PRIME - PREFERENCE RATIOS IN MULTIATTRIBUTE EVALUATION

Ahti A. Salo and Raimo P. Hämäläinen
Technical Research Centre of Finland
02044 VTT, Finland
Systems Analysis Laboratory, Helsinki University of Technology
Otakaari 1 M, 02150 Espoo, Finland

December, 1997

Abstract

This paper develops the PRIME method which elicits the decision maker's preferences through possibly imprecise statements about value differences. In PRIME, the alternatives need not have numerical measurement scales, and because the elicitation procedure can be carried in several different ways, PRIME offers a number of approaches for supporting multiattribute problems under certainty. In terms of its use of ratio comparisons, PRIME resembles the AHP, but because the ratio comparisons relate to differences on measurable value functions, it avoids some of the problems for which the AHP has been criticized. An example is developed to illustrate the method.

1. Introduction

The tradeoff method of Keeney and Raiffa (1976) elicits a multiattribute value (MAV) function through indifference judgments where the decision maker (DM) adjusts missing attribute levels of partially specified consequences until these become equally preferred to other, completely specified consequences. This theoretically sound method is well-founded in conjoint measurement (Krantz et al. 1971), but does require that the attributes have continuous scales on which the performance levels of the alternatives can be indicated. In the absence of such scales, it becomes necessary to introduce proxy variables which are highly correlated with the attributes and thus suitable for measuring the performance of the alternatives.

Yet proxy variables are not always readily available, and even if they can be found, the DM may find difficulty in giving precise replies to the tradeoff questions. In consequence, more straightforward methods such as SMART (Edwards 1977) and the analytic hierarchy process (AHP) (Saaty 1980) have won popularity in applications of multiattribute decision analysis. A common feature in both SMART and the AHP is that they rely on ratio comparisons about the "relative importance" of attributes, although the resulting weights are not explicitly linked to the values of unit changes in the component value functions. Indeed, this is one of the reasons for the recommendation that the tradeoff method should be employed instead of ratio-based comparisons (see, e.g., Hobbs 1980; Weber, Eisenführ and von Winterfeldt 1988; von Nitzsch and Weber

1993; Weber and Borcherding 1993; Pöyhönen and Hämäläinen 1996, 1997; Pöyhönen, Vrojlik and Hämäläinen 1997).

The PRIME method of the present paper seeks to strike a balance between the theoretical validity of the tradeoff method and the functionality of decomposed ratio judgments. Towards this end, the elicitation of preferences is founded on the comparison of preference differences in pairs of consequences. Such comparisons may be specified either as exact point estimates or, more typically, as interval judgments which impose linear constraints on the single-attribute scores of the alternatives. In this way, the preference model becomes increasingly specific whereby more definitive dominance results can gradually be established. The possibility to work with imprecise preference statements may be particularly suitable for group decision support since the DMs' conflicting views can be combined into an aggregate preference model (Hämäläinen, Salo and Pöysti 1992; Hämäläinen and Pöyhönen 1996).

The remainder of this paper is structured as follows. Section 2 gives a summary of related techniques for the elicitation of imprecisely specified value representations. Section 3 considers ratio judgments in the assessment of preference information and describes the dominance concepts employed in PRIME. Section 4 focuses on criteria weights in hierarchical value tree representations and discusses PRIME in relation to the AHP. Section 5 illustrates the properties of PRIME in the context of an example.

2. Additive Preference Representations

Multiattribute choice problems under certainty can be characterized by the consequence space $X = X_1 \times \ldots \times X_n$ where X_i denotes the set of possible consequences on the i-th attribute. The DM's preferences for the consequences are modeled through a binary relation such that $\mathbf{x} \succsim \mathbf{y}$ only if $\mathbf{x}$ is at least as preferred as $\mathbf{y}$. Under well-known conditions (see, e.g., Debreu 1960; Fishburn 1970; Krantz et al. 1971; Wakker 1989), the preference relation $\succsim$ has and an additive numerical representation such that

$$\mathbf{x} \succsim \mathbf{y} \iff \sum_i v_i(x_i) \geq \sum_i v_i(y_i). \tag{1}$$

This representation is unique up to affine positive transformations, and it induces a well-defined preference relation $\succsim_i$ on X_i through $x_i \succsim_i y_i \iff v_i(x_i) \geq v_i(y_i)$.

The *additive measurable value functions* of Dyer and Sarin (1979) capture the DM's preferences for positive differences between pairs of consequences. In particular, if the DM's preference relation $\succsim^*$ (defined on pairs of positive differences between consequences $\mathbf{x}, \mathbf{x}'$ such that $\mathbf{x} \succsim \mathbf{x}'$) satisfies conditions such as difference consistency and difference independence, the representation in (1) has the property

$$(\mathbf{x}, \mathbf{x}') \succsim^* (\mathbf{y}, \mathbf{y}') \iff \sum_i [v_i(x_i) - v_i(x_i')] \geq \sum_i [v_i(y_i) - v_i(y_i')]. \tag{2}$$

The above equivalence justifies the use of strength of preference statements in the estimation of the value representation. For instance, single-attribute value functions can be determined from indifference judgments in which the DM specifies missing attribute levels until exchanges between the achievement levels become equally preferred. As a result, the resulting value functions are defined on an interval scale even though references to other attributes are not made (see, e.g., von Winterfeldt and Edwards 1986).

In (1), the worst achievement levels on the attributes are customarily fixed to zero through $v_i(x_i^\circ) = 0$. Assuming that x_i^* denotes the best achievement level on the i-th attribute, the additive representation can thus be written as

$$v(\mathbf{x}) = \sum_i w_i v_i^N(x_i), \qquad (3)$$

where the normalized value functions $v_i^N(\cdot) = [v_i(\cdot) - v_i(x_i^\circ)]/[v_i(x_i^*) - v_i(x_i^\circ)]$ map the achievement levels onto the range $[0,1]$ and the scaling constants $w_i = v_i(x_i^*) - v_i(x_i^\circ)$, usually called attribute weights, relate unit increases in normalized value functions to increases in the overall value. By convention, the attribute weights add up to one because the ideal consequence $(x_1^*, \ldots, x_n^*)$ has the value one.

A variety of methods have been proposed for the elicitation of the additive representation (1) or its normalized version (3). The trade-off method (Keeney and Raiffa 1976) requests the DM to specify missing attribute levels until a partially specified consequence becomes equally preferred to another consequence. More direct methods, such as direct rating and swing weighting, ask for numerical inputs that do not involve choices between pairs of consequences. In this regard, the direct methods have a weaker theoretical foundation; nevertheless, they have found extensive use in applications, perhaps because they are generally easier to apply. Summaries of the major assessment techniques are given e.g. by Fishburn (1967) and von Winterfeldt and Edwards (1986).

The numerical estimates in preference assessment should be quantitatively meaningful in the sense that the estimates should not change even if the representation (1) is replaced by another, equally valid representation of the underlying preference relation (French 1986). This fundamental requirement implies, for example, that the DM should not be asked for the specific values that a particular value representation attaches to individual consequences. More generally, it implies that ratio comparisons should be phrased in terms of comparisons of positive preference differences because the ratio

$$\frac{v(\mathbf{x}) - v(\mathbf{x}')}{v(\mathbf{y}) - v(\mathbf{y}')} = \frac{\sum_i [v_i(x_i) - v_i(x_i')]}{\sum_i [v_i(y_i) - v_i(y_i')]} \tag{4}$$

does not change if the value function $v(\cdot)$ is replaced by $v'(\cdot) = \alpha v(\cdot) + \beta, \ \alpha > 0$.

Judgments about ratios of preference differences can be criticized on the grounds that they, like strength of preference statements, rely on the DM's introspection and are not actionable as choices between naturally occurring options. Nevertheless, Sarin (1983) convincingly argues for strength of preference judgments which he considers helpful as a means of facilitating the elicitation process. The axiomatization of Vansnick (1984), on the other hand, provides a theoretical foundation for the type of measurement needed in the estimation of the ratio (4). Farquhar and Keller (1989) discuss these and other relevant results which give further justification for the use of strength of preference statements.

Apart from theoretical considerations, a practical motivation for ratio estimates about preference differences is the popularity of related judgments in widely applied techniques of direct measurement. In direct rating, for example, the DM evaluates alternatives by indicating their position relative to the worst and best alternatives at the endpoints of the [0,100] scale. Such assignments are essentially ratio comparisons, for an estimate about the ratio $[v_i(x_i) - v_i(x_i^\circ)]/[v_i(x_i^*) - v_i(x_i)]$ determines the position of $v_i(x_i)$ on the [0,100] scale. The judgments needed in SWING weighting reduce to ratio comparisons much in the same way, whereas direct value assignments, as employed for instance in the HOPIE method (Weber 1985), can be regarded as ratio statements about value improvements from the worst achievement level to the achievement levels of the actual alternatives.

Traditional decision analyses are typically based on an exhaustive specification of the DM's preferences. Yet, in many cases the results of the analyses are insensitive to elicited model parameters, which in turn makes it hard to justify the effort spent in the assessment of a precise preference description.

Several methods have been developed in order to alleviate problems in the elicitation of complete preference models. For instance, Hazen (1985) discusses concepts of dominance and potential optimality, and presents optimization techniques for identifying dominated alternatives under partial preference information. The rank ordering algorithm of Kirkwood and Sarin (1985) is applicable to problems with imprecise attribute weights and known single-attribute scores.

Weber's (1985) HOPIE technique accepts both holistic and decomposed preference statements and synthesizes these into dominance results through linear programming. The MCRID method of Moskowitz, Preckel and Yang (1992) associates value intervals with the alternatives by processing holistic comparisons and possibly imprecise ratio judgments about the relative importance of the attributes. The PAIRS technique (Salo and Hämäläinen 1992) processes imprecise statements of relative importance into two dominance relations for the alternatives; it also provides explicit support for maintaining the consistency of the preference model. Barron (1992) suggests that the centroid of the feasible weight be used in the computation of overall values when the DM merely ranks the attributes in the order of relative importance.

A major difference between PRIME and approaches such as AHP, SMART, MCRID, and PAIRS is that the ratio comparisons are explicitly linked to the ranges of the alternatives so that problems arising from the vague notion of "relative importance" can be avoided (see, e.g., von Nitzsch and Weber 1993). Furthermore, PRIME is more flexible than its predecessors in that it offers explicit guidance for consistent refinements to the preference model and supplies intermediate dominance results throughout the analysis. In contrast to PAIRS, PRIME is also capable of handling holistic judgments in which consequences are compared against attributes on the higher levels of the value tree; it is in this sense that PRIME can be regarded as a generalisation of PAIRS.

3. Preference Elicitation and Synthesis

Score Assessment

Under each attribute, the relative magnitudes of the scores can be derived from ratio comparisons in which the consequences differ from each other on one attribute only. The general format of such comparisons can be written as

$$\frac{v_i(x_i) - v_i(x_i')}{v_i(y_i) - v_i(y_i')}, \tag{5}$$

where x_i, x_i' and y_i, y_i' are pairs of alternatives' distinct achievement levels on the i-th attribute. The elicitation of these ratios can be begun by asking the DM first to rank order the alternatives with respect to this attribute. As the next step, the DM can be requested for a ranking of the positive value differences between adjacent achievement levels. After these initial comparisons, ratio statements about these or any other value preference differences may be entered. Possible anchoring effects can be countered by not employing the same value difference across most of the comparisons, even though this might improve the consistency of the elicited jugdments.

With m alternatives, there can be as many as $\frac{1}{2}m(m-1)$ positive value differences under each attribute. Because any two distinct differences can be compared, the maximum number of ratio comparisons under each attribute is $\frac{1}{8}m(m-1)[m(m-1)-2]$, while only $m-2$ exact ratio statements are needed to determine the scores in relation to the worst and best achievement levels. Thus there is room for redundant comparisons that can be used to check the consistency of the DM's statements.

The scores can be elicited either 1) through exact but possibly conflicting ratio estimates or 2) by using the DM's judgments as increasingly tight constraints on the alternatives' scores. The second approach leads to an interactive analysis in which precise scores need not be known before weight assessment is begun.

In the case of exact ratio estimates, the DM can be asked to make precise and exhaustive statements about preference differences from x_i° to the achievement levels of the alternatives. The resulting ratio estimates $[v_i(x_i) - v_i(x_i^\circ)]/[v_i(x_i') - v_i(x_i^\circ)]$ define a reciprocal comparison matrix from which the normalized scores $v_i^N(\cdot) = [v_i(\cdot) - v_i(x_i^\circ)]/[v_i(x_i^*) - v_i(x_i^\circ)]$ can be estimated with the logarithmic least squares method or the eigenvector solution of the AHP. Details about these and other techniques of averaging the entries of a reciprocal matrix are given for instance by Fichtner (1986) (see also Yu 1985).

However, while the matrix representation is computationally convenient, it is not flexible enough to accommodate ratio judgments between arbitrarily chosen value differences. These can be handled by observing that the estimate $r(x_i x_i'; y_i y_i')$ for (5) corresponds to the constraint

$$r(x_i x_i'; y_i y_i') = \frac{v_i(x_i) - v_i(x_i')}{v_i(y_i) - v_i(y_i')}. \qquad (6)$$

Thus, ratio judgments lead to an overdetermined set of equations from which normalized scores can be obtained by solving the minimization problem

$$\min \quad \sum \frac{[r(x_i x_i'; y_i y_i')\Delta(y_i, y_i') - \Delta(x_i, x_i')]^2}{r(x_i x_i'; y_i y_i')^2 + 1} \qquad (7)$$

$$\sum_i v_i^N(x_i^*) = 1,$$

where the sum is taken over all the estimates the DM has specified and $\Delta(x_i, x_i')$ denotes the difference $v_i^N(x_i) - v_i^N(x_i')$. This approach is analogous to the geometric least squares method of Islei and Lockett (1988) in that the error term of the ratio judgment in (6) (and that of its reciprocal $r(y_i y_i'; x_i x_i')$) is the Euclidean distance in the plane defined by the nominator and denumerator.

Fur purposes of interactive analysis, the relationship (6) can be used to model both exact and imprecise ratio statements as constraints on the underlying scores. For example, if the ratio $[v_i(x_i) - v_i(x_i')]/[v_i(y_i) - v_i(y_i')]$ is judged to belong to the range

$[L, U]$, the scores that are consistent with this statement must satisfy the inequalities

$$\begin{cases} [v_i(x_i) - v_i(x_i')] \geq L[v_i(y_i) - v_i(y_i')], \\ [v_i(x_i) - v_i(x_i')] \leq U[v_i(y_i) - v_i(y_i')], \end{cases} \tag{8}$$

The elicitation of such ratio comparisons can be supported through *consistency bounds* found by minimizing and maximizing the ratio in (5) through algorithms of linear fractional programming. Such bounds illustrate the implications of the earlier statements and help the DM preserve the consistency of the preference model.

Weight Assessment

Tradeoff information about the DM's preferences can be elicited through ratio judgments in which the achievement levels of the consequences differ on at least two attributes. Although such judgments are linear in terms of the scores in (4), they do not imply linear constraints on the attribute weights of the normalized representation

$$\frac{v(\mathbf{x}) - v(\mathbf{x}')}{v(\mathbf{y}) - v(\mathbf{y}')} = \frac{\sum_i w_i[v_i^N(x_i) - v_i^N(x_i')]}{\sum_i w_i[v_i^N(y_i) - v_i^N(y_i')]} \tag{9}$$

unless the normalised scores in (9) have already been precisely determined. As a result, the usual normalized representation leads in the modeling of imprecision to nonlinearities which are avoided in the sequel by working directly with the model (4).

The elicitation of attribute weights can be simplified by restricting the comparisons to consequence pairs such that the achievement levels are identical except on two attributes, in which case the ratio (4) can be written as

$$\frac{v_i(x_i) - v_i(x_i')}{v_j(y_j) - v_j(y_j')}, \quad x_i \succ_i x_i', y_j \succ_j y_j'. \tag{10}$$

As an application of this representation, the DM can discriminate between the two alternatives $\mathbf{x}$ and $\mathbf{x}'$ by ranking them under each attribute and by supplying estimates r_i for the ratios $[v_i(x_i) - v_i(x_i')]/[v_1(x_1) - v_1(x_1')]$ (here the achievement levels x_1, x_1' are assumed not to be equally preferred). Based on these estimates, the preferred alternative is given by the sign of the difference

$$\begin{aligned} v(\mathbf{x}) - v(\mathbf{x}') &= \sum_{i=1}^{n}[v_i(x_i) - v_i(x_i')] \\ &= [v_1(x_1) - v_1(x_1')][\sum_{i=1}^{n} \frac{v_i(x_i) - v_i(x_i')}{v_1(x_1) - v_1(x_1')}] \\ &= [v_1(x_1) - v_1(x_1')] \sum_{i=1}^{n} r_i. \end{aligned}$$

There is no need to perform all comparisons relative to the first attribute; in fact, it may be advisable to elicit ratio judgments between other attributes as well in order to counter possible anchoring effects.

Even if ratio judgments are limited to the type in (10), the DM has still many possibilities of restricting the range of the attribute weights. With m alternatives, there can be as many as $\frac{1}{2}m(m-1)$ positive value differences under each attribute, and because any two of these can be compared, there are up to $\frac{1}{4}m^2(m-1)^2$ different ways of specifying a ratio judgment for a pair of attributes. On the other hand, if the scores are exactly known, only one such judgment is needed to determine the relative magnitudes of the two attribute weights.

As in score elicitation, the attribute weights can be derived either 1) by synthesizing the inconsistencies into an overspecified system of exact judgments or 2) by converting all judgments into linear constraints on the underlying scores.

With precisely determined score information, ratio judgments about (10) lead to a system of linear equations through

$$w_i/w_j = r(x_i x_i'; y_j y_j')\Delta(y_j, y_j')/\Delta(x_i, x_i')$$

where $r(x_i, x_i'; y_j, y_j')$ is the DM's estimate for the ratio (6) with the value difference in the denominator now being taken on the j-attribute. By analogy to (7), the attribute weights can be computed from the least-squares minimization problem where the sum is again taken over all comparisons that the DM has entered.

The second approach, i.e., the conversion of ratio judgments into linear constraints provides support for interactive preference analysis. For example, if the DM states that the ratio (10) belongs to the range $[L, U]$, the scores must satisfy the inequalities

$$\begin{cases} [v_i(x_i) - v_i(x_i')] \geq L[v_j(y_j) - v_j(y_j')], \\ [v_i(x_i) - v_i(x_i')] \leq U[v_j(y_j) - v_j(y_j')]. \end{cases} \tag{11}$$

Before the assessment of each new ratio comparison, the ratio (10) can be minimized and maximized subject to the constraints of the earlier judgments. As in the case of score elicitation, the resulting consistency bounds assist the DM in maintaining the consistency of the preference model.

Like many other techniques (see, e.g., Kirkwood and Sarin 1985; Hazen 1986; Weber 1985, 1987; Moskowitz, Preckel and Yang 1992; Salo and Hämäläinen 1992), PRIME converts the DM's imprecise preference model into dominance relations by examining the values that feasible scores associate with the alternatives. The more restrictive of its two dominance concepts, called *absolute dominance*, is related to the alternatives' value intervals so that alternative $\mathbf{x}$ dominates $\mathbf{x}'$ only if the smallest feasible value of $\mathbf{x}$ exceeds the largest feasible value of $\mathbf{x}'$, i.e.,

$$\mathbf{x} \succ_A \mathbf{x} \iff \min \sum_i v_i(x_i) > \max \sum_i v_i(x_i'). \qquad (12)$$

Results about absolute dominance can be conveniently displayed through the value intervals $V(\mathbf{x}) = [\underline{v}(\mathbf{x}), \overline{v}(\mathbf{x})] = [\min \sum_i v_i(x_i), \max \sum_i v_i(x_i)]$, the bounds of which are computed from the linear programs in (12).

The set of dominated alternatives is determined by the second criterion, called *pairwise dominance*, according to which alternative $\mathbf{x}$ is preferred to $\mathbf{x}'$ only if for all feasible scores the value of $\mathbf{x}$ exceeds that of $\mathbf{x}'$, i.e.,

$$\mathbf{x} \succ_P \mathbf{x}' \iff \min[v(\mathbf{x}) - v(\mathbf{x}')] = \min \sum_i [v_i(x_i) - v_i(x_i')] > 0. \qquad (13)$$

Pairwise dominance need be explicitly computed only if the value intervals satisfy the inequalities $\overline{v}(\mathbf{x}) > \overline{v}(\mathbf{x}') \geq \underline{v}(\mathbf{x}) > \underline{v}(\mathbf{x}')$. In other cases, the dominance either follows from or is excluded by the value intervals.

The Impact of Imprecise Judgments on Dominance Results

With imprecise judgments, it is hard to make definite predictions as to how many judgments are needed before the most preferred alternative can be identified. This is because the pace at which the set of nondominated alternatives becomes smaller depends not only on the structure of the problem, but also on the precision, type, and the relative order of the DM's judgments.

This notwithstanding, we performed a Monte Carlo simulation study with randomly generated problem instances in order to investigate the completeness of the dominance relations under various types of preference statements. In this study the number of attributes n varied from one to seven, and the number of alternatives m was in the range from two to five. The attribute weights were generated from a uniform distribution over the set $W = \{(w_1, \ldots, w_n) \mid \sum_{i=1}^{n} w_i = 1, w_i \geq 0\}$. The worst and best alternatives under each attribute were picked at random so that every alternative had an equal probability of assuming either one of these positions. For the remaining alternatives the scores were taken from the uniform distribution over the unit interval. This procedure defined a completely specified value structure for each of the fifty problem instances that were generated for every combination of n and m. As a preliminary step to the subsequent simulations, the alternatives were rank ordered under each of the attributes.

In the first round of simulations, $n(m-1)-1$ imprecise judgments were entered by multiplying the actual ratios of randomly chosen value differences by the lower and upper bounds of $\frac{4}{5}$ and $\frac{5}{4}$, respectively. For score elicitation $m-1$ value differences were selected under each of the attributes, and an additional value difference was selected in order to make the $n-1$ comparisons needed for weight assessment. Throughout constraint generation, all value differences had an equal chance of becoming selected subject to the restriction that every single-attribute score was required to be included in at least one of the comparisons for score elicitation. After the introduction of constraints, typically more than half of the alternatives became dominated in problems involving only a few attributes (see Table 1).

Table 1. Number of nondominated alternatives after ratio judgments.

The second round of simulations was run by rank ordering all the value differences between the alternatives. Because the overall number of differences (i.e., $1/2nm(m-1)$) grows quickly with the number of alternatives and attributes, the larger problem instances involved more constraints in proportion to n and m; this explains why the number of nondominated alternatives decreases as the number of attributes increases (see Table 2). On the whole, the results indicate that ordinal comparisons of value differences are helpful in reducing the set of nondominated alternatives.

Table 2. Number of nondominated alternatives after ordinal judgments.

4. Hierarchical Attribute Representations

Meaning of Attribute Weights

It is often helpful to structure the relevant objectives as a hierarchical value tree in which higher level attributes are decomposed into an exhaustive set of mutually preferentially independent attributes (Keeney and Raiffa 1976). In such a tree, the higher level attributes can be thought of as aggregates of the twig-level attributes so that each attribute a_k can be identified with a subset of $\{1, \ldots, n\}$ such that $i \in a_k$ only if the ith twig-level attribute is in the subtree of which a_k is the root. Thus, for instance, the entire set $\{1, \ldots, n\}$ corresponds to the topmost attribute while the singleton sets $\{i\}$ stand for twig-level attributes.

PRIME permits holistic comparisons with respect to the topmost or other higher level attributes. From the computational viewpoint, such comparisons constrain scores in the same way as decomposed ratio comparisons; for example, the inequality $\sum_{i \in a_k} [v_i(x_i) - v_i(x_i')] > 0$ must hold if the DM prefers $\mathbf{x}$ to $\mathbf{x}'$ when considering the attribute a_k. In the same way, ordinal and ratio judgments at the higher levels lead to linear constraints on the underlying scores.

Given a higher-level attribute a_k, the additive value representation can be written in the form

$$v(\mathbf{x}) = w(a_k) \sum_{i \in a_k} \frac{[v_i(x_i) - v_i(x_i^\circ)]}{\sum_{i \in a_k} [v_i(x_i^*) - v_i(x_i^\circ)]} + \sum_{i \notin a_k} [v_i(x_i) - v_i(x_i^\circ)] \tag{14}$$

where $w(a_k) = \sum_{i \in a_k} [v_i(x_i^*) - v_i(x_i^\circ)]$ is the weight of attribute a_k. This emphasizes that the weight of a_k should be proportional to the value increase brought by exchanging the worst consequence for the hypothetical consequence $\mathbf{x}'$ such that $x_i' = x_i^\circ,\ i \notin a_k$ and $x_i' = x_i^*,\ i \in a_k$. Furthermore, this equality demonstrates that even at the higher levels of the value tree, the attribute weights are indeed linked to the value differences associated with the best and worst achievement levels (see also Keeney and Raiffa 1976).

The attributes' weight intervals

$$W(a_k) = [\underline{w}(a_k), \overline{w}(a_k)] = [\min \sum_{i \in a_k} v_i(x_i^*), \max \sum_{i \in a_k} v_j(x_i^*)]$$

indicate to what extent the value differences between extreme achievement levels may vary in relation to the unit improvement from the worst to the ideal consequence. These intervals tend to become wider when moving upwards in the tree (with the exception of the topmost attribute whose weight is one by definition); the weight interval of any attribute is, however, tighter than that obtained by summing the corresponding bounds at the respective subattributes.

Stages in Interactive Analysis

In the interactive analysis supported by PRIME, the DM's preference statements are translated into linear constraints on the single-attribute scores. Since the addition these such tends to reduce the number of non-dominated alternatives, the DM gets gradually more specific results as he goes on making holistic judgments or ratio comparisons about value differences. The main steps in the interactive analysis are summarized in the following list.

Step 1. Construct the value tree and verify the independence conditions.

Step 2. Rank the alternatives with respect to the lowest level attributes.

Step 3. Choose the next comparison, observe the corresponding consistency bounds, elicit the judgment either as a precise point estimate or an interval judgment, and enter the corresponding constraints into the preference model.

Step 4. Check absolute and pairwise relations for dominance.

Step 5. Iterate through steps 3 and 4 to reduce the set of nondominated alternatives.

Steps 3 and 4 leave plenty of room for the choice of the particular comparison about which the next estimate is requested. Because of this, PRIME can be thought of as a family of alternative methods which differ from each other in the details concerning the order of the elicitation process. The above list is, therefore, only an outline of a general template which applies to all variants within the family.

Although PRIME imposes no particular restrictions on the order of the comparisons, it is advisable to begin with score elicitation since this helps to clarify the alternatives and attributes to the DM. After Step 2, detailed score information can be obtained for instance through the direct ordered metric technique (see, e.g., Fishburn 1967) so that the DM ranks the value differences between adjacent alternatives in the attribute-specific rank orders. The DM may then go on to ratio statements about these value differences, or to comparisons which relate them to the ranges between the extreme achievement levels. If necessary, further constraints may be assessed through statements about any actual or artificially constructed alternatives.

The elicitation of attribute weights may, to some extent, rely on ordinal comparisons of alternatives with regard to higher level attributes. However, because the purpose of the analysis is to support the overall evaluation of alternatives, these comparisons should be restricted to those that do not pose any difficulties to the DM, either because he definitely prefers one alternative to the other or because the holistic comparison involves attributes that are on the lower levels of the value tree.

Because the dominance results are determined as solutions to linear programs, standard sensitivity results serve to identify the judgments the tightening of which would contribute to a particular dominance relation between any two alternatives. Further support can be obtained by adding the dominance relation as an temporary constraint to the preference model (i.e., as the inequality (13)), whereafter any given ratio of value differences can be minimized and maximized subject to the modified constraint set. The results of such an analysis indicate which values (if any) of the ratio would give rise to the dominance relation under consideration.

Some techniques for the management of imprecision advocate procedures that eliminate alternatives which have not yet become dominated according to the pairwise dominance criterion. For example, Barron (1992) picks the centroid vector to represent the DM's ordinal perceptions about the relative importance of attributes, and Moskowitz, Preckel and Yang (1992) assign implicit probabilities to the parameters that the DM could arrive at if the elicitation of preferences were to be continued. Both approaches run the risk of drawing premature conclusions by extrapolating from the DM's earlier statements.

In PRIME, the recommendation is to elicit additional preference statements until the DM is either unwilling or unable to enter any further constraints. Only at such a stage, if more than one nondominated alternative still remains, should one consider probabilistic inferences in which a triangular or truncated normal distribution is assumed on the feasible value differences between the alternatives. Such distributions can, in principle, be analyzed for dominance statements using the procedures of the HOPIE method (Weber 1985).

Relationship to the Analytic Hierarchy Process

Saaty's analytic hierarchy process (AHP) (Saaty 1980; Saaty 1994) is a well-known method of hierarchical weighting which elicits preferences through ratio comparisons alone. At each attribute, the DM's replies to the pairwise comparisons of relative importance (or performance) are inserted into a reciprocal comparison matrix. From these matrices, overall weights for the alternatives are found by dividing the unit weight of the topmost attribute in proportion to the components of the respective right principal eigenvectors.

¿From the perspective of the multiattribute value analysis, the standard AHP appears deficient on several accounts. In particular, the elicitation of scores through questions such as "Which of the alternatives Mercedes and Honda is better with respect to Performance ? By how much ?" is based on the assumption that the values associated with to various levels of performance are measurable on a ratio scale (Watson and Freeling 1982). Yet, as we argued in the presentation of the value function framework in Section 2, only value differences - and not values as such - can be legitimately compared on a ratio scale. This observation in turn suggests that the AHP can be given a theoretical foundation by interpreting them as implicit references to underlying reference points (for an extensive coverage, see the discussion paper Salo and Hämäläinen (1997) and its rejoinders in the same issue).

The AHP allows the attribute weights to be elicited independently of the set of alternatives, although (14) shows that the weights should be proportional to the value differences between the worst and best achievement levels. This, together with the enforced normalization of local priority vectors, occasionally causes the addition of a new alternative to reverse the relative rank order of the other alternatives (Belton and Gear 1983). In contrast, the elicitation of preferences in PRIME is so constructed that the attribute weights are automatically tied to the ranges between the extreme achievement levels, which in turn eliminates the possibility of rank reversals.

In summary, PRIME can used as a revised AHP where the controversial employment of ratio comparisons in the AHP is eliminated. Another advantage of PRIME is the interactiveness which arises from the possibility of working with imprecise preference statements.

5. An Example

This section illustrates the use of PRIME in supporting a company to to choose one of three alternative sites for a manufacturing facility. In the first step, the value tree in Figure 1 is constructed by grouping the seven twig-level attributes under three higher level attributes. Among these, *Staff* has been divided into the subattributes *Competence* and *Salary expenditure*, and *Industrial policy* is defined by *Support measures* and *Taxation* rules. The subattributes under *Logistics* are concerned with access by *Sea, Rail* and *Air*.

Figure 1. Value tree for siting a facility.

The alternative sites are denoted by the capital letters A, B, and C, and the lowest level attributes are indexed from one to seven in the order in which they appear from top to bottom in Figure 1. Thus the first attribute refers to salary, the second to cost of living etc. The higher level attributes are identified with subsets of the twig-level attributes as discussed in Section 4.

In step 2, it is assumed that company ranks the alternatives on the twig-level attributes as shown in Table 3. On the first attribute, for instance, the ranking implies the inequalities $v_1(A_1) > v_1(C_1) > v_1(B_1) = 0$.

Table 3. Rankings on the twig-level attributes.

In the third step, assume that the company makes the following judgments about the alternatives within the three main attributes groups:

1. The competence improvements from B to C and from C to A are equal.
2. The competence improvement from B to A is at least seven but no more than nine times larger than the difference on salary expenditure.
3. The difference between A and C on taxation is twice as large as that on support measures.
4. With respect to the two-attribute group of industrial policy, the alternatives A and B are equally preferred.
5. The difference between B and A on air logistics is at most three times larger than their difference on rail logistics, but the differences on sea and rail are equal.
6. With regard to air logistics, the improvements from A to B and from B to C are equal.

The first and last judgments are strength of preference statements, the third is a precise ratio statement, and the fourth is a holistic comparison on a higher level attribute. The second and fifth statements, on the other hand, are imprecise ratio statements. Yet they all lead to linear constraints to score increments: the first one, for example, is equivalent to $v_1(C_1) - v_1(B_1) = v_1(B_1) - v_1(A_1)$.

Before each new judgment the consistency bounds can be determined to assist the company in specifying judgments that are compatible with the earlier ones. For example, for the comparison of alternatives A and B on industrial policy, the bounds $\hat{L}^{43}_{BA;AB} = 0 \leq [v_4(B_4) - v_4(A_4)]/[v_3(A_3) - v_3(B_3)] \leq 2 = \hat{U}^{43}_{BA;AB}$ show that the value difference of these alternatives on taxation can be at most twice as large as that on support measures. In Figure 2, these consistency bounds are displayed as the endpoints of a darker shaded area on a horizontal bar which can be divided into two parts whose lengths reflect the perceived ratio of the value differences $v_4(B_4) - v_4(A_4)$ and for $v_3(A_3) - v_3(B_3)$. Any division which falls into the darker shaded area corresponds to a precise and consistent ratio judgment, and imprecise ratio judgments correspond to intervals on this horizontal bar. As a result, further consistent judgments can be entered by restricting the range of possible ratios within the consistency bounds.

Figure 2. Consistency bounds on a horizontal bar.

Once the above judgments have been transformed into linear constraints, the linear programs in (12) in step 4 give the value intervals $V(A) = [0, 0.9], V(B) = [0.1, 0.75], V(C) = [0.55, 0.88]$ for the alternatives. At this point absolute dominance does not hold for any pair of alternatives since all three intervals overlap. These intervals exclude the possibility of pairwise dominance for alternatives other than C and B, but because the minimum for $v(C) - v(B)$ is zero, there exists a set of feasible scores for which C and B have the same value and thus therefore pairwise dominance does not hold.

Going back to step 3, assume that DM makes the following judgments:

1. The difference between B and A on taxation is three to five times larger than that on salary expenditure.
2. The overall difference between C and A on logistics as a whole is larger than that on taxation.

After these judgments the revised value intervals become

$$V(A) = [0, 0.62], \quad V(B) = [0.31, 0, 75], \quad V(C) = [0.65, 0.88].$$

Because the lower bound of C lies above the upper bound of A, the alternative C dominates A according to both dominance concepts (see Figure 4). Checking the possible dominances between B and the other alternatives from (13) shows that the differences $v(C) - v(B)$ and $v(B) - v(A)$ have non-positive minima so that pairwise dominance does not hold yet.

Since the alternatives A and B have the widest value intervals, these alternatives can be compared with respect to the first level attributes staff considerations and logistics whose weight intervals are given in Table 2. The consistency bounds for this comparison indicate that in view of the earlier statements, the value of the competence difference between A and B cannot be more than 4.2 times larger than on logistics. If the value function for competence has been elicited, this consistency bound can be converted into the least decrease in the level of competence for alternative A that could offset the improvement in exchanging the logistic characteristics of A for the better characteristics associated with alternative B.

Finally, assume that the company restricts the ratio of value differences between A and B on the competence and logistics attributes from the previously implied $[0, 4.2]$ to $[1.33, 3]$. Adding the new judgment to the set of constraints leads to the value intervals

$$V(A) = [0.44, 0.6], \quad V(B) = [0.34, 0.45], \quad V(C) = [0.65, 0.73].$$

Although the two first value intervals overlap, alternative A does dominate B because the minimum for their value difference in (13) is positive, i.e. $0.03 > 0$. Still, the most preferred of alternative is C which is preferred to the two other ones according to both dominance concepts. Figure 3 illustrates the evolution of value intervals and pairwise dominance in the analysis.

Figure 3. Evolution of results.

---

6. Conclusion

The PRIME technique seeks to reduce the elicitation effort in multiattribute evaluation under certainty by accepting and analyzing imprecise preference statements. Such statements can be either holistic comparisons between actual or hypothetical alternatives, ordinal strength of preference judgments, or ratio comparisons about preference differences. Regardless of their type, the DM's statements are modeled as linear constraints on the alternatives, and as new judgments and refinements to the earlier ones are introduced, these constraints become more restrictive and gradually produce more informative dominance results. Throughout the interactive refinement process, the logical implications of the earlier statements are shown in order to help the DM enter consistent judgments which contribute to the identification of the preferred alternative(s).

Because the inputs of PRIME do not require that the attributes have explicit numerical measurement scales attached to them, the proposed technique appears particularly useful in problems where the dimensions of evaluation are highly qualitative or for other reasons lack suitable proxy variables. Problems of group decision support provide another promising area of application, because the DMs seem to be more capable of moving towards a consensus solution once their conflicting judgments have been combined into an imprecise aggregate model. This, and other behavioral aspects of applying PRIME and related techniques to decision problems calls for empirical and comparative research.

Acknowledgment

This work has been supported by the Research Council of Technology of the Academy of Finland.

References

BARRON, F. H. 1992. Selecting a Best Multiattribute Alternative with Partial Information about Attribute Weights. Acta Psych. 80, 1-103.

BELTON, V. AND T. GEAR. 1983. On a Short-coming of Saaty's Method of Analytic Hierarchies. Omega 11, 228-230.

DEBREU, G. 1960. Topological Methods in Cardinal Utility Theory. In K. J. Arrow, S. Karlin, P. Suppes (Eds.), *Mathematical Methods in the Social Sciences, 1959.* Stanford University Press, Stanford, 16-26.

DYER, J. S. AND R. K. SARIN. 1979. Measurable Multiattribute Value Functions. Oper. Res. 27, 810-822.

EDWARDS, W. 1977. How to Use Multiattribute Utility Measurement for Social Decisionmaking. IEEE Trans. Systems, Man, Cybernet. 7, 326-340.

FARQUHAR, P. H. AND L. R. KELLER. 1989. Preference Intensity Measurement. Ann. Oper. Res. 19, 205-217.

FICHTNER, J. 1986. On Deriving Priority Vectors from Matrices of Pairwise Comparisons. Socio-Economic Planning Sci. 20, 341-345.

FISHBURN, P. C. 1967. Methods of Estimating Additive Utilities. Management Sci. 13, 435-453.

FISHBURN, P. C. 1970. *Utility Theory for Decision Making*, John Wiley & Sons, New York.

FRENCH, S. 1986. *Decision Theory: An Introduction to the Mathematics of Rationality.* Ellis Horwood Ltd., Chichester.

HÄMÄLÄINEN, R. P. AND M. PÖYHÖNEN. 1996. On-Line Group Decision Support by Preference Programming in Traffic Planning. Group Decision and Negotiation 5, 485-500.

HÄMÄLÄINEN, R. P., A. A. SALO AND K. PÖYSTI. 1992. Observations about Consensus Seeking in a Multiple Criteria Environment. Proceedings of the 25th Hawaii International Conference on System Sciences, Vol. IV, January 1992, 190-198.

HAZEN, G. B. 1986. Partial Information, Dominance, and Potential Optimality in Multiattribute Utility Theory. Oper. Res. 34, 296-310.

HOBBS, B. F. 1980. A Comparison of Weighting Methods in Power Plant Siting. Decision Sci. 11, 725-737.

ISLEI, G. AND A. G. LOCKETT. 1988. Judgemental Modelling Based on Geometric Least Square. European J. Oper. Res. 36, 27-35.

KEENEY, R. L. AND H. RAIFFA. 1976. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* John

Wiley & Sons, New York.

KIRKWOOD, C. W. AND R. K. SARIN. 1985. Ranking with Partial Information: A Method and an Application. Oper. Res. 33, 38-48.

KRANTZ, D. H., R. D. LUCE, P. SUPPES AND A. TVERSKY. 1971. *Foundations of Measurement*, Vol. I. Academic Press, New York.

MOSKOWITZ, H., P. V. PRECKEL AND A. YANG. 1992. Multiple- Criteria Robust Interactive Decision Analysis (MCRID) for Optimizing Public Policies. European J.Oper. Res. 56, 219-236.

VON NITZSCH, R. AND M. WEBER. 1993. The Effect of Attribute Ranges on Weights in Multiattribute Utility Measurement. Management Sci. 39, 937-943.

PÖYHÖNEN, M. A. AND R. P. HÄMÄLÄINEN. 1996. Notes on the Weighting Biases in Value Trees. Helsinki University of Technology, Systems Analysis Laboratory, Research Reports A63.

PÖYHÖNEN, M. A. AND R. P. HÄMÄLÄINEN. 1997. On the Convergence of Multiattribute Weighting Methods. Helsinki University of Technology, Systems Analysis Laboratory, Research Reports A66.

PÖYHÖNEN, M. A., H. VROJLIK AND R. P. HÄMÄLÄINEN. 1997. Behavioral and Procedural Consequences of Structural Variation in Value Trees. Systems Analysis Laboratory, Research Reports A69.

SAATY, T. L. 1980. *The Analytic Hierarchy Process*, McGraw-Hill, New York.

SAATY, T. L. 1980. How to Make a Decision: The Analytic Hierarchy Process. Interfaces 24, 19-43.

SALO, A. A. AND R. P. HÄMÄLÄINEN. 1992. Preference Assessment by Imprecise Ratio Statements. Oper. Res. 40, 1053-1061.

SALO, A. A. AND R. P. HÄMÄLÄINEN. 1997. On the Measurement of Preferences in the Analytic Hierarchy Process. J. Multi-Crit. Dec. Anal. 6, 309-319.

SARIN, R. K. 1983. Measurable Value Function Theory: Survey and Open Problems. In P. Hansen (Ed.), *Essays and Surveys in Multiple Criteria Decision Making*, Lecture Notes in Economics and Mathematical Systems 209, Springer-Verlag, Berlin, 337-346.

VANSNICK, J. C. 1984. Strength of Preference: Theoretical and Practical Aspects. In J. P. Brans (Ed.), *Operational Research'84* J. B. Brans (Ed.), North-Holland, Amsterdam.

WAKKER, P. P. 1989. *Additive Representations of Preferences: A New Foundation of Decision Analysis.* Kluwer Academic Publishers, Dordrecht.

WATSON, S. R. AND A. N. S. FREELING. 1982. Assessing Attribute Weights. Omega 10, 582-585.

WEBER, M. 1985. A Method of Multiattribute Decision Making with Incomplete Information. Management Sci.

34, 431-445.

WEBER, M. 1987. Decision Making with Incomplete Information. European J. Oper. Res. 28, 44-57.

WEBER, M., F. EISENFÜHR AND D. VON WINTERFELDT. 1988. The Effects of Splitting Attributes on Weights in Multiattribute Utility Measurement. Management Sci. 34, 431-445.

WEBER, M. AND K. BORCHERDING. 1993. Behavioral Influences on Weight Judgments in Multiattribute Decision Making. European J. Oper. Res. 67, 1-12.

VON WINTERFELDT, D. AND W. EDWARDS. 1986. *Decision Analysis and Behavioral Research.* Cambridge University Press, New York.

YU, P.-L. 1985.*Multiple-Criteria Decision Making : Concepts, Techniques, and Extensions.* Plenum Publishing Corp., New York.

Table 1
Mean number of nondominated alternatives
after $nm - 1$ judgments

Alternatives	Attributes (n)						
m	1	2	3	4	5	6	7
2	1.00	1.00	1.06	1.12	1.24	1.24	1.38
3	1.08	1.18	1.34	1.60	1.46	1.92	2.06
4	1.12	1.44	1.94	2.06	2.38	2.54	2.72
5	1.36	1.84	2.10	2.54	3.04	2.86	3.52

Table 2
Mean number of nondominated alternatives
after ordinal comparisons of value differences

Alternatives	Attributes (n)						
m	1	2	3	4	5	6	7
2	1.00	1.40	1.52	1.64	1.60	1.62	1.68
3	1.48	1.62	1.90	2.02	1.94	1.88	2.10
4	1.96	1.82	2.02	1.92	1.88	1.82	1.84
5	2.02	1.96	2.16	1.92	1.58	1.48	1.40

Table 3
Rankings on the twig-level attributes

Attribute	Ranking
Competence	$A \succ C \succ B$
Salary	$B \sim C \succ A$
Support	$A \succ C \succ B$
Taxation	$B \sim C \succ A$
Sea	$B \sim C \succ A$
Rail	$B \succ A \sim C$
Air	$C \succ B \succ A$