# MACBETH

*Carlos A.* ***Bana e Costa***[1,2], *Jean-Marie* ***De Corte***[3], *Jean-Claude* ***Vansnick***[3]

[1] *Department of Operational Research, London School of Economics*
*Houghton Street, London WC2A 2AE, U.K. (c.bana@lse.ac.uk)*

[2] *Centre for Management Studies of Instituto Superior Técnico, Technical University of Lisbon*
*Av. Rovisco Pais, 1049-001 Lisbon, Portugal (carlosbana@netcabo.pt)*

[3] *Centre de Recherche Warocqué, Université de Mons-Hainaut*
*Place du Parc, 20, 7000 Mons, Belgium (DeCorte@umh.ac.be / Vansnick@umh.ac.be)*

**Abstract: This paper presents an up-to-date comprehensive overview of MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique). MACBETH is a multicriteria decision analysis approach that requires only qualitative judgements about differences of value to help a decision maker, or a decision-advising group, quantify the relative attractiveness of options. The approach, based on the additive value model, aims to support interactive learning about the evaluation problem and the elaboration of recommendations to prioritise and select options in individual or group decision making processes. We revise in detail the theoretical foundations of MACBETH and present a simple example that illustrates the use of the M-MACBETH decision support system (www.m-macbeth.com). It permits the structuring of value trees, the construction of criteria descriptors, the scoring of options against criteria, the development of value functions, the weighting of criteria, and extensive sensitivity and robustness analyses about the relative and intrinsic value of options. Reference is also made to some successful real-world consulting applications and a historical survey is also included, which describes the key stages, with their corresponding publications, in the development of MACBETH since early the 1990's.**

***Keywords:*** *MACBETH, Qualitative value judgements, Multicriteria decision analysis.*

LSE OR Working Paper 03.56 ISBN 0 7530 1520 X

## 1. Introduction

MACBETH is an approach to multicriteria decision aid whose development was set in motion in the early 1990's by C.A. Bana e Costa and J.-C. Vansnick. The team expanded in subsequent years when it was joined by J.-M. De Corte. The brief historical survey of the MACBETH research offered in section 4 shows that, on a technical level, MACBETH has evolved through the course of theoretical research conducted about it, and also through numerous practical applications (see section 5).

Its essential characteristics, however, have never changed; in fact MACBETH has always remained consistent with the ideas that led to its creation as an humanistic, interactive and constructive approach to the problem of how to build a quantitative model of values based on qualitative (verbal) difference judgements, that facilitates the path from ordinal to cardinal preference modelling, namely analysing judgmental inconsistence and offering suggestions to move the process forward.

MACBETH is:

- ✓ Humanistic in the sense that it should be used to help decision-makers ponder, communicate, and discuss their value systems and preferences.
- ✓ Interactive because we are convinced that this reflection and learning process can best spread through socio-technical facilitation sustained by straightforward questioning-answering protocols. From a practical viewpoint, this suggests that such interaction would greatly benefit from an extremely efficient and user-friendly decision support system, as it is actually the case of the M-MACBETH software.
- ✓ Constructive because MACBETH rests on the idea that full-bodied convictions about the kind of decision to make do not (pre-)exist in the mind of the decision maker, nor in the mind of each of the members of a decision advising group, but that we can provide them with help to form such convictions and to build robust (shared) preferences concerning the different possible options to solve the problem.

Let us point out that this type of construction does not have to be, and should not be, anarchic in nature. In order to ensure the coherence of the construction process, it is advisable that it be done within the framework of a well defined model of aggregating ("partial") preferences established at the level of each of the multiple criteria. The MACBETH approach adopted the additive value aggregation model as the reference scope that guarantees the coherence of the help provided in the course of a multicriteria process of building ("global") preferences. Here are some of the reasons that led us to choose the additive model:

- ✓ It is particularly simple
- ✓ it is well known
- ✓ its technical parameters have a clear and easily explicable substantive interpretation
- ✓ it allows the processing of the difficult problem of relative importance of criteria in a precise way
- ✓ it permits the avoidance of the difficulties that are inherent in every ordinal aggregation (Condorcet's Paradox, Arrow's Theorem).

The additive value model has some attractive assets but its practical implementation is sometimes difficult because it requires the availability of very rich information, concerning the decision makers' preferences, in order to produce mathematically significant results. Within the framework of MACBETH, we did not perceive this necessity to be a drawback. Instead, we considered it to be an opportunity to compel the decision makers to think about their preferences at a greater depth than usual and thus get a clearer vision of the decisional situation. Furthermore, this difficulty associated with the process of additive aggregation led us to think profusely about how to get the decision makers to provide such rich information. The first publications regarding MACBETH were about this very issue (see section 4) and we consider that one of the strengths of MACBETH is the original and in practice well calculated manner in which this crucial aspect is treated; this subject is presented in section 2.

In section 3, we will tackle the multicriteria aspect of the MACBETH approach; more specifically, we will present how the technique described in section 2 can be used, not only for the determination of the value scales at the level of each criterion but also for the determination of the "weights" of these criteria. We will also refer to some of the numerous functionalities of M-MACBETH, in particular the different sensitivity and robustness analyses that this software allows one to make. These capabilities are another strong point of the MACBETH approach; they are liable to greatly assist the emergence of convictions that enable to move forward the decision making process.

## 2. How to obtain cardinal information

### 2.1. Concept of ordinal information

Let X be a finite set of elements. We will say that we have "ordinal information" (regarding the attractiveness of the elements of X) when these elements are ranked (with equals or not) in accordance with decreasing attractiveness. In such a case, it is possible to associate to each element x of X a number v(x) that satisfies the following measurement conditions (ordinal measurement conditions):

CM1) $\forall$ x, y $\in$ X: [x is more attractive than y $\Leftrightarrow$ v(x) > v(y)]

CM2) $\forall$ x, y $\in$ X: [x is as attractive as y $\Leftrightarrow$ v(x) = v(y)].

A numerical scale v: X $\rightarrow \Re$: x $\rightarrow$ v(x) satisfying the ordinal measurement conditions is unique up to a strictly increasing transformation, that means that:

1) if v satisfies CM1 and CM2, then, for each strictly increasing function $\phi: \Re \rightarrow \Re$, ($\phi \circ v$) satisfies also CM1 and CM2
2) if v* and v** are two functions X $\rightarrow \Re$ satisfying CM1 and CM2, then, there exists a strictly increasing function $\phi^* : \Re \rightarrow \Re$ such that $v^{**} = \phi^* \circ v^*$.

Such a scale is called an ordinal scale.

Let us point out that, if we write P the binary relation defined on X by

[$\forall$ x,y $\in$ X: xPy $\Leftrightarrow$ x is more attractive than y],

then one can prove that P determines an ordinal information

if and only if P is asymmetric and negatively transitive

if and only if P is asymmetric and transitive and I is transitive

where I is the binary reflexive relation defined on X by

[$\forall$ x, y $\in$ X: xIy $\Leftrightarrow$ $x \not P y$ and $y \not P x$ )].

## 2.2. Concept of cardinal information

We will say that we have "cardinal information" (regarding the attractiveness of the elements of X) when

1) we have ordinal information (regarding this attractiveness)
2) $\forall$ (x,y), (z,w) $\in$ P, we have a strictly positive number that we will write R((x,y)/(z,w)) ("measuring" the ratio of differences of attractiveness between x and y on the one hand, and z and w on the other hand), these numbers satisfying the following conditions:
   a) $\forall$ (u,w), (w,x), (s,t) $\in$ P: R((u,w)/(s,t)) + R((w,x)/(s,t)) = R((u,x)/(s,t))
   b) $\forall$ (u,w), (x,y), (s,t) $\in$ P: R((u,w)/(s,t)) $\times$ R((s,t)/(x,y)) = R((u,w)/(x,y)).

When we have such information available, it is possible to associate a number v(x) to each element of X satisfying the following measurement conditions (cardinal measurement conditions):

CM1) $\forall$ x, y $\in$ X: [xPy $\Leftrightarrow$ v(x) > v(y)]
CM2) $\forall$ x, y $\in$ X: [xIy $\Leftrightarrow$ v(x) = v(y)]
CM3) $\forall$ (x,y), (z,w) $\in$ P: R((x,y)/(z,w)) = [v(x)-v(y)] / [v(z)-v(w)].

A numerical scale v: X $\rightarrow \Re$: x $\rightarrow$ v(x) satisfying the cardinal measurement conditions is unique up to a positive affine transformation, that means that:

1) if v satisfies CM1, CM2 and CM3, then, $\forall$ $\alpha$, $\beta \in \Re$ with $\alpha$>0, $\alpha$.v + $\beta$ also satisfies CM1, CM2 and CM3
2) if v* and v** are two functions X $\rightarrow \Re$ satisfying CM1, CM2 and CM3, then two real numbers exist, $\alpha$* and $\beta$* with $\alpha$*>0, such that $\forall$ x$\in$X: v**(x) = $\alpha$*.v*(x) + $\beta$*.

Such a scale is called an interval scale.

Let us observe that a way to give cardinal information (regarding the attractiveness of the elements of a finite set X) consists of placing the elements of X on a vertical axis so that:

1) $\forall$ x, y $\in$ X : xPy $\Leftrightarrow$ x is positioned above y on the vertical axis
2) $\forall$(x,y),(z,w) $\in$ P, R((x,y)/(z,w)) = d(x,y) / d(z,w), where $\forall$ s, t $\in$ X, d(s, t) is the measure, in some unit, of the distance that separates the elements s and t on the vertical axis.

The use of the additive aggregation model, in the MACBETH approach, as reference scope to ensure the coherence of the process of building preferences requires that cardinal information concerning the attractiveness of the elements of a finite set be obtained from the decision makers.

## 2.3. Regarding the transition from ordinal information to cardinal information

### *2.3.1. Introduction*

The transition from ordinal information to cardinal information constitutes a considerable jump in terms of the wealth of the information. When we have ordinal information concerning the attractiveness of the elements of a finite set X, entering the cardinal world reveals the origin of the notion of strength of preference, that we designate in MACBETH by "difference of attractiveness" to avoid any misunderstanding with any other approach that uses another type of cardinal preference information, such as ratio judgements. Even if people feel as though they easily understand the notion of difference of attractiveness, it is rare for them to think about it in any depth. Therefore, if we want the collected information to be reliable, in the sense that it accurately represents what the person feels, it makes little sense to ask for cardinal information without some previous preparation. We must, therefore, develop stages to bridge the long way between ordinal and cardinal information.

*2.3.2. MACBETH questioning and MACBETH preference information*

In order to facilitate the transition from ordinal to cardinal, we decided to begin by asking the evaluator J questions whose formulation used only two elements of X; our goal was to make the first "cardinal" questions, those regarding the notion of difference of attractiveness, as simple and easy as possible.

After various experiences in practical cases, we decided upon the following questioning mode. When xPy (x more attractive than y), what do you feel is the difference of attractiveness between x and y: "very weak", "weak", "moderate", "strong", "very strong" or "extreme"? If unsure about this difference of attractiveness, J is allowed to choose several successive categories.

Note that the name "MACBETH approach" comes from this mode of questioning: **M**easuring **A**ttractiveness by a **C**ategorical **B**ased **E**valuation **T**ec**h**nique.

By definition, we will say that we have "MACBETH preference" information (concerning the attractiveness of the elements of X) when:

1) we have ordinal information concerning this attractiveness (asymmetric and negatively transitive binary relation P on X)
2) $\forall$ (x,y) $\in$ P, the couple (x,y) has been allocated to one of the six categories of difference of attractiveness "very weak", "weak", "moderate", "strong", "very strong", "extreme" or to the union of several successive categories among these.

*2.3.3. Concept of precardinal information*

We will say that the MACBETH preference information (concerning the attractiveness of the elements of X) is "precardinal" when it is compatible with cardinal information, that is when it is possible to extend it to cardinal information, without modifying it but only by refining it.

The following proposition is easy to prove.

Proposition:
The MACBETH preference information (concerning the attractiveness of the elements of X) is precardinal if and only if it is possible to associate to each element x of X a number v(x) satisfying the following conditions:

CM1) $\forall$ x, y $\in$ X: [xPy $\Leftrightarrow$ v(x) > v(y)]
CM2) $\forall$ x, y $\in$ X: [xIy $\Leftrightarrow$ v(x) = v(y)]
C3) $\forall$ (x,y), (z,w) $\in$ P, if it results from the MACBETH information that the difference of attractiveness between x and y is greater than the difference of attractiveness between z and w,
then v(x) – v(y) > v(z) – v(w).

A numerical scale v: $X \rightarrow \Re$: $x \rightarrow v(x)$ satisfying the conditions CM1, CM2 and C3 is called a precardinal scale.

Regarding uniqueness, the notion of precardinal scale is between those of ordinal scale and cardinal scale in the sense that: if v* satisfies CM1, CM2 and C3, it is possible to generate the set of scales satisfying CM1, CM2 and C3 by composing v* with each element of a set $\mathcal{J}$ of functions from $\Re$ in $\Re$ so that

$$G_1 \subsetneq \mathcal{J} \subsetneq G_2$$

where $G_1$ is the set of strictly increasing functions from $\Re$ in $\Re$ and $G_2$ is the set of positive affine functions from $\Re$ in $\Re$ [functions as $f(r) = \alpha.r + \beta$ where $\alpha, \beta \in \Re$ and $\alpha > 0$]. Note that, even though $G_1$ and $G_2$ are groups for the composition rule, that is not the case for $\mathcal{J}$.

## 2.4. Process to obtain cardinal information in MACBETH

### *2.4.1. Introduction*

In order to help a person (or a group of persons) J to provide cardinal information concerning the attractiveness of the elements of a finite set X, the “MACBETH philosophy” consist of getting J to, through an interactive process of reflection and learning, progressively refine his (or their) judgements. Generally, this process is managed in accordance with the great lines exposed hereafter. Let us mention that there are other possibilities of interaction already explored in practical cases (see section 3.3.2); as a matter of fact, the M-MACBETH software includes several options (for example, the possibility of accepting answers such as “I don’t know” when comparing two actions) that assist the reflection process even in extremely difficult cases – see [44] for a deep presentation and discussion of all those options.

### *2.4.2. Obtaining ordinal information*

We ask J for a ranking of the elements of X by order of decreasing attractiveness. When we find J to have difficulties giving such a ranking, we propose instead that J compare the elements two at a time: is one of the two elements more attractive than the other and if yes, which one? For each of J’s answers, the M-MACBETH software will test the compatibility of the collected information with ordinal information. If an incompatibility is detected, J will receive a warning message (”no ranking”) and a discussion will begin. To facilitate such a discussion, the software allows the source of the problem to be graphically displayed (see Figure 1) and even provides suggestions of the judgement modifications that would make the pairwise comparison information compatible with ordinal information.

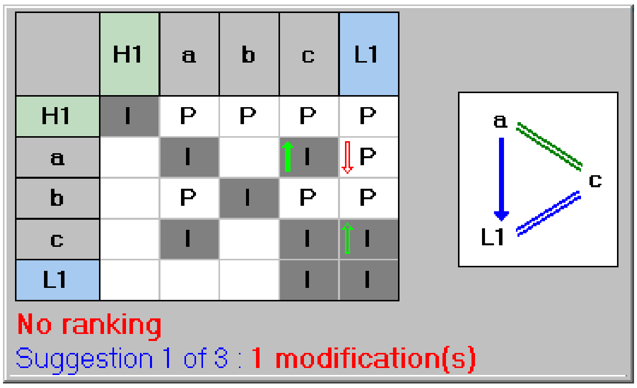



Figure 1. Example of ordinal incompatibility. ($H_1$ and $L_1$ are reference actions which role will be explained in section 3.1.2)

### *2.4.3. Obtaining a precardinal MACBETH preference information*

When we have ordinal information, the entrance into the cardinal world is made through MACBETH questioning. The questions may be asked in any order and can be stopped at any moment: if xPy and if questioning were stopped before J answered the MACBETH questioning concerning the couple (x,y), we would consider that the MACBETH preference information relating to this couple belongs to the union of all categories of difference of attractiveness (it is a borderline answer but it is acceptable and appears in the M-MACBETH software with the mention: “positive” difference of attractiveness).

For each of J’s answers, the software tests the compatibility of the information collected with cardinal information. If an incompatibility is detected, J receives a warning message (“inconsistent judgements”) and a discussion will begin. To facilitate such a discussion, the software allows the source of the problem to be graphically displayed and even provides suggestions of the judgement modifications that would make the information precardinal (see Figure 2). The software has the ability to display the source of an incompatibility and provide all of the possible ways to resolve it.

| | H1 | b | a | c | L1 |
|---|---|---|---|---|---|
| H1 | no | very weak | strong | positive | positive |
| b | | no | weak-mod | moderate | positive |
| a | | | no | weak | positive |
| c | | | | no | strong |
| L1 | | | | | no |

Inconsistent judgements
Suggestion 1 of 4 : 1 modification(s)

| Problems | Diff. | Couples | | Couples | Diff. |
|---|---|---|---|---|---|
| 1 | strong | H1 - a | > | b - c | moderate |
| | weak | a - c | > | H1 - b | very weak |

Figure 2. Suggestion of judgement modifications.

### *2.4.4. Presentation of a precardinal scale*

Upon demand, the M-MACBETH software can, at any moment in which the MACBETH preference information is precardinal, determine an example of a scale that satisfies the conditions CM1, CM2 and C3 (see section 2.3.3). This scale can be presented in a numerical (see Figure 3) or graphical ("thermometer") way with the ability, in the latter case, to choose whether or not to display the numerical value associated with each element of X (see Figure 4).

| | H1 | b | a | c | L1 | Macbeth transformed | Macbeth basic |
|---|---|---|---|---|---|---|---|
| H1 | no | very weak | mod-strg | positive | positive | 100.00 | 9.00 |
| b | | no | weak-mod | strong | positive | 88.89 | 8.00 |
| a | | | no | weak | v. strong | 66.67 | 6.00 |
| c | | | | no | strong | 44.44 | 4.00 |
| L1 | | | | | no | 0.00 | 0.00 |

Consistent judgements

Figure 3. Numerical display of a precardinal scale. (The "MACBETH transformed scale" is a direct linear conversion of the "MACBETH basic scale"; by default, the M-MACBETH software assigns "transformed" scores of 100 and 0 to the references $H_1$ and $L_1$, respectively; if references are not explicitly defined, the 100 and 0 scores are automatically assigned to the most and least attractive actions, respectively.)

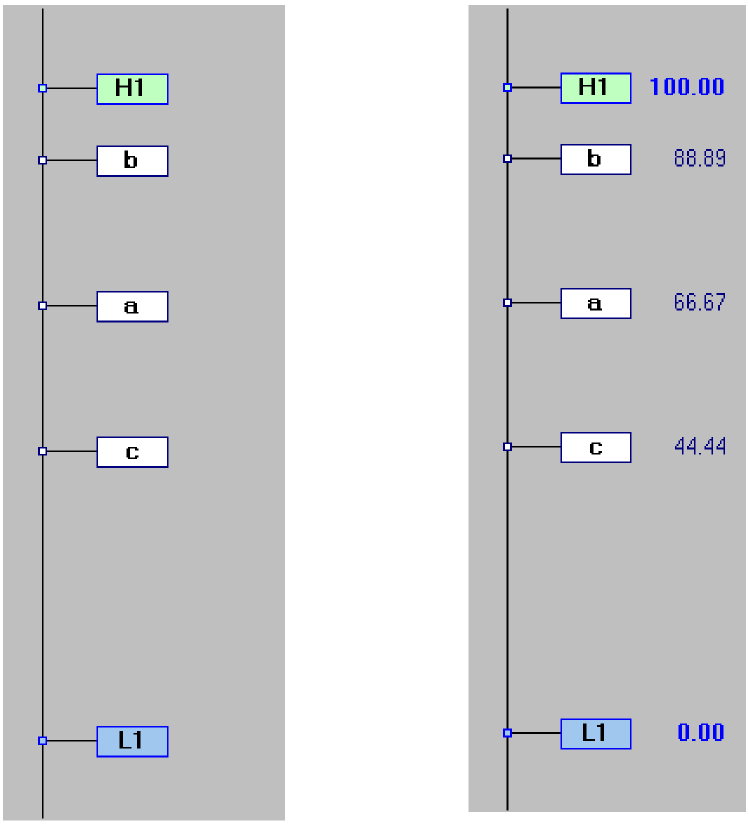



Figure 4. Graphical displays of a precardinal scale.

The basic MACBETH scale can be obtained by solving the following linear program (PL-MACBETH):

Min $v(x^+)$
under the constraints:

$\forall\, x, y \in X$: $xPy \Rightarrow v(x) \geq v(y) + 1$

$\forall\, x, y \in X$: $x\,I\,y \Rightarrow v(x) = v(y)$

$\forall\, (x,y), (z,w) \in P$, if the result from the MACBETH preference information is that the difference of attractiveness between x and y is greater than the difference of attractiveness between z and w,
then $v(x) - v(y) \geq v(z) - v(w) + 1 + \delta(x,y,z,w)$

$v(x^-) = 0$;

where

- P is the asymmetric and negatively transitive binary relation defined on X that models the ranking of the elements of X by order of decreasing attractiveness
- I is the binary relation defined on X by, $\forall\, x, y \in X$: $x\,I\,y \Leftrightarrow x \not P y$ and $y \not P x$
- $x^+$ is an element of X so that, $\forall\, x \in X$: $x^+(P \cup I)x$
- $x^-$ is an element of X so that, $\forall\, x \in X$: $x(P \cup I)x^-$
- $\delta(x,y,z,w)$ is the minimal number of categories of difference of attractiveness between the difference of attractiveness between x and y and the difference of attractiveness between z and w.

The optimal solution of this linear program is not necessarily unique; multiple optimal solutions may exist, with different values for at least one element $x \in X \backslash \{x^-, x^+\}$. To mathematically ensure the uniqueness of the basic MACBETH scale, supplementary technical linear programs are used in the software, after running PL-MACBETH [44]. Obtaining this precardinal scale is not an end in itself, but a step forward in the learning process towards a cardinal representation of J's preferences.

We call every scale obtained by the composition of the basic MACBETH scale with a positive affine function from $\Re$ in $\Re$ a transformed MACBETH scale. In Figure 3, the transformed MACBETH scale has been obtained by imposing that this positive affine transformation must be such that $v(H_1) = 100$ and $v(L_1) = 0$ (possible in the software). Of course, every transformed MACBETH scale is also a precardinal scale.

*2.4.5. From precardinal to cardinal: discussion around a scale*

The precardinal MACBETH scale is graphically presented to J and is commented upon; we explain, in particular, to J that the elements of X have been positioned on a vertical axis so that:

(1) $\forall\, x, y \in X$: x is positioned above y $\Leftrightarrow$ J has declared that x is more attractive than y

(2) $\forall\, x, y, z, w \in X$ with x more attractive than y and z more attractive than w,
if the result from the MACBETH information is that the difference of attractiveness between x and y is greater than the difference of attractiveness between z and w,
then the distance between x and y is greater than the distance between z and w;

and these assertions are verified in some cases.

We point out to J that many other possible placements of the elements of X exist which can graphically represent the verbal information that J has provided and we show to J that:

- ✓ when J selects an element (with the mouse), an interval (in red in the software) appears around this element
- ✓ by moving the mouse, J can modify the position of the selected element but only within that interval
- ✓ by modifying the position of an element inside the interval, J obtains a new positioning of the elements of X compatible with J's verbal judgements – in the sense that it verifies conditions (1) and (2)
- ✓ an attempt to position the selected element outside of the interval without modifying the position of at least one other element of X, would make the resulting positioning of the elements no longer verify conditions (1) and (2).

J has the opportunity to modify the positioning of the elements of X, should J want to, so that, not only conditions (1) and (2) are satisfied but also, the relative distances which appear between the elements reflect the relative differences of attractiveness that J deems to exist between these elements (see example in Figure 5).

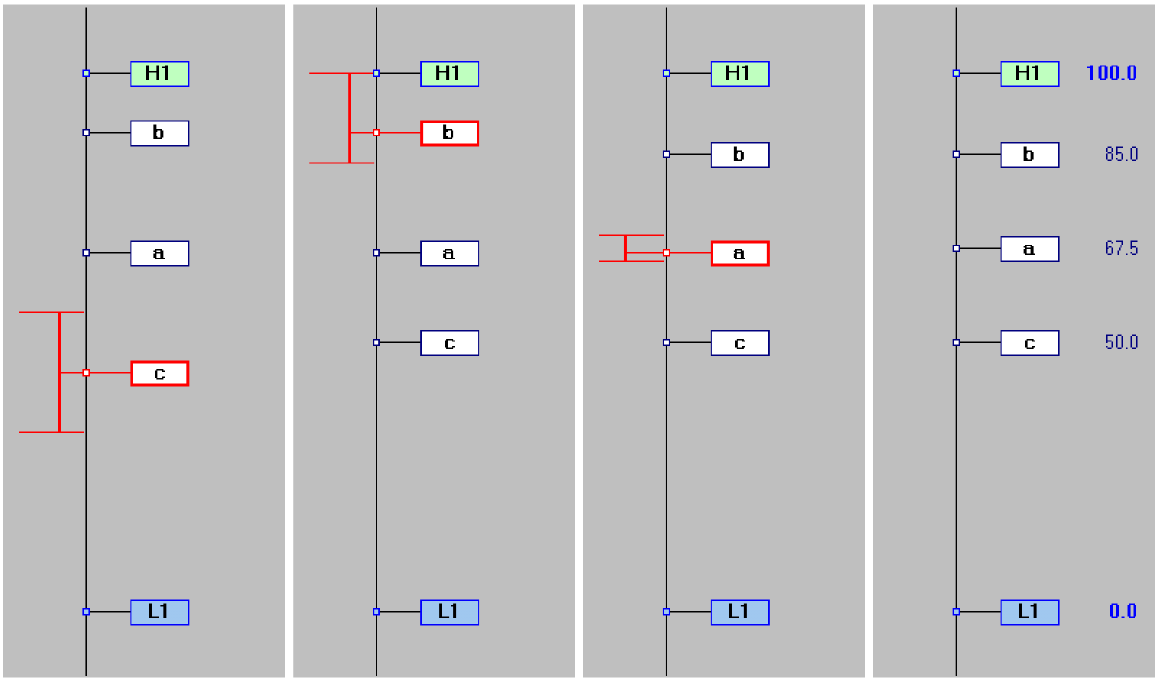



Figure 5. Adjusting the scale.

We then make the cardinal information concerning the attractiveness of the elements of X available, so as to allow the determination of an interval scale on X "measuring" the attractiveness of the elements of X (for J).

Note that, if J absolutely wants to position an element outside its interval, the software will permit J to do so (not by moving the mouse), provided that the new position of the element allows, by modifying the positions of the other elements, the compliance with conditions (1) and (2) – this is graphically represented by a new interval (in green in the software) around the considered element (this new interval will always include the first interval within it, as shown in Figure 6).

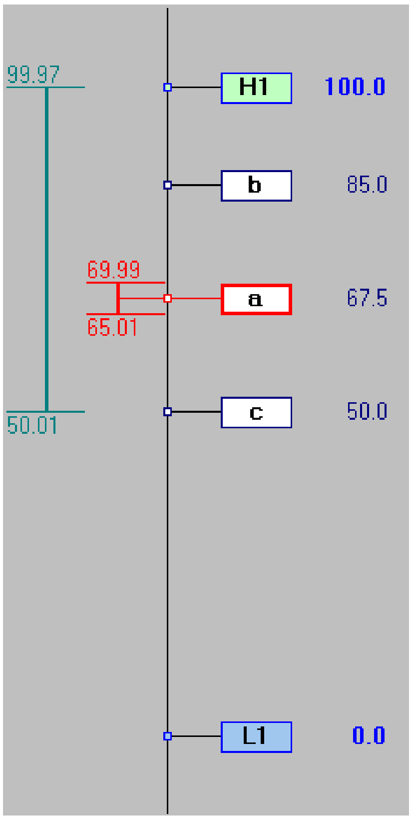


Figure 6. The green interval.

**3. The MACBETH multicriteria decision aid process**
3.1. Model structuring
*3.1.1. Criteria*

The multicriteria analysis process usually starts with a discussion of the different points of view about how potential actions should be evaluated. A "point of view" (or a "concern") is any aspect that (1) emerges during the discussion as relevant for evaluating potential actions, (2) in the perspective of at least one actor, and (3) has a value meaning that is well defined and understood by everybody so as to avoid ambiguity and misunderstanding. The role of the facilitator during the initial stage of structuring a model of values consists of stimulating the reflection process of the actors in order to progressively:

- make all sorts of points of view emerge, clarify their meaning, and analyse why and in what way they are relevant, in order to
- identify and agree on a shared family of "fundamental points of view" ($PV_1$, $PV_2$, …, $PV_K$), each of these being an individual point of view, or a cluster of several of them, that will serve as a decision "criterion" in terms of which the actors agree to a separate evaluation of potential actions (note that this implies preferential independence).

The points of view are commonly organized in a tree form, usually referred to as a "value tree". This provides a useful visual overview of the structure of the points of view in several levels of increasing specification (this is why they are sometimes alternatively called "hierarchies", although this designation implies the existence of some form of "subordination" to parent nodes, that may not be present). Some criteria (fundamental points of view) can appear immediately after the top ("overall") node; other criteria are peers within an area of concern represented by a parent node. When the tree's purpose is to emphasise only the criteria, it can be structured in such a way that each of the criteria appears at the end of a branch. In other cases, the tree can be structured so as to provide a detailed overview of the value issues, in which case the criteria can appear at any level, rather than only at the terminating nodes.

*<u>Example</u>*
*Person J, responsible for the purchase of an inkjet printer, has made a short list A of five models with similar prices:*

*A = {PH, Conan, Espon, Sister, Nomark}*

*J started the MACBETH multicriteria process by introducing these five potential actions (options of choice) into the M-MACBETH software. The software was then used to develop the value tree shown in Figure 7, in which J identified three fundamental points of view (the decision criteria, shown in bold characters in the tree):*

*$PV_1$: black printing quality (abbreviated: Black)*
*$PV_2$: colour printing quality (abbreviated: Colour)*
*$PV_3$: black printing speed (abbreviated: Speed).*

*Note that the point of view "Printing quality" is not a decision criterion, it is an area of concern specified by two criteria.*

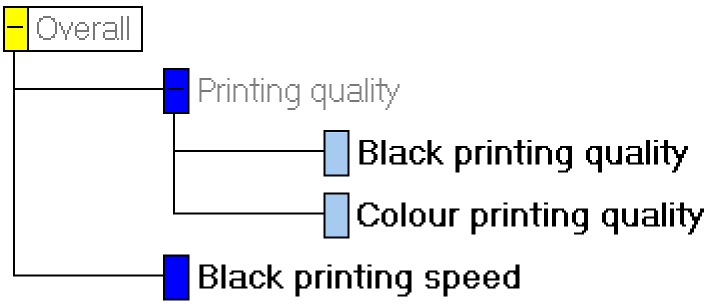



Figure 7. Value tree. (The structuring component of the M-MACBETH software was designed with the purpose of being flexible enough to allow for all sorts of value trees; that is why, each time a point of view is inserted in the tree, the software requires the user to specify whether or not it is to be taken as fundamental, that is, as a decision criterion.)

### *3.1.2. Explicit references*

In the MACBETH decision aid process, the assessor J is asked to specify, for each criterion $PV_i$, two particular "references" $H_i$ and $L_i$ (where $H_i$ is more attractive than $L_i$), which make sense to J because of their concrete nature (if they are described as particular levels of performance), or symbolic nature (if they are defined as particular levels of intrinsic value), or a combination of these. In accordance with the theory, fixing the value scores of these references permits one to obtain a unique numerical representation of cardinal intra-criterion preference information. However, their principal interest is that $\{(L_i, H_i) \mid i \in \{1, \ldots, K\}\}$ constitutes a particularly solid and well adapted basis for the assessment of inter-criteria preference information for weighting the criteria, as we will describe in section 3.3.2.

In the traditional multicriteria analysis literature, $H_i$ and $L_i$ are almost always implicitly defined as either the limits of a range of plausible performances or the performances of the "most attractive" and the "least attractive" actions in each criterion $PV_i$, for all i = 1, …, K. The M-MACBETH software allows the user to follow this traditional approach, because, when no explicit references are introduced in the software, such implicit references are automatically assumed. In our opinion, however, it is better to invite J to specify two concrete reference levels *a priori*, because not only will they make clear the value meaning of the considered criterion but also the difference of preference between such levels will be well understood. Our practical experience has led us to believe that the result of this *a priori* specification will even be better than that if, in addition to a concrete description of two reference levels, J makes an explicit statement about the intrinsic value of one or both of these references. For instance, we have observed that it is sometimes appropriate in practice to let $L_i$ correspond to the "status quo" or the "do nothing" option, as in [5], and let $H_i$ correspond to an option judged to be strongly more attractive than $L_i$, or a satisfactory aspiration level or a benchmark in $PV_i$. We have also often recommended [6, 11, [12, 15, 30] the explicit definition of $H_i$ as an intrinsically "good" (undoubtedly satisfying) level of performance and the explicit definition of $L_i$ as an intrinsically "neutral" (neither satisfying nor unsatisfying) level of performance, in each criterion $PV_i$. Experience has revealed that the effort required to identify the references $good_i$ and $neutral_i$ contributes significantly to the intelligibility of the criterion: it is one thing to say that an option is better than another in printing quality, for example, yet, it is quite another to specify what is meant by a good quality or a neutral quality. It has been observed [11, 12] that this very much facilitates the acquisition of cardinal preference information. Furthermore, an explicit statement regarding the nature of "neutral" enables one to use the notion of the "intrinsic" attractiveness of an action to distinguish, at the local level of each criterion $PV_i$, between

*an unattractive option on $PV_i$*, if it is less attractive than $neutral_i$. and

*an attractive option on $PV_i$*, if it is more attractive than $neutral_i$,.

If, moreover, an explicit statement regarding the nature of "good" is also available, one can also distinguish

*a very attractive option on $PV_i$*, if it is at least as attractive as "$good_i$".

Therefore, fixing a "neutral" reference for each criterion enables the highlighting of the "pros" and "cons" of each action by simply comparing its impacts with the "neutral" references in the various criteria. Similarly, fixing the reference "good" for each criterion enables to one to determine if an action is outstanding. In the evaluation phase, once an aggregation model is constructed, the references "good" and "neutral" enable one to use the notion of the overall intrinsic attractiveness of an action to assign to it one of the following categories:

- *very attractive (or outstanding) option*, when it is at least as attractive as the profile "good all over";
- *simply attractive option*, if it is more attractive than the profile "neutral all over", but less attractive than the profile "good all over";
- *neutral option*, if it is as attractive as the profile "neutral all over", and
- *unattractive option*, if it is less attractive than the profile "neutral all over".

A great advantage associated with defining the references intrinsically, rather than only in relative terms, is that it avoids situations in which an inappropriate action is chosen, simply because it is the best option in a set of unattractive options. Moreover, the comparison of the value associated with a change from neutral to good in one criterion with the value associated with a change from neutral to good in another criterion, rather than the changes or “swings” from least attractive to most attractive, makes it possible not only to derive weights for the criteria before actions are known [15], but also to avoid forcing J to consider unrealistic tradeoffs [11]. Last but not least, starting by defining neutral and good references can significantly facilitate the construction of descriptors, as will be detailed in the next section.

*Example (continued)*
*J defined references “good” and “neutral” in concrete terms, for each of the three criteria. Figures 8 and 9 show these references for “Black printing quality” and “Colour printing quality”, respectively; 10 and 6 pages per minute were fixed as indicating good and neutral black printing speeds, respectively (Figure 10).*

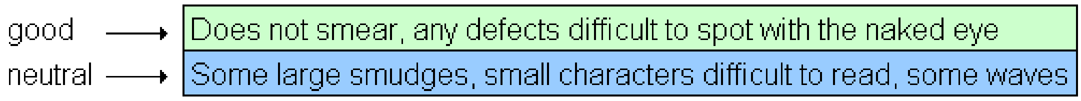



Figure 8. References of “Black printing quality”.

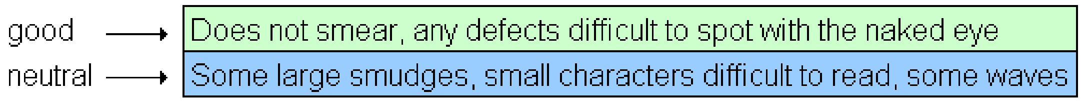



Figure 9. References of “Colour printing quality”.

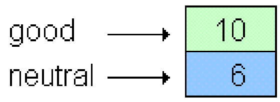



Figure 10. References of “Black printing speed” (pages/min).

### *3.1.3. Descriptors*

On the basis of a family of criteria it is already possible to use MACBETH to evaluate actions, by comparing them in terms of difference of attractiveness in each one of the criteria. While this may be valid in many cases, the transparency of the result can be limited. Alternatively to working directly with the actions, the introduction of descriptors of “impacts” (plausible performances, or consequences of potential actions against the criteria) will contribute decisively to well informed judgements as well as to a justified and transparent evaluation. A “descriptor” is an ordered set $D_i$ of plausible impact levels associated with a criterion (fundamental point of view) $PV_i$. It is intended to:

1. Operationalise the appraisal of the impact of each action in the criterion.
2. Make possible the objective description (insofar as this is possible) of the impacts of actions with respect to that criterion. The higher the level of objectivity associated with the appraisal of the impacts, the better understood (less ambiguous) and therefore the better accepted (less controversial) will the evaluation model be.
3. Better frame the evaluation model, by restricting (whenever appropriate) the range of impact levels to a plausible domain (from a most attractive or desirable level to a least attractive level). This impact-range can be defined by screening out impacts or options that are non-admissible or out-of-context.
4. Verify the ordinal independence of the corresponding criterion. If dependence is detected in this phase, feedback is necessary to re-structure the family of

criteria, so that “all other things being equal” (*ceteris paribus*) comparisons of actions may be individually made against each criterion.

The contextual operational requirement imposed to a descriptor will permit a criterion to be described in such a way that clarifies why it is important for the decision maker, and, when several actors are involved, each one will better understand the other ones’ values. This is also enhanced by the ordinal requirement imposed upon a descriptor, which gives to the descriptor the structure of an ordinal preference scale (either continuous or discrete). Note that this preference is limited to the criterion under consideration; hence, it is called “local” attractiveness, to distinguish it from “overall” attractiveness (i.e. at the level of the entire family of criteria).

Discussing the descriptors for the various criteria can also provide insights which will lead to a better definition of the value tree. For example, disagreement between actors arising from the ordinal requirement imposed upon a descriptor should be clarified at this stage, not postponed. A couple may agree that the size of the car they will both use is an important evaluation aspect, but, for instance, the husband may privilege a bigger car while his wife stands for a small one. Inspection of the reasons behind these contradictory preferences will eventually reveal different hidden fundamental points of view. The family of criteria then needs to be restructured, to include the new criteria instead of the size of the car. On the other hand, the discussion of alternative descriptors for the same criterion, in order to select the most appropriate in a specific decision context, can help to avoid redundancy in the model.

Depending on the context, a descriptor may be more or less directly related with the criterion. It is then useful to distinguish between direct and indirect descriptors:

- ✓ The levels of a “direct (or natural) descriptor” directly reflect effects (such as the number of people affected by respiratory diseases).
- ✓ The levels of an “indirect (or proxy) descriptor” indicate causes more than effects (such as the degrees of concentration of air pollutants that cause respiratory diseases).

It is not always certain that a natural descriptor is necessarily more adequate than a proxy one, namely for reasons of lack of information (see section 3.1.4). For similar reasons, or when the criterion has a subjective or intangible nature (image, for example), or it is a cluster of several interrelated elementary aspects that are judgementally dependent, neither a direct nor an indirect descriptor exists or can be assessed. In such cases, a “constructed descriptor” has to be developed within the specific context. The levels of a constructed descriptor can be qualitative or quantitative (or mixed) and several types can be conceived, like:

- ✓ “One-dimension qualitative scales”: finite sets of verbal descriptions of one-dimension plausible impacts.
- ✓ “Multidimensional descriptors”: sets of multidimensional impact levels, each one for instance describing a plausible consequential scenario involving several elementary aspects. They can be multidimensional (discrete) scales or indices. The former are constructed by individual descriptions of each of their multidimensional impact levels, the latter are defined comprehensively by a formula that combines several dimensions analytically.
- ✓ “Pictorial descriptors”: finite set of reference visual representations of impacts (like sketches, drawings, pictures, videos, computer simulations, etc.).

It is not always an easy task to develop a constructed descriptor. As its construction inherently requires value judgments, conflicts of viewpoints may arise in group-structuring; different actors may tend to privilege different descriptors, even if all of them agree on the criterion. Unavailability of input data can further complicate the task. It may also be difficult for J to directly rank the levels of a complex qualitative descriptor in terms of relative attractiveness. In such cases, the MACBETH process for obtaining ordinal information described in section 2.4.2 can be of great help.

Our experience in applying multicriteria decision analysis to public policy decision contexts has made it clear that the difficultness associated with constructing a descriptor for a criterion $PV_i$ can be very much facilitated if one starts by taking the two fixed references $good_i$ and $neutral_i$ as anchors to develop additional levels. One can then stimulate reflection by asking J simple questions like: Can you conceive a very attractive impact, that is, better than good? What could an attractive impact that is worse than good be? Can you accept an unattractive impact, that is, worst than neutral, and if so what could an example of such an impact be?

The M-MACBETH software can deal with qualitative or quantitative descriptors. In the case of a continuous descriptor (that is, when $D_i$ is an interval of real numbers to which information of relative local attractiveness is associated with regards to the respective criterion $PV_i$) the software requires the impact range to be split up into a few intervals and takes the limits of these intervals (not the intervals) as a reference set of quantitative levels $S = \{d_1, d_2, \ldots, d_q\} \subset \Re$. The reason for this is to enable these levels to be, subsequently, compared by using the MACBETH interactive process described in section 2, and obtain a cardinal scale $v_i$: $S \to \Re$: $d \to v_i(d)$; this scale will be extended to a piecewise linear value function, simply defining a $D_i \to \Re$ scale by linear interpolation.

The M-MACBETH software always requires the user to define a descriptor for each criterion, but it offers the possibility of taking the set of potential actions introduced into the software as the levels of the descriptor. Therefore, the software can deal with cases in which one wants to compare the options directly. Note that a descriptor defined as an ordered set of actions is a limit type of constructed descriptor that is perfectly admissible when the actions are familiar to the decision maker. In this and all other cases, each of the two references $H_i$ and $L_i$ can, but do not have to, be a level of the descriptor $D_i$ of the criterion $PV_i$; in any case, we will always write $X_i = D_i \cup \{H_i, L_i\}$.

*Example (continued):*
*Figures 11 and 12 show the qualitative descriptors associated with "Black printing quality" and "Colour printing quality", respectively. The quantitative descriptor associated with "Black printing speed" is the number x of pages printed per minute, with x ranging from 11 (most attractive) to 5 (least attractive).*

| | Description of the impact levels |
|---|---|
| Black1 | Quality comparable to that of a laser printer |
| Black2 | Does not smear, any defects difficult to spot with the naked eye |
| Black3 | Smudges are rare and minuscule, some slight defects (almost unnoticeable) |
| Black4 | Some smudges (not very significant), few or no "waves" |
| Black5 | Some large smudges, small characters difficult to read, some "waves" |
| Black6 | Gutenberg could have done as well |

Figure 11. Descriptor of “Black printing quality” (abbreviated: “Black”).

| | Description of the impact levels |
|---|---|
| Colour1 | Printed pictures looks like developed at a professional studio |
| Colour2 | Excellent colours, can be used to reproduce pictures |
| Colour3 | Smudges are rare and minuscule, some slight defects (almost unnoticeable) |
| Colour4 | Colour quality is good enough to print drawings but not pictures |
| Colour5 | Colour quality just enough to be used to enhance text |
| Colour6 | The blue is not blue, the red is not red, the green is not green |

Figure 12. Descriptor of “Colour printing quality” (abbreviated: “Colour”).

### *3.1.4. Impacts of a potential action*

When selecting or constructing a descriptor, one has to bear in mind that it must be operational for the estimation and analysis of the impacts of actions in the respective criterion. This assumes that, for each criterion, it will be possible to appraise the impact of implementing each potential action from among the levels of its descriptor. Why and for what purpose would one select or laboriously define a descriptor that it is not ultimately suitable to appraise the impacts of the options, for instance, due to lack of information, or time, or any other necessary resources? For example, in the case of the comparison of alternative packages of new roads in the Lisbon Region [5], a natural descriptor of “air pollution effects” was the “increase in number of people affected by pollutant emissions” after the construction of the roads. However, the use of this descriptor to estimate impacts would require a rigorous and expensive mathematical model that could not be afforded. That descriptor had to be abandoned and, alternatively, a rough (but effective) qualitative descriptor had to be constructed, {“decrease in emissions”, “no significant change in emissions”, “increase in emissions”}. Expert judgements were then used to appraise the qualitative impacts of the packages. In conclusion, the definition of a descriptor is only complete when a procedure for estimating the impacts of the options is also associated with it. Such a procedure can consist of the use of a mathematical (optimisation, simulation, etc.), econometric, or technical model or formula, an empirical rule, a panel of experts, a survey technique, an experimental process, etc.

Mathematically, the “impact” of an action a on criterion $PV_i$ is a sub-set of $D_i$. The definition of the impact of an action in the M-MACBETH software always requires one to begin by choosing one and only one level of $D_i$, which we will write $\pi_i(a)$. We call $\pi(a) = [\pi_1(a), \pi_2(a), \ldots, \pi_K(a)]$ the “impacts-profile” of action a and “impact table” the organised collection of the $\pi_i(a)$, for all $a \in A$ (A being the set of actions) and for all $PV_i$ ($i = 1, \ldots, K$).

*<u>Example</u> (continued)*
*The second to sixth lines in Figure 13 represent the impact profiles of the five printers, which together constitute the impact table of the example. For instance, J defined for "PH":*

$$\pi(PH) = [Black3, Colour2, 7 \text{ pages per minute}]$$

*The two bottom lines in Figure 13 show the impact profiles of [good] and [neutral], each one formed by the references "$good_i$" and "$neutral_i$", respectively, already established in each criterion $PV_i$. Thus, [good] and [neutral] can be seen as the impact profiles of the two fictitious reference options "good all over" and "neutral all over", respectively.*

| Printers | Black printing (qualitative) | Colour printing (qualitative) | Speed (pages/min) |
|---|---|---|---|
| PH | Black3 | Colour2 | 7 |
| Conan | Black1 | Colour4 | 9 |
| Espon | Black5 | Colour2 | 8.5 |
| Sister | Black6 | Colour3 | 10.5 |
| Nomark | Black3 | Colour3 | 9 |
| [good] | Black2 | Colour2 | 10 |
| [neutral] | Black5 | Colour5 | 6 |

Figure 13. Impact table and reference profiles

When no uncertainty is involved in the procedure of impact estimation, an impact table provides an overview of the impacts of the options in the criteria identified in the structuring process. This is, of course, the most desirable case. By comparing the impacts of options on each criterion one can conclude whether or not one option is better than another. At the same time, the comparison of impacts to the references $neutral_i$ and $good_i$ serves to determine whether an option is very attractive, simply attractive or unattractive with regard to criterion $PV_i$. These are ordinal results that provide qualitative knowledge about the local attractiveness of the actions. They are not, however, sufficient to specify whether one option is "better" or "much better" than another, nor how attractive or unattractive an option is in each criterion. These types of cardinal results require deeper investigations of preferences and can only be obtained using methods that appraise preference strengths of one impact over another. This is the subject of quantitative value analysis (see section 3.3).

*<u>Example</u> (continued)*
*Figure 14 shows the rankings of the printers by decreasing order of local attractiveness in each of the criteria, derived directly from Figure 13. Figure 14 offers J interesting information, both of relative and intrinsic nature. Conan provides an example of the former type, as it simultaneously has the best black and the worst colour printer quality among all the printers considered by J. Conan also provides an example of the latter type, as it is a very attractive printer (because it is better than good) in terms of black printing quality, but it is simply attractive in terms of both colour printer quality and printing speed.*

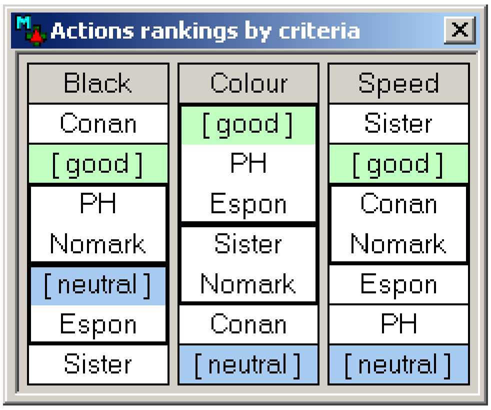



Figure 14. Rankings (“indifference” is indicated by a border around two or more designations).

Several types of uncertainty may, however, substantially affect the estimation of precise impacts. Bouyssou [39] offers an interesting discussion of what he considers to be the four main sources of uncertainty: “the ‘map’ is not the ‘territory’” (meaning that an impacts-profile of an option can not be taken as the option in itself), “the ‘future’ is not a ‘present’ to come” (referring to the classical notion of uncertainty in Decision Analysis), “the data are not the result of exact measurement”, and “the model is not the description of a real entity independent of the model” (highlighting, in a constructive perspective, that the way the modelling process is conducted influences the model in itself). The extent to which it is worthwhile to formalise and detail the modelling of uncertainty phenomena is indissoluble from what a “requisite model” [52] is in the specific decision-context: that is, what is the necessary and sufficient information that should be incorporated in the evaluation model to resolve the issues at hand.

When J considers that the uncertainty relating to the determination of the impact of an action a on criterion $PV_i$ is great, the M-MACBETH software lets J specify, in addition to $\pi_i(a)$, two levels of $X_i$ between which J is sure to see the expression of the effect of action a on $PV_i$. This “interval” will be used in the robustness analysis of the results provided by the additive aggregation model (see section 3.5).

*Example (continued)*
*J had doubts about the black printing quality of “Conan”, hesitating between the impact-levels “quality comparable to that of a laser printer” (Black1), which seemed to be closer to what J observed (see Figure 13) and “does not smear, any defects difficult to spot with the naked eye” (Black2). This uncertainty information was therefore introduced into the software as shown in Figure 15.*

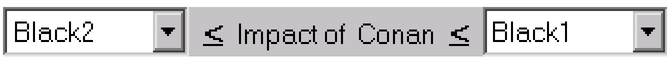



Figure 15. “Interval” of impact uncertainty.

### 3.2. Dominance analysis

The analysis of the impacts-profiles of the actions can yield, by itself, important insights about the relative and intrinsic attractiveness of the actions, even if only in ordinal terms. Relative analysis consists of comparing impacts of different actions, namely using the concept of “dominance”. Analysis of intrinsic attractiveness can be technically reduced to the relative analysis that consists of the individual comparison of the impacts-profile of each action with reference profiles.

Let a and b be any two potential actions. One says that:

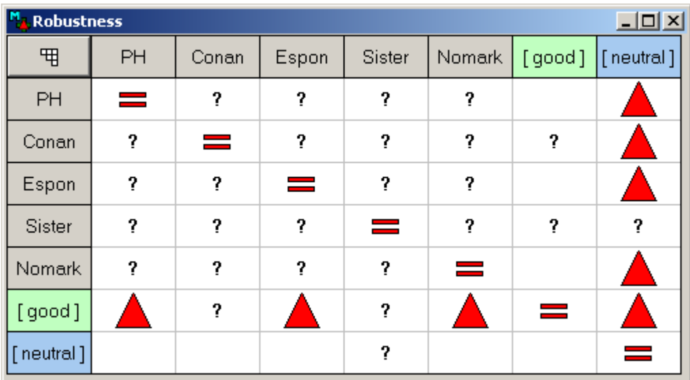
a (strictly) dominates b if and only if

- ✓ the impact of b is not more attractive than the impact of a in any criterion and
- ✓ the impact of a is more attractive than the impact of b in at least one criterion.

The results of dominance analysis can be synthesised in a "dominance table".

<u>Example</u> *(continued)*

*The dominance table in Figure 16, produced by the M-MACBETH software for the impact table in Figure 13, reveals that dominance does not hold between any two printers (this is represented by a question mark). There are, however, several dominance situations (represented by triangles in Figure 16) involving a printer and one of the two reference profiles [good] and [neutral]. They offer several pieces of relevant information for J; for instance, from the last column of Figure 16, it is robust to assert that, considering only the impacts of the printers and the reference levels defined, "PH", "Conan", "Espon" and "Nomark" are globally attractive printers (note that the "?" between "Sister" and [neutral] does not mean that "Sister" is globally unattractive; it only indicates that, with the information available at this stage, the assertions "Sister is an attractive printer" and "Sister is an unattractive printer" are not robust.*




| | PH | Conan | Espon | Sister | Nomark | [good] | [neutral] |
|---|---|---|---|---|---|---|---|
| PH | = | ? | ? | ? | ? | | ▲ |
| Conan | ? | = | ? | ? | ? | ? | ▲ |
| Espon | ? | ? | = | ? | ? | | ▲ |
| Sister | ? | ? | ? | = | ? | ? | ? |
| Nomark | ? | ? | ? | ? | = | | ▲ |
| [good] | ▲ | ? | ▲ | ? | ▲ | = | ▲ |
| [neutral] | | | | ? | | | = |

Figure 16. Dominance analysis.

Dominance analysis is the simplest type of (global) robustness analysis (see section 3.5). Unfortunately, its results are usually too poor to derive a complete ranking of the actions in terms of overall attractiveness or a most preferred option. This will require, in almost all cases, the collection of cardinal preference information, from J, both at the local and global levels.

## 3.3. Building a cardinal value model

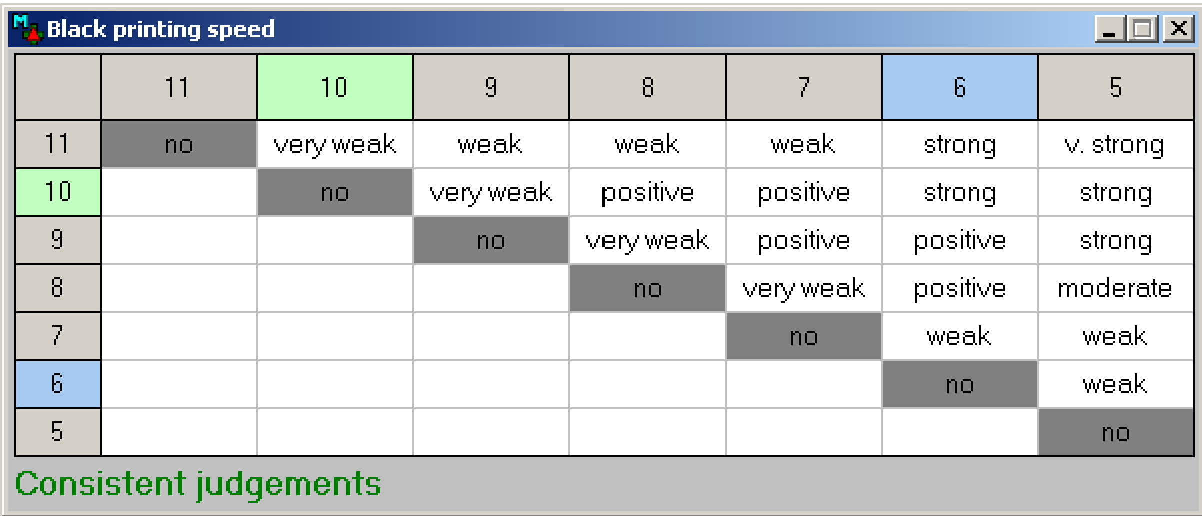
### *3.3.1. Acquisition of local preference information*

<u>Example</u> *(continued)*

*Observe, in the last column of Figure 13, the black printing speeds estimated for the printers. For instance, the difference of speed between Sister and Conan (10.5-9=1.5 pages/min) is equal to the difference of speed between Espon and PH (8.5-7=1.5 pages/min). Nonetheless there is no evidence that suggests that, with regard to this* *criterion only, when confronted with such input data J would necessarily feel that the difference of attractiveness between Sister and Conan is the same as the difference of attractiveness between Espon and PH. Such a statement can only be made after J has provided cardinal preference information. (That is cardinal information regarding the attractiveness of speed levels, not cardinal information about speed.)*

Cardinal (intra-criterion) information regarding the local attractiveness (relating to $PV_i$ only) of the elements of $X_i$ (for each $i \in \{1, 2, \ldots, K\}$) can be acquired from J, in accordance with the MACBETH learning process described in section 2. We will hereafter write $v_i$ a numerical scale defined on $X_i$ that satisfies the cardinal measurement rules CM1, CM2 and CM3: $v_i$ is an interval scale which "measures" the local attractiveness (in regards to $PV_i$) of the elements of $X_i$.

<u>*Example*</u> *(continued)*
*For "Black printing speed", J has provided cardinal information concerning the local attractiveness (in regards to "Speed") of levels "5", "6", "7", "8", "9", "10" and "11" pages/min. J's MACBETH judgements are shown in the matrix of Figure 17.*


Black printing speed

| | 11 | 10 | 9 | 8 | 7 | 6 | 5 |
|---|---|---|---|---|---|---|---|
| 11 | no | very weak | weak | weak | weak | strong | v. strong |
| 10 | | no | very weak | positive | positive | strong | strong |
| 9 | | | no | very weak | positive | positive | strong |
| 8 | | | | no | very weak | positive | moderate |
| 7 | | | | | no | weak | weak |
| 6 | | | | | | no | weak |
| 5 | | | | | | | no |

Consistent judgements

Figure 17. J's matrix of judgements for the criterion "Black printing speed".

*The MACBETH questioning-answering process took place with the direct interactive support of the M-MACBETH software, as follows:*

1. *It began with the comparison of the good and neutral speeds (10 and 6 pages/min, respectively). J qualitatively judged the respective difference of attractiveness as being "strong" (see Figure 17).*
2. *J was then asked to compare the two extreme levels of impact-range of "Black printing speed", that is, J compared the most attractive speed and the least attractive speed (11and 5 pages/min, respectively), followed by the second most attractive level (10 pages/min) with the least attractive, and so on, thereby completing (from top to bottom) the last column of the matrix; this step implicitly used the least attractive level (5 pages/min) as a fixed reference.*
3. *The most attractive level was then compared to each of the other levels, in order of increasing attractiveness, thereby completing (from right to left) the first row of the matrix, now taking as the fixed reference the most attractive level (11 pages/min).*

4. *The next step consisted of comparing the most attractive level with the second most attractive level, the second most attractive with the third, and so on, thereby completing the diagonal border of the upper triangular portion of the matrix.*

Comments:

- ✓ This sequence is not at all compulsory. For instance, step 1 is often skipped in practice, as it was the case described in [12], in which an additional final step was added:

  5. To assess the remaining judgements.

  It is not necessary to complete the matrix for a MACBETH scale to be created, however, the more preference information provided the greater the scale's level of accuracy.
- ✓ An alternative questioning order would be to fill in the MACBETH matrix diagonally (see example in Figure 3), starting with the comparison of consecutive levels (step 4 above). This is actually the way we used to develop the process in the first applications of MACBETH. From our practical experience, however, we tended to infer that the diagonally process is less judgementally appealing for the assessors.

<u>*Example*</u> *(continued)*
*As each judgement was given, the software automatically verified the matrix's consistency. No inconsistency was detected in this particular case. Should any consistency occur, the software would immediately suggest judgement modification(s) (see section 2.4.3) and the detected inconsistency would be fixed by J before moving forward.*

*From the consistent matrix of judgments in Figure 17, the M-MACBETH software created the precardinal scale (see section 2.4.4) that was then discussed. J considered that it adequately represented the relative magnitude of the judgements, that is, the scale could be considered as an interval scale. As described in section 2.4.5, this could have required the adjustment of numerical scores by J. The interval scale defined on {"5", "6", "7", "8", "9", "10", "11"} has been extended by the software to the full range of the continuous descriptor of speed by linear interpolation, giving rise to the piecewise linear value-function $v_3$ whose graph is represented in Figure 18. So, for example, we would have:*

$$v_3(8.5) = v_3(8) + \frac{8,5-8}{9-8} \,.\, (v_3(9) - v_3(8)) = 70.5.$$

*Note that the value function is anchored in the good and neutral speeds, to which the scores of 100 and 0 were arbitrarily assigned, that is $v_3(10) = 100$ and $v_3(6) = 0$, thereby giving any other score a substantive meaning as a numerical representation of the relative attractiveness of the respective speed level. Thus, for instance, given that the estimated black speed printing of Espon is 8.5 pages/min, its value of 70.5 can be interpreted by J as meaning that the added benefit, with reference to a neutral speed, of purchasing Espon is 70.5% of the added benefit that J would obtain if Espon had a good printing speed, all other things being equal. Moreover, fixing the value of neutral equal to 0, implies that a positive score directly indicates a (locally) attractive printer, and a negative score directly indicates an unattractive printer; similarly, fixing the value of good equal to 100, a score greater or equal to 100 indicates a very attractive printer.*

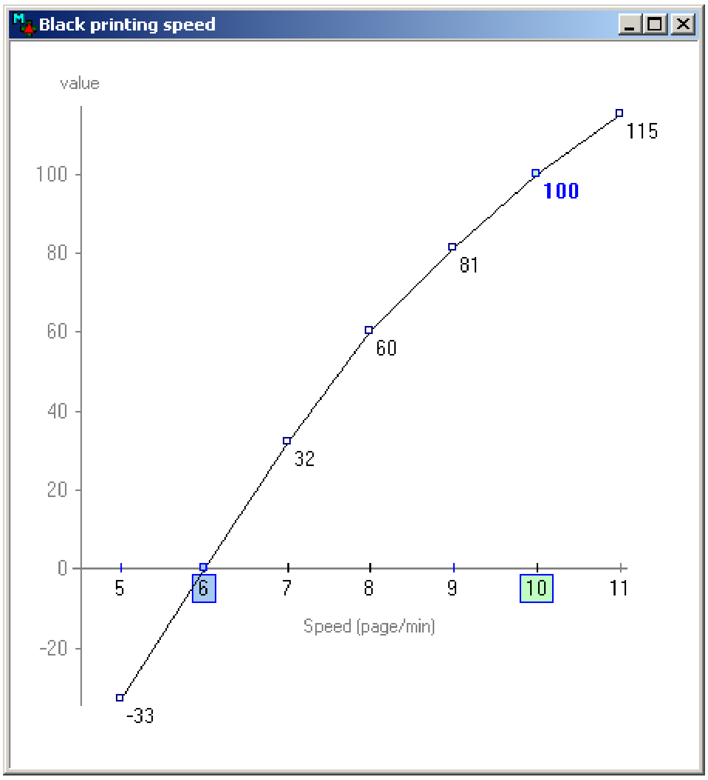



Figure 18. Piecewise linear value function of black printing speed ($v_3$).

*A similar process was then followed to create a value scale for each of the two qualitative criteria, "Black printing quality" and "Colour printing quality (see Figure 19).*

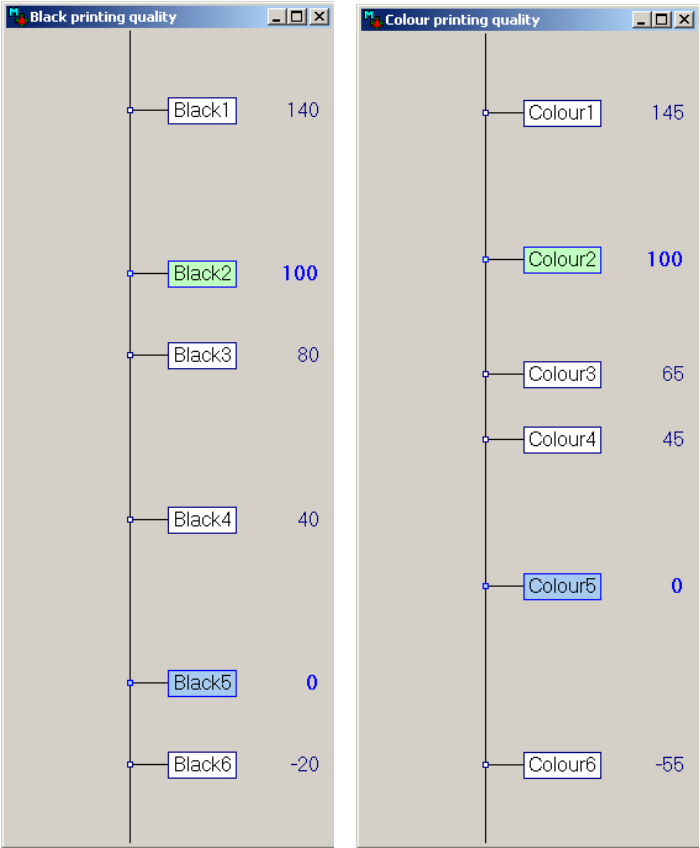



Figure 19. Value scales of black and colour printing quality ($v_1$ and $v_2$, respectively).

*The local values of each of the printers in the three criteria could then be calculated from $v_1$, $v_2$ and $v_3$. J was presented with this information in the different forms offered by the M-MACBETH software, either numerically, as the "value table" at the left in Figure 20, or graphically, as the "value profile" of Sister at the right in the same figure. J could also observe the differences of local values between any two printers graphically, as exemplified by the "difference profile" in Figure 21. This visual representation is very useful for an easy appraisal of the extent to which one option is more, or less, attractive in each criterion, although it does not yet offer which of them, if any, is globally more attractive than the other one, and by how much. Indeed, an overall (or global) comparison of the printers requires one to first acquire inter-criteria preference information from J.*

Actions scores

| Actions | Overall | Black | Colour | Speed |
|---|---|---|---|---|
| PH | ? | 80.00 | 100.00 | 32.00 |
| Conan | ? | 140.00 | 45.00 | 81.00 |
| Espon | ? | 0.00 | 100.00 | 70.50 |
| Sister | ? | -20.00 | 65.00 | 107.50 |
| Nomark | ? | 80.00 | 65.00 | 81.00 |
| [ good ] | ? | 100.00 | 100.00 | 100.00 |
| [ neutral ] | ? | 0.00 | 0.00 | 0.00 |
| Scaling constants: | | ? | ? | ? |

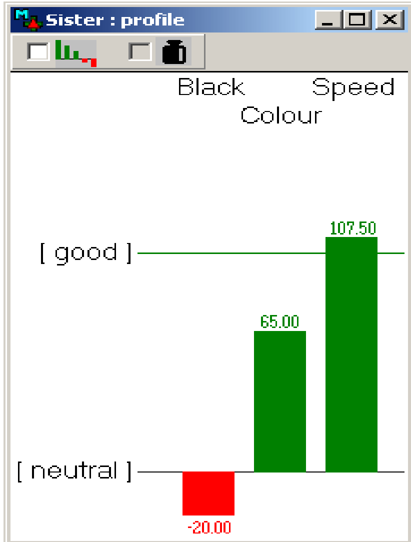



Figure 20. Value table and value profile of "Sister".

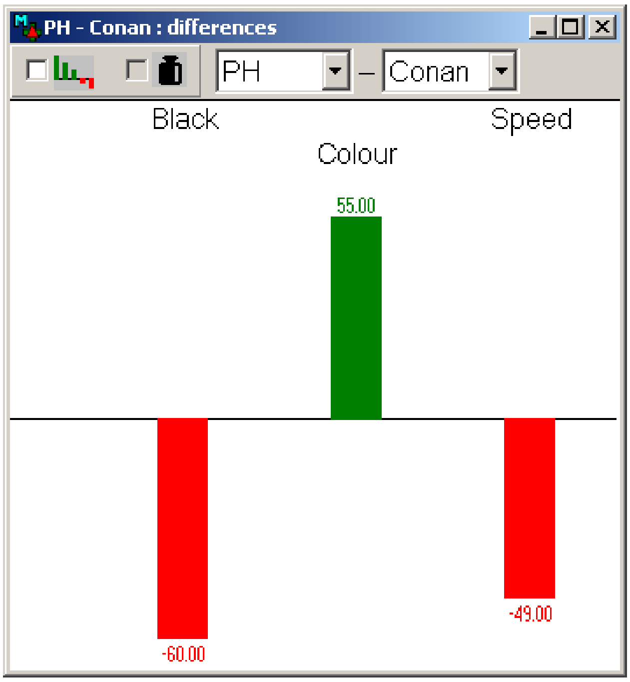



Figure 21. Difference profile between "PH" and "Conan".

*3.3.2. Acquisition of inter-criteria preference information*

This stage consists of acquiring preference information from J concerning differences of attractiveness taken for each criterion. As this stage is difficult, it is important that the differences of attractiveness chosen for each criterion are clear and well understood by J. This is why, as announced in section 3.1.2, the set of references $\{(L_i, H_i) \mid i \in \{1, \ldots, K\}\}$ constitutes a particularly solid and well adapted foundation for that purpose. They allow one to define the (K+1) following reference profiles:

$$[L] = [L_1, L_2, L_3, \ldots, L_{K-1}, L_K]$$
$$[PV_1] = [H_1, L_2, L_3, \ldots, L_{K-1}, L_K]$$
$$[PV_2] = [L_1, H_2, L_3, \ldots, L_{K-1}, L_K]$$
$$\ldots$$
$$[PV_K] = [L_1, L_2, L_3, \ldots, L_{K-1}, H_K].$$

Let us observe that the difference in (overall) attractiveness between $[PV_i]$ and [L], $i \in \{1, \ldots, K\}$, corresponds to the added value of the "swing" from $L_i$ to $H_i$ and that the comparison of the relative attractiveness of the profiles $[PV_i]$ and $[PV_j]$, $i \neq j \in \{1, \ldots, K\}$, corresponds to the comparison of the added values of the "swings" from $L_i$ to $H_i$ and from $L_j$ to $H_j$.

In terms of these reference profiles, it is possible to correctly state the aim of the stage of "acquisition of inter-criteria preference information": obtaining cardinal information regarding the (overall) attractiveness of [L], $[PV_1]$, $[PV_2]$, …, $[PV_K]$. We will hereafter write $v_0$ a numerical scale defined on $X_0 = \{[L], [PV_1], [PV_2], \ldots, [PV_K]\}$ that satisfies the cardinal measurement rules CM1, CM2 and CM3: $v_0$ is then an interval scale which "measures" the overall attractiveness of the reference profiles.

*<u>Example</u> (continued)*
*As both neutral and good levels had been previously determined (see Figures 8 to 10), the first question asked to J was phrased as follows: "Imagine an option exists that is neutral in every criterion; how much would a swing from neutral to good black printing quality increase its overall attractiveness?" Once again J was invited to respond with a MACBETH qualitative judgement. J answer was: that increase is very strong. A similar question was subsequently asked for each of the other two criteria, thus completing the last column of the "judgement weighting matrix" (Figure 22).*

Comment:
Let us point out that these MACBETH judgements of J allowed to rank the reference profiles

[neutral] = [$\text{neutral}_1$, $\text{neutral}_2$, $\text{neutral}_3$],
[Black] = [$\text{good}_1$, $\text{neutral}_2$, $\text{neutral}_3$],
[Colour] = [$\text{neutral}_1$, $\text{good}_2$, $\text{neutral}_3$],
[Speed] = [$\text{neutral}_1$, $\text{neutral}_2$, $\text{good}_3$],

by order of decreasing overall attractiveness: [Black], followed by [Colour], followed by [Speed] and finally [neutral]. As announced in section 2.4.1, this is an example of another possible interaction mode than the classical one presented in section 2.4.

*<u>Example</u> (continued*
*The next step was to elicit qualitative judgments from J regarding the difference of attractiveness between swings. It began with the comparison of the most attractive swing to the second most attractive swing, by asking: "How much more attractive is a*

*swing from neutral to good black printing quality than a swing from neutral to good black printing colour?" A similar comparison was subsequently made between [Black] and [Speed] and between [Colour] and [Speed], thus completing the MACBETH matrix of judgements (see Figure 22). As before, consistency checks were automatically made each time a judgement was entered into the matrix. Using the procedure described in section 2, we finally obtained, after the adjustment by J of the MACBETH scale initially proposed by the software, the following scale:*

$$v_0([\text{Black}]) = 60$$
$$v_0([\text{Colour}]) = 25$$
$$v_0([\text{Speed}]) = 15$$
$$v_0([\text{neutral}]) = 0.$$

Overall

| | [Black] | [Colour] | [Speed] | [neutral] |
|---|---|---|---|---|
| [Black] | no | strong | strong | v. strong |
| [Colour] | | no | very weak | moderate |
| [Speed] | | | no | weak |
| [neutral] | | | | no |

Consistent judgements

Figure 22. J's matrix of judgements concerning overall attractiveness of reference profiles.

### *3.3.3. Construction of an overall attractiveness scale*

From the scales $v_0, v_1, v_2, \ldots, v_K$, we build an overall attractiveness scale Att:

$$\forall\, x = (x_1, x_2, \ldots, x_K) \in X_1 \times X_2 \times \ldots \times X_K,\ \text{Att}(x) = \sum_{i=1}^{K} \frac{v_0([PV_i]) - v_0([L])}{v_i(H_i) - v_i(L_i)} \cdot v_i(x_i).$$

We recognise, in this formula, the additive aggregation model

$$\sum_{i=1}^{K} p_i \cdot v_i(x_i) \quad \text{in which} \quad p_i = \frac{v_0([PV_i]) - v_0([L])}{v_i(H_i) - v_i(L_i)}.$$

It's easy to prove that Att is an interval scale when the scales $v_0, v_1, v_2, \ldots, v_K$ are interval scales.

Comments:

- ✓ A better notation for Att would be

  $$\text{Att}_{(v_0, v_1, \ldots, v_K)}$$

  because it emphasises that the overall attractiveness scale Att depends on the scales $v_0, v_1, \ldots, v_K$; this is particularly important if one intents to meaningfully compare Att(x) and Att(y), where $x, y \in X_1 \times X_2 \times \ldots \times X_K$.
- ✓ If

  $$v_0([L]) = 0 \text{ and, } \forall\, i \in \{1, 2, \ldots, K\},\ v_i(L_i) = 0$$

  then

  $$\text{Att}([PV_i]) = v_0([PV_i]),$$

which shows that the overall attractiveness scale Att build on $X_1 \times X_2 \times \ldots \times X_K$ extends the overall attractiveness scale $v_0$ obtained on the subset $X_0$ of $X_1 \times X_2 \times \ldots \times X_K$ from the cardinal preference information given by J.

✓ When, moreover, as it is common in practice,

$$\forall\, i \in \{1, 2, \ldots, K\},\ v_i(H_i) = 100 \text{ and } \sum_{i=1}^{K} v_0([PV_i]) = 100$$

then

$$Att(x) = \sum_{i=1}^{K} \frac{v_0([PV_i])}{100} . v_i(x_i)$$

and $\frac{v_0([PV_i])}{100}$ is called "weight" of criterion $PV_i$; these weights are represented in the M-MACBETH software in the form of a "histogram" (see Figure 23).

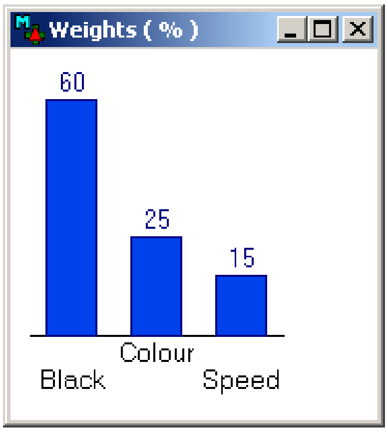



Figure 23. Histogram of weights.

Hereafter, we will write "Att(a)" instead of "Att(π(a))", for all a ∈ A (the set of actions).

*<u>Example</u> (continued)*
*Figure 24 shows in the overall column the overall attractiveness of the printers and of the profiles [good] and [neutral] (ranked by order of decreasing attractiveness). In the last three columns, the scales $v_1$, $v_2$ and $v_3$ are presented. The respective weights of the criteria appear in the last line.*

Actions scores

| Actions | Overall | Black | Colour | Speed |
|---|---|---|---|---|
| Conan | 107.40 | 140.00 | 45.00 | 81.00 |
| [ good ] | 100.00 | 100.00 | 100.00 | 100.00 |
| PH | 77.80 | 80.00 | 100.00 | 32.00 |
| Nomark | 76.40 | 80.00 | 65.00 | 81.00 |
| Espon | 35.58 | 0.00 | 100.00 | 70.50 |
| Sister | 20.38 | -20.00 | 65.00 | 107.50 |
| [ neutral ] | 0.00 | 0.00 | 0.00 | 0.00 |
| Scaling constants: | | 0.60 | 0.25 | 0.15 |

Figure 24. Value table.

3.4. Sensitivity analysis

The M-MACBETH software allows numerous sensitivity analyses to be performed. We will not describe all of them here but instead report that every change of a value, "textual" (keyboard) or "graphical" (mouse), is instantaneously reflected upon all other dependent values (and graphics).

*Example (continued)*
*If we increase the weight of Speed or if we decrease the score of Black3 on "Black printing quality (for instance, by dragging the mouse), we can immediately see that the global score of Nomark decreases (see Figure 25).*

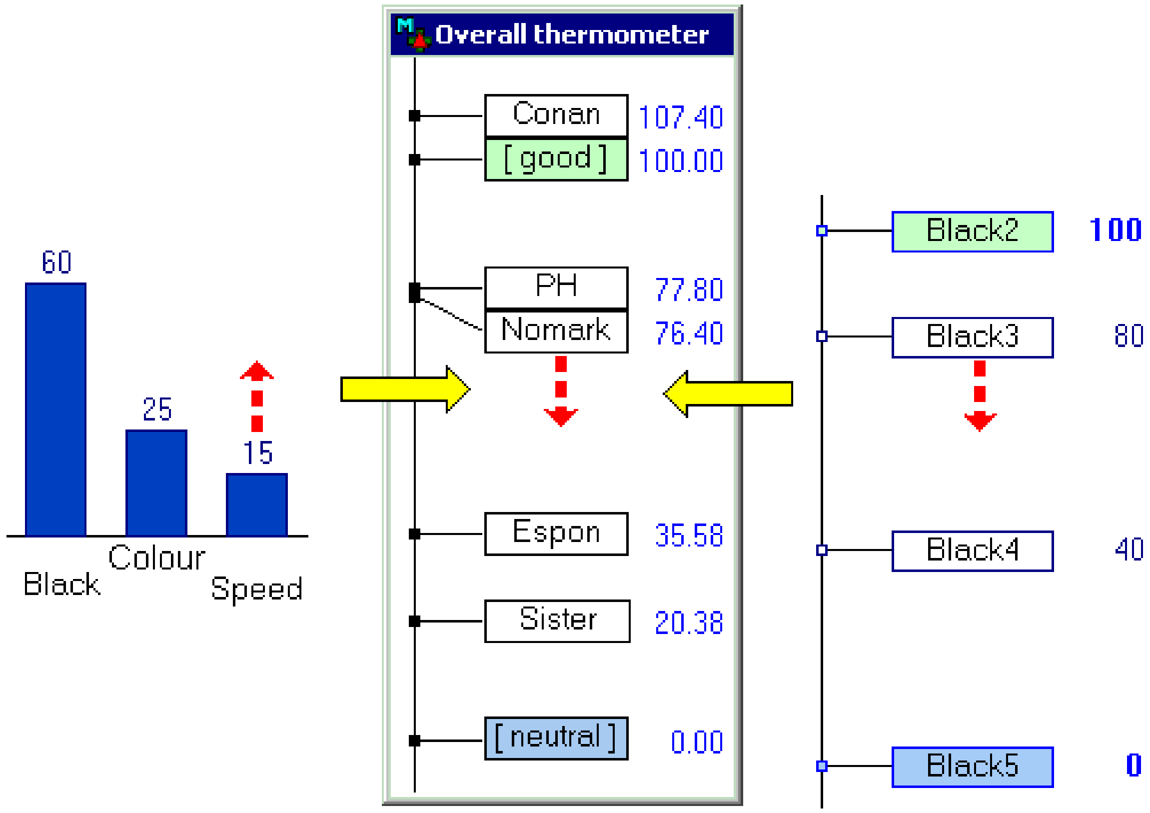



Figure 25. Example of sensitivity analysis.

A window in the software is dedicated to the performance of sensitivity analysis on weight: it permits the transformation of the global scores of the potential actions to be observed in case the weight of a chosen criterion $PV_i$ is changed.

*Example (continued)*
*Figure 26 shows, among other things, that:*

- ✓ *the current weight (on horizontal axis) of Colour is 25.00;*
- ✓ *when the weight of Colour increases, the global score (on vertical axis) of PH increases and the global score of Nomark decreases;*
- ✓ *the global score of Nomark would be greater than the global score of PH (and greater than 76.9) if the weight of Colour was lower than 21.9.*

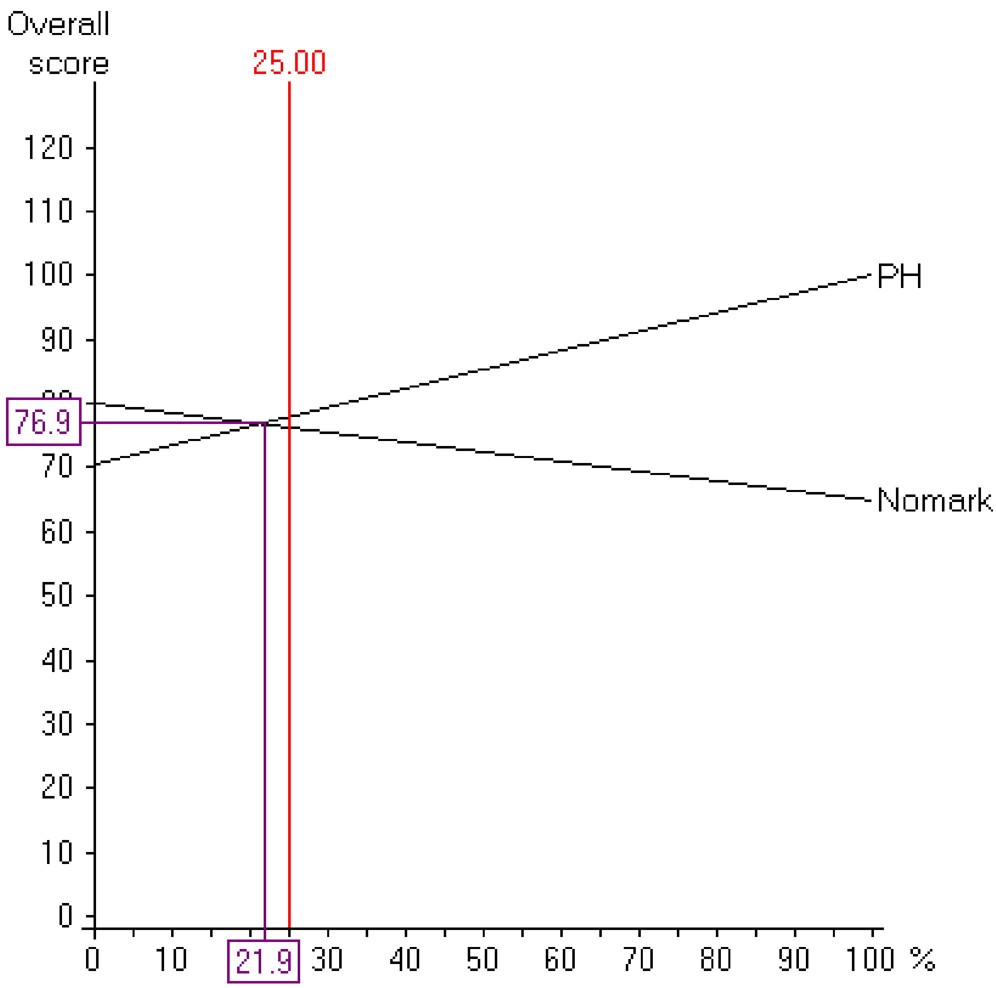



Figure 26. Example of sensitivity analysis on the weight of “Colour”

3.5. Robustness analysis

Let a and b be any two potential actions. One says that:

*a is globally more attractive than b considering some preference information (ordinal, precardinal or cardinal)* when the statement

$$Att_{(v_0,v_1,\dots,v_K)}(a) > Att_{(v_0,v_1,\dots,v_K)}(b)$$

is true for all the numerical scales $v_0$, $v_1,\dots,v_k$ which are, taking into account the conditions CM1, CM2, CM3 and C3, admissible representations of the selected preference information (conditions CM1 and CM2 in the case of ordinal information, conditions CM1, CM2 and C3 in the case of precardinal information, conditions CM1, CM2 and CM3 in the case of cardinal information).

A table of global comparisons of the actions is integrated in the M-MACBETH software. It displays all couples (a,b) such as “a dominates b” (a triangle on line a and column b) or “a is globally more attractive than b considering the selected preference information ” (a cross on line a and column b). Supposing that Att(a) > Att(b) with cardinal information, the table of global comparisons allows one to answer questions such as: “if we take into account only ordinal or precardinal information on some criteria and/or for the comparison of reference profiles, could we always maintain that a is globally more attractive than b?”

*Example* (continued)

*In the table of global comparisons of the Figure 27, it can be seen that, when the selected preference information is the precardinal (MACBETH) information for the three criteria (local information) and the ordinal information for the comparison of reference profiles (global information), then Conan is no longer globally more attractive than all other printers. However, if we take into account precardinal (MACBETH) information not only for the local information but also for the global information (see Figure 28), it is enough for Conan to be globally more attractive than PH, Espon, Sister and Nomark.*

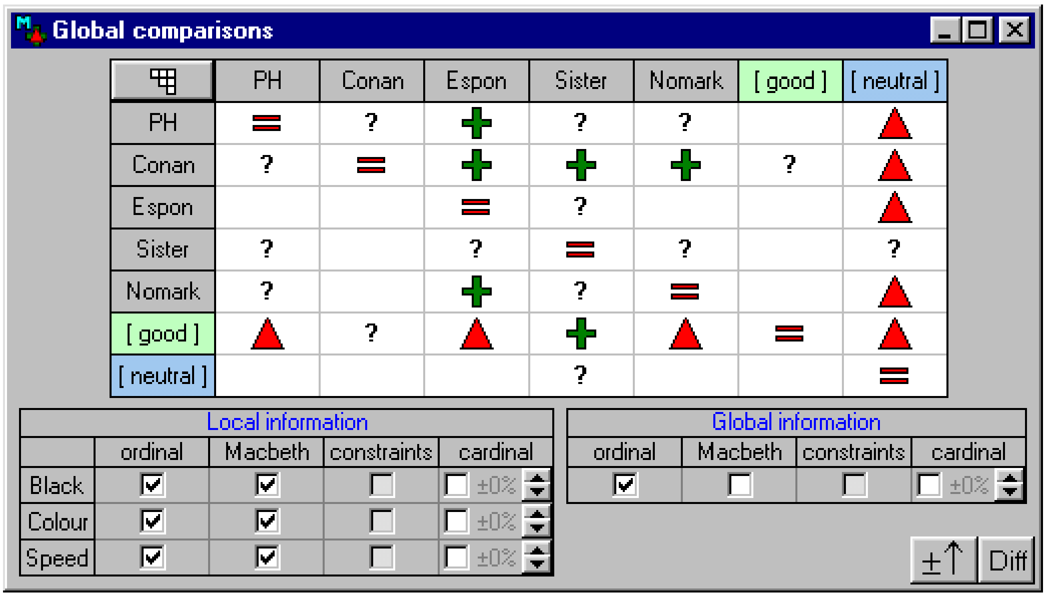



Figure 27. First example of table of global comparisons.

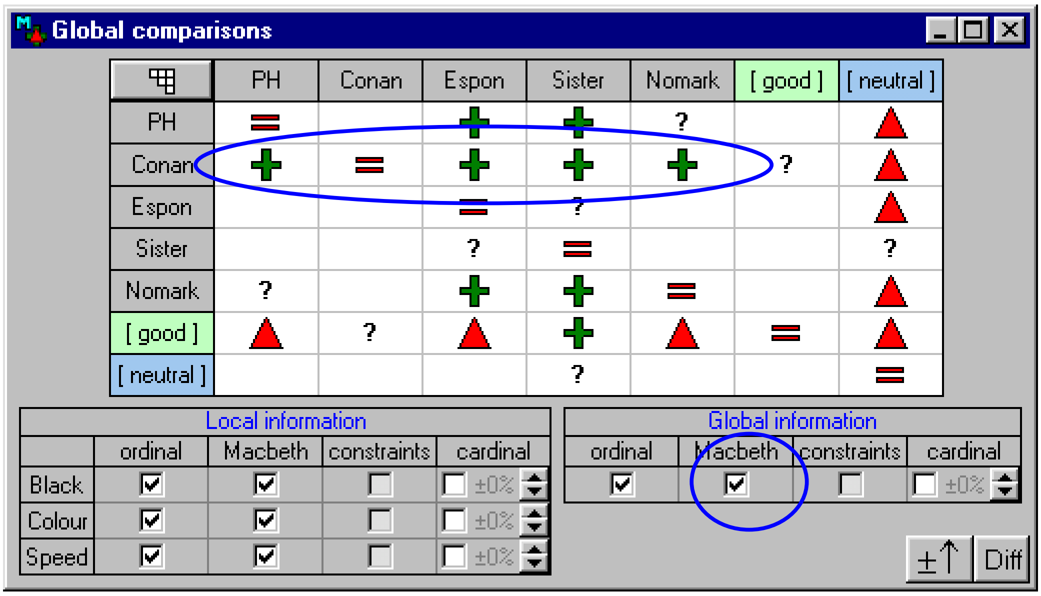



Figure 28. Second example of table of global comparisons

It is also possible, with the table of global comparisons, to see the effect of taking into account uncertainty on the actions' impacts.

*Example (continued)*
*If J decides to take into account the uncertainty specified for the impact of Conan on the criterion “Black printing quality” (“Black”) (see the “Impacts precision” window at the right end corner of Figure 29), the table of global comparisons of Figure 29 shows that Conan would no longer be globally more attractive than PH and Nomark.*

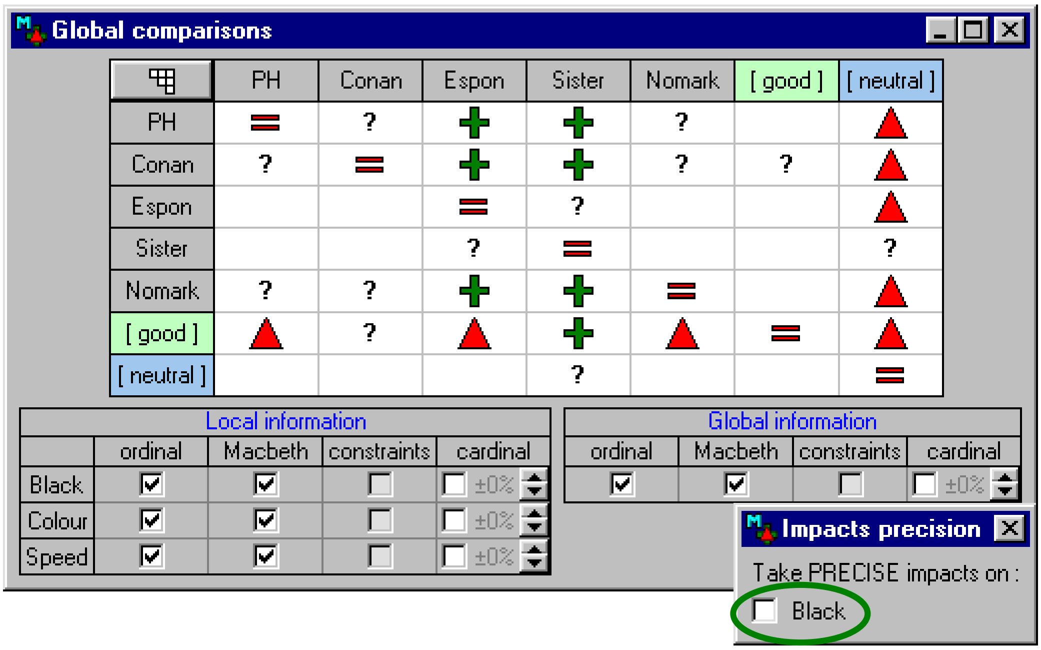



Figure 29. Taking into account the uncertainty on the impacts on criterion “Black” and third example of table of global comparisons.

*Figure 30 shows that, when taking into account the uncertainty, Conan is globally more attractive than all other printers if we select the cardinal information on the three criteria (keeping a precardinal (MACBETH) global information).*

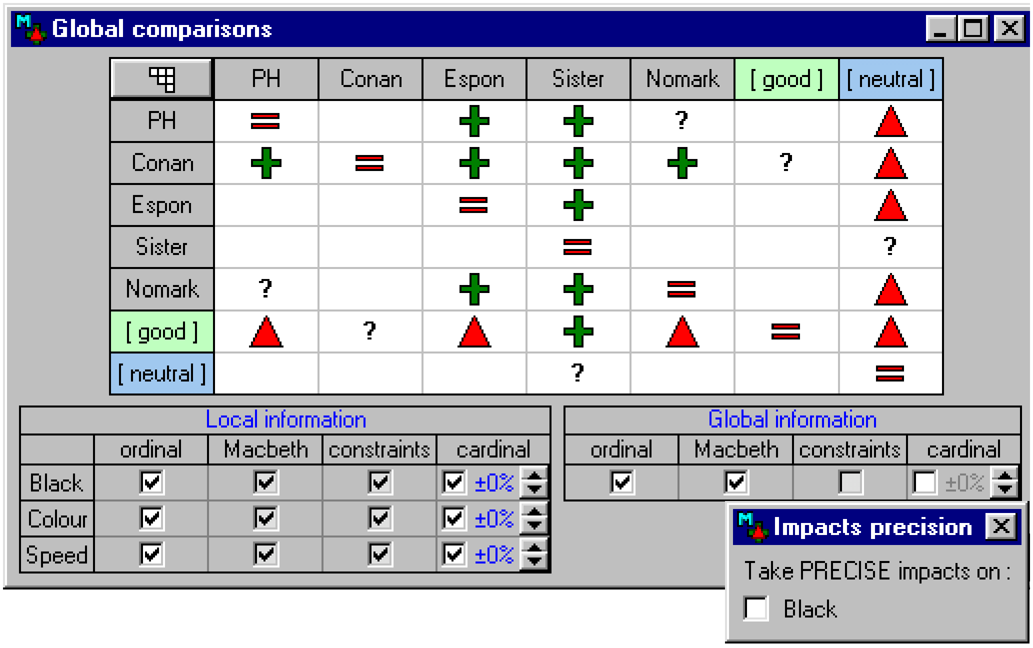



Figure 30. Fourth example of table of global comparisons.

*Let us recall that "take into account the cardinal information, for instance, on Speed" means that:*

$$\text{if} \quad \begin{cases} v_3(10)=100 \\ v_3(6)=0 \end{cases}, \quad \text{then} \quad \begin{cases} v_3(11)=115 \\ v_3(9)=81 \\ v_3(8)=60 \\ v_3(7)=32 \\ v_3(5)=-33 \end{cases}$$

*But, is it really necessary to select "all" of the cardinal information on Speed criterion for Conan to be globally more attractive than all other printers? The answer is no, as can be seen in the Figure 31.*

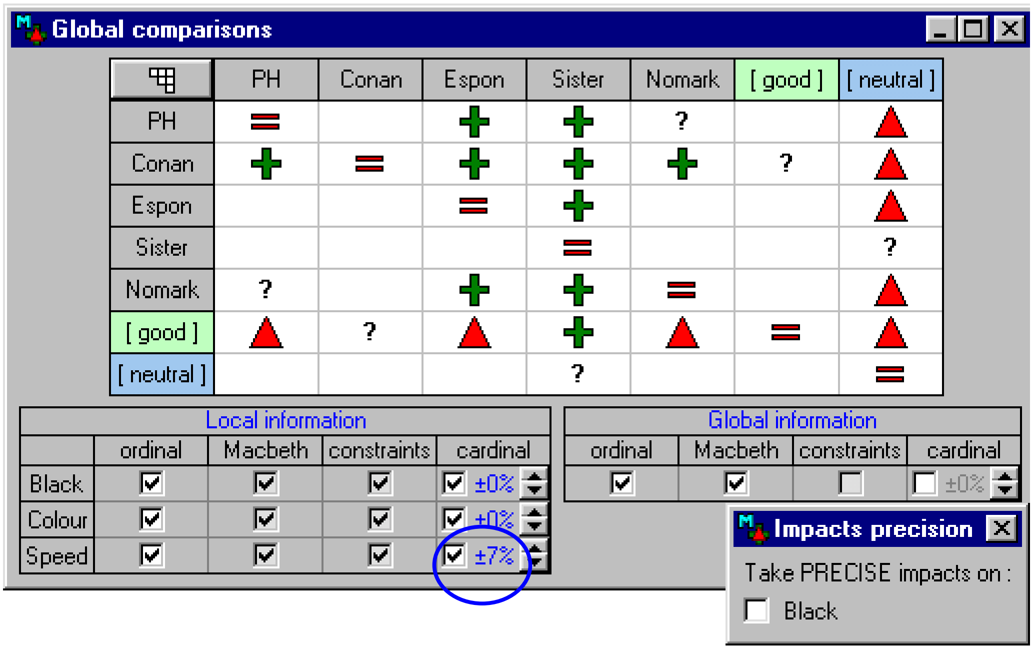



Figure 31. Fifth example of table of global comparisons.

*What does this "±7%" mean? Simply that:*

$$\text{if} \quad \begin{cases} v_3(10)=100 \\ v_3(6)=0 \end{cases}, \quad \text{then} \quad \begin{cases} 108 = 115-7 \le v_3(11) \le 115+7 = 122 \\ 74 = 81-7 \le v_3(9) \le 81+7 = 88 \\ 53 = 60-7 \le v_3(8) \le 60+7 = 67 \\ 25 = 32-7 \le v_3(7) \le 32+7 = 39 \\ -40 = -33-7 \le v_3(5) \le -33+7 = -26 \end{cases}$$

*This information is clearly less rich than "strict" cardinal information, which can help J to be convinced of the validity of his final decision.*

3.6. Diagram of the MACBETH approach

We can now summarise the MACBETH approach with the diagram of Figure 32.

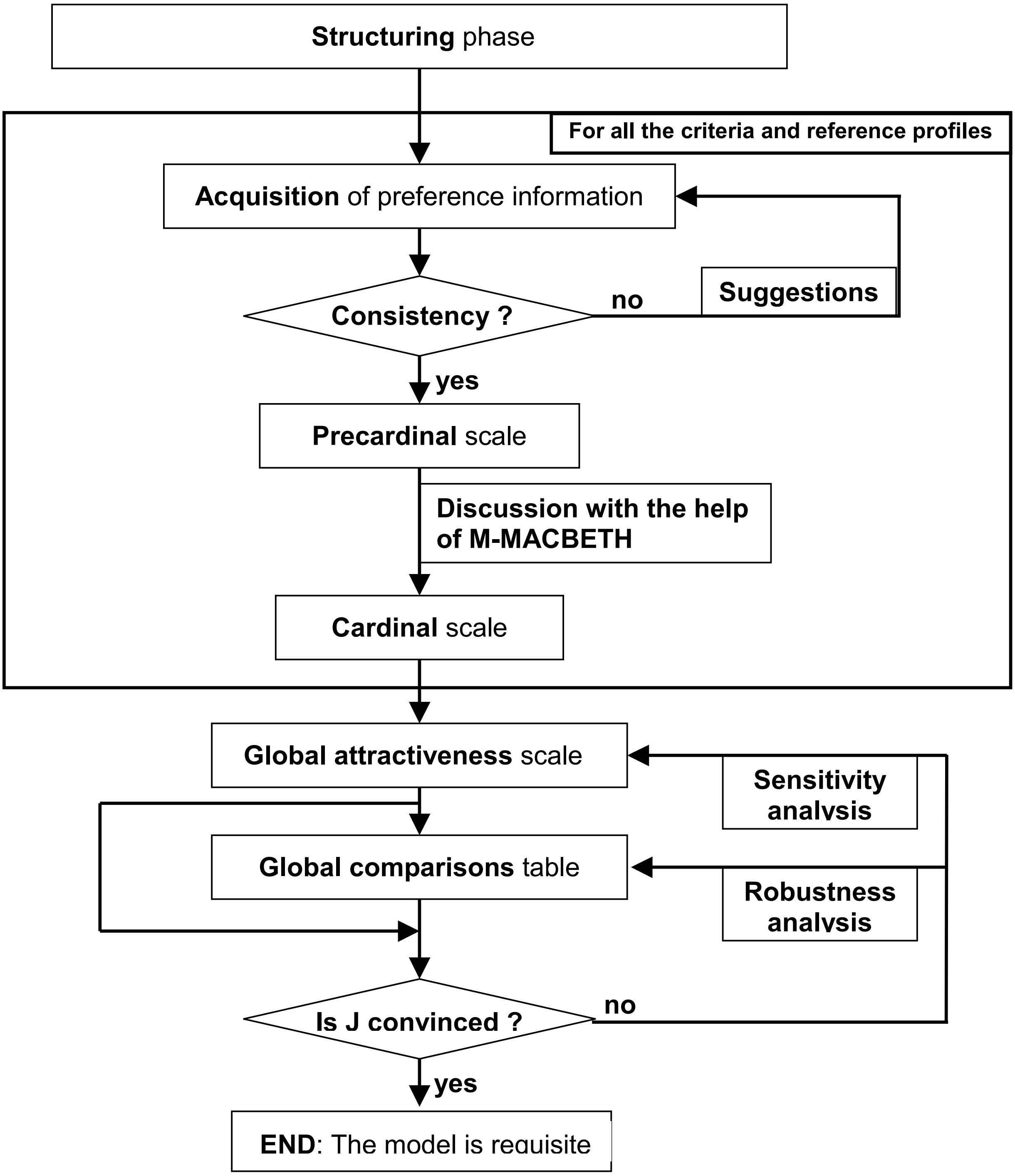



Figure 32. Summary of the MACBETH decision aid process.

## 4. Brief historical survey of the MACBETH research

The initial research that gave rise to the MACBETH approach was conducted by C.A. Bana e Costa and J.-C. Vansnick, in the early 1990's [7, 32, 37], as a response to the question: how to build an interval scale of preferences on a set of options without forcing evaluators to produce direct numerical representations of their preferences? They conceived the MACBETH pairwise questioning procedure that compares two options at a time requesting only a qualitative judgment about the difference of attractiveness between options. The six semantic categories "very weak", "weak", "moderate", "strong", "very strong" or "extreme" should be represented by non-overlapping (disjoint) intervals of real numbers whose limits should be determined together with numerical value scores for the options. Of course, for each particular set of judgements, one should start by previously testing the existence of such intervals. Research was then conducted to find a mathematical programming answer to that problem [7, chapter IV], which gave rise to the formulation of a chain of four linear programs [34, 35, 36, 37] conceived to be used in practice as illustrated in Figure 33. Programs MC1 to MC4, implemented in GAMS, were used in the first real-world applications of MACBETH as a decision support tool to derive value scores and weights in the framework of an additive aggregation model [30, 42, 43].

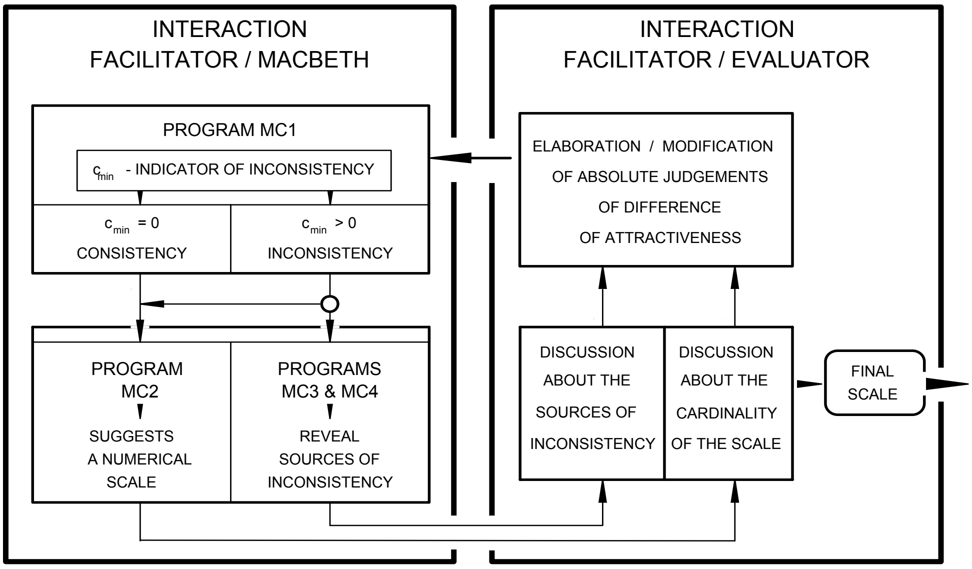



Figure 33. Initial conception of the "interactive Scheme of MACBETH" (adapted from [34]).

Theoretical research conducted meanwhile, and first presented in 1994 at the XIth International Conference on MCDM [31], demonstrated the equivalence of the approach by thresholds and the approach by measurement conditions.

In 1997, the first version of the MACBETH software was developed by J.-M. De Corte, following essentially the diagram shown in Figure 33, although the objective functions used in MC1 and MC2 programs were modified on the basis of a simple principle [27, 28], that makes possible the determination "by hand" of the basic MACBETH scale in simple cases [33]. The simplest hand procedure, applicable to full matrices of judgements in which no hesitation exists, consists of:

1) to assign to each ordered pair of elements (x,y), corresponding to a cell of the diagonal border of the upper triangular portion of the matrix of judgements, a number n(x,y) as follows:
   n(x,y) = 0   if there is no difference of attractiveness between x and y
   n(x,y) = 1   if the difference of attractiveness between x and y is “very weak”
   n(x,y) = 2   if the difference of attractiveness between x and y is “weak”
   n(x,y) = 3   if the difference of attractiveness between x and y is “moderate”
   n(x,y) = 4   if the difference of attractiveness between x and y is “strong”
   n(x,y) = 5   if the difference of attractiveness between x and y is “very strong”
   n(x,y) = 6   if the difference of attractiveness between x and y is “extreme”;
2) to assign to each other ordered pair (z,w), with z more attractive than w, a number n(z,w) as follows:
   $$n(z,w) = n(z,x_1) + n(x_1,x_2)+ \dots + n(x_{k-1},x_k) + n(x_k,w)$$
   where $(z,x_1)$, $(x_1,x_2)$, … , $(x_{k-1},x_k)$ and $(x_k,w)$ are ordered pairs of elements corresponding to the diagonal border of the upper triangular portion of a matrix of judgements.

If these numbers verify:

<u>Condition δ</u>: For all elements x, y, z, w with x more attractive than y and z more attractive than w, if the difference of attractiveness between x and y is greater than the difference of attractiveness between z and w, then
$$n(x,y) \geq n(z,w) + 1 + \delta(x,y,z,w)$$
where δ(x,y,z,w) is the minimal number of categories of difference of attractiveness between the difference of attractiveness between x and y and the difference of attractiveness between z and w,

then, the basic MACBETH scale v is given by the rule:
- ✓ for each less attractive element $x^-$ : $v(x^-) = 0$
- ✓ for each element x more attractive than $x^-$: $v(x) = v(x^-) + n(x,x^-) = n(x,x^-)$.

For example, in the case of the matrix of judgements in Figure 22, one has:
1) n([Black],[Colour]) = 4, n([Colour],[Speed])= 1 and n([Speed],[neutral]) = 2
2) n([Black],[Speed]) = n([Black],[Colour]) + n([Colour],[Speed]) = 4+1 = 5
   n([Black],[neutral]) = n([Black],[Colour]) + n([Colour],[Speed]) + n([Speed],[neutral]) = 4+1+2 = 7
   n(Colour],[neutral]) = n([Colour],[Speed]) + n([Speed],[neutral]) = 1 + 2 = 3;

as Condition δ is satisfied, the basic MACBETH scale is:
   v([neutral]) = 0
   v([Speed]) = n([Speed],[neutral]) = 2
   v([Colour]) = n([Colour],[neutral]) = 3
   v(Black]) = n([Black],[neutral]) = 7.

When Condition δ is not verified, the procedure for obtaining the basic MACBETH scale is more complex [33].

Enabling interactivity was the key concern behind the development of the first MACBETH software: it allowed judgements to be encoded in a user-friendly way (something absent in the initial GAMS implementation) and offered a graphical representation of the MACBETH value scale, which allowed the direct adjustment (by way of the mouse) of the value score of an option within an interval of variation respecting measurement conditions related to the matrix of judgements. Moreover, in the insurgence of an inconsistency, this first software informed the user, in a visual way, of the possible changes that would make the matrix of judgements consistent, thereby greatly facilitating interactivity in judgmental revision. Based upon measurement conditions, we, at the time, made the distinction between (several cases of) “semantic inconsistency” and “incoherence”, although we did not have complete procedures to address and manage all cases. This is why a compromise scale was still available in case of an inconsistency.

Even if successfully used in practice [2, 3, 16, 19, 20, 22, 28, 29, 30, 35, 38, 45 to 49, 51, 55], this first MACBETH software imposed several limitations to the practical development of an efficient, interactive and learning model-building process. Indeed:

1) The determination of suggestions was still heuristic and did not guarantee the minimal number of changes necessary to achieve consistency;
2) it was not possible for the evaluator to hesitate between several semantic categories when expressing judgements. It, therefore, did not enable one to facilitate the management of group judgemental disagreements;
3) it forced the evaluator to first give all judgements before running any procedure and, consequently, judgemental inconsistency could only be detected for a full matrix of judgements. Therefore, suggestions of changes to bypass inconsistency could only then be discussed.

In 1999, the drawbacks of the first MACBETH software led us to make the decision to abandon it and to develop a totally new software, now indeed a decision support system, M-MACBETH (see http://www.m-macbeth.com): the first “M” of the name emphasises its ability to deal with multicriteria problems. Successively ameliorated, M-MACBETH could take advantage of the results of the theoretical research conducted since 2000, namely those reported in [44] and [50], allowing inconsistency to be treated in a solid mathematical way. This was the turning point in the path to find a more interactive formulation. First, working with incomplete matrices, we found that it was in practice meaningless to distinguish between incoherence and inconsistency. Then, leaving this distinction in two cases, we implemented the automatic detection of “inconsistency”, even for incomplete matrices. The objective of abandoning the suggestion of a compromise scale could finally be achieved for the origin of inconsistency could now be easily found (detection of elementary incompatible systems) and explained to the evaluator. We can guarantee that the M-MACBETH software finds the minimal number of changes to be made and that we have all of the possibilities for any number of changes not greater than five. We can also give suggestions of multiple category changes, where a “k categories change” is considered to be equivalent to k “1 category changes”.

In what concerns the structuring phase of the decision aid process, M-MACBETH offers tools to easily build a value tree and descriptors, to define an impact table (see section 3.1.4), to show value profiles of actions (see section 3.3.1), etc.

We also began to work around the concept of “robustness” at that time, and the first “global comparisons table” appeared in 2000 (see section 3.5). In parallel, the extensive use of M-MACBETH in decision-conferencing consulting work (see section 5 and many other applications, for example [1, 11, 25, 26, 40]), led to the addition, in 2001, of several useful tools regarding sensitivity analyses (see section 3.4 and [12]). Moreover, in 2002, we added the possibility of defining additional constraints and the option of moving away, to a certain extent, from strict numerical (cardinal) representations (see section 3.5).

In 2003, the possibility of

- ✓ dealing with imprecision (uncertainty) about impacts of options (see sections 3.1.4 and 3.5) either for qualitative or quantitative descriptors
- ✓ incorporating at any time reference levels for a criterion
- ✓ graphically representing comparisons of potential actions on any two groups of criteria

was added to the M-MACBETH software.

Also in 2003, a version of the MACBETH software limited to scoring and weighting has been embedded in the HIVIEW3 software package [41].

## 5. Some consulting projects using MACBETH

- Evaluation of bids in international public calls for tenders and contractors' choice [4, 6, 8, 9, 10, 13, 15, 53].
- Evaluation of measures of European structural programs [28, 30, 42, 43].
- Resource allocation to the construction of new inter-municipal roads in the Region of Lisbon [5].
- Design of a new railway link to the Port of Lisbon [20, 21, 30].
- Assigning priorities for maintenance, repair and refurbishment in managing the Lisbon's housing stock [22, 23, 24].
- Design of a total quality system for the Lisbon Gas Company [14].
- Decision support systems for the evaluation of Lisbon Gas Company's service and material suppliers [51].
- Prioritisation of the municipalities of the State of Ceará, Brazil, in terms of dryness effects [54].
- Development of the Strategic Town Plan of Barcelos [17, 18].
- Evaluation of the competitive capacity of small and medium sized textile firms in the State of Santa Catarina, Brazil [16].
- Evaluation of alternative strategies to attach foreigner investment to the Brazilian micro-electronic sector.
- Development of a Vision for Puerto Rico 2025.
- Evaluation of military strategies.
- Analysis of the operational performance of electricity distributors in Brazil.
- Performance evaluation of managers of the Portuguese Railway Company and directors of the Portuguese Gas Company.
- Evaluation of alternative sites for the new International Airport of Lisbon.
- Analysis of a new railway link through the Central Pyrenean.
- Establishing priorities for intervention in repairing Portuguese monuments.
- Policy analysis and resource allocation in Porto Development Plan 1998-2001.

## 6. References


[1] A.C.Z.B. Andrade and C.A. Bana e Costa. Measuring security value: a multicriteria decision analysis approach. *CEG-IST Research Paper* 20/2001, Lisbon, 2001.

[2] P. Antão da Silva, C.A. Bana e Costa and F.N. Correia. Avaliação multicritério das incidências ambientais de medidas de controlo de cheias: aplicação à bacia hidrográfica da Ribeira do Livramento. In *Trabalhos Técnicos do 1º Congresso sobre Aproveitamentos e Gestão de Recursos Hídricos em Países de Idioma Português* (CD-ROM edition), Rio de Janeiro: 210-217, 2000.

[3] P. Antão da Silva, C.A. Bana e Costa and F.N. Correia. Multicriteria evaluation of flood control measures: the case of Ribeira do Livramento. *CEG-IST Research Paper* 22/2001, Lisbon, 2001.

[4] C.A. Bana e Costa. Issues in facilitating bid evaluation in public call for tenders. In F. Khosrowshahi (ed.), *Proceedings of the 3rd International Conference on Decision Making in Urban and Civil Engineering* (CD-ROM edition), London, 2002.

[5] C.A. Bana e Costa. The use of multicriteria decision analysis to support the search for less conflicting policy options in a multi-actor context: case-study. *Journal of Multi-Criteria Decision Analysis*, 10(2): 111-125, 2001.

[6] C.A. Bana e Costa. O modelo de apoio à avaliação de propostas nos concursos do Metro do Porto. *FER XXI*, XVIII: 111-115, 1999.

[7] C.A. Bana e Costa. Structuration, Construction et Exploitation d'un Modèle Multicritère d'Aide à la Décision. *PhD Thesis*, Technical University of Lisbon, 1992.

[8] C.A. Bana e Costa and J.A. Antunes Ferreira. Concursos de obras públicas: metodologia de avaliação de propostas. Aplicação a uma grande empreitada de infraestruturas de transportes. In *Actas do Congresso 94, Engenharia Portuguesa na Viragem do Século,* Ordem dos Engenheiros, Lisbon: 47-w/9pages, 1994.

[9] C.A. Bana e Costa, J.A. Antunes Ferreira and E.C. Corrêa. Metodologia multicritério de apoio à avaliação de propostas em concursos públicos. In C.H. Antunes, L.V. Tavares (eds.), *Casos de Aplicação da Investigação Operacional*, McGraw-Hill, Lisbon: 336-363, 2000.

[10] C.A. Bana e Costa, J.A. Antunes Ferreira and J.C. Vansnick. Avaliação multicritério de propostas: o caso de uma nova linha do Metropolitano de Lisboa. *Revista de Transporte e Tecnologia*, VII(14): 31-65, 1995.

[11] C.A. Bana e Costa, L. Barroso and J.O. Soares. Qualitative modelling of credit scoring: a case study in banking. *Journal of European Research Studies*, V(1-2): 37-51, 2002.

[12] C.A. Bana e Costa and M.P. Chagas. A career choice problem: an example of how to use MACBETH to build a quantitative value model based on qualitative value judgments. *European Journal of Operational Research* (in press). (Preprint: *Working Paper LSEOR* 02.53, London School of Economics, 2002.)

[13] C.A. Bana e Costa and E.C. Corrêa. Framing public calls for tenders through a multicriteria approach. In D.K. Despotis, C. Zapounidis (eds.), *Proceedings of 5th International Conference of the Decision Sciences Institute (Volume II*), Athens: 1353-1355, 1999.

[14] C.A. Bana e Costa and E.C. Corrêa. Construction of a total quality index using a multicriteria approach: the case of Lisbon Gas Company. *CEG-IST Research Paper* 10/2000, Lisbon, 2000.
[15] C.A. Bana e Costa, E.C. Corrêa, J.M. De Corte and J.C. Vansnick. Facilitating bid evaluation in public call for tenders: a socio-technical approach. *OMEGA: The International Journal of Management Science*, 30(3): 227-242, 2002.
[16] C.A. Bana e Costa, E.C. Corrêa, L. Ensslin and J.C. Vansnick. Decision support systems in action: integrated application in a multicriteria decision aid process. *European Journal of Operational Research*, 113(2): 315-335, 1999.
[17] Bana e Costa, M.L. Costa-Lobo and I.A.J. Ramos. Contributo da metodologia multicritério na elaboração do Plano Estratégico de Barcelos. *Sociedade e Território*, 24: 102-115, 1997.
[18] C.A. Bana e Costa, M.L. da Costa-Lobo, I.A. Ramos and J.C. Vansnick. Multicriteria approach for strategic town planning: the case of Barcelos. In D. Bouyssou, E. Jacquet-Lagrèze, P. Perny, R. Słowiński, D. Vanderpooten and P. Vincke (eds.), *Aiding Decisions with Multiple Criteria. Essays in Honor of Bernard Roy*, Kluwer Academic Publishers, Boston: 429-456, 2002. (Preprint: *Working Paper LSEOR* 01.36, London School of Economics, 2001.)
[19] C.A. Bana e Costa, L. Ensslin and E. Holz. Avaliação ambiental em situação de conflito de interesses – Um estudo de caso. In IX SILUBESA, *Simpósio Luso-Brasileiro de Engenharia Sanitária e Ambiental* (CD-ROM edition), ABES/APRH, Porto Seguro (BA): 1914-1923, 2000.
[20] C.A. Bana e Costa and F. Nunes da Silva. Concepção de uma 'boa' alternativa de ligação ferroviária ao porto de Lisboa: uma aplicação da metodologia multicritério de apoio à decisão e à negociação. *Investigação Operacional*, 14(2): 115-131, 1994.
[21] C.A. Bana e Costa, F. Nunes da Silva and J.C. Vansnick. Conflict dissolution in the public sector: a case-study. *European Journal of Operational Research*, 130(2): 388-401, 2001.
[22] C.A. Bana e Costa and R.C. Oliveira. Assigning priorities for maintenance, repair and refurbishment in managing a municipal housing stock. *European Journal of Operational Research*, 138(2): 380-391, 2002.
[23] C.A. Bana e Costa, R.C. Oliveira and M.L. Alvarez. SMAI: sistema multicritério de avaliação de intervenções num parque habitacional municipal. In A.A. Bezelga, A.B. Leitão (eds.), *A Avaliação do Imobiliário*, APAE, Lisbon: 109-122, 1997.
[24] C.A. Bana e Costa, R.C. Oliveira and M.L. Alvarez. Concepção de um sistema multicritério de definição de prioridades de intervenção/conservação do património habitacional da Câmara Municipal de Lisboa. In *Comunicações ao 2º ENCORE – Encontro sobre Conservação e Reabilitação de Edifícios (Vol. II)*, LNEC, Lisbon: 925-934, 1994.
[25] C.A. Bana e Costa and J.O. Soares. A multicriteria model for portfolio management. *The European Journal of Finance*, 9: 1-14, 2003.
[26] C.A. Bana e Costa and J.F. Thomaz. Locating centres of information and recruitment of volunteers for the Portuguese Armed Forces: a decision-analysis case-study. In *Proceedings of the 42nd Annual Conference of the International Military Testing Association* (CD-ROM edition), IMTA, Edinburgh: 173-180, 2000. (Preprint: *Working Paper LSEOR* 01.38, London School of Economics, 2001.)

[27] C.A. Bana e Costa and J.C. Vansnick. Cardinal value measurement with MACBETH. In S.H. Zanakis, G. Doukidis and C. Zopounidis (eds.), *Decision Making: Recent Developments and Worldwide Applications*, Kluwer Academic Publishers, Dordrecht: 317-329, 2000.
[28] C.A. Bana e Costa and J.C. Vansnick. The MACBETH approach: basic ideas, software and an application. In N. Meskens and M. Roubens (eds.), *Advances in Decision Analysis*, Kluwer Academic Publishers, Dordrecht: 131-157, 1999.
[29] C.A. Bana e Costa and J.C. Vansnick. Preference relations and MCDM. In T. Gal, T.J. Stewart and T. Hanne (eds.), *Multicriteria Decision Making: Advances in MCDM Models, Algorithms, Theory and Applications*, Kluwer Academic Publishers, Boston: 4-1-4-23, 1999.
[30] C.A. Bana e Costa and J.C. Vansnick. Applications of the MACBETH approach in the framework of an additive aggregation model. *Journal of Multi-Criteria Decision Analysis*, 6(2): 107-114, 1997.
[31] C.A. Bana e Costa and J.C. Vansnick. A theoretical framework for measuring attractiveness by a categorical based evaluation technique (MACBETH). In J. Climaco (ed), *Multicriteria Analysis*, Springer-Verlag, Berlin: 15-24, 1997.
[32] C.A. Bana e Costa and J.C. Vansnick. The MACBETH approach: basic ideas. In *Proceedings of the International Conference on Methods and Applications of Multicriteria Decision Making*, FUCAM, Mons: 86-88, 1997.
[33] C.A. Bana e Costa and J.C. Vansnick. L'approche MACBETH: Note technique de presentation. *Research Paper of CESUR*, 779, Lisbon, 1997.
[34] C.A. Bana e Costa and J.C. Vansnick. General overview of the MACBETH approach. In P.M. Pardalos, Y. Siskos and C. Zopounidis (eds.), *Advances in Multicriteria Analysis*, Kluwer Academic Publishers, Dordrecht: 93-100, 1995.
[35] C.A. Bana e Costa and J.C. Vansnick. Uma nova abordagem ao problema de construção de uma função de valor cardinal: MACBETH. *Investigação Operacional*, 15(1): 15-35, 1995.
[36] C.A. Bana e Costa and J.C. Vansnick. MACBETH - An interactive path towards the construction of cardinal value functions. *International Transactions in Operational Research*, 1(4): 489-500, 1994.
[37] C.A. Bana e Costa and J.C. Vansnick. Sur la quantification des jugements de valeur : l'approche MACBETH. *Cahiers du Lamsade*, 117, Paris, 1993.
[38] C.A. Bana e Costa, A.C. Vitorino. Construção multicritério da 'matriz de custos' de um modelo de afectação para gestão de stands de aeronaves. *Investigação Operacional*, 17(2): 179-199, 1997.
[39] D. Bouyssou. Modelling inaccurate determination, uncertainty, imprecision using multiple criteria. In A.G. Lockett and G. Islei (eds.), *Improving Decision Making in Organisations*, Springer Verlag, Heidelberg: 78-87, 1989.
[40] F.J.P. Cardoso, L.F.A.M. Gomes, M. Sanglard Júnior. Administração das operações de telecomunicação: uma análise de decisão. *Revista Portuguesa e Brasileira de Gestão*, 2(2): 91-103, 2003.
[41] Catalyze Ltd. *HIVIEW3 Starter Guide*, Catalyze Ltd, London, 2003.
[42] C3E. Applying the Multi-criteria Method to the Evaluation of Structural Programmes. *MEANS Handbook*, 4, European Commission, DG XVI/02, Brussels, 1995.
[43] C3E. Evaluation Pilote Multicritère du Hainaut. *Rapport d'Expérience*, European Commission, DG XVI/02, Brussels, 1995.

[44] J.M. De Corte. Un logiciel d'exploitation d'informations préférentielles pour l'aide à la décision. Bases mathématiques et algorithmiques. *PhD Thesis*, University of Mons-Hainaut, Mons, 2002.
[45] L. Ensslin, A. Dutra and S. Ensslin. MCDA: a constructivist approach to the management of human resources at a governmental agency. *International Transactions in Operational Research*, 7(1): 79-100, 2000.
[46] L. Ensslin and S. Ensslin. Elaboração de um modelo construtivista para identificação de oportunidades de aperfeiçoamento de docentes do EPS-UFSC. *Revista Produto & Produção*, 2(3): 143-149, 1998.
[47] L. Ensslin, G.N. Montibeller and M.V.A. Lima. Constructing and implementing a DSS to help evaluate perceived risk of accounts receivable. In Y. Y. Haimes and R. E. Steuer (eds.), *Research and Practice in Multiple Criteria Decision Making*, Lecture Notes in Economics and Mathematical Systems, Vol. 487, Springer-Verlag, Berlin: 245-259, 2000.
[48] L. Ensslin, G.N. Montibeller and S.M. Noronha. *Apoio à Decisão: Metodologias para Estruturação de Problemas e Avaliação Multicritério de Alternativas*, Editora Insular, Florianópolis (SC), 2001.
[49] J.C. Mello E.G. Gomes, M.E. Lins and L.M. Vieira. Um caso de estudo de integração SIG-DEA-MCDA: a influência de uma instituição de ensino superior em vários municípios do estado do Rio de Janeiro. *Investigação Operacional*, 21(2): 171-190, 2001.
[50] C. Mousset and J.C. Vansnick. About the representation of a precardinal information (to appear).
[51] R.C. Oliveira and J.C. Lourenço. A multicriteria model for assigning new orders to service suppliers. *European Journal of Operational Research*, 139(2): 390-399, 2002.
[52] L.D. Phillips. A theory of requisite decision models. *Acta Psychologica*, 56: 29-48, 1984.
[53] J.L. Porto. Analyse multicritère dans le cadre des appels d'offres pour la construction de travaux publics et privés: le cas du Métro de Porto au Portugal. *Newsletter of the European Working Group "Multicriteria Aid for Decisions*, 15: 1-2, 1999. (http://www.inescc.pt/~ewgmcda/OpPorto.html)
[54] J.S.V. Silva and C.A. Bana e Costa. Especificação de uma ferramenta de apoio à decisão para gestão pública em regiões semi-áridas. *CEG-IST Research Paper* 23/2001, Lisbon, 2001.
[55] V. Vieira, C.S. Oliveira and C.A. Bana e Costa. Estruturação de um modelo de avaliação do grau de importância estratégica (MAGIE) de viadutos face à vulnerabilidade sísmica: aplicação à cidade de Lisboa. In C.A.P. Martins (ed.), *Sísmica 99 – 4º Encontro Nacional sobre Sismologia e Engenharia Sísmica / 2ème Rencontre en Génie Parasismique des Pays Méditerranéens*, Faro: 717-726, 1999.

Acknowledgments
The authors thank Manuel Pedro Chagas and two anonymous referees for their valuable comments on an earlier version of this paper.