Chapter 7
Searching for the "Technical Compromise Solution": Solving the Discrete Multi-Criterion Problem in an SMCE Framework

7.1 Pair-Wise Comparison of Alternatives

Given a set of evaluation criteria $G = \{g_m\}$, $m = 1, 2, \ldots, M$, and a finite set $A = \{a_n\}$, $n = 1, 2, \ldots, N$ of potential alternatives (actions), let us start with the simple assumption that the performance (i.e. the criterion score) of an alternative a_n with respect to a judgement criterion g_m is based on an *interval or ratio* scale of measurement. For simplicity's sake, it is assumed that a higher value of a criterion is preferred to a lower one (i.e. the higher, the better). The pair-wise comparison of alternatives proposed here is a preference modelling structure based on the so-called threshold model and fuzzy preference relations.

As shown in Chap. 4, by introducing a positive indifference threshold q the resulting preference model is the *threshold model:*

$$\left\{\begin{array}{lcl} a_j P a_k & \Leftrightarrow & g_m(a_j) > g_m(a_k) + q \\ a_j I a_k & \Leftrightarrow & \left| g_m(a_j) - g_m(a_k) \right| \le q \end{array}\right\} \tag{7.1}$$

where a_j and a_k belong to the set A of alternatives and g_m to the set G of evaluation criteria.

The *double threshold model* is a preference relation where indifference and preference thresholds have been introduced, i.e.:

$$\left\{\begin{array}{lcl} a_j P a_k & \Leftrightarrow & g_m(a_j) > g_m(a_k) + p\big(g_m(a_k)\big) \\ a_j Q a_k & \Leftrightarrow & g_m(a_k) + p\big(g_m(a_k)\big) \ge g_m(a_j) > g_m(a_k) + q\big(g_m(a_k)\big) \\ a_j I a_k & \Leftrightarrow & \left\{\begin{array}{l} g_m(a_k) + q\big(g_m(a_k)\big) \ge g_m(a_j) \\ g_m(a_j) + q\big(g_m(a_j)\big) \ge g_m(a_k) \end{array}\right. \end{array}\right\} \tag{7.2}$$

for any $m = 1, 2, \ldots, M$, p being a positive preference threshold.

G. Munda, *Social Multi-Criteria Evaluation for a Sustainable Economy*.

A *pseudo-order structure* is a double threshold model upon which the following consistency condition is imposed

$$g_m(a_j) > g_m(a_k) \Leftrightarrow \begin{cases} g_m(a_j) + q\big(g_m(a_k)\big) > g_m(a_k) + q\big(g_m(a_k)\big) \\ g_m(a_j) + p\big(g_m(a_k)\big) > g_m(a_k) + p\big(g_m(a_k)\big) \end{cases} \quad (7.3)$$

A problem is that the modelling procedure based on the notion of a pseudo-criterion may display a serious lack of stability. Such undesirable discontinuities make a sensitivity analysis (or robustness analysis) necessary; however, this important analysis step is very complex in its execution because of the combinatorial nature of the various sets of data (Saltelli et al., 2004). One should combine variations of two thresholds (indifference and preference) and k possible scores of the M criteria. A solution to this problem may be found in the concept of *valued preference relations*, that is a preference relation in which it is necessary to assign to each ordered pair of alternatives (a_j, a_k) a value $v(a_j, a_k)$ representing the "strength" or the "degree of preference" (Fishburn, 1970, 1973a; Roubens and Vincke, 1985; Ozturk et al., 2005).

In this framework, an interesting concept is the one of a *fuzzy preference relation* (Kacprzyk and Roubens, 1988). If A is assumed to be a finite set of N alternatives, a *fuzzy preference relation* is an element of the $N\times N$ matrix $R = (r_{jk})$, i.e.:

$$r_{jk} = \mu_R(a_j, a_k) \qquad \text{with } j,k = 1,2,\ldots,N \text{ and } 0 \le r_{jk} \le 1 \quad (7.4)$$

$r_{jk} = 1$ indicates the maximum credibility degree of the preference of a_j over a_k; each value of r_{jk} in the open interval (0.5, 1) indicates a definite preference of a_j to a_k (a higher value means a stronger credibility); $r_{jk} = 0.5$ indicates the indifference between a_j and a_k. This definition implies that fuzzy preference relations can be used as mathematical models of intensity of preference.
Usually, fuzzy preference relations are assumed to satisfy two properties:

(1) Reciprocity, i.e. $r_{jk}+r_{kj} = 1$
(2) Max–min transitivity, i.e. if a_i is preferred to a_j and a_j is preferred to a_k, then a_i should be preferred to a_k with at least the same credibility degree, i.e.:

$$r_{ij} \ge 0.5, \quad r_{jk} \ge 0.5 \Rightarrow r_{ik} \ge \min\ (r_{ij}, r_{jk}) \quad (7.5)$$

Since small variations of input data (scores and thresholds) are modelled by means of a continuous membership function, by using fuzzy preference modelling as developed in (7.6), the combinatorial drawbacks of the pseudo-criterion model can be avoided.

Let us now consider any criterion g_m belonging to the set G and any pair of alternatives a_j and a_k belonging to the set A. The criterion scores $g_m(a_j)$ and $g_m(a_k)$ are measured on an interval or ratio scale. Let p_m be a constant preference threshold and q_m a constant indifference threshold for the criterion g_m. Then the credibility degree μ of preference (P) and indifference (I) relations between a_j and a_k can be computed as follows:

$$\begin{cases} \mu(a_j\, P\, a_k) = \left[1 + c_{pm} \left(g_m(a_j) - g_m(a_k)\right)^{-2}\right]^{-1} \\ \mu(a_j\, I\, a_k) = e^{-c_{qm}\left|g_m(a_j) - g_m(a_k)\right|} \\ \mu(a_k\, P\, a_j) = \left[1 + c_{pm} \left(g_m(a_k) - g_m(a_j)\right)^{-2}\right]^{-1} \end{cases} \tag{7.6}$$

where $\mu(a_j\, I\, a_k)\ \forall g_m(a_j)$ *and* $g_m\,(a_k)$ and

$$\mu(a_j\, P\, a_k) \qquad \textit{if} \quad g_m(a_j) - g_m(a_k) > 0 \tag{7.7}$$

$$\mu(a_k\, P\, a_j) \qquad \textit{if} \quad g_m(a_j) - g_m(a_k) < 0 \tag{7.8}$$

In (7.6) the parameters (c_{pm}) and (c_{qm}) are derived in function of the cross-over point, i.e. the value of the difference between two criterion scores where the credibility degree of the corresponding indifference/preference relation is equal to 0.5; see Figs. 7.1 and 7.2 for an example[1] (in these figures in the *y*-axis the credibility degrees and in the *x*-axis the thresholds are represented respectively). The relations $\mu(a_j P\ a_k)$and $\mu(a_k P\ a_j)$ are derived from values satisfying the condition of additive transitivity, thus it is trivial to prove that all these relations are max-min transitive (Kacprzyk and Roubens, 1988); however the property of reciprocity does not hold, thus these are not fuzzy preference relations in a strict sense. The relation $\mu(a_j I\ a_k)$ is a *resemblance relation*, which is reflexive and symmetrical but no transitivity is implied (thus the Luce paradox cannot occur).

It has to be admitted that the shape of the function representing the credibility degrees of the preference and indifference relations is arbitrary. However, there do exist some consistency requirements e.g. that the functions be continuous and monotonic and that $p_m > q_m$ thereby reducing considerably the degree of arbitrariness.

---

[1] Algebraically the parameters are the solution of this equation

$$\begin{cases} 0.5 = \left[1 + c_{pm} \left(p_m\right)^{-2}\right]^{-1} \\ 0.5 = e^{-c_{qm}\left|q_m\right|} \\ 0.5 = \left[1 + c_{pm} \left(p_m\right)^{-2}\right]^{-1} \end{cases}$$

where p_m and q_m are the preference and indifference thresholds respectively.

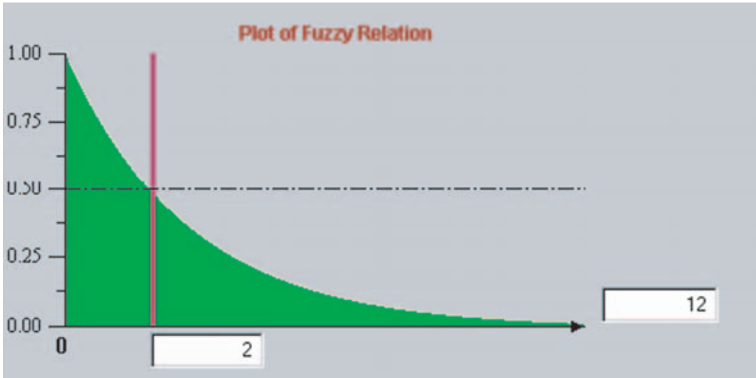

Fig. 7.1 Example of credibility degrees of a fuzzy indifference relation

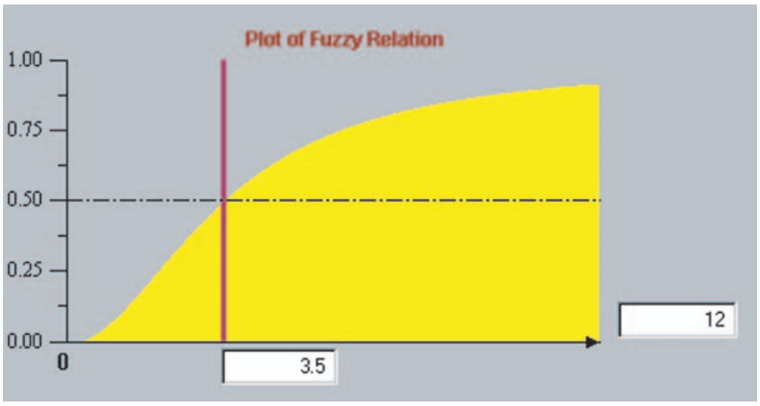

Fig. 7.2 Example of credibility degrees of a fuzzy preference relation

7.2 Extensions: The Case of Mixed Information on Criterion Scores

Ideally the information available for a policy problem should be precise, certain, exhaustive and unequivocal. But in real-world situations, it is often necessary to use information which lacks these characteristics and thus to deal with uncertainty of a stochastic and/or fuzzy nature in the data. Let us then introduce a more realistic assumption, i.e. that the set of evaluation criteria $G = \{g_m\}$, $m = 1, 2,\ldots, M$, on the set $A = \{a_n\}$, $n = 1, 2,\ldots, N$ of potential alternatives may include either crisp (i.e. impacts measured on interval or ratio scales), stochastic and fuzzy criterion scores.

The treatment of mixed information on criterion scores proposed here is mainly based on the semantic distance I developed some years ago (Munda, 1995, Chap. 6). This because this semantic distance allows us to deal consistently with an impact (or evaluation) matrix which may include crisp, stochastic or fuzzy measurements of the performance of an alternative with respect to an evaluation criterion. As a consequence the multi-criterion problem is considered in its more general form (the next section will show that ordinal criterion scores can also be considered). The only restriction is in the case of fuzzy information, when continuous, convex membership functions allowing for a definite integration are required.

Let us start with the case of fuzzy criterion scores (to complete the axiomatic system in Appendix 7.1 it is proved that this distance satisfies the property of triangle inequality):

if $\mu_1(x)$ and $\mu_2(x)$ are two fuzzy numbers, one can write (see Ragade and Gupta, 1977 for a formal proof):

$$f(x) = k_1\ \mu_1(x) \quad \text{and} \quad g(y) = k_2\ \mu_2(x) \tag{7.10}$$

where $f(x)$ and $g(y)$ are two functions obtained by rescaling the ordinates of $\mu_1(x)$ and
$\mu_2(x)$ through k_1 and k_2, such that

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty} g(y)dy = 1 \tag{7.11}$$

The distance between all points of the membership functions is computed as follows:

If $f(x)$ is defined on $X = [x_L, x_U]$ and $g(y)$ is defined on $Y = [y_L, y_u]$
where sets X and Y can be non-bounded from one or either sides, then

$$S_d\big(f(x),\ g(y)\big) = \iint_{x,\,y} |x-y| f(x)g(y)dydx \tag{7.12}$$

If the intersection between the two membership functions is empty, it is $x > y$ $\forall x \in X$ *and* $\forall y \in Y$, it follows that a continuous function in two variables is defined over a rectangle. Therefore the double integral can be calculated as iterated single integrals; the result is

$$S_d\big(f(x),\ g(y)\big) = |E(x) - E(y)| \tag{7.13}$$

where $E(x)$ and $E(y)$ are the expected values of the two membership functions.

When the intersection between two fuzzy sets is not empty, their distance is greater than the difference between the respective expected values, since $|x - y|$ is always greater than $(x - y)$. In this case one finds:

$$S_d\big(f(x),\ g(y)\big) = \int_{-\infty}^{+\infty}\int_{x}^{+\infty} (y-x) f(x)g(y)dydx \ + \ + \int_{-\infty}^{+\infty}\int_{-\infty}^{x} (x-y) f(x)g(y)dydx \tag{7.14}$$

This is the case of a double integral over a general region; since this is not vertically or horizontally simple, its computation is not possible by means of iterated integration; it is necessary to take the limit of the Rieman sum. This problem can easily be overcome by means of numerical analysis (in Munda, 1995 a Monte Carlo type numerical algorithm for the computation of this distance was developed. This is presented in Appendix 7.2).

As an example, we will compute the semantic distance between a symmetrical and a LR fuzzy number. Let us assume:

$$\tilde{A} = e^{-0.077(x-143)^2} \qquad \text{and}$$

$$\tilde{B} = \begin{cases} 1 - e^{-1500\left(\frac{x-140}{140}\right)^2} & if\ 140 \le x < 147 \\ 1 & if\ x = 147 \\ \left[e^{-470\left(\frac{x-147}{147}\right)^2} \right]^2 & if\ 147 < x < +\infty \end{cases}$$

Then we have

$$S_d(\tilde{A}, \tilde{B}) = \int_{147}^{+\infty} \int_{-\infty}^{+\infty} |x-y| \frac{\left[e^{-470\left(\frac{x-147}{147}\right)^2} \right]^2}{7.965} \bullet \frac{e^{-0.077(y-143)^2}}{6.349} dydx +$$

$$+ \int_{140}^{147} \int_{-\infty}^{+\infty} |x-y| \frac{1 - e^{-1500\left(\frac{x-140}{140}\right)^2}}{7.965} \bullet \frac{e^{-0.077(y-143)^2}}{6.349} dydx \cong 5$$

while the difference between their expected values is about 4.368.

From a theoretical point of view, the following main conclusions can be drawn:

(1) The absolute value metric is a particular case of the semantic distance
(2) The comparison between a fuzzy number and a crisp number is equal to the difference between the expected value of the fuzzy number and the value of the crisp number
(3) Stochastic information can also be taken into account

In sum the semantic distance allows us to deal with fuzzy numbers, probability distributions and crisp numbers with the theoretical guarantee that all these sources of information are tackled equivalently, thus solving an open problem for multi-criteria methods dealing with mixed information. Of course, this search for an equivalent treatment of available information implies a trade-off with precision. For example, if stochastic information only is available, a stochastic dominance approach is more effective (see e.g. Markowitz, 1989; Martel and Zaras, 1995), or if fuzzy numbers only have to be compared, Matarazzo and Munda (2001) present a more sophisticated approach based on area comparison. However, in the case of mixed information in a multi-criteria framework, the semantic distance illustrated here is probably the best available compromise solution between generality and precision. Moreover, the use of this semantic distance allows a homogeneous preference modelling on all the criteria, otherwise impossible; this can be illustrated as follows.

Going back to the pair-wise comparison of alternatives, let us assume $f(x) = g_m(a_j)$ and $g(y) = g_m(a_k)$, where g_m is any criterion belonging to the set G and a_j and a_k any pair of alternatives belonging to the set A. The criterion scores $g_m(a_j)$ and $g_m(a_k)$ are *fuzzy or stochastic* in nature. Let p_m be a preference threshold and q_m an indifference threshold for the criterion g_m. Then we have:

$$\begin{cases} \mu(a_j\, P\, a_k) = \left[1 + c_{pm}\left(\iint\limits_{x,\,y} (x - y)\; f(x)g(y)\; dydx\right)^{-2}\right]^{-1} \\[2ex] \mu(a_j\, I\, a_k) = e^{-c_{qm}\left(\iint\limits_{x,\,y} |x - y|\, f(x)g(y)\, dydx\right)} \\[2ex] \mu(a_k\, P\, a_j) = \left[1 + c_{pm}\left(\iint\limits_{y,\,x} (y - x)\; f(x)g(y)\; dydx\right)^{-2}\right]^{-1} \end{cases} \tag{7.15}$$

where

$$\mu(a_j\, I\, a_k) \quad \forall \quad x, y \qquad \text{and}$$

$$\mu(a_j\, P\, a_k) \;\; if \quad \iint\limits_{x,\,y} (x - y) f(x)g(y)\, dydx > 0 \tag{7.16}$$

$$\mu(a_k P\, a_j) \;\; if \quad \iint\limits_{x,\,y} (x - y) f(x)g(y)\, dydx < 0 \tag{7.17}$$

One should note that the comparison between the criterion scores of each pair of actions is carried out by means of the semantic distance. Since the absolute value metric is a particular case of this distance, fuzzy, stochastic and crisp criterion scores are dealt with equivalently.

7.3 Extensions: Introducing Weights as Importance Coefficients

At this point, a very delicate step has still to be tackled, i.e. the exploitation of the inter-criteria information in the form of weights. Let us then assume the existence of a set of criterion weights $W = \{w_m\}$, $m = 1, 2, \ldots, M$, with $\sum_{m=1}^{M} w_m = 1$ derived as importance coefficients. The problem is the theoretical guarantee that weights are really treated as importance coefficients and not as trade-offs. The point is that no connection can be made between criterion weights and the corresponding criterion intensity of preference. Our objectives are then:

(1) To find a way to combine weights with credibility degrees without a direct interpretation of the latter as intensity of preference
(2) To divide each criterion weight into two parts proportionally to the credibility degrees of the indifference and preference fuzzy relations. In doing so, the requirement that $\sum_{m=1}^{M} w_m = 1$ should not be lost.

Let us define μ_p as the fuzzy preference relation between a pair of alternatives and μ_I as the fuzzy indifference relation between the same pair. Let us put $\mu_{min} = \min(\mu_p, \mu_I)$ and $\mu_{max} = \max(\mu_p, \mu_I)$.

Clearly, it is $\mu_p = \mu_{min}$ on the left of the intersection point between the indifference and the preference fuzzy relations and vice versa on the right. I propose that a criterion weight w_m be divided proportionally to μ_p and μ_I according to (7.18).

$$\begin{cases} w_{m1} = w_m \dfrac{\mu_{\min}}{\mu_{\max} + \mu_{\min}} \\ \\ w_{m2} = w_m \dfrac{\mu_{\max}}{\mu_{\max} + \mu_{\min}} \end{cases} \tag{7.18}$$

(7.18) presents the following properties:

$$w_{m1} + w_{m2} = w_m \tag{7.19}$$

$$if\ \mu_{\min} = 0 \Rightarrow w_{m2} = w_m \tag{7.20}$$

$$if\ \mu_{\min} = \mu_{\max} = 0 \Rightarrow w_m = 0 \tag{7.21}$$

$$if\ \mu_{\min} = \mu_{\max} \Rightarrow w_{m1} = w_{m2} = \frac{1}{2} w_m \tag{7.22}$$

As a consequence (7.18) meets our objective of keeping the sum of weights perfectly equal to one. Moreover, in (7.18) no direct use of the concept of intensity of preference is made; as a result we can be sure that criterion weights are being used consistently with their nature as importance coefficients. Finally if a criterion score is *ordinal in nature*, it can be considered a particular case where $\mu_{\min} = 0$. Again *the treatment of crisp, fuzzy, stochastic and ordinal criterion scores is perfectly* *equivalent*. Moreover, when indifference and preference thresholds are not used, the corresponding criteria can be dealt with as ordinal criteria[2], where

$$\begin{cases} a_j P a_k \Leftrightarrow g_m(a_j) > g_m(a_k) \\ a_j I a_k \Leftrightarrow g_m(a_j) > g_m(a_k) \end{cases} \tag{7.23}$$

Now an $N \times N$ matrix E can be built, where any generic element e_{jk} with $j \neq k$ is the result of the pair-wise comparison between alternatives j and k according to all the M criteria. Such a global pair-wise comparison is obtained by means of (7.24):

$$e_{jk} = \sum_{m=1}^{M} \left(w_m(P_{jk}) + \frac{1}{2} w_m(I_{jk}) \right) \tag{7.24}$$

where $w_m\ (p_{jk})$ and $w_m\ (I_{jk})$ are derived from μ_P and μ_I through (7.18). It is

$$e_{jk} + e_{kj} = 1 \tag{7.25}$$

Property 7.25 is very important since it allows us to consider matrix E as a *voting matrix* i.e. a matrix where instead of using criteria, alternatives are compared by means of voters' preferences (on the principle of one agent, one vote). This analogy between the multi-criterion and the social choice problem, as noted by Arrow and Raynaud (1986), is very useful for tackling the step of ranking the N alternatives in a consistent axiomatic framework: a Condorcet consistent rule can now be used to exploit the pair-wise comparisons to order alternatives.

7.4 Ranking of Alternatives in a Complete Pre-Order

The issue here is whether it is possible to find a ranking algorithm consistent with the desirable properties of social multi-criteria evaluation. And conversely, given the results of Arrow's impossibility theorem (Arrow, 1963), whether it is possible to ensure that no essential property is lost. Both social choice literature and multi-criteria decision theory agree that whenever the majority rule can be operationalized,

---

[2] If criterion scores are used with an ordinal meaning only, as we saw in Chap. 6, the following axiomatic conditions must be added (adapted from Arrow and Raynaud, 1986, p. 81–82).

Axiom 1: Diversity. Each criterion is a total order on the finite set A of alternatives to be ranked, and there is no restriction on criteria; they can be any total order on A.

Axiom 2: Symmetry. Since criteria have incommensurable scales, the only preference information they provide is the ordinal pair-wise preferences they contain.

Axiom 3: Positive Responsiveness. The degree of preference between two alternatives ***a*** and ***b*** is a strictly increasing function of the number and weights of criteria that rank ***a*** before ***b***.

it should be applied. However, majority rule often produces undesirable intransitivities, thus "more limited ambitions are compulsory. The next highest ambition for an aggregation algorithm is to be Condorcet" (Arrow and Raynaud, 1986, p. 77). As we have discussed in Chap. 6, in the framework of SMCE the C–K–Y–L ranking procedure seems the most appropriate.

According to this ranking procedure, the maximum likelihood ranking of alternatives, in a social multi-criterion framework, is that ranking supported by the maximum number of criteria for each pair-wise comparison, summed over all pairs of alternatives. More formally, the C–K–Y–L ranking procedure can be adapted to a multi-criteria framework as follows.

All the $N(N-1)$ pair-wise comparisons compose the matrix E, in which we remember that $e_{jk} + e_{kj} = 1$, with $j \neq k$. Let us call R the set of all the $N!$ possible complete rankings of alternatives: $R = \{r_s\}$, $s = 1, 2,\ldots, N!$. For each r_s, let us compute the corresponding score φ_s as the sum of e_{jk} over all the $\binom{N}{2}$ pairs jk of alternatives, i.e.

$$\varphi_s = \sum e_{jk} \tag{7.26}$$

where $j \neq k$, $s = 1, 2, \ldots N!$ and $e_{jk} \in r_s$

The final ranking (r^*) is that[3] which maximizes (7.26), which is:

$$r_* \Leftrightarrow \varphi_* = \max \sum e_{jk} \qquad \text{where } e_{jk} \in R \tag{7.27}$$

A final issue to be discussed is the matter of ties, i.e. the case that in some individual profiles alternatives can be ranked in the same position. This does not constitute a problem since such an event can easily be taken into account in the concordance index used for the construction of an outranking matrix (see (7.24)).

However, if some ties occur in the outranking matrix E, this might sometimes create a problem for the interpretation of the final results. In this case, one has to choose a tie-breaking rule, thus neutrality is necessarily lost. The question is then: which is the probability of finding ties in the outranking matrix E? Proposition 7.1 states that this probability is approximately zero; as a consequence ties in the outranking matrix are not a serious problem. The proof of this proposition can be found in Appendix 7.1.

Proposition 7.1 *In the outranking matrix E the event of obtaining ties, that is,* $e_{jk} = e_{kj} = \frac{M}{2}$, *is possible but its probability is approximately zero.*

[3] It is important to remember that sometimes the final ranking is not unique. This is a desirable property since it can be considered a measure of the *robustness* of the results provided.

7.5 Introducing the Minority Principle: A Borda Approach

At this point, we refer to the normative tradition in political philosophy, which also has an influence in modern social choice (Moulin, 1981) and public policy (Mueller, 1978). The fundamental idea is that any coalition controlling more than 50% of the votes may be converted into an actual dictator. As a consequence, the "remedy to the tyranny of the majority is the minority principle, requiring that all coalitions, however small, should be given some fraction of the decision power. One measure of this power is the ability to veto certain subsets of outcomes.… "(Moulin, 1988, p. 272). The introduction of a veto power in a multi-criteria framework can be further justified in the light of the so-called "prudence" axiom (Arrow and Raynaud, 1986, p. 95), whose principle is that it is not prudent to accept alternatives whose degree of conflictuality is too high (and thus might make the final decision very vulnerable[4]). The point is then how to implement this idea of veto power in a multi-criteria framework.

Historically, the first attempt was made by Roy (1985, 1996) in the so-called ELECTRE methods. Basically, Roy proposed that for any pair of alternatives one should look at the majority principle expressed as a concordance index and to the minority principle in the form of the discordance index. The discordance index is calculated according to the intensity of preference any single criterion has against the concordance coalition. This means that for each single criterion a veto threshold must be defined.

In my opinion, the implementation of veto power in an SMCE framework presupposes three desirable characteristics:

1. To be independent of arbitrary ad hoc thresholds.
2. To consider the global opposition to the final ranking and not to a pair of alternatives (Roy's approach), or any specific possible ranking (Paelinck, 1978).
3. No specific intensity of preference should be considered (if a weight is combined with a veto threshold for each criterion, the resulting concept of criterion importance also depends on the intensity of preference; this means that weights probably can no longer be considered importance coefficients).

It is interesting to note that an approach meeting these requirements can again be found in classical social choice theory, in particular, in the Borda approach. The Borda rule is normally used to find a Borda winner, where the winner is the alternative which receives the highest score in favour (an alternative receives $N-1$ points if it ranks first, and so on until *0* score if it ranks last on a given criterion). In the same way, a *Borda loser* can be defined as the alternative which receives the highest score against (where $N-1$ points are assigned to the last alternative in the ranking and so on until 0 points are given to the option which ranks first).

[4] It should be noted that mitigating the vulnerability of the C–K–Y–L ranking procedure is very important since this is one of the main criticisms of the method.

Formally, the procedure I am proposing can be described as follows by taking inspiration from the concept of frequency matrices (Hinloopen et al., 1983; Matarazzo, 1988). Let us call F the matrix where any element f_{ij} means that a given criterion g_m scores alternative a_j in the i-th ordinal position. Now it is possible to define the $N \times N$ matrix Φ where any element ϕ_{ij} represents the summation of the weights of criteria which score alternative j at the i-th position; that is

$$\phi_{ij} = \sum_{m \in G_i} w_m \tag{7.28}$$

$$where \;\; G_i = \left\{ g_m : g_m(a_j) = f_{ij} \right\} \quad with \quad G_i \subset G \tag{7.29}$$

$i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, N$
It is obviously:

$$\sum_{i=1}^{N} \phi_{ij} = 1 \quad \forall \; a_j \in A \quad \text{and} \tag{7.30}$$

$$\sum_{j=1}^{N} \phi_{ij} = 1 \quad with \;\; j = 1, 2, \ldots, N \tag{7.31}$$

Now for any alternative a_j let us apply the Borda rule in search for the Borda loser, i.e.

$$B(a_j) = \sum_{i=1}^{N} (\phi_{ij} \times b_i) \tag{7.32}$$

$$where \quad b_i = N-1, \; N-2, \ldots, 0 \quad with \; i = N, \; N-1, \ldots, 1$$

The vetoed alternative $\overline{a}_j$ is the Borda loser, i.e. the a_j for which $B(a_j) = \max$.

One should note that by means of this procedure weights are never combined with intensities of preference and no *ad hoc* parameter is needed. Consistently with the Borda approach, only one alternative, considered the one with the greatest opposition, is selected to be vetoed. It must be remembered that the Borda procedure respects all the properties of the C–K–Y–L, except local stability. This is the main reason why Borda consistent rules are more appropriate for the selection of one alternative only and not for the generation of rankings.

Finally a question to be answered is: do Borda and Condorcet rules normally lead to different solutions? One might in fact believe that the divergence of solutions is a very special case and thus the value added of introducing the Borda loser is very limited. As we have seen in Chap. 6, this question can be answered very easily. Fishburn (1973b) proves the following theorem: *there are profiles where the Condorcet winner exists and it is never selected by any scoring method*. Moulin (1988, p. 249) proves that "a Condorcet winner (loser) cannot be a Borda loser (winner)". In other words, Condorcet consistent rules and scoring voting rules are deeply different in nature.

7.6 Numerical Examples

Let us consider an evaluation problem concerning three types of publicly provided goods (A, B, C). Let us assume that it has been agreed that these goods have to be evaluated by taking into account three dimensions, i.e. economic, social and environmental, and that each dimension has the same weight.

These dimensions are operationalized by means of the following evaluation criteria:

1. ***Financial cost*** (economic dimension), weight = 0.167; its criterion scores are in millions of Euro measured in crisp terms, indifference threshold = €250,000, preference threshold = €500,000 (see Fig. 7.3).
2. ***Employment*** (economic dimension), weight = 0.167; its criterion scores are in number of persons/year, measured by means of symmetric fuzzy numbers, indifference threshold = 30 persons/year, preference threshold = 50 persons/year (see Fig. 7.4).
3. ***Avoidance of social exclusion*** (social dimension), weight = 0.333; its criterion scores are qualitative, measured by means of an ordinal scale of measurement (good better than moderate).

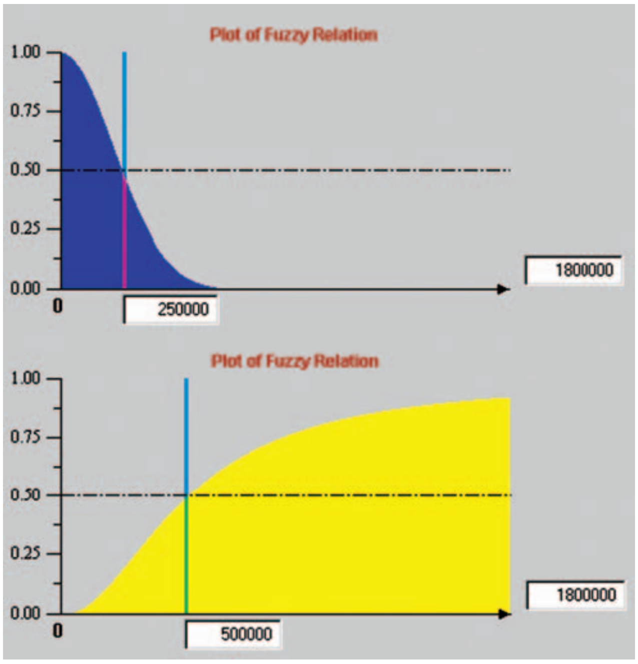

Fig. 7.3 Indifference and preference fuzzy relations on the criterion "financial cost"

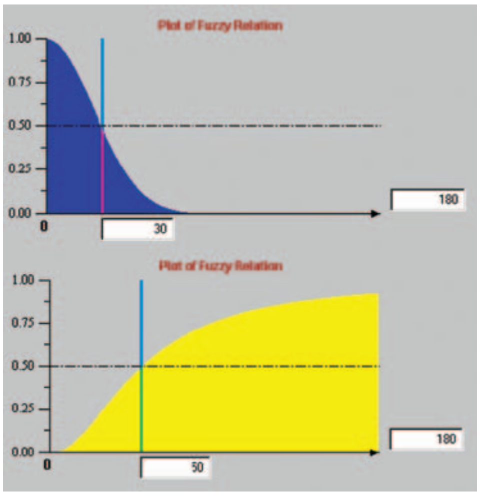

Fig. 7.4 Indifference and preference fuzzy relations on the criterion “employment”

Table 7.1 Evaluation matrix of a hypothetical public policy problem

Alternatives	A	B	C
Criteria			
financial cost	13.2	13.5	15
employment	approx. 100	approx. 135	approx. 200
avoidance of social exclusion	Very Good	Moderate	Good
environmental impact	2-nd	1-st	3-rd

Table 7.2 Values of the semantic distance for the criterion “financial cost”

	Semantic distance
(A,B)	0.30
(A,C)	1.80
(B,C)	1.50

4. ***Environmental impact*** (environmental dimension), weight = 0.333; its criterion scores are qualitative, measured by means of an ordinal scale of measurement (1° better than 2°).

This policy problem can be summarized in the evaluation matrix described in Table 7.1.

Let us now compare each pair of alternatives according to each criterion. For the ordinal criterion scores the comparison is obvious. For the other criteria, let us apply the semantic distance. The results are presented in Tables 7.2 and 7.3.

Now it is possible to compute the fuzzy preference and indifference relations. Values are given in Tables 7.4 and 7.5.

Table 7.3 Values of the semantic distance for the criterion "employment"

	Expected value difference	Semantic distance
(A,B)	–35.0081	56.2391
(A,C)	–99.9977	107.9972
(B,C)	–64.9896	87.8337

Table 7.4 Values of the fuzzy relations for the criterion "financial cost"

	μ_P	μ_I
(A,B)	0.2647	0.5946
(A,C)	0.9284	0.0442
(B,C)	0.9000	0.0743

Table 7.5 Values of the fuzzy relations for the criterion "employment"

	μ_P	μ_I
(B,A)	0.3290	0.3827
(C,A)	0.8000	0.1570
(C,B)	0.6282	0.2224

Through (7.18), let's now compute the relative weights on each criterion for any pair of alternatives.

Criterion 1: "Financial Cost".

$$w_1(P)_{(A,B)} = \frac{0.2647}{0.2647 + 0.5946} \times 0.167 = 0.051$$

$$w_1(I)_{(A,B)} = \frac{0.5946}{0.2647 + 0.5946} \times 0.167 = 0.116$$

$$w_1(P)_{(A,C)} = \frac{0.9284}{0.9284 + 0.0442} \times 0.167 = 0.159$$

$$w_1(I)_{(A,C)} = \frac{0.0442}{0.9284 + 0.0442} \times 0.167 = 0.0075$$

$$w_1(P)_{(B,C)} = \frac{0.9000}{0.9000 + 0.0743} \times 0.167 = 0.154$$

$$w_1(I)_{(B,C)} = \frac{0.0743}{0.9000 + 0.0743} \times 0.167 = 0.013$$

Criterion 2: "Employment".

$$w_2(P)_{(B,A)} = 0.0762$$
$$w_2(I)_{(A,B)} = 0.08872$$
$$w_2(P)_{(C,A)} = 0.1379$$
$$w_2(I)_{(A,C)} = 0.0270$$
$$w_2(P)_{(C,B)} = 0.1218$$
$$w_2(I)_{(B,C)} = 0.04319$$

Criterion 3: “Avoidance of Social Exclusion”.

$$w_3(P)_{(A,B)} = 0.333$$
$$w_3(P)_{(A,C)} = 0.333$$
$$w_3(P)_{(C,B)} = 0.333$$

Criterion 4: “Environmental Impact”.

$$w_4(P)_{(B,A)} = 0.333$$
$$w_4(P)_{(A,C)} = 0.333$$
$$w_4(P)_{(B,C)} = 0.333$$

By applying (7.24) the following results are obtained (see Table 7.6):

By applying the C–K–Y–L rule to the *3!* possible rankings it is:

$$\text{ABC}\,\varphi_1 = 0.485 + 0.513 + 0.841 = 1.83$$
$$\text{BCA}\,\varphi_2 = 0.513 + 0.159 + 0.515 = 1.18$$
$$\text{CAB}\,\varphi_3 = 0.159 + 0.485 + 0.487 = 1.13$$
$$\text{ACB}\,\varphi_4 = 0.841 + 0.487 + 0.485 = 1.81$$
$$\text{BAC}\,\varphi_5 = 0.515 + 0.841 + 0.513 = 1.86$$
$$\text{CBA}\,\varphi_6 = 0.487 + 0.515 + 0.159 = 1.16$$

The final ranking r^* is then BAC

Let us now look for the Borda loser. Matrix F is presented in Table 7.7:

Computing the elements ϕ_{ij} of matrix Φ, we obtain:

Table 7.6 Matrix E of a hypothetical public policy problem

$$E = \begin{bmatrix} & A & B & C \\ A & 0 & 0.485 & 0.841 \\ B & 0.515 & 0 & 0.513 \\ C & 0.159 & 0.487 & 0 \end{bmatrix}$$

Table 7.7 Matrix *F* of a hypothetical public policy problem

Alternatives	A	B	C
Criteria			
financial cost	1-st	2-nd	3-rd
employment	3-rd	2-nd	1-st
avoidance of social exclusion	1-st	3-rd	2-nd
environmental impact	2-nd	1-st	3-rd

$$\phi_{1A} = 0.167 + 0.333 = 0.5$$
$$\phi_{2A} = 0.333$$
$$\phi_{3A} = 0.167$$
$$\phi_{1B} = 0.333$$
$$\phi_{2B} = 0.167 + 0.167 = 0.33333$$
$$\phi_{3B} = 0.333$$
$$\phi_{1C} = 0.167$$
$$\phi_{2C} = 0.333$$
$$\phi_{3C} = 0.333 + 0.167 = 0.5$$

Then matrix Φ is the following $\Phi = \begin{bmatrix} 0.5 & 0.333 & 0.167 \\ 0.333 & 0.333 & 0.333 \\ 0.167 & 0.333 & 0.5 \end{bmatrix}$

By applying(7.32), we have:
$B(\mathrm{A}) = 0.333 \times 1 + 0.167 \times 2 = 0.666$
$B(\mathrm{B}) = 0.333 \times 1 + 0.333 \times 2 = 1$
B(C) $= 0.333 \times 1 + 0.5 \times 2 =$ **1.333**

The Borda loser is alternative C which, in this case, is also the C–K–Y–L loser.

Let us now look at a completely ordinal example. As discussed in Box 4.1, with composite indicators it is essential to use weights as importance coefficients. Moreover, indifference and preference thresholds would increase the degree of arbitrariness too much, thus a proper ranking procedure for composite indicators should be ordinal in nature (Munda and Nardo, 2003). Let us then apply the C–K–Y–L ranking procedure to the urban sustainability assessment example with the criterion weights illustrated in Sect. 4.5. The corresponding outranking matrix is presented in Table 7.8.

The 24 possible rankings and the corresponding scores φ_s are the following (where A is Budapest, B is Moscow, C is Amsterdam and D is New York):

Table 7.8 Weighted outranking matrix

	Budapest	Moscow	Amsterdam	New York
Budapest	0	0.3	0.4	0.4
Moscow	0.7	0	0.5	0.6
Amsterdam	0.6	0.5	0	0.3
New York	0.6	0.4	0.7	0

B	**D**	**C**	**A**	**3,6**	B	C	A	D	2,9
D	B	C	A	3,5	C	B	A	D	2,9
D	C	B	A	3,5	A	B	D	C	2,9
B	D	A	C	3,5	B	A	C	D	2,8
D	B	A	C	3,4	A	D	B	C	2,8
B	A	D	C	3,3	A	D	C	B	2,8
B	C	D	A	3,2	C	D	A	B	2,7
C	B	D	A	3,2	C	A	B	D	2,6
D	C	A	B	3,2	C	A	D	B	2,5
C	D	B	A	3,1	A	B	C	D	2,5
D	A	B	C	3,1	A	C	B	D	2,5
D	A	C	B	3,1	A	C	D	B	2,4

In comparison with the results obtained by applying the linear aggregation rule, without any criterion weights as described in Chap. 1, Moscow is still in top position, but this time Budapest is at the bottom. New York again scores better than Amsterdam.

Note that the use of weights and the improvement of the mathematical aggregation procedure (in comparison with the simple linear aggregation rule) do not change the results spectacularly. The structuring process, and in this case above all, the input information used for the indicator scores clearly determine the final ranking. "*Garbage in, garbage out*" phenomena are almost impossible to avoid (Funtowicz and Ravetz, 1990). This is a fundamental lesson to bear in mind in real-world applications of SMCE. Good mathematical algorithms guarantee consistency with the problem structuring and nothing else. Of course, *ceteris paribus*, the mathematical properties of a ranking algorithm may make an important difference.

Let us conclude by examining thereafter examples from the field of composite indicators (Munda and Nardo, 2003). Let us take into consideration a simple hypothetical example with three countries (A, B, C) to be ranked according to a composite sustainability indicator. Let us assume that three dimensions have to be considered, i.e. economic, social and environmental, and that each dimension should have the same weight, i.e. 0.3333.

The following individual indicators are used:

Economic dimension

Indicator: GDP per capita. Weight: 0.167. Objective: maximization of economic growth. Variable: US dollar per year.

Indicator: Unemployment rate. Weight: 0.167. Objective: minimization of unemployed people. Variable: percentage of population.

Environmental dimension

Indicator: Solid waste generated per capita. Weight: 0.333. Objective: minimization of environmental impact. Variable: tons per year.

Social dimension

Indicator: Income disparity. Weight: 0.167. Objective: minimization of distributional inequity. Variable: Q5/Q1.

Indicator: Crime rate. Weight: 0.167. Objective: minimization of criminality. Variable: robberies per 1000 inhabitants.

The impact matrix described in Table 7.9 can then be constructed.

The pair-wise comparison results can be summarized in the following outranking matrix:

$$E = \begin{bmatrix} & A & B & C \\ A & 0 & 0.666 & 0.333 \\ B & 0.333 & 0 & 0.333 \\ C & 0.666 & 0.666 & 0 \end{bmatrix}$$

By applying the C–K–Y–L rule to the 3! possible rankings we obtain:

$$\begin{aligned}
&\text{ABC}\,\varphi_1 = 0.666 + 0.333 + 0.333 = 1.333\\
&\text{BCA}\,\varphi_2 = 0.333 + 0.333 + 0.666 = 1.333\\
&\textbf{CAB}\,\varphi_3 = 0.666 + 0.666 + 0.666 = 2\\
&\text{ACB}\,\varphi_4 = 0.333 + 0.666 + 0.666 = 1.666\\
&\text{BAC}\,\varphi_5 = 0.333 + 0.333 + 0.333 = 1\\
&\text{CBA}\,\varphi_6 = 0.666 + 0.666 + 0.333 = 1.666
\end{aligned}$$

The final ranking r^* is then CAB.

Note that using one of the standard ways to produce a composite indicator would produce a different result. If the composite indicator for each country is calculated in terms of the difference from the group leader (which assigns 100 to the leading country and ranks the others in percentage points away from the leader), the impact matrix becomes as shown in Table 7.10.

The index will be calculated by averaging each indicator (with the same weights as in the multi-criterion matrix), obtaining $\boldsymbol{I}_A = 69.8$, $\boldsymbol{I}_B = 79.7$, and $\boldsymbol{I}_C = 81.9$. The ranking would be CBA, different from the ranking obtained with the other algorithm.

Table 7.9 Impact matrix of the illustrative numerical example

Indicators	GDP	Unemp. rate	Solid waste	Inc. dispar.	Crime rate
Countries					
A	22,000	0.17	0.4	10.5	40
B	45,000	0.09	0.45	11.0	45
C	20,000	0.08	0.35	5.3	80

Table 7.10 Impact matrix: distance from the leader

Indicators	GDP	Unemp. rate	Solid waste	Inc. dispar.	Crime rate
Countries					
A	48.9	47.05	87.5	50.5	100
B	100	88.9	77.8	48.2	88.9
C	44.4	100	100	100	50

Table 7.11 Performance in the knowledge-based economy: a tentative indicator

Indicators	BERD	Researchers	Patents	HT
Countries				
Australia	23.1	22.0	8.3	64.5
Austria	38.7	27.5	24.4	76.3
Belgium	46.5	40.2	32.0	22.3

Consider another example, the *composite indicator of Industrial innovation* (OECD, 2003). This composite indicator is based on four sub-indicators: Business enterprise R&D as percentage of GDP (*BERD*), the number of business researchers per 10,000 labour force (*Researchers*), the number of patents per million population (*Patents*), and the share of firms having introduced at least one new or improved product or process on the market *(HT)*. For sake of simplicity let us take the first three countries of OECD classification (see Table 7.11) (see OECD, 2003, p. 19)[5]:

The composite index is the simple average of indicators (thus we have a case of equal weights): 41.7 for Austria, 35.2 for Belgium, and 29.5 for Australia.

However, the outranking matrix with weight equal to ¼ for each index is as follows:

$$E = \begin{bmatrix} & AU & A & B \\ AU & 0 & 0 & 0.25 \\ A & 1 & 0 & 0.25 \\ B & 0.75 & 0.75 & 0 \end{bmatrix}$$

From the comparison of 3! possible combinations it turns out that the one with the highest score is (B, A, AU) with 2.25[6]. Again the ranking produced with the standard methods of summing up normalized variables is different from that produced with the C–K–Y–L ranking procedure.

[5] The indicators in the matrix shown in Table 7.11 have been normalized with the min-max method which ranks each country with respect to the global maximum (the leader = 100) and the global minimum (the laggard = 0). The index is calculated as: (actual value – minimum value)/(maximum value – minimum value)*100. Note that none of the countries chosen is either a maximum or a minimum.

[6] The outranking matrix is the same for the original data and for the normalized indicators.

Let us conclude with a real-world example which shows the importance of a computational algorithm: the "Environmental Sustainability Index" (ESI). The index for 2005 was produced by a team of environmental researchers from Yale and Columbia Universities, in co-operation with the World Economic Forum and the Joint Research Centre of the European Commission.

The aim of the ESI is to benchmark the ability of 146 nations to protect the environment over the next decades, by integrating 76 data sets into 21 indicators of environmental sustainability (see Esty et al., 2005). The database used to construct the ESI covers a wide range of aspects of environmental sustainability ranging from the physical state and stress of the environmental systems (like natural resource depletion, pollution, ecosystem destruction) to the more general social and institutional capacity to respond to environmental challenges. Poverty, short-term thinking and lack of investment in capacity and infrastructure committed to pollution control and ecosystem protection thus compete to determine the measure of a country's sustainability.

Although the official ESI ranking is based upon the linear aggregation of 21 equally weighted indicators, an attempt has been made, in the methodological appendix, to apply the non-compensatory approach presented in this chapter, in order to tackle the issues of weights as "importance measure" and the compensability of different and crucial dimension of environmental sustainability (see the Methodological Appendix in Esty et al., 2005).

Figure 7.5 compares the ranking obtained by means of the non-compensatory aggregation rule with that of the ESI2005. In both cases all 21 indicators are equally weighted. From this figure it is clear that the aggregation method used affects principally the middle-of-the-road and, to a lesser extent, the leader and the laggard countries. Overall, for the set of 146 countries, the assumption underlying the

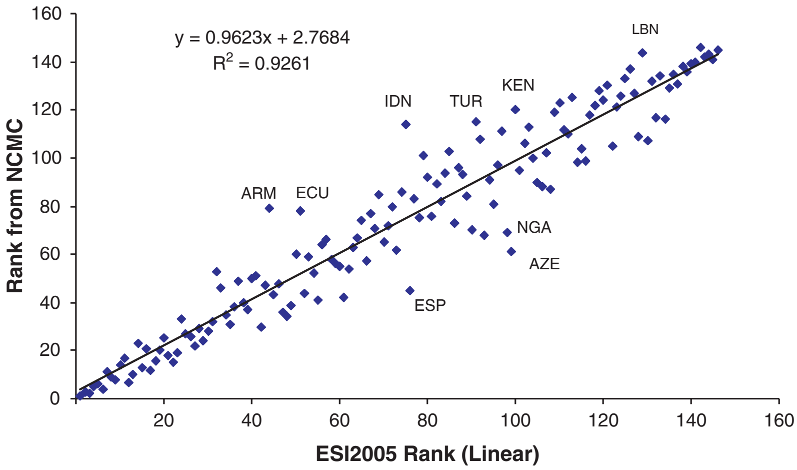

Fig. 7.5 Comparison of rankings obtained by the linear aggregation (ESI2005 on the *x*-axis) and the non-linear/non-compensatory –C–K–Y–L– (NCMA on the *y*-axis) rules

Table 7.12 ESI rankings obtained by linear aggregation (LIN) and the C–K–Y–L ranking procedure: countries that greatly improve or greatly worsen their rank position

	Aggregation	ESI rank with LIN	Rank with NCMC	Change in rank
Improvement	Azerbaijan	99	61	38
	Spain	76	45	31
	Nigeria	98	69	29
	South Africa	93	68	25
	Burundi	130	107	23
Deterioration	Indonesia	75	114	39
	Armenia	44	79	35
	Ecuador	51	78	27
	Turkey	91	115	24
	Sri Lanka	79	101	22
	Average change over 146 countries			8

aggregation scheme has an average impact of eight ranks and a rank-order correlation coefficient of 0.962. In particular, while the top 50 countries move on average only five positions, the next 50 countries on average move twelve positions and the remaining 46 countries eight positions.

It is important to underline that although both aggregation schemes seem to produce consistent rankings (the R^2 is 0.92), those rankings do not nevertheless coincide. Using the non-compensatory approach, 43 out of 146 countries experience a change in rank greater than ten positions (none before the 30th ESI rank). When compensability among indicators is not allowed, countries with very poor performance in some indicators, such as Indonesia or Armenia, worsen their rank with respect to the linear yardstick, whereas countries that have less extreme values improve their ranking, such as Azerbaijan or Spain. Table 7.12 shows the countries with the largest variation in their ranks.

7.7 Conclusion

This chapter has presented a new mathematical aggregation convention for the solution of the so-called discrete multi-criterion problem in a SMCE context. This multi-criterion aggregation convention can be divided into two main steps:

- Pair-wise comparison of alternatives
- Ranking of alternatives in a complete pre-order

Throughout the pair-wise comparison step it is guaranteed that ordinal, crisp, stochastic and fuzzy criterion scores are tackled equivalently. The double threshold model, generating a pseudo-order structure, is used for preference modelling; as a consequence the so-called Luce paradox is avoided. To deal with the lack of stability of the pseudo-order structure, valued preference relations modelled by means of fuzzy preference relations are introduced. Given the requirement of consistency between indifference and preference thresholds, the functional form of these fuzzy relations looks descriptively reasonable. Weights are never combined with intensities of preference; as a consequence the theoretical guarantee that they are importance coefficients holds. Given that the sum of weights is equal to one, the pair-wise comparisons can be synthesized in a matrix, which can be interpreted as a voting matrix. Thanks to Proposition 7.1, it is known that ties are possible but that the probability of coming across one is approximately zero; neutrality is then in general respected.

The information contained in the voting matrix is exploited to rank all alternatives in a complete pre-order by using a Condorcet consistent rule. The Condorcet tradition has been chosen for four main reasons:

- Non-compensability is implied, since intensities of preference are never used.
- Manipulation rules of weights guarantee that they are importance coefficients.
- It is the most consistent approach for generating a complete ranking.
- There is a low probability of obtaining rank reversals.

A problem connected with the use of Condorcet consistent rules is the occurrence of cycles. A cycle-breaking rule normally demands some arbitrary choices, such as eliminating the cycle with the lowest support, and so on. In search of a non-arbitrary cycle breaking rule the Condorcet–Kemeny–Young–Levenglick ranking procedure was chosen; no arbitrary choice is called for with this procedure. Given the fact that criterion weights are used, anonymity is necessarily lost. However, given that Arrow's impossibility theorem forces us to make trade-offs between decisiveness and anonymity, the loss of anonymity in favour of decisiveness in our framework is a positive feature. An important advantage of the C–K–Y–L procedure is that its properties are completely known and meet the requirements of social multi-criteria evaluation. A problem connected with the C–K–Y–L procedure is its computational complexity. Given that this problem can be solved by the numerical algorithm presented, its implementation in a multi-criteria framework is possible without any restriction on the number of alternatives considered. Consistently with the normative tradition in political philosophy and following the prudence axiom, the minority principle is introduced by means of a veto power. A vetoed alternative, the Borda loser, is found by means of the original Borda approach, implemented through a frequency matrix. This approach has been chosen because:

- It is independent of arbitrary ad hoc thresholds.
- It considers the global opposition to the final ranking.
- No specific intensity of preference is considered, thus weights continue to be importance coefficients.

The issue that makes multi-criterion aggregation conventions intrinsically complex, is the fact they are simultaneously *formal, descriptive* and *normative* models (Munda, 1993). As a consequence, the properties of an approach have to be evaluated at least in the light of these three dimensions. In the framework of the debate on the maximization assumption in microeconomics, Musgrave (1981) made a very useful classification of the assumptions made in economic theory. He makes a distinction between *negligibility assumptions, domain assumptions and heuristic assumptions*. The first type is required to simplify and focus on the essence of the phenomena studied. The second type of assumption is needed when applying a theory to specify the domain of applicability. The third type is needed either when a theory cannot be directly tested or when the essential assumptions give rise to such a complex model that successive approximation is required.

Let us then try to clarify the properties of the approach I am proposing in the light of these considerations.

Descriptive domain assumptions:

- Mixed information is tackled in the form of ordinal, crisp, stochastic and fuzzy criterion scores.
- The preference model is a pseudo-order structure with constant indifference and preference thresholds.
- The most useful result for policy-making is a complete ranking of alternatives.

Normative domain assumptions:

- Simplicity is desirable and means the use of as few ad hoc parameters as possible.
- Weights are meaningful only as importance coefficients and not as trade-offs. As a consequence, complete compensability cannot be implemented.
- A minority principle must be implemented for ethical and prudential reasons.

Formal domain assumptions:

- Unanimity
- Monotonicity
- Neutrality
- Reinforcement

Heuristic descriptive assumptions:

- Criteria can always be derived from the higher dimensions to which they univocally belong.
- Valued preference relations (in the form of fuzzy relations) are useful for solving the problem of lack of stability of a pseudo-order structure.

Heuristic formal assumptions:

- Local stability
- Cycle-breaking without losing neutrality.
- Semantic distance as a compromise solution between generality and precision

Negligibility formal assumptions:

- Anonymity
- Independence of irrelevant alternatives

In conclusion, we may state that the main characteristic of the multi-criterion aggregation convention I am proposing is that all the steps are fully justified and all the properties made explicit. Of course this is not to imply that it is the "best" possible approach to the discrete multi-criterion problem. It is a "reasonable" approach based on theoretical and empirical grounds, all of them explicit and thus open to evaluation in relation to a particular purpose.

Appendix 7.1

Proposition 7.1 *In the outranking matrix E the event of finding ties, that is* $e_{jk} = e_{kj} = \dfrac{M}{2}$ *is possible but its probability is approximately zero.*

Proof

The probability that $e_{jk} = e_{kj} = \dfrac{M}{2}$ always depends at least on the number of criteria (voters or individual indicators) in favour of a_j and a_k. Let us also assume the existence of a set of criterion weights $W = \{w_m\}$, $m = 1, 2,\ldots,M$, with $\sum_{m=1}^{M} w_m = 1$, which is very common in multi-criteria analysis and in the construction of composite indicators. Let us then look at the specific probabilities of each single factor.

Given S criteria in favour of a_j and T criteria in favour of a_k, the probability of a specific combination is $\dfrac{1}{\binom{M}{S}}$.

A specific value on the vector of weights depends on how many numbers we are using for the definition of each weight. Let us assume that each weight is defined by two integer numbers (e.g. 0.02, 0.10, 0.25,….,..), then the probability of a specific weight is $\dfrac{1}{10^2}$. Since the value 0.00 does not make sense, the probability is $\dfrac{1}{99}$. Thus a specific vector has a probability of $\left(\dfrac{1}{99}\right)^M$.

At this point it is possible to compute the probability $\pi(v)$ of finding a tie on the outranking matrix E. The probability is:

$$\pi_v = \frac{1}{\binom{M}{S}} \times \left(\frac{1}{99}\right)^M$$

It is evident that $\pi_v \simeq 0$.

Let us now assume that no criterion weight is used. In this case, to have a tie, it is necessary that the number of criteria M be even. The probability is

$\pi_v = \dfrac{1}{\binom{M}{M/2}} \times \left(\dfrac{1}{2}\right)$, where $\dfrac{1}{2}$ is the probability of having an even number of criteria (on the grounds of Laplace insufficient reason principle). In this case too, the probability is very low.[7]

One should note finally that the probability to get ties is always close to zero in the case where indifference and preference thresholds are used in the preference modelling.

In fact, given the vector of intensities I_m, of criteria g_m, the values $\mu(P)_m$ and $\mu(I)_m$ depend on the thresholds q_m and p_m defined on each criterion g_m. Let us denote with $\pi(q_m)$ and $\pi(p_m)$ the respective probabilities of getting a precise value of q_m and p_m on a criterion g_m. Then the probability of getting a specific vector of values on q_m and p_m is $\bigcap_{m=1}^{M} \pi\,(q)_m$ and $\bigcap_{m=1}^{M} \pi\,(p)_m$. The question now concerns the values of $\pi(q_m)$ and $\pi(p_m)$? In theory, the thresholds may vary a priori on any point of the intensity of preference I_m, I_m being a set which is not finite or countable. Thus, it is[8]

$$\pi(q_m) = \pi(p_m) = \lim_{n \to +\infty} \frac{1}{n} = 0$$

At this point it is possible to compute the probability $\pi(v)$ of having a tie in the matrix E in the most general case, i.e. where thresholds and weights are defined on all criteria. It is

$$\pi_v = \frac{1}{\binom{M}{S}} \times \left(\frac{1}{99}\right)^M \times \pi(q)_m{}^M \times \pi(p)_m{}^M$$

Let us put $\pi(q)_m = \pi(p)_m = \lambda$, then

[7] In the case with the smallest M, which makes sense with a ranking problem, i.e. four criteria, the probability is $\pi_v = 1/2 \simeq 0.083$; for $M = 6$, it is: $\pi_v = 1/40 \simeq 0.025$.

[8] One could argue that, from a descriptive point of view, it is not very realistic to assume that all thresholds have the same probability along the set I_m. Let us make the VERY optimistic assumption that only ten thresholds of each type are possible for each criterion and four criteria exist. In this case the probability of a specific vector on q_m and p_m is $\left(\frac{1}{10}\right)^8 = 0.00000001$. As one can see, even in this optimistic case the probability is very close to zero.

$$\pi_v = \frac{1}{\binom{M}{S}} \times \left(\frac{1}{99}\right)^M \times \lambda^{2M}$$

It is evident that $\pi_v \simeq 0$

Appendix 7.2

It is a trivial matter to prove that the semantic distance satisfies the properties of non-negativity and symmetry: the fulfilment of the property of triangle inequality can be proven as follows (Munda, 1995). Let us assume three functions:

$$f(x): X \to R^+,\ g(y): Y \to R^+,\ h(z): Z \to R^+$$

Let's also assume that $X \cap Y \cap Z \neq \varnothing$
We first prove that $\forall\ x \in X$, $\forall\ y \in Y$ *and* $\forall\ z \in Z$; the relationship
$|x-y|+|y-z| \geq |x-z|$ is always true.

The total number of possible cases is *3!*

$$\begin{aligned}
&x \geq y \geq z \to (x-y)+(y-z)-(x-z)=0\\
&x \geq z \geq y \to (x-y)+(-y+z)-(x-z)= 2(z-y) \geq 0\\
&y \geq x \geq z \to (-x+y)+(y-z)-(x-z)= 2(y-x) \geq 0\\
&y \geq z \geq x \to (-x+y)+(y-z)-(-x+z)= 2(y-z) \geq 0\\
&z \geq x \geq y \to (x-y)+(-y+z)-(-x+z)= 2(x-y) \geq 0\\
&z \geq y \geq z \to (-x+y)+(-y+z)-(-x+z)=0
\end{aligned}$$

therefore

$$|x-y|+|y-z| \geq |x-z| \geq 0 \quad \forall\, x \in X,\ \forall\, y \in Y \text{ and } \forall\, z \in Z.$$

Since $f(x) \geq 0$, $g(y) \geq 0$ and $h(z) \geq 0$, it is:

$$\int_x \int_y \int_z \left[|x-y|+|y-z|-|x-z|\right] f(x)\, g(y)\, h(z)\, dz\, dy\, dx \geq 0$$

This integral can be decomposed as follows:

$$\begin{aligned}
&\int_x \int_y \int_z |x-y|\, f(x)\, g(y)\, h(z)\, dz\, dy\, dx + \int_x \int_y \int_z |y-z|\, f(x)\, g(y)\, h(z)\, dz\, dy\, dx -\\
&\int_x \int_y \int_z |x-z|\, f(x)\, g(y)\, h(z)\, dz\, dy\, dx = \int_x \int_y |x-y| f(x)\, g(y)\, dy\, dx +\\
&+\int_y \int_z |y-z|\, g(y)\, h(z)\, dz dy - \int_x \int_z |x-z| f(x)\, h(z)\, dz\, dx
\end{aligned}$$

This is because the triple integrals can be computed by means of iterated integrals and because it is: $\int_x f(x)\,dx = \int_y g(y)\,dy = \int_z h(z)\,dz = 1$

Therefore, it is:

$$S_d\left[f(x), g(y)\right] + S_d\left[g(y), h(z)\right] - S_d\left[f(x), h(z)\right] \geq 0 \qquad \text{or}$$
$$S_d\left[f(x), g(y)\right] + S_d\left[g(y), h(z)\right] \geq S_d\left[f(x), h(z)\right]$$

When both variables x and y are defined in the same interval, i.e. $X = Y = [x_L, x_U] = [y_L, y_U]$, for reasons of consistency it is necessary to prove that the value of the semantic distance is other than zero. For simplicity let us say $x_L = y_L = a$ and $x_U = y_U = b$ where $a < b$. Let us now consider the following algebraic expression: $|x-y|\,f(x)\,g(y)$. This product assumes the value of zero if at least one of its three elements is zero. Concerning *f(x)* and *g(y)*, the following two assumptions are made:

$$(1)\begin{cases} f(x) > 0 & \forall\, x \in (a,b) \\ f(x) = 0 & \text{if } x = a \; or \; x = b \end{cases}$$

$$(2)\begin{cases} g(y) > 0 & \forall\, y \in (a,b) \\ g(y) = 0 & \text{if } y = a \text{ or } y = b \end{cases}$$

For the factor $|x-y|$ it is:

$$(3)\begin{cases} |x-y| > 0 & \forall\, x \neq y \in [a,b] \\ |x-y| = 0 & \text{if } x = y \end{cases}$$

On the basis of these assumptions it is possible to conclude that

$$\begin{cases} |x-y|\,f(x)\,g(y) > 0 & \forall\, x \neq y \in (a,b) \\ |x-y|\,f(x)\,g(y) = 0 & \text{elsewhere} \end{cases}$$

If we take into consideration the sum of all

$$|x-y|\,f(x)\,g(y) \qquad \forall\, x \in [a,b] \text{ and } \forall\, y \in [a,b], \quad i.e.$$
$$\int_x\int_y |x-y|\,f(x)\,g(y)\,dy\,dx, \text{ given that it is not } |x-y|\,f(x)\,g(y) = 0,$$
$$\forall\, x \in [a,b] \text{ and } \forall\, y \in [a,b], \quad \text{it is}$$
$$\int_x\int_y |x-y|\,f(x)\,g(y)\,dy\,dx > 0$$

Note that

$$S_d(f(x), g(y)) = 0 \;\; iff \;\; x = y \;\forall\, x \in [a,b] \;\; and \;\; \forall\, y \in [a,b],$$
$$i.e. \; iff \; a = x = y = b.$$

This is true only if x and y are two equal, crisp real numbers.

Finally, one should note that if the distance between a fuzzy number and itself is computed by definition, the condition $S_d\,(f(x), f(x)) = 0$ has to be imposed.

Appendix 7.3

To make the semantic distance presented in Sect. 7.2 operational, a Monte Carlo type numerical algorithm is required (Munda, 1995).

The initial *assumptions* are:

$$(1)\quad \begin{array}{l} f(x): X = [x_L, x_U] \to \mathrm{M} \\ g(y): Y = [x_{L'}, x_{U'}] \to \mathrm{M} \end{array}$$

where M is the membership space.

(2) All $x \in X$ and all $y \in Y$ can be obtained by means of a random generator that supplies uniformly distributed numbers $r \in [0,1]$.

We have:

$$x = rx_L + (1-r)x_U \qquad \text{and}$$
$$y = rx_{L'} + (1-r)x_{U'}$$

(3) The probability of obtaining a point p inside e.g. $f(x)$, whose value on the x-axis is x_0 depends on the shape of the function. An auxiliary variable Z, with $z \in [0, \max f(x)]$, is then introduced by means of a random generator.

Now the *procedure* is as follows:

STEP 1: draw a random number r_0
STEP 2: $x_0 = r_0 x_L + (1 - r_0)\, x_U$
STEP 3: draw a random number Z_0
STEP 4: if $Z_0 \le f(x_0)$ then go to next step
if $Z_0 > f(x_0)$ then return to step 1.
STEP 5: draw a random number r_1
STEP 6: $y_1 = r_1 x_{L'} + (1 - r_1) x_{U'}$
STEP 7: draw a random number Z_1
STEP 8: if $Z_1 \le g(y_1)$ then compute $|x_0 - y_1|$
if $Z_1 > g(y_1)$ then return to step 5.

By repeating this procedure N times, N values of $|x_i - y_i|$ are obtained. The semantic distance between two fuzzy sets is approximately equal to the arithmetic mean of all the points bounded by their respective membership functions obtained by drawing random numbers. In more formal terms it is:

$$S_d\big(f(x),\ g(y)\big) = \iint_{x,\,y} |x-y| f(x) g(y)\, dy\, dx \simeq \frac{\sum_{i=1}^{N} |x_i - y_i|}{N}$$

with $i = 1, 2, \ldots, N$.

Of course, the greater N, the more precise the computation.