A Critical Survey on the Status of Multiple Criteria Decision Making Theory and Practice

TJ Stewart
University of Cape Town
South Africa

OMEGA Int. J. of Mgmt Sci., Vol. 20, No. 5/6, pp. 569–586, 1992
Printed in Great Britain. All rights reserved

0305-0483/92 $5.00 + 0.00

A Critical Survey on the Status of Multiple Criteria Decision Making Theory and Practice

TJ STEWART

University of Cape Town, South Africa

(Received December 1991)

This paper seeks to review and to contrast the main streams of thought in Multiple Criteria Decision Making (MCDM) theory and practice, without attempting to review all MCDM methods in detail. The main purpose is to identify pitfalls in the usage of various approaches, and to suggest approaches which are most robustly and effectively useable, especially by the non-expert in MCDM methodology. Problem areas in MCDM still requiring further research are also discussed.

Key words—multicriteria, decision support system, utility, goal programming

1. MULTIPLE CRITERIA DECISION MAKING PROBLEMS

THE SUCCESSES of early OR/MS projects were at least partially due to the fact that they addressed operational problems, such as production scheduling and inventory control, for which more-or-less well defined objectives could be identified with little controversy (e.g. minimize cost). As the sphere of application of quantitative management science moved from these operational decision making situations to higher level managerial planning and decision making, well defined problems gave way to what Ackoff [1] (in his by-now famous phrase) termed 'messes'. One consequence of this shift is that decision making goals became increasingly imprecise. The key philosophical departure point defining Multiple Criteria Decision Making (MCDM) as a formal approach to types of problem solving (or mess reduction), lies in attempting to represent such imprecise goals in terms of a number of individual (relatively precise, but generally conflicting) *criteria*. Over the past two decades, MCDM has developed into a discipline in its own right, with specialized conferences or specialized streams at OR/MS conferences, and with the first MCDM journal starting publication in 1992. As MCDM thus stands on the threshold of adulthood, our purpose with this paper is to attempt a critical look at what the MCDM field has achieved over the two decades, what practical tools are available, and what has still to be achieved. This is by no means a comprehensive review of MCDM theory and approaches, but rather an assessment of broad streams of thought and their impact on real-world problem solving (actual or potential). Our assessment of the methodologies will be based on the operational usefulness of each (admittedly a somewhat personalistic judgement), where 'operationally useful' requires at the very least (a) ease of use by non-experts, (b) transparency of the logic of the method to the decision maker, and (c) freedom from ambiguity regarding interpretation of inputs required from the decision maker, For more comprehensive reviews, the reader is referred to [2, 9, 16, 41].

Let us first introduce some nomenclature. We define A to be the set of decision *alternatives*, from which the *decision maker* (*DM*) has to make a selection of an $a \in A$. Identification of the set A by no means a trivial task, and it is not even clear that A can be pre-defined before the start of analysis (see for example [55] for an in-depth discussion of the manner in which the stresses generated during the search for the most preferred element of A can have the synergistic effect of bringing new decision alternatives into play). Nevertheless, it still seems useful to consider methods of analysis applicable to a fixed decision space A, as this is what formal MCDM methodologies and algorithms in fact purport to do, even though we are fully cognizant of the fact that this analysis may form just one cycle (of many repeated cycles) in the creation or discovery of a course of action which will resolve or reduce the 'mess'. A differentiation is often made in MCDM theory [18] between cases in which A is defined explicitly by a finite list of alternative actions (sometimes termed multiattribute decision making), and those in which A is defined implicitly by a mathematical programming structure (sometimes termed multiobjective optimization theory). We shall not make such a differentiation unless essential to the discussion.

A further defining feature of the MCDM model is of course the set of criteria by which elements of A are to be compared. Once again the identification of the criteria relevant to a particular context is a non-trivial task. Furthermore the selection of such criteria is part of the modelling and problem formulation process (i.e. has no objective existence), a fact that is frequently under-emphasized. Criteria are commonly developed in a hierarchical fashion, starting from some general but imprecise goal statement, which is refined into more precise sub- and sub-sub goals. Terminology is not entirely consistent between authors, but a useful general definition of a *criterion* is that of Bouyssou [10] as a tool allowing comparison of alternatives according to a *particular significance axis or point of view*. As with the generation of alternatives, we shall not here discuss the identification of criteria, but refer the reader to [21, 34, 50] to name a few who discuss *inter alia* a variety of desirable properties which should be satisfied by a set of decision criteria (such as independence, in the sense to be discussed in Section 3, operational clarity of definition, and avoidance of double counting of issues). In this context, it is worth noting the paper of Weber *et al.* [52] who show how variable levels of detail in different parts of the hierarchical tree of criteria can affect the results obtained in the subsequent analysis.

It is generally assumed that each criterion can be represented by a surrogate measure of performance, represented by some measurable *attribute* of the consequences arising from implementation of any particular decision alternative. For convenience of exposition, we shall also adopt this assumption, viz. that with each alternative $a \in A$ we can associate a vector of attributes $\mathbf{z}^a = (z_1^a, z_2^a, \ldots, z_p^a)$, where p is the number of criteria, and z_i^a is the attribute representing the outcome of decision alternative a as it affects criterion i. We note however, that this can be an oversimplification, particularly where a criterion represents a specific group interest, as then it may be difficult to associate value to the group with a single objectively measurable attribute. For the most part in this paper, we assume that the vector $\mathbf{z}^a$ does not depend on exogenous and/or random events, and can thus be determined at least approximately for each $a \in A$. We comment briefly in Section 8 on the problems which arise when the attributes are stochastically generated. *For the purposes of exposition, and without loss of generality,* we shall suppose that the attributes are defined in an increasing sense, i.e. that the DM prefers larger to smaller values of each z_i, all other things being equal.

If for two alternatives a and b, $z_i^a \geqslant z_i^b$ for all $1 \leqslant i \leqslant p$, without total equality, then we say that the alternative represented by $\mathbf{z}^a$ *dominates* $\mathbf{z}^b$. Alternatives which are not dominated by any other are also termed *Pareto optimal* or *efficient*. It is often useful to standardize attribute values in terms of the *ideal* and *nadir* values for each attribute, i.e. the best and worst available values defined by $z_i^* = \mathrm{Max}_{a \in A} \{z_i^a\}$ and $z_i^0 = \mathrm{Min}_{a \in A} \{z_i^a\}$ respectively.

In Section 2 we comment briefly on the general aims and structure of MCDM analysis, following which in Sections 3 and 4 we review theoretical and practical principles of the two best known classes of approach, viz. value function (sometimes termed utility function or scoring) approaches, and goal programming and variations thereof. In Section 5 we give attention to the rather different class of methods using 'outranking' concepts, and the role which they can play in MCDM. We briefly consider the 'fuzzy set' approach in Section 6. The use of descriptive multivariate statistical techniques, which has received relatively little attention in the MCDM literature, is discussed in Section 7. We comment on the problem of uncertainty in Section 8, before drawing together the discussion in terms of conclusions regarding what is available to the practitioner today, and what areas of research in MCDM still need attention.

2. THE GOALS OF FORMAL MCDM TECHNIQUES

Multicriteria decision making is a human, managerial task. It can and never will be automated by tools, techniques or algorithms. The aim of any MCDM technique is thus to provide help and guidance to the DM in discovering his or her most desired solution to the problem (in the sense of that course of action which best achieves the DM's long-term goals). Some philosophical debate is conducted from time to time regarding whether or not the preferences of the DM which would define best achievement of goals pre-exist in the DM's mind prior to the analysis, or whether it is formed in the process of seeking a solution to the decision problem. From a practical point of view, the distinction is, however, often not so important. Either way round, any MCDM technique seeks to make the DM's search as effective and efficient as possible, maintaining some degree of consistency in the search, or at least warning the DM of inconsistencies as they arise, without imposing undue and unjustifiable structure on the decision maker.

MCDM methods may be used in two contexts. In the first, the DM would be either a single individual or an essentially homogenous group, seeking to make a decision which does not seriously impact, or require justification to, other parties. In this case, methods can be relatively informal, and the rationale behind the decision reached does not need substantial documentation. This contrasts with the context in which the DM (individual or group) has to make decisions on behalf of a much larger group or community, or has in fact only to generate a short-list of alternatives for consideration elsewhere. This might occur with managers in large corporations, or with public servants. In such cases, the rationale for choices must be clearly documented, and justice must be seen to be done, in the sense that criteria might refer to different members of the community being served, and proper consideration of each interest must be demonstrated. This requirement necessitates the use of rather more formal methods of analysis, even where these may be less efficient, and/or may impose structures (of 'rationality', for example) which may not strictly be justifiable.

The context of any particular case needs to be taken into consideration when selecting a particular MCDM method for use. One distinction between methods which is particularly relevant in this context is that between 'prior' and 'progressive' articulation of preferences [16, 18], where methods in the latter category are sometimes termed 'interactive'. Methods of prior articulation of preferences require the DM to specifiy value judgements somewhat in isolation from the particular choices at hand, and these are then translated into the particular choice, or choices, from A which are consistent with these preferences. This approach is relevant particularly in contexts where full justification and rationale for decisions are required. Methods of progressive articulation of preference allow the DM to explore the decision space systematically, without having to specify any prior preferences. This is more efficient, and requires less sweeping assumptions regarding preference structures, but is also more open to manipulation by skilled users, and is thus less defensible when solutions have to be justified and/or rationalized.

In the sections which follow, we attempt to illustrate the extent to which each MCDM technique can be implemented in either the prior or progressive articulation mode (or both).

3. VALUE OR UTILITY BASED APPROACHES

3.1 Basic principles of preference measurement

For the quantitatively oriented manager or management scientist, there is great appeal in being able to establish some means of associating a numerical score or value with each decision alternative, after which choice of the optimal alternative becomes automatic. This gives a sense of objectivity to the process, and certainly helps to focus discussion on the 'borderline' choices, and to defuse some of the emotion, in a group decision making setting. In particular, and in the interests of simplicity and parsimony, it is common to want to seek a *value function* based on a simple addition of *scores* representing goal achievement according to each criterion, i.e. to represent the total score or value of the alternative described by the attribute vector $\mathbf{z}$ as follows:

$$V(\mathbf{z}) = \sum_{i=1}^{p} v_i(z_i) \quad (1)$$

where $v_i(z_i)$ is the *score* associated with a level of performance for criterion i represented by the attribute value z_i.

At a fundamental level, $V(\mathbf{z})$ need not be endowed with any properties other than that of preference ordering: $\mathbf{z}^a$ is preferred to $\mathbf{z}^b$ if and only if $V(\mathbf{z}^a) > V(\mathbf{z}^b)$. The behavioural assumptions necessary to ensure the *existence* of such a value function preserving preference ordering are well-known (cf. [21, 50]), and yet frequently seem to be overlooked. Keeney and Raiffa [21] term the assumptions *preferential independence* of the criteria; in essence this means that a decision maker can state whether an offered trade-off between any two attributes, representing two criteria, is or is not acceptable, "all other things being equal" (i.e. no other attributes varying), without requiring knowledge of performance levels achieved for the other attributes. It is in fact evident that the concept of trade-off is central to the interpretation of an *additive* value function. In particular, an increase of one unit in $v_i(z_i)$ will always precisely compensate for a loss of one unit in $v_j(z_j)$, "all other things being equal", irrespective of the performance levels achieved for any attribute (including i and j). One implication is that unit increments in $v_i(z_i)$ have the same marginal value to the decision maker, in terms of trade-offs with other criteria, irrespective of the value of z_i, i.e. $v_i(z_i)$ represents an *interval scale* of preference. The scores implied by the additive form (1) thus represent a very specific form of currency in which all criteria can be measured, and are not any arbitrary ordinal preference measure.

Keeney and Raiffa ensure (as far as is practically possible) that the scores they generate satisfy this constant trade-off condition by construction. Their 'lock-step' and 'mid-value splitting' methods, build up the $v_i(.)$ functions precisely by considering pairwise trade-offs between which the decision maker is indifferent, or increments judged to be of equal importance. The problem is that these approaches can be rather tedious and mystifying to the decision maker, and in consequence it is very appealing to represent scores in the form:

$$v_i(z_i) = w_i u_i(z_i) \quad (2)$$

where:

(a) $u_i(z_i)$ is a *marginal utility function* assessed separately for each criterion (without reference to trade-offs between criteria), and normalized to some convenient scale (e.g. 0–100);

(b) w_i represents a *weight* associated with the importance of the criterion i, and is in effect used to scale the individual scores for each criterion on to a commensurate scale.

The marginal utilities may be evaluated indirectly by the standard decision analysis technique of ascertaining certainty equivalents for two-point lotteries, or by direct scoring (e.g. [50], Section 7.3). In direct scoring, the ideal and nadir are given scores of for example 100 (or 1) and 0 respectively, after which other values are fitted into this range in such a way that equal score differences correspond to equal strengths of preference for one over the other. The weights can be determined directly by evaluation of the importances of the swings from best to worst on each criterion, relative to either the most or least important (e.g. [50], p. 286), i.e. the relative importances of the ranges represented by $[z_i^0, z_i^*]$ are assessed directly for each attribute relative to the others. It seems natural to assess these importance weights in ratio terms, i.e. to assess directly the ratios w_i/w_j. But it is important to note that the additive score function imposes a very specific meaning on to the weights. Suppose for purposes of argument that the marginal utilities are scaled on to the [0–1] interval, and that $w_i < w_j$; then the ratio w_i/w_j is precisely the maximum proportion of the utility range on criterion j that would be sacrificed, in order to increase attribute i from z_i^0 to z_i^*. The fundamental principle is that the weights have a natural implicit trade-off interpretation, that can only be assessed *in the context of the ranges of options available*. (For example, the 'importance' of risk to human life relative to cost savings can only be assessed in the context of a particular range of possible outcomes.) The 'swing-weighting' approach of von Winterfeldt and Edwards [50, p. 286] uses this principle explicitly. Whether simpler procedures such as setting the weight of the most important attribute to 100, and scoring others relatively, satisfy the trade-off property or not is a moot point; but even in such a procedure, the range of values being assumed must be made evident explicitly for the approach to have any justification at all.

In principle, any requirement that the value function preserves *strengths of preferences* as expressed by the DM between different gains or losses in one or more attributes, imposes a somewhat stronger requirement on the value function than the simple need to preserve order. On theoretical grounds this demands either stronger behavioural assumptions than mere preferential independence in order to justify the simple additivity in (2), or a need to aggregate the individual marginal utility functions in ways other than additive, for example in the multiplicative form:

$$1 + kU(z) = \prod_{i=1}^{p} [1 + kw_i u_i(z_i)] \qquad (3)$$

Fortunately, it seems that little is lost in retaining the assumption of additivity, particularly when outcomes are deterministic (but see Section 8). Any modelling imprecisions are likely to be outweighed by the high degree of imprecision in measuring the $u_i(z_i)$ and w_i (cf. [50], Section 7.1). Provided, therefore, that adequate care is taken in capturing non-linearities in the $u_i(z_i)$, and in ensuring that the importance weights w_i are assessed in the proper context, there is good reason not to attempt any more esoteric preference models than that given by (2). This form of additive model is well-justified theoretically, and is easily understood, in that the connections between the inputs provided by the decision maker and the outputs obtained are not hidden behind a screen of complex mathematical manipulation. Belton [4] or Belton and Vickers [7] give a useful description of the practicalities of this approach, with reference to the sensitivity analysis which should accompany it.

We have dealt with the principles underlying the use of additive scores or value functions at some length, because anecdotal evidence suggests some form of additive scoring is a very wide-spread approach to resolving multi-criteria decision problems in practice. This might not be evident from the literature, which tends only to report relatively sophisticated applications of multi-attribute value or utility theory. It is the author's experience however, that when decision makers, particularly in a group or committee setting, are confronted with difficult multi-criteria choices, they almost instinctively turn to some form of simple scoring. In fact, our experience is that more time is spent in convincing clients *not* to use some inappropriate scoring technique, or in warning of the dangers of interpretation of such scoring, than in presenting alternative MCDM approaches. If this practice is so wide-spread, it is probably most effective to retain the natural and easily understood additive scoring, instead of more sophisticated methods, but to ensure that the definition of criteria, and the scoring methods used, are properly justified and understood by those providing the inputs.

3.2 The Analytic Hierarchy Process (AHP)

Although, as we have indicated, our recommendation is to assess additive score functions in a very simple straightforward manner, we do need, for completion, to comment on the Analytic Hierarchy Process (AHP) methodology introduced by Saaty [34], which also makes use of an additive value function of the form given by (2) (although the terms used in the previous section are not conventionally used in AHP). There has developed an AHP school quite distinct from the rest of MCDM thinking. A variety of applications have been reported [48]. Many decision makers seem to find AHP an appealing tool to use, yielding results which are at least plausible in many settings, although this may at least partially be due to the use of the instinctively natural additive aggregation, as we have discussed above. In spite of this, a number of writers, many of them after extensive experience with AHP (e.g. [5, 6, 14, 20, 24, 25, 38] have criticized various fundamental aspects of AHP, which criticism we now attempt to understand in the light of the principles discussed above.

Saaty, and other proponents of AHP, have strived to emphasize that the axiomatic basis of AHP is fundamentally different from that of utility theory (e.g. [35]). In fact, there seems to be a substantial divide between the AHP school and the remainder of the MCDM field, in spite of pleas for more cross-fertilization (e.g. [14, 24]). The differences in philosophy notwithstanding, however, the basic trade-off implications of the additive preference scoring function, as we have discussed them, apply whatever the axiomatic justifications, and it is here that the first cause for concern arises. The basis of AHP, is that the decision maker provides assessments of the ratios of (in our nomenclature) $u_i(z_i^a)/u_i(z_i^b)$ for each attribute i, and for every pair of available alternatives a and b. The resulting values for $u_i(z_i)$ would appear to be on a *ratio scale* of preferences, and not on the interval scale as implied by the additive form. Consider, for example, two criteria i and j having equal importance weights, and two alternatives a and b differing only on criteria i and j, and such that $u_i(z_i^a)/u_i(z_i^b) = u_j(z_j^b)/u_j(z_j^a)$. It will not in general be true that a and b have equal scores on (1); and yet when a decision maker expresses equal ratios in this context (i.e. that preferences for a over b on criterion i is of the same importance as preference for b over a on criterion j), it is difficult to see that the DM has anything else in mind but that the loss on one criterion when moving from alternative a to b is exactly compensated for by the gain in the other. Lootsma [24] argues in effect from this observation that it is the ***logarithm*** of Saaty's marginal value scores which should be used in the additive function, and this does seem to be more consistent with our earlier arguments.

A second point of contention regarding AHP derives in effect from the practice in AHP of normalizing the marginal utility functions such that:

$$\sum_{a \in A} u_i(z_i^a) = 1 \tag{4}$$

for each criterion i. Since the range of values for $u_i(\,)$ differs from criterion to criterion, it follows that the magnitude of the effect of a change in z_i is determined partly by this range, and partly by the weight w_i. This makes the meanings of the weights more difficult to interpret than with swing weights. This leads to the phenomenon of rank reversal in AHP, first pointed out by Belton and Gear [5]. They showed that the relative rankings of two alternatives according to (1) can change depending on what other 'irrelevant' alternatives are available, unless the criterion weights are at the same time modified (cf. also [4]). If the other alternatives do not affect the ranges of responses, however, it is unlikely that a DM would want to change these weights. This phenomenom led Dyer [14], with perhaps some hyperbole, to conclude that the AHP rankings are arbitrary. Saaty and Vargas [37] argued in response to Belton and Gear that human preferences *are* affected by the presence or absence of other alternatives, due either to changed ranges of options available, or to a realization that a previously unrecognized criterion should also be taken into consideration. These effects need, however, to be modelled explicitly, and can then be accommodated into any MCDM approach. As it stands, there is no evidence, or even *prima fa^ie* motivation, as to why the AHP-induced rank reversal is in any sense a good model of either of these effects, and without such motivation the rank reversal must be seen as a disturbing property of a normative decision-aiding procedure. Saaty [36] has pointed out that rank-reversal can be avoided by applying what he terms the 'absolute measurement mode' of assessment, viz. identifying a fixed set of possible outcomes for each criterion i, z_i^1, $z_i^2, \ldots, z_i^m$ say, and assessing $u_i(z_i^r)/u_i(z_i^s)$ across these m possibilities independently of the actual set of alternatives currently under consideration. This would seem to be the better approach in most cases, although it is not commonly emphasized in discussions on AHP.

For completion, we should also mention three further practical problems in applying AHP. Firstly, the manner in which AHP is often implemented encourages users to assess importance weights in isolation from the specific ranges of options available. Weights cannot be expressed in the absence of context, and the result can easily be that the users assess weights relevant to a context different to the current one. Secondly, the usual form of input required by AHP is not the numerical ratio described above, but rather a preference statement on a nominal nine-point scale *which is interpreted as a ratio*. (Thus for example, a value of 5 on Saaty's scale corresponds to the nominal description '... more important than ...', but is used as if the decision maker has asserted that $u_i(z_i^a)/u_i(z_i^b) = 5$ or $w_i/w_j = 5$, as the case may be.) Justification for this quantitative interpretation of a nominal scale is anecdotal, and has been questioned (cf. [24]). The final practical problem relates to the insistence in AHP applications on using the eigenvector procedure, for estimating the marginal values and criterion weights from the ratios provided. If all the ratios are fully consistent, then the eigenvector procedure recovers the correct individual values, but then so do many other procedures. Alternatives such as logarithmic least squares (e.g. [24]), and what Islei and Lockett [20] term geometric least squares have been suggested on other theoretical grounds, and have a major practical advantage over the eigenvector method, in that they do not require that all pairwise comparisons be provided, a task which can quickly become extremely tedious and time-consuming.

In spite of the above criticisms, the rankings generated by AHP may still be useful in some contexts, especially when applied under the guidance of a skilled facilitator. Nevertheless, AHP needs to be used with considerable caution, and use of the simpler procedures of Section 3.1 may be more justifiable (although, of course, *any* scoring function can only give guidance to the decision maker, and cannot be viewed prescriptively). A further advantage of the direct scoring methods of Section 3.1 is that the processes involved are under the direct control of, and quite transparent to the decision maker, in contrast to the more esoteric processing occurring in AHP (i.e. the translation of the nominal into a ratio scale, and the eigenvector analysis).

3.3 Use of value functions in interactive mode

The problems with *a priori* articulation of value functions as discussed above, are that the value functions must be valid over wide ranges, and that the DM must be able to express global preferences (e.g. the importance weight of the full range of change in one criterion). Methods of progressive articulation of preference can also be based on the existence of a value function such as that of (1), or a more general form, but with the advantage that any assumptions and/or value judgements need only apply over restricted ranges. Perhaps the first such attempt is the method of Geoffrion *et al.* [15]. They require the decision maker to provide trade-offs in the vicinity of a particular feasible solution, which is equivalent to giving ratios of partial derivatives:

$$\frac{\partial V(\mathbf{z})/\partial z_i}{\partial V(\mathbf{z})/\partial z_j} \tag{5}$$

These are sufficient to provide the search directions for standard non-linear programmng algorithms, although the step length has also to be judged interactively by the decision maker. The Geoffrion–Dyer–Feinberg algorithm has become a standard of comparison in the literature for assessing the performance of other algorithms. It is however not very efficient in using preference information, as trade-offs from one iteration are entirely discarded before the next, and there has been little serious practical use reported of the approach. One of the challenges to research in MCDM is to develop interactive procedures in which trade-offs are assessed at various points in the search for a solution, and are used to constrain later search even though trade-offs may change during the course of this search.

Specifically within the context of multiple objective linear programming, some interactive methods have developed, based on a full linearization of (1) into the form:

$$V(\mathbf{z}) \approx \sum_{i=1}^{p} \lambda_i z_i \tag{6}$$

Although such a linearization must of necessity only have local validity, the methods will converge to the true optimum, provided that the true value function satisfies the relatively mild condition of pseudo-concavity. This follows from the fact that if the set of feasible attribute vectors **z** is convex (as must be true in multiple objective linear programming), then the true most preferred solution will maximize (6) for *some* set of λ_i weights. Two approaches of this type are those of Steuer [41], Sections 13.4 and 13.5, and of Zionts and Wallenius [59, 60]. Both work on the basis of placing increasing restrictions on the allowable range of values which the λ_is may take on, based on preference judgements expressed by the decision maker. In the case of Zionts–Wallenius, the decision maker is required to compare a sequence of pairs of attribute vectors, and to select the most preferred in each such pair. Suppose therefore that the decision maker states that vector $\mathbf{z}^a$ is preferred to vector $\mathbf{z}^b$: this implies that the λ_is must satisfy a constraint of the form:

$$\sum_{i=1}^{p} [z_i^a - z_i^b]\lambda_i \geqslant 0 \tag{7}$$

With careful selection of the pairwise comparisons to be made, the above constraints quickly reduce the range of feasible weights to the point where all generate the same optimal basic solution to the LP. This strictly only generates the most preferred basic solution; a more preferred solution may be found on adjoining facets of the LP, which would have to be found during a secondary search, as described in [60], but this is unlikely to have serious consequences in LPs of realistic size. Stewart [43, 44] has suggested an extension of the Zionts–Wallenius scheme which does evaluate additional non-basic solutions as well.

The Steuer approach starts by generating a small number of sets of weights over a wide range, and computing the corresponding solutions maximizing (6). The decision maker indicates which of these is least preferred, and this leads to excluding part of the range of possible weights (according to a somewhat heuristic algorithm, details of which can be found in the original reference). Here, too, eventually a point is reached where all remaining sets of weights generate the same solution to the LP.

Both the Steuer and the Zionts–Wallenius methods are most conveniently applied in the multiple objective linear programming context, but can in principle be applied whenever the space of feasible attribute vectors is convex. In fact this extends directly to the discrete case, provided that the attribute vectors corresponding to the discrete alternatives lie on a convex surface. The Zionts–Wallenius idea has been extended to more general discrete problems by Korhonen *et al.* [22], but the efficiency of this in terms of use of decision maker's time is not clear. The extension discussed in Stewart [43, 44] does apply equally well to general discrete problems.

Although the Steuer and Zionts–Wallenius ideas appear to be simple and efficient to implement, there is relatively little practical implementation reported in the literature, apart from that reported by the authors themselves.

4. GOALS AND REFERENCE POINTS

While the use of some form of scoring or value function may be appealing to the quantitatively oriented, and in fact widely used in the quest for 'objectivity', the earliest formal MCDM methods were of the goal programming form (attributed generally to Charnes and Cooper [12]). Useful reviews of the development of goal programming (GP) are found in [17, 19, 54], while [32] gives a comprehensive bibliography of work up to 1982. In some senses, GP can be seen as an operationalization of Simon's 'satisficing' concept. According to this model (cf. [39], pp. 272–273), the natural decision making heuristic is to concentrate initially on improving what appears to be the most critical problem area (criterion), until it has been improved to some satisfactory level of performance. Thereafter, attention is shifted to the next most important issue, and so on. Goal programming formalizes this heuristic, although we should note that Simon did not view this heuristic as necessarily desirable, but merely that it is a response to bounded rationality. (In the introduction to the third edition of Simon [39], he refers to: "... human beings who *satisfice* because they have not the wits to *maximize*".)

We suppose therefore that for each attribute i (representing a particular criterion of evaluation), the decision maker can specify some desirable goal or target level of achievement, T_i, say. If the targets T_i are realistically specified, then there may exist only one (or at most a small number) of elements of A such that $z_i \geqslant T_i$ for all attributes i (recalling our convention that each attribute is defined to be increasing in preference).

In general, however, the decision maker may find it extremely difficult to know what are realistic targets to set; there may then either be no alternative, or very large numbers of alternatives, which satisfy the goals. Traditional GP assumes that there are some absolute target levels, which can be specified almost context-free, above which the decision maker will always be satisfied. It is implicitly assumed to be very unlikely that any solution exists satisfying all the goals; but if there are any, then any such solution will be entirely satisfactory, and there is no further problem. Since, by assumption, however, such a fully satisfying solution will probably not exist, the aim of GP is to find a solution which is as near as possible to the target. This requires some definition of 'distance', or measure of discrepancy from the target, which is always of a somewhat *ad hoc* form (in the sense that it is generally more difficult to link such distance measures to behavioural assumptions, than it is for the value function scores discussed in Section 3). Three possible forms of discrepancy have been proposed, which can be used either singly, or in combination with each other:

(a) ***Archimedean***: Weights w_i are associated with each attribute, and the alternative $a \in A$ is selected which minimizes:

$$\sum_{i=1}^{p} w_i \cdot \text{Max}[0; T_i - z_i^a] \quad (8)$$

Clearly this is nothing more than a variation of the additive preference function given by (1), with little check on whether the assumptions of the additive function are satisfied. Nevertheless, when discrepancies are small it is a plausible distance measure (between a desired point and the feasible region). When discrepancies are large, however, the results can be quite at variance with reasonable expectations. No cognizance is taken of the fact that large deviations are likely to be of greater concern than small; it is for example quite possible that the 'optimal solution' will have small or zero deviations from target on most criteria, but very large deviations for a few, where these few may not necessarily be by any means the least important. This problem is ameliorated by using a metric other than L_1 in (8), for example L_2, but this introduces a further *ad hoc* element. The special case of L_∞, however, is of particular interest, and is discussed under (c) below.

(b) ***Pre-emptive***: According to Hannan [17], this discrepancy measure was introduced in order to circumvent the difficulty of specifying the weights w_i in (8), and certainly became at one stage almost the standard approach. In the simplest form of pre-emptive approach, the criteria are first ordered from most ($i = 1$) to least ($i = p$) important. For the first criterion, define $z_1^+ = \text{Min}[T_1, z_1^*]$, and restrict the set A to those alternatives a for which $z_1^a \geqslant z_1^+$, adjusting the ideals z_i^* for all other criteria accordingly. This process is repeated for each criterion in turn, until such time as A is reduced to a single element. More generally, criteria may be classified into ordered importance classes, and the pre-emptive rule applied only between classes using an Archimedean measure of discrepancy from the target within each class as a surrogate criterion representing the class. The pre-emptive approach led to quite elegant variations on the simplex method when applied to multiple objective linear programming, but as a preference model is questionable, unless the goals are very realistically set. The main weakness is that no trade-off is allowed between importance classes, even if major gains are achievable on one attribute with infinitesimal losses on another. This appears to be massively at variance with human preference structures (e.g. many increase their risk of death every day for the minor convenience of using their own cars rather than public transport, or walking!). We would concur with Hannan [17, p. 539] in his comment "Thus preemptive GP should be used only when the priorities truly are preemptive, and not as a surrogate for problems with commensurable goals just to avoid specifying weights for the goals". In our experience it is rare for priorities to be 'truly preemptive'.

(c) ***Tchebycheff, or Min-Max***: This is a useful alternative to the traditional discrepancy measures, i.e. Archimedean with the L_1 norm or pre-emptive, but retains their simplicity (e.g. LP methods can still be applied to the MOLP porblem using this approach). The idea is to minimize the maximum weighted deviation, i.e. to select the alternative $a \in A$ which yields the smallest value of $\text{Max}_{\{1 \leqslant i \leqslant p\}} w_i[T_i - z_i^a]$ (truncated at zero if $z_i^a \geqslant T_i$ for all i). Although in a limiting sense this remains an Archimedean measure, based on the L_∞ norm, it turns out to be a much more robust measure. For any reasonably plausible target values, solutions are generally quite acceptable, in the sense that some of the silly answers generated by the other two discrepancy measures cannot occur, and often not

> overly sensitive to choice of weights. In fact, the Tchebycheff procedure seems to be closer to the spirit of the satisficing heuristic, by ensuring that the most urgent criterion always receives maximum attention. There is one technical problem however: it is possible that ties may occur (i.e. different alternatives giving the same minimum discrepancy), in which case the algorithm may select an alternative which is dominated. We could avoid this by ensuring that the set A is Pareto optimal at the outset, but if this is not practical, then the problem can be overcome by modifying the measure to one of minimizing:
>
> $$\mathrm{Max}_{\{1 \leq i \leq p\}} w_i[T_i - z_i^a] + \epsilon \cdot \sum_{i=1}^{p} w_i z_i^a \qquad (9)$$
>
> where ϵ is a small positive constant. Expression (9) has been motivated especially in the context of interactive goal programming ideas (cf. [53]), but appears to be equally appropriate to the standard GP approach.

GP has a real advantage over value/utility based approaches when the number of criteria becomes large (greater than about 10, perhaps), in that the construction of trade-offs and/or value functions can become increasingly tedious. This is particularly true when the numbers of values possible on each criterion is also large (or perhaps a continuum). Of course, with large numbers of criteria, GP also suffers from the difficulties entailed in establishing weights subjectively. Nevertheless, especially with the Tchebycheff norm, GP is probably the method of choice for the purpose of screening a large (or infinite) number of alternatives down to a short-list, when the number of criteria is large. Weights will not then play an overly important role, and can be chosen purely on considerations of scaling (e.g. ensuring that the ranges between T_i and z_{i^*} have equal weight in each case, as in [28]).

One problem in applying GP with purely 'prior articulation of preferences' is that it can be difficult for DMs to specifiy goals meaningfully *a priori*. It seems therefore that, where possible, GP should be used in an interactive ('progressive articulation of preferences') mode, and in fact lends itself easily to such an approach. An initial set of goals can be specified (perhaps even simply the ideals for each criterion), and the GP solution found. This solution is presented to the DM who may then wish to modify the goals in the light thereof. The process can be repeated as often as needed. It is in fact difficult to believe that GP is ever used in any other way!

The above interactive GP approach can become extremely unstructured or 'hit-and-miss', although one attempt at introducing more structure within essentially the standard GP structure above may be found in Masud and Hwang [28]. Other authors have, however, suggested variations of the broad GP theme specifically for interactive implementation, but very often these are correctly classified as *reference point* methods to use a more general term. Within this more general context, the 'reference point' need not necessarily be aspiration levels for each criterion (beyond which the DM is 'satisfied'), but could be either lower bounds on desirable options for each, or more simply an estimate of a likely acceptable compromise position. One of the earliest of these variations is the Step Method (STEM) of Benayoun *et al.* [8]. The distinguishing feature of STEM is that at each phase of interaction, the DM, after observing the last solution obtained, indicates the maximum amount he would be prepared to sacrifice on each criterion (possibly zero for some, but not for all criteria). The maximum sacrifices allowed are converted into hard *lower bounds* on performance for each criterion, and further analysis is restricted to those criteria for which no sacrifice was accepted. STEM thus maintains two sets of reference points, viz. the ideals which serve as goals, and lower bounds imposed as hard constraints.

The Interactive Multiple Goal Programming (IMGP) procedure of Spronk [40] also works in terms of two sets of reference levels which converge towards each other. The nadir values are set as lower bounds, and the ideals as 'potentials' representing the best that can be achieved. At each iteration, the DM examines the lower bounds (rather than a specific alternative), and indicates which of these should be improved first. A tentative increase in the lower bound for this criterion is implemented, and the effect on the potentials with this tightened constraint on acceptable values is calculated and displayed to the DM. If the DM accepts that the loss in potential is acceptable, the new lower bound is made definitive, and the next phase of interaction starts; otherwise the tentative increase is reduced systematically until the losses in potential are acceptable. In some practical experience [45], we have found this procedure to be relatively slow, but to produce very satisfying results. One particularly useful property of IMGP is that the DM is not required to sacrifice anything which he may perceive he has already gained. This is important in that people are loathe to accept sure losses, and may tend to be too cautious in lowering aspirations in more conventional goal programming, leading to early termination at a poor solution. A fairly similar approach to IMGP, but which allows a degree of backtracking is the PRIAM method [23].

It was Wierzbicki [53], however, who really formalized the concept of the reference point approach. He clearly perceived of the reference point as neither an ultimate aspiration level (beyond which lies satisfaction), nor merely a minimum necessary level of performance. In fact different decision makers may differ considerably in the degree of caution with which they specify goals or reference levels. It is necessary therefore not only to minimize measures of underachievement, but also to maximize overachievement as far as is possible in a balanced manner. In effect, he defines a *scalarizing function* to be optimized, which is really a surrogate value function, but which is defined so as to give first preference to improving the worst underperformances relative to the reference point. A typical form of scalarizing function is essentially that given in (9), but without truncation to zero (as in standard GP) when $z_i > T_i$, and with the weights proportional to the reciprocal of the range from nadir to ideal. The procedure is for the DM simply to continue modifying his expectations as represented by the reference levels until no further gains are perceived. Our experience has been that variations of this approach are particularly well suited for use in the Decision Support System framework (cf. [26, 46]).

5. THE OUTRANKING CONCEPT

Implicit in both the value function and goal/reference point approaches, are two assumptions, viz. (i) that there is always scope for some form of 'compensation' between attributes (i.e. that no matter how important one attribute is, a sufficiently large gain in a lesser attribute will eventually compensate for a small loss in the more important, cf. [10] for discussion); and (ii) that there exists a 'true' ordering of the alternatives (and by implication a 'best' alternative) representative of the DM's preferences, which needs to be 'discovered'. It is difficult to envisage a situation in practice of no compensation, and assumption (i) seems relatively mild therefore. No serious decision analyst would, on the other hand, view assumption (ii) as being a precise and accurate representation of the real world. Preferences are not constant in time, are not unambiguous, and are not independent of the process of analysis. Nevertheless, both the value function and goal programming approaches place a structure on the process of learning and discovery which ensures that the solution found does satisfy some internal consistency properties, which many find appealing as at least a desired feature of 'rational' decision making. The use of value functions or of goal programming techniques does not of course preclude the possibility of re-examination of the analysis after the solution is obtained, in order to evaluate whether the values and preferences expressed earlier are still valid in the light of what has been learnt subsequently. If not, another round of analysis can be conducted, until convergence is reached in the sense that the solution is optimal relative to the final 'true' preference orderings revealed. (Care is needed to ensure that the re-evaluations are not just an attempt to juggle inputs to obtain a *post hoc* justification for a decision already reached, however!). Thus although assumption (ii) is unlikely to be a true representation of reality, it does provide a normative guiding principle, in searching for an acceptable solution.

The 'outranking' class of approach to MCDM, which is particularly popular in Europe, arises out of an attempt to avoid assumptions (i) and (ii) as far as possible. One definition of outranking may be that alternative $\mathbf{z}^a$ *outranks* $\mathbf{z}^b$ ($\mathbf{z}^a S \mathbf{z}^b$) if there is a "sufficiently strong argument" [33] in favour of the assertion that $\mathbf{z}^a$ is at least as good as $\mathbf{z}^b$, from the decision maker's point of view. This is a relatively fuzzily defined notion, and we can well have situations in which neither of a pair of alternatives outranks the other ('incomparable'), and those in which each outranks the other (argument in favour of 'indifference'). A popular manner of rendering this concept operational for practical MCDM (for other methods, see Chapter II of [2]) is through the ELECTRE approaches (e.g. [33]). ELECTRE is based on the so-called *concordance* and *discordance* indices, which represent respectively the 'arguments' for and against the outranking relationship. For any specific pair of alternatives $\mathbf{z}^a$ and $\mathbf{z}^b$, the concordance index is typically defined by something like:

$$C(a, b) = \sum_{\{i: z_i^a \geq z_i^b\}} w_i$$

i.e. the relative weight of attributes on which $\mathbf{z}^a$ is preferred, where generally the weights are normalized to sum to one. These weights are not seen in trade-off terms, and are rather relative voting weights accorded to the respective criteria. This intrinsic, or non-compensatory, interpretation of weights is seen by the proponents of outranking methods to be one of its main advantages. On the other hand, it is not at all clear whether decision makers can or would perceive weights in any terms other than at least the relative importance of the trade-offs offered by the available ranges of values on each attribute. (Certainly this writer, on introspection, finds it difficult to evaluate weights in any other terms.) Where weights do represent some form of trade-off desirability, the legitimacy of their use as voting weights is not immediately evident, and in any case their use in this way represents a discarding of important preference information.

The discordance index *for criterion i* is positive only if $z_i^b > z_i^a$, and represents the extent to which the difference $z_i^b - z_i^a$ is sufficient of its own to contradict ('veto') the assertion that $\mathbf{z}^a$ outranks $\mathbf{z}^b$, and is commonly defined as follows:

$$d^i(a, b) = \max\left\{\frac{z_i^b - z_i^a - p_i}{s_i}, 0\right\}$$

where p_i is a threshold of importance for differences, and s_i is a scaling factor for the attribute values z_i.

The earlier versions of ELECTRE declared $\mathbf{z}^a$ as outranking $\mathbf{z}^b$ if $C(a,b) \geqslant c^*$ and $d^i(a,b) \leqslant d^*$ for some suitable threshold values c^* and d^*. The outranking relationships can be displayed either in matrix form or as a graph with nodes representing alternatives and directed arcs representing outrankings. In addition, a partial ordering of the alternatives can be derived from the outranking relationships. Some quite complex algorithms have been suggested for this purpose, but a simple and effective approach is simply to score each alternative by the number of alternatives which it outranks less the number which outrank it. For greatest ease of interpretation of these results, it is desirable that the number of outranking relationships be neither too large nor too small. This is critically dependent upon choice of c^* and d^*: if c^* is too large and/or d^* too small, then there will be no outrankings, while in the contrary case a situation may be reached where virtually every alternative outranks every other. In practice, it is necessary to experiment with a variety of threshold levels until a suitable density of outranking relationships is obtained. A systematic manner of achieving this is suggested by Vetschera [49].

This writer's experience is that ELECTRE used in this way is particularly valuable as a descriptive device when the number of alternatives remaining under consideration is small (e.g. 6 or less). The outranking graphs help DMs to focus attention on critical issues and to gain insight into their own preference structures. It can also assist in understanding how and why other MCDM methods generate the answers they do. For larger numbers of alternatives, however, much of this insight is lost through the overwhelming level of detail, while the partial rankings may do little more than to group alternatives into 2 or 3 equivalence classes, on a basis which contains some rather *ad hoc* elements (e.g. the precise algorithm used for establishing the ranking).

Later versions of ELECTRE (version III) attempt to give some measure of the degree of outranking, by using the concordance index moderated by a function of the discordances. This assists in deriving a ranking of alternatives, but introduces further *ad hoc* functional forms, which although intuitively appealing are difficult to verify empirically as models of human preferences.

6. FUZZY SET THEORY

The problem of multi-criteria decision making is fundamentally one of imprecision in human preferences: we cannot say precisely whether one alternative is preferred to another because one is better on some criteria, and the second on other criteria. Zadeh's fuzzy set theory (see [57] for a review in the operations research context) has thus been suggested as a means of resolving MCDM problems; in fact the first such suggestion goes back more than 20 years to Bellman and Zadeh [3].

In fuzzy set terms, each alternative would have some degree of membership $\mu_i(z_i)$ in the fuzzy set of good (or acceptable, or satisfactory) solutions for each criterion taken in turn. The alternative's membership in the fuzzy intersection of all these single-criterion fuzzy sets then indicates the strength of its claim to being good (or acceptable, or satisfactory) overall. In fact, we might well then rank order alternatives according to the degree of their membership in the fuzzy intersection. There are, however, two practical problems associated with this approach.

The first problem relates to the assessment of the membership functions $\mu_i(z_i)$. In principle this should represent the extent to which an attribute value of z_i satisfies achievement of criterion i. Operational procedures for assessing this are poorly defined in the literature, if at all; and where defined, it is difficult to see that the DM would understand what is required in terms of anything other than those of value functions. Reported applications of fuzzy sets to MCDM problems tend to propose rather *ad hoc* measures for $\mu_i(z_i)$, such as a linear function increasing from 0 to 1 over some range of values (possibly just the range from ideal to nadir values), e.g. [47, 56]. The algorithmic effect of such functions is no different to that of value functions or of achievement or scalarizing functions in goal programming, without the advantage afforded by the discipline, especially of value function theory, of specific behavioural assumptions regarding DM preferences which can be checked.

Secondly, the definition of the fuzzy intersection itself is not without argument. The original Zadeh definition requires that the fuzzy intersection of the degrees of achievement on two criteria i and k be given by the fuzzy set with membership function $\mathrm{Min}\{\mu_i(z_i), \mu_k(z_k)\}$. Zimmermann and Zysno [58] have shown that human preferences are considerably more complex than this representation indicates, and in any case, on *prima facie* grounds the Min function seems an unlikely candidate for MCDM at least, as it allows for no compensation between criteria. Alternative representations of the fuzzy intersection have been suggested, such as the product $\mu_i(z_i) \cdot \mu_k(z_k)$, but these too are *ad hoc*, and not clearly grounded in specific and testable behavioural assumptions.

Our view, therefore, is that at this stage, attempts to apply fuzzy set theory to MCDM in any operational manner leads to models which are effectively either goal programming or value function models, but with the inputs required from the DM obscured behind a language which may seem more 'natural', but which gives greater scope for misunderstanding between analyst and DM than the simpler clear-cut language of goal levels, aspiration levels, trade-offs or relative values. It is of little concern to practical goal programming or value function analysis that the DM may be imprecise in specifying the necessary inputs: these are easily handled by relevant sensitivity analyses.

7. DESCRIPTIVE METHODS

All of the MCDM methods discussed thus far are at least partially normative in the sense that the aim is to provide some form of guidance or recommendation regarding a full or partial rank ordering of decision alternatives. In this sense, MCDM can be seen as having evolved from optimization theory. We have previously suggested [42] that an alternative view is that of MCDM as a problem of multivariate statistical analysis. We can view the set of decision alternatives, represented by the set of attribute vectors $\mathbf{z}^a$ for all $a \in A$, as an $n \times p$ matrix $\mathbf{Z}$, where each row is the attribute vector for one alternative, and n is the number of alternatives in A. The analysis is then directed towards an examination of the relationships between the attributes ('variables' in statistical terminology), so as to develop an understanding of what can realistically be achieved, and what are the constraints on performance imposed by the current decision set A. This understanding can assist the DM in formulating his goals realistically, and in *constructing* a desired solution to the decision problem.

A more-or-less direct application of this multivariate statistical approach is demonstrated in [42]. *Factor analysis* is used to identify linear combinations of the attributes which tend to vary together across the set of decision alternatives. Each such linear combination is used as an axis in a co-ordinate system, on which the alternatives can be plotted. In this way, a large proportion of the differences which exist between alternatives can be represented in a low dimensionality plot (where often as few as two dimensions suffice). By projecting the ideal on to the same plot, the points representing the alternatives can be seen as diverging from the ideal in a number of different directions, each direction representing a particular set of trade-offs which are actually offered by the alternatives in A. The specific plots produced by the factor analysis in [42] were at that time seen to be of limited value, in retrospect possibly due to the separation of positive and negative factor loadings into separate 'cost' and 'benefit' axes, and to the fact that the direct implications for each individual criterion were not directly displayed. The alternative suggested in [42] was to use *correspondence analysis*, in which a similar set of plots were obtained, but on which preference axes from best to worst were shown for each criterion, all on the same two dimensional plot of the alternatives.

Brans and Mareschal [11] and Mareschal and Brans [27], in an extension of their 'PROMETHEE' method (an outranking approach) termed 'GAIA', used *principal components* analysis in much the same way as above (recalling that factor analysis is in effect merely a rotation of the 'principal component' axes), but plotting both the alternatives, and the contributions of each attribute to each component, on the same set of axes. This yields a plot allowing similar interpretation to that of the 'correspondence analysis' plots proposed by Stewart [42]. It should be noted that Brans and Mareschal do not apply principal components to the Z matrix directly. They first establish by interaction with the DM an 'intensity of preference' measure between each pair of alternatives according to each criterion in turn. The pairwise intensities of preference between any one alternative and all others are consolidated into a single 'nett flow' representing the value of this alternative relative to the others, in terms of each criterion. The principal components analysis is applied to the matrix of these 'nett flows', rather than to the original z_i values. Simple experimentation reveals, however, that the plots produced by analysis of Z directly (without the need for the DM to specify intensities of preferences for various differences in attribute values) are nearly indistinguishable from those produced by Brans and Mareschal, cf. the example below.

Perhaps the multivariate statistical approach is better understood by means of a simple example. A useful example is that used by Brans and Mareschal [11], in which choice is to be made between alternative sites (different European countries) for an electric power plant. Six criteria have been identified as follows:

z_1: manpower required (minimize);
z_2: power generated in MW (maximize);
z_3: construction costs in 10^6 US dollars (minimize);
z_4: annual maintenance costs in 10^6 US dollars (minimize);
z_5: number of villages to be evacuated (minimize);
z_6: safety level on a nominal scale (maximize).

Table 1 summarizes the relevant data. The first two principal components explain close to 80% of the variation between the six alternatives. These two components are respectively:

$$-0.15z_1 - 0.79z_2 + 0.94z_3 - 0.58z_4 + 0.82z_5 - 0.80z_6$$

and

$$-0.93z_1 + 0.52z_2 + 0.25z_3 + 0.33z_4 - 0.15z_5 - 0.43z_6$$

where the signs of the z_i have been reversed in the case of the minimizing criteria (so that, according to our convention, increasing values are always preferred), and where the z_i have been standardized to a mean of 0 and a standard deviation of 1. By computing these component scores for each alternative, and plotting these in the plane defined by these components, we have effectively a projection of each alternative on to the plane of maximum variation between alternatives. This is done in Fig. 1, on which is also shown the projection of the 'ideal' and the directions of deviation from the mean corresponding to good performance on each criterion. (Note that Fig. 1 corresponds very closely to the figure on p. 241 of Brans and Mareschal [11], apart from a mirror image reflection and some rotation of the axes.) Certain observations can immediately be made from

Table 1

	z_1	z_2	z_3	z_4	z_5	z_6
Italy	80	90	600	54	8	5
Belgium	65	58	200	97	1	1
Germany	83	60	400	72	4	7
United Kingdom	40	80	1000	75	7	10
Portugal	52	72	600	20	3	8
France	94	96	700	36	5	6

Fig. 1. Firstly, good performances on criteria z_3 and z_5 tend to occur together and at the expense of the other criteria, except perhaps for z_1 which is somewhat orthogonal to the remaining criteria (cf. the horizontal axis in Fig. 1 and/or the positive coefficients for z_3 and z_5 in the above expressions). Similarly, z_2 and z_4 tend to perform together. Then we notice that the alternatives differ from the mean, and from the projection of the ideal in three main directions, with two alternatives lying in each direction; these represent the trade-off directions which are available. Finally we see that Portugal lies much closer to the projection of the ideal than does any other alternative, suggesting that it is a strong candidate unless there is one criterion that is really substantially more important than all others. One word of warning, however, is that the two-dimensional plot is only an approximate representation, and in particular a third axis or component may reveal features not evident with only two axes. Nevertheless, the descriptive approach can be a very valuable tool for understanding a complex MCDM problem.

Rivett [29–31] also suggested a form of graphical display based on multivariate statistical ideas. He did not, however, base this analysis on the $\boldsymbol{Z}$ matrix, or any variant thereof. In fact, he specifically allowed for consideration of cases in which the attribute values were not readily available explicitly. He rather set up an $n \times n$ matrix of dissimilarities between alternatives, based on holistic judgements by the DM as to which alternatives were most similar. The technique of *multidimensional scaling* is applied to this matrix in order to produce a two-dimensional map on which the distances between alternatives are maximally consistent with the dissimilarities. Rivett demonstrates that the major axis of this plot correlates well with preference orders obtained by more complex analyses, but without requiring the more difficult judgemental tasks of assessing strengths of preference in value function fitting for example. The Rivett approach does require the DM to carry out quite a large number of comparisons between alternatives, and although Clarke and Rivett [13] demonstrate that it is not necessary for the DM to compare all pairs of alternatives in this way, this can become a constraining factor as the number of alternatives increases.

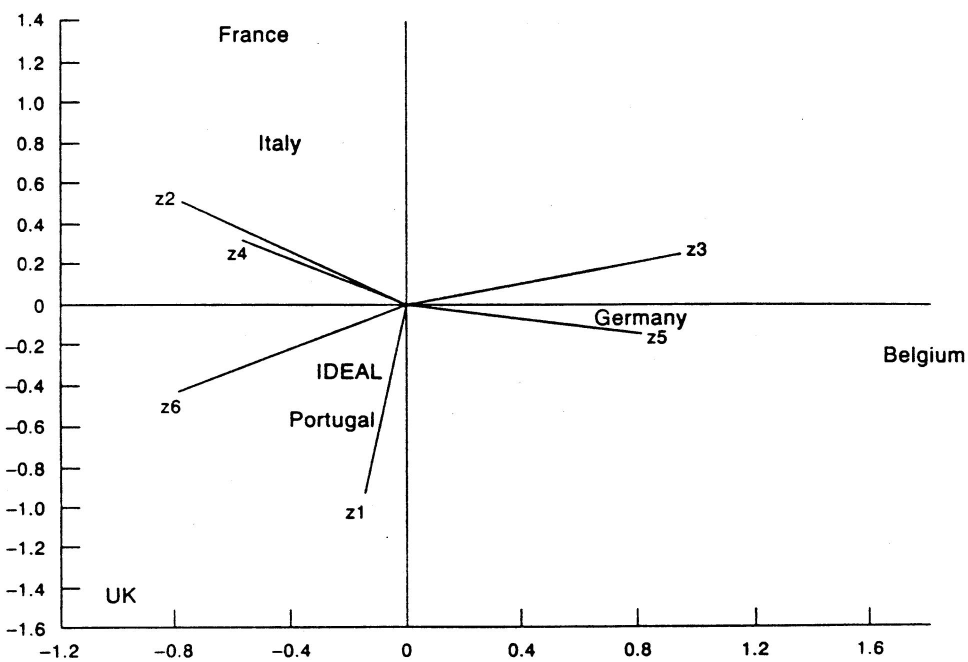

Fig. 1. Principal component plot for six alternatives and six criteria.

8. THE PROBLEM OF UNCERTAINTY

In this paper we have assumed that the consequences of a decision can be assessed deterministically, even if somewhat imprecisely. A much more difficult problem arises when the consequences themselves are stochastic. Keeney and Raiffa [21] do give considerable attention to this problem, and derive conditions under which a multi-attribute utility function can be found such that the DM's preferences between alternatives can be modelled in terms of maximization of expected utility. This approach is equivalent to combining to value function approach of Section 3 with von Neumann–Morgenstern expected utility theory. Certainly there are no good reasons to expect that expectation of a value function derived per Section 3 will be a particularly useful criterion for MCDM under uncertainty. Little additional results have been reported since the Keeney and Raiffa book, and in fact a lot of literature has appeared concerning weaknesses in the von Neumann–Morgenstern theory itself, as a predictive model of human behaviour, even in the case of a single criterion (cf. [51] for a review and discussion).

Where uncertainties can be represented in terms of a small number of exogenous scenarios, independent of actions being taken, the performance in terms of each deterministic criterion for each scenario can be viewed as a 'super-criterion' in its own right, but quite clearly the number of super-criteria can explode so quickly that anything more than (say) around three possible scenarios would be practically impossible to handle. The problem of uncertainty in MCDM must therefore remain a major open problem.

9. SOME PERSONAL CONCLUSIONS

Where now do we stand regarding solution aids for MCDM problems? For the deterministic case at least, we have a plethora of approaches, some (but not all) of which we have reviewed above. Some of these are *ad hoc*, and largely unjustified on theoretical and/or empirical grounds. In selecting an appropriate method to use, the following desiderata can be identified:

(i) The inputs required from the DM should be operationally meaningful and free from ambiguities of meaning.

(ii) The translation of these inputs into partial or complete recommendations should be consistent with the inputs used and with reasonable behavioural assumptions, and should be as far as possible transparent to the DM.

(iii) The method should be simple and efficient to use.

It may be evident that the present author's view is that there are a relatively small number of generally quite simple approaches which satisfy these desiderata. Within this view, we could conclude as follows:

(a) Goal programming in as fully an interactive mode as possible, perhaps relying on the reference point ideas of Wierzbicki [53], is a valuable means of understanding the structure of the problem at an early stage of analysis, when the number of potential decision alternatives is large or even infinite. This can be used to narrow the search for the 'best' solution quite considerably.

(b) Descriptive methods (Section 7) and/or outranking methods (Section 5) are useful once the number of alternatives under consideration has been reduced to a relatively small number (certainly less than about 15–20). These approaches assist in understanding and visualizing the key hard judgemental choices which have to be made, and can generate tentative partial orderings of the alternatives. Sometimes this on its own is sufficient to give the DM confidence in making a final choice directly.

(c) If the number of alternatives cannot be reduced far enough to allow use of descriptive or outranking methods, and/or if these methods do not generate a final result, and/or (very importantly) if the rationale for the final choice has to be defended in the public arena, then the value function approaches need to be used. In this case however, as we have stressed in Section

> 3, it is vital that the methodology used is such that the implicit assumptions made are both justifiable and easily understood by the DM and the public concerned.

For the future, the field of MCDM needs urgently to give attention to three issues at least, viz.

(1) The empirical validation and testing of the various approaches which are available (which all too often are justified by anecdotal 'success stories' which may reflect more the personality of the analysts than anything about the methodology!).

(2) The extension of MCDM decision aids into the group decision making situation, especially where there are considerable value-conflicts between group members.

(3) The treatment of uncertainty in MCDM, as discussed in Section 8.

REFERENCES

1. Ackoff RL (1979) The future of operational research is past. *J. Opl Res. Soc.* **30**, 93–104.
2. Bana e Costa CA (Ed.) (1990) *Readings in Multiple Criteria Decision Aid.* Springer, Berlin.
3. Bellman RE and Zadeh LA (1970) Decision making in a fuzzy environment. *Mgmt Sci.* **17**, B141–B164.
4. Belton V (1986) A comparison of the analytic hierarchy process and a simple multi-attribute value function. *Eur. J. Opl Res.* **26**, 7–21.
5. Belton V and Gear T (1982) On a shortcoming of Saaty's method of analytic hierarchies. *Omega* **11**, 228–230.
6. Belton V and Gear T (1985) The legitimacy of rank reversal—A comment. *Omega* **13**, 143–144.
7. Belton V and Vickers S (1990) Use of a simple multi-attribute value function incorporating visual interactive sensitivity analysis for multiple criteria decision making. [2], pp. 319–334.
8. Benayoun R, de Montgolfier J, Tergny J and Laritchev O (1971) Linear programming with multiple objective functions: Step method (STEM). *Math. Prog.* **1**, 366–375.
9. Bogetoft P and Pruzan P (1991) *Planning with Multiple Criteria.* Elsevier–North Holland, Amsterdam.
10. Bouyssou D (1990) Building criteria: A prerequisite for MCDA. [2], pp. 58–80.
11. Brans JP and Mareschal B (1990) The PROMETHEE methods for MCDM; the PROMCALC, GAIA and BANKADVISER software. [2], pp. 216–252.
12. Charnes A and Cooper WW (1961) *Management Models and Industrial Applications of Linear Programming.* Wiley, New York.
13. Clarke D and Rivett BHP (1978) A structural mapping approach to complex decision-making. *J. Opl Res. Soc.* **29**, 113–128.
14. Dyer JS (1990) Remarks on the analytic hierarchy process. *Mgmt Sci.* **36**, 249–258.
15. Geoffrion AM, Dyer JS and Feinberg A (1972) An interactive approach for multicriterion optimization with an application to the operation of an academic department. *Mgmt Sci.* **19**, 357–368.
16. Goicoechea A, Hansen DR and Duckstein L (1982) *Multiobjective Decision Analysis with Engineering and Business Applications.* Wiley, New York.
17. Hannan El (1985) An assessment of some criticisms of goal programming. *Computs Opns Res.* **12**, 525–541.
18. Hwang CL, Paidy SR, Yoon K and Masud ASM (1980) Mathematical programming with multiple objectives: A tutorial. *Computs Opns Res.* **7**, 5–31.
19. Ignizio JP (1983) Generalized goal programming: An overview. *Computs Opns Res.* **10**, 277–289.
20. Islei G and Lockett AG (1988) Judgemental modelling based on geometric least square. *Eur. J. Opl Res.* **36**, 27–35.
21. Keeney RL and Raiffa H (1976) *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* Wiley, New York.
22. Korhonen P, Wallenius J and Zionts S (1984) Solving the discrete multiple criteria problem using convex cones. *Mgmt Sci.* **30**, 1336–1345.
23. Levine P and Pomerol JC (1986) PRIAM, an interactive program for choosing among multiple attribute alternatives. *Eur. J. Opl Res.* **25**, 272–280.
24. Lootsma FA (1990) The French and the American school in multi-criteria decision analysis. *Recherche opérationelle/Opns Res.* **24**, 263–285.
25. Lootsma FA (1991) Scale sensitivity and rank preservation in a multiplicative variant of the AHP and SMART. Report 91-67, Faculty of Technical Mathematics and Informatics, TU Delft.
26. Lotfi V, Stewart TJ and Zionts S (1992) An aspiration-level interactive model for multiple criteria decision making. *Computs Opns Res.* **19**(7), 671–681.
27. Mareschal B and Brans JP (1988) Geometrical representations for MCDA. *Eur. J. Opl Res.* **34**, 69–77.
28. Masud AS and Hwang CL (1981) Interactive sequential goal programming. *J. Opl Res. Soc.* **32**, 391–400.
29. Rivett BHP (1977) Policy selection by structural mapping. *Proc. R. Soc. London A* **354**, 407–423.
30. Rivett BHP (1977) Multidimension scaling for multiobjective policies. *Omega* **5**, 367–379.
31. Rivett BHP (1978) Structural mapping applied to single value policies. *Omega* **6**, 407–417.
32. Romero C (1986) A survey of generalized goal programming (1970–1982). *Eur. J. Opl Res.* **25**, 183–191.
33. Roy B (1990) The outranking approach and the foundations of ELECTRE methods. [2], pp. 155–183.
34. Saaty TL (1980) *The Analytic Hierarchy Process.* McGraw-Hill, New York.
35. Saaty TL (1990) An exposition of the AHP in reply to the paper "Remarks on the analytic hierarchy process". *Mgmt Sci.* **36**, 259–268.
36. Saaty TL (1990) How to make a decision: The analytic hierarchy process. *Eur. J. Opl Res.* **48**, 9–26.
37. Saaty TL and Vargas LG (1984) The legitimacy of rank reversal. *Omega* **12**, 513–516.
38. Schoner B and Wedley WC (1989) Ambiguous criteria weights in AHP: Consequences and solutions. *Decis. Sci.* **20**, 462–475.
39. Simon HA (1976) *Administrative Behaviour*, 3rd edn. The Free Press, New York.

40. Spronk J (1981) *Interactive Multiple Goal Programming: Applications to Financial Planning*. Martinus Nijhoff Publishing, Boston.
41. Steuer RE (1986) *Multiple Criteria Optimization: Theory, Computation and Application*. Wiley, New York.
42. Stewart TJ (1981) A descriptive approach to multiple criteria decision making. *J. Opl Res. Soc.* **32**, 45–53.
43. Stewart TJ (1987) Pruning of decision alternatives in multiple criteria decision making, based on the UTA method for estimating utilities. *Eur. J. Opl Res.* **28**, 79–88.
44. Stewart TJ (1987) An interactive multiple objective linear programming method based on piecewise linear additive value functions. *IEEE Trans. Sys. Man Cyber.* **SMC-17**, 799–805.
45. Stewart TJ (1988) Experience with prototype multicriteria decision support systems for pelagic fish quota determination. *Naval Res. Logist.* **35**, 719–731.
46. Stewart TJ (1991) A multi-criteria decision support system for R&D project selection. *J. Opl Res. Soc.* **42**, 17–26.
47. Tapia CG and Murtagh BA (1991) Interactive fuzzy programming with preference criteria in multiobjective decision-making. *Computs Opns Res.* **18**, 307–316.
48. Vargas LG (1990) An overview of the analytic hierarchy process and its applications. *Eur. J. Opl Res.* **48**, 2–8.
49. Vetschera R (1988) An interactive outranking system for multi-attribute decision making. *Computs Opns Res.* **15**, 311–322.
50. Von Winterfeldt D and Edwards W (1986) *Decision Analysis and Behavioral Research*. Cambridge University Press.
51. Weber M and Camerer C (1987) Recent developments in modelling preferences under risk. *OR Spektrum* **9**, 129–151.
52. Weber M, Eisenführ F and von Winterfeldt D (1988) The effects of splitting attributes on weights in multiattribute utility measurement. *Mgmt Sci.* **34**, 431–445.
53. Wierzbicki AP (1980) The use of reference objectives in multiobjective optimization. In *Multiple Criteria Decision Making Theory and Application* (Edited by G Fandel and T Gal). Springer, Berlin.
54. Zanakis SH and Gupta SK (1985) A categorized survey of goal programming. *Omega* **13**, 211–222.
55. Zeleny M (1982) *Multiple Criteria Decision Making*. McGraw–Hill, New York.
56. Zimmermann HJ (1978) Fuzzy programming and linear programming with several objective functions. *Fuzzy Sets Syst.* **1**, 45–55.
57. Zimmermann HJ (1983) Using fuzzy sets in operational research. *Eur. J. Opl Res.* **13**, 201–216.
58. Zimmermann HJ and Zysno P (1979) Latent connectives in human decision making. *Fuzzy Sets Sys.* **2**, 37–51.
59. Zionts S and Wallenius (1976) An interactive programming method for solving the multiple criteria problem. *Mgmt Sci.* **22**, 652–663.
60. Zionts S and Wallenius (1983) An interactive multiple objective linear programming method for a class of underlying nonlinear utility functions. *Mgmt Sci.* **29**, 519–529.

ADDRESS FOR CORRESPONDENCE: *Professor TJ Stewart, Department of Statistical Sciences, University of Cape Town, Private Bag, Rondebosch 7700, South Africa.*