

# A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid\*

Michel GRABISCH\*\*<sup>1</sup>, Christophe LABREUCHE<sup>2</sup>

<sup>1</sup> Centre d'Économie de la Sorbonne  
106-112, Bd. de l'Hôpital, 75103 Paris, France  
e-mail: michel.grabisch@univ-paris1.fr

<sup>2</sup> Thales Research & Technology  
RD 128, 91767 Palaiseau Cedex, France  
e-mail: christophe.labreuche@thalesgroup.com

Version of June 26, 2009

**Abstract** The main advances regarding the use of the Choquet and Sugeno integrals in multi-criteria decision aid over the last decade are reviewed. They concern mainly a bipolar extension of both the Choquet integral and the Sugeno integral, interesting particular submodels, new learning techniques, a better interpretation of the models and a better use of the Choquet integral in multi-criteria decision aid. Parallel to these theoretical works, the Choquet integral has been applied to many new fields, and several softwares and libraries dedicated to this model have been developed.

MSC Codes: 90B50, 90B99, 91A12, 91B06, 91C05.

---

**Keywords:** Choquet integral, Sugeno integral, capacity, bipolarity, preferences

## 1 Introduction

Despite its birth more than fifty years ago, the Choquet integral (Choquet, 1953) has not been studied and applied in decision making under uncertainty till the end of the eighties (see the pioneering work of Schmeidler (1986)), and till the beginning of the nineties for multi-criteria decision aid (MCDA).

---

\* This is an updated version of the paper that appeared in 4OR (2008) 6, 1-44.

\*\* Corresponding author.

If in decision under uncertainty, the use of the Choquet integral immediately received a firm theoretical basis through providing axiomatic characterizations (Schmeidler, 1989), the first works in MCDA remained on a rather intuitive and experimental level. Perhaps the idea of Murofushi of using the Shapley value as an importance index (Murofushi, 1992), and later his proposal of interaction index (Murofushi and Soneda, 1993), were the starting point for a more theoretical basis for the use of Choquet integral in MCDA. The survey paper of Grabisch (1996) provides a good snapshot of the situation in 1996, where these ideas of Shapley value and interaction index are developed, and perhaps has led people to use more and more the Choquet integral in MCDA.

More than one decade has passed since this first survey paper, and the field has been growing incredibly during this short period of time. The aim of the present survey paper is precisely to make a new snapshot of the situation, in order to make clear what are the main advances, and what are the theoretical and practical tools we have now at our disposal for tackling in a better way practical problems in MCDA.

The content and structure of the paper are dictated by the new advances obtained. It also draws on the analysis of the gaps that were identified in 1996. To this purpose, let us first detail the situation at that time, as given in (Grabisch, 1996):

- Properties for aggregation of the Choquet and Sugeno integrals were well known, as well as their relationship with classical aggregation operators.
- Interaction among criteria and global importance of criteria were well explained through the concepts of interaction index and Shapley value.
- Automatic learning methods were known, to determine an optimal model given a set of input/output data.
- Already many applications used these tools (mainly conducted in Japan).

What was lacking is the following:

1. Even if the semantics attached to capacities and the Choquet integral model was almost clear, the actual construction of the model was limited to either a blind and rather restrictive optimization method, or to ad hoc manipulations limited to toy problems. Moreover, the determination of the utility functions in a way that is consistent w.r.t. the Choquet integral was missing.
2. The exponential complexity of the model made its practical use very difficult. It was necessary to look for simpler models, since most of the time the full complexity of the model was not needed. Moreover, the interpretation in terms of decision making of these simpler models should be made clear.
3. Most of the time, scores to be aggregated were considered as nonnegative quantities, with 0 being the least possible score (unipolar model). But

many studies in psychology show that models in decision making should be bipolar, with value 0 indicating a neutral level demarcating “good” from “bad” scores. Considering this fact, is the definition of the Choquet integral still valid or rich enough?

4. The Sugeno integral was considered as the ordinal version of the Choquet integral, but its properties related to decision making were not well known, nor its exact situation in the realm of ordinal aggregation functions. Also, it was not possible to consider negative scores for aggregation (see the previous item).
5. If the Shapley value and the interaction index are useful tools for an introspective analysis of the model, are they sufficient? Can they explain all aspects? In particular, the degree to which a criterion is a veto or a favor is also needed.
6. The use of the Choquet integral in MCDA requires an in-depth analysis of the result of this model on several alternatives. The recommendation proposed by the multi-criteria model must be explained to the decision maker (DM). Moreover, an important question in practice is to know on which criterion an improvement should be done in order to get the maximal possible improvement of the overall score.
7. Lastly, powerful softwares and real-size applications were lacking.

Amazingly, all of the above aspects have been addressed and almost completely solved in the past decade, as this survey (see also Grabisch and Labreuche (2008)) will make clear: Section 2 addresses the first point (construction of the model), Section 2.7 presents important non-exponential submodels, Section 3 presents bipolar cardinal models, Section 4 addresses the ordinal case, Section 5 presents tools for the analysis of models, Section 6 presents methods for the exploitation of the models on several options, and lastly Section 7 presents some applications and software tools.

Size limitation forbid us to fully develop the above topics (this would take a whole book). We will try instead to focus on the main points and give references for further details.

Throughout the paper, the set of criteria is denoted by  $N := \{1, \dots, n\}$ . Min and max are denoted by  $\wedge, \vee$  respectively. For convenience, subsets of  $N$  will be denoted by uppercase letters, e.g.  $A \subset N$ , and their cardinality by the corresponding lowercase, e.g.,  $a = |A|$ .

## 2 Construction of a model based on the Choquet integral

Throughout this section, we suppose that scores, utility or value functions, etc., are nonnegative (unipolar case). The real-valued case will be addressed in Section 3 (bipolar case).

## 2.1 General framework and measurement scales

We consider a set  $X = X_1 \times \dots \times X_n$  of potential alternatives (e.g., candidates, cars, etc.), each alternative  $x := (x_1, \dots, x_n)$  being described by a vector of  $n$  descriptors or attributes (e.g., technical ability, purchase price, performances, etc.) taking values in sets  $X_1, \dots, X_n$ . The decision maker is supposed to have a preference over  $X$ , expressed by a binary relation  $\succeq$ , that is reflexive and transitive (possibly complete). The fundamental problem of decision theory is to build a numerical representation of  $\succeq$ . In the framework of this paper, this representation has the form

$$x \succeq y \Rightarrow F(u_1(x_1), \dots, u_n(x_n)) \geq F(u_1(y_1), \dots, u_n(y_n)), \quad (2.1)$$

where  $F$  is the Choquet integral (or the Sugeno integral in the ordinal case),  $u_i : X_i \rightarrow S$  ( $i = 1, \dots, n$ ) are called *utility functions* or *value functions* (the latter term will be used in the sequel, since the former refers more to decision under uncertainty and risk), and  $S \subseteq \mathbb{R}_+$  is a common scale<sup>1</sup> on which the preferences of the DM are represented.

We call the pair  $(X_i, u_i)$  a *criterion* (abusing terminology,  $X_i$  is also sometimes called a criterion), and  $N := \{1, \dots, n\}$  is the index set of criteria. For convenience, we define the *overall value function*  $U : X \rightarrow S$  by  $U(x) := F(u_1(x_1), \dots, u_n(x_n))$ .

More generally, when  $F$  is any increasing function from  $S^n$  to  $S$ , equation (2.1) is the so-called *decomposable model* of measurement theory (Krantz et al., 1971), and  $F$  is called an *aggregation function*. For details about the justification of the use of Choquet integral for  $F$ , see Grabisch and Labreuche (2005b); Grabisch (2005).

We can distinguish two types of scales when dealing with nonnegative real numbers:

- **Bounded unipolar scale:** this is the case when  $S$  is a closed bounded interval, e.g.,  $[0, 1]$ . Two typical examples of such scales are the scales of credibility of an event (belief or certainty degree, probability, etc.), and the scale of membership degree of an element to a fuzzy set (see, e.g., the excellent synthesis of fundamental aspects of fuzzy sets in Dubois et al. (2000a)). The boundaries 0 and 1 represent respectively the absence of the property (no credibility, no membership), and the total satisfaction of the property (full credibility, that is, the event is true; full membership, that is, the element belongs to the set in the classical sense). By contrast to bipolar scales (see Section 3), there is no opposite nor symmetric notion to the considered property.

---

<sup>1</sup> Employing the usual informal terminology. In measurement theory, the scale is a homomorphism between the set of objects to be measured and, e.g., the real numbers. Hence, the word “scale” should refer to the mapping  $u_i$  instead of  $S$ .

Coming back to our framework of MCDA, saying that  $u_i$  is a bounded unipolar scale implies the existence in  $X_i$  of two elements denoted by  $\mathbf{U}_i$  and  $\mathbf{P}_i$ , which have an absolute meaning:  $\mathbf{U}_i$  is an element of  $X_i$  which is thought by the DM as completely unsatisfactory relatively to his concerns w.r.t. criterion  $i$ , and  $\mathbf{P}_i$  is an element of  $X_i$  that is considered as perfectly satisfactory (Labreuche and Grabisch, 2003; Grabisch and Labreuche, 2005b; Grabisch, 2005). We set for convenience  $u_i(\mathbf{U}_i) = 0$  and  $u_i(\mathbf{P}_i) = 1$ .

- **Unbounded unipolar scale:** in this case  $S$  is no more bounded from above, hence  $S$  is taken to be  $\mathbb{R}_+$ . Typical examples are the scales of priority and importance (e.g., of obligations, laws, things to do, etc.). These notions are unipolar because no opposite notion exists. Moreover, it is always possible to find obligations more priority or more important than a given one, so that no upper bound exists.

In the framework of MCDA, the difference with the previous bounded case is that the element  $\mathbf{P}_i$  does not exist in  $X_i$ . Instead we assume the existence in  $X_i$  of an element denoted by  $\mathbf{S}_i$ , which the DM considers as good and completely satisfying if he could obtain it on criterion  $i$ , even if more attractive elements could exist. This special element corresponds to the *satisficing level* in the theory of bounded rationality of Simon (1956). We set for convenience  $u_i(\mathbf{S}_i) = 1$ .

In the rest of this section, any of these two scales can be used, but the same type of scale must be used for all criteria. For ease of notation, the unsatisfactory element on attribute  $X_i$  will be denoted by  $\mathbf{0}_i$ , and the notation  $\mathbf{1}_i$  will indicate either the upper bound  $\mathbf{P}_i$  (bounded unipolar scale) or the satisficing element  $\mathbf{S}_i$  (unbounded unipolar scale).

We introduce the following convenient notation: for two alternatives  $x, y \in X$  and a subset  $A \subseteq N$ , the compound alternative  $z := (x_A, y_{-A})$  is defined by  $z_i := x_i$  if  $i \in A$ , and  $z_i := y_i$  otherwise.

## 2.2 From the weighted sum to the Choquet integral

Most MCDA methods use as aggregation function the weighted arithmetic mean (weighted sum), i.e.,  $F(a_1, \dots, a_n) := \sum_{i=1}^n w_i a_i$ , with  $w_i \geq 0$  and  $\sum_{i=1}^n w_i = 1$ . It is well known however that in many situations, the weighted sum cannot represent the preferences of the decision maker. Let us try to illustrate this, and to motivate the definition of the Choquet integral.

*Example 2.1.* Let  $a, b, c$  be three alternatives evaluated on 2 criteria as follows:

$$\begin{aligned} u_1(a) &= 0.4, u_1(b) = 0, u_1(c) = 1 \\ u_2(a) &= 0.4, u_2(b) = 1, u_2(c) = 0, \end{aligned}$$

where scores are given in  $[0, 1]$ . Suppose that the decision maker (DM) says  $a \succ b \sim c$ . Let us find  $w_1, w_2$  so that the weighted sum represents the preference. We get:

$$\begin{aligned} b \sim c &\Leftrightarrow w_1 = w_2 \\ a \succ b &\Leftrightarrow 0.4(w_1 + w_2) > w_2 \end{aligned}$$

equivalent to  $0.8 w_2 > w_2$ , which is impossible.

To understand the underlying reason of this failure, we should notice that for the weighted sum,  $w_1$  is the overall score achieved by an alternative having a totally satisfactory score on the first criterion (1), and not acceptable on the others (0). Obviously, our DM is more attracted by alternatives being well balanced on the two criteria. It would be possible to take into account this preference if one allows to define weights not only on each criterion, but also on groups of criteria. In our very simple example, this amounts to defining a weight  $w_{12}$  on both criteria, which represents the score assigned to an alternative being totally satisfactory on both criteria. This alternative being obviously the best one, it is natural to assign to it the maximal score 1, hence  $w_{12} = 1$ . In order to model the fact that the DM considers alternatives being satisfactory only on one criterion as not attractive, we may set, e.g.,  $w_1 = w_2 = 0.3$ . Let us try to rewrite the weighted sum, taking into account this new weight  $w_{12}$ . Keeping in mind the interpretation of weights, we are led to the following computation:

- $a$  has equal scores on both criteria, which corresponds to the situation depicted by  $w_{12}$ , up to the factor 0.4. Supposing the model to be homogeneous, we may put  $U(a) = 0.4 w_{12} = 0.4$ .
- $b$  and  $c$  correspond respectively to the situations depicted by  $w_2, w_1$ , hence  $U(b) = w_2 = 0.3$ , and  $U(c) = w_1 = 0.3$ .

The model indeed represents the preference of the DM. It is easy to see that by choosing appropriate values for  $w_1, w_2, w_{12}$ , *any* preference among  $a, b, c$  can be represented this way.

The above example works well because the alternatives  $a, b, c$  fit exactly to the situations depicted by the weights. What if this is no more the case, for example considering an alternative  $d$  such that  $u_1(d) = 0.2$  and  $u_2(d) = 0.8$ ? We may consider that the DM prefers  $d$  to  $b$  and  $c$ , and  $a$  to  $d$ . To solve the problem, we consider that  $d$  is the sum of two fictitious alternatives  $d', d''$  defined by:

$$\begin{aligned} u_1(d') &= 0.2, u_1(d'') = 0 \\ u_2(d') &= 0.2, u_2(d'') = 0.6. \end{aligned}$$

Supposing that our model is additive for such alternatives, the overall score of  $d$  is the sum of the overall scores of  $d'$  and  $d''$ . But it is possible to compute

them, because  $d', d''$  correspond to situations depicted by the weights. We obtain:

$$\begin{aligned} U(d') &= 0.2 w_{12} = 0.2 \\ U(d'') &= 0.6 w_2 = 0.18 \\ U(d) &= U(d') + U(d'') = 0.38. \end{aligned}$$

Observe that we obtain the desired ranking:  $a \succ d \succ b \sim c$ .

This method of computing the overall score is in fact nothing else than the Choquet integral, and the weights on groups of criteria define a capacity or fuzzy measure.

### 2.3 Basic definitions

**Definition 2.2.** 1. A function  $\nu : 2^N \rightarrow \mathbb{R}$  is a game if it satisfies  $\nu(\emptyset) = 0$ .  
2. A game  $\mu$  which satisfies  $\mu(A) \leq \mu(B)$  whenever  $A \subseteq B$  (monotonicity) is called a capacity (Choquet, 1953) or fuzzy measure (Sugeno, 1974). The capacity is normalized if in addition  $\mu(N) = 1$ .

A capacity is *additive* if for all disjoint  $A, B \subseteq N$ , we have  $\mu(A \cup B) = \mu(A) + \mu(B)$ . A capacity is *symmetric* if for any subsets  $A, B$ ,  $|A| = |B|$  implies  $\mu(A) = \mu(B)$ . The *conjugate* or *dual* of a capacity  $\mu$  is a capacity  $\bar{\mu}$  defined by

$$\bar{\mu}(A) := \mu(N) - \mu(A), \quad \forall A \subseteq N.$$

**Definition 2.3.** Let us consider  $f : N \rightarrow \mathbb{R}_+$  (or equivalently a vector in  $\mathbb{R}_+^n$ ). The Choquet integral of  $f$  w.r.t. a capacity  $\mu$  is given by

$$\mathcal{C}_\mu(f) := \sum_{i=1}^n [f_{\sigma(i)} - f_{\sigma(i-1)}] \mu(\{\sigma(i), \dots, \sigma(n)\}),$$

where  $f_i$  stands for  $f(i)$ ,  $\sigma$  is a permutation on  $N$  such that  $f_{\sigma(1)} \leq \dots \leq f_{\sigma(n)}$ , and  $f_{\sigma(0)} := 0$ .

The above definition is also valid if  $\mu$  is a game. A fundamental property is:

$$\mathcal{C}_\mu(1_A, 0_{-A}) = \mu(A), \quad \forall A \subseteq N. \quad (2.2)$$

Two particular cases are of interest.

- If  $\mu$  is additive, then the Choquet integral reduces to a weighted arithmetic mean:

$$\mathcal{C}_\mu(f) = \sum_{i \in N} \mu(\{i\}) f_i.$$

- If  $\mu$  is symmetric, the Choquet integral reduces to the so-called Ordered Weighted Average (OWA) introduced by Yager (Yager, 1988):

$$C_\mu(f) = \sum_{i \in N} (\mu_{n-i+1} - \mu_{n-i}) f_{\sigma(i)}$$

with  $\mu_i := \mu(A)$ , such that  $|A| = i$ , and  $\sigma$  is defined as before.

**Definition 2.4.** *Let us consider  $f : N \rightarrow \mathbb{R}_+$ . The Sugeno integral (Sugeno, 1974) of  $f$  w.r.t. a capacity  $\mu$  is given by*

$$S_\mu(f) := \bigvee_{i=1}^n [f_{\sigma(i)} \wedge \mu(\{\sigma(i), \dots, \sigma(n)\})],$$

with same notation as above.

Note that the above definition also works if  $f, \mu$  are valued on some ordinal scale (possibly finite).

We introduce two important linear transformations over capacities.

**Definition 2.5.** *Let  $\nu$  be a game on  $N$ .*

1. *The Möbius transform of  $\nu$ , denoted by  $m^\nu$ , is the unique solution of the equation*

$$\nu(A) = \sum_{B \subseteq A} m^\nu(B), \quad \forall A \subseteq N,$$

given by

$$m^\nu(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} \nu(B).$$

2. *The interaction transform of  $\nu$ , denoted by  $I^\nu$ , is defined by*

$$I^\nu(A) := \sum_{B \subseteq N \setminus A} \frac{(n-b-a)!b!}{(n-a+1)!} \sum_{K \subseteq A} (-1)^{|A \setminus K|} \nu(B \cup K), \quad \forall A \subseteq N. \quad (2.3)$$

The meaning of the interaction transform will be detailed in Section 5.1. All details concerning these transformations as well as others can be found in Grabisch et al. (2000); Denneberg and Grabisch (1999); Grabisch (1997a) (for summaries see Grabisch (2000, 2002)). The value of  $I^\nu$  for singletons plays a special role, and is called the *Shapley value* (Shapley, 1953), usually denoted by  $\phi_i(\nu)$ ,  $i \in N$ :

$$\phi_i(\nu) := I^\nu(\{i\}) = \sum_{A \subseteq N \setminus i} \frac{(n-a-1)!a!}{n!} [\nu(A \cup i) - \nu(A)].$$

## 2.4 Construction of a model based on the MACBETH approach

This section is based on Labreuche and Grabisch (2003). Our main ingredient for the construction of the model is the existence of the two special elements  $\mathbf{0}_i$  and  $\mathbf{1}_i$  on each  $X_i$  representing the scale  $S$  (see Section 2.1).

The MACBETH approach (Bana e Costa and Vansnick, 1994, 1999) is a MCDA methodology enabling the construction of value functions and weights in the weighted sum in a meaningful way in the sense of measurement theory. It is based on the existence of two reference levels, such as the two previously presented ones. The construction of the Choquet model is obtained by a generalization of the MACBETH approach and is done in two steps: the construction of value functions  $u_1, \dots, u_n$  by using intra-criterion information, and the construction of  $F$ , or more precisely of the capacity  $\mu$ , by using inter-criteria information.

*Construction of value functions.* In the literature, the value function on an attribute  $X_i$  is often constructed by asking questions directly regarding the preference of the DM on the set  $X_i$  rather than his preferences on  $X$ . This is completely justified when the aggregation function is a weighted sum since each criterion can be isolated thanks to the preferential independence property. This is no longer true with the Choquet integral since criteria interact together. The construction of the value function on  $X_i$  must then be performed by only asking information regarding elements of  $X$ . We consider a particular subset of  $X$ :

$$X_i] := \{(\mathbf{0}_1, \dots, \mathbf{0}_{i-1}, x_i, \mathbf{0}_{i+1}, \dots, \mathbf{0}_n \mid x_i \in X_i)\}.$$

Using our notation, elements of  $X_i]$  are denoted more conveniently by  $(x_i, \mathbf{0}_{-i})$ . The MACBETH approach allows, by assuming that the DM is able to give information using intensity of preference, to build an interval scale  $v_i$  encoding the attractiveness of elements of  $X_i]$ . Since interval scales are determined up to a shift and dilation, i.e.,  $v'_i := \alpha v_i + \beta$  with  $\alpha > 0$  encodes the same information as  $v_i$ , assuming  $(\mathbf{1}_i, \mathbf{0}_{-i}) \succ \mathbf{0}$ , we choose the unique  $v_i$  satisfying

$$v_i(\mathbf{0}) = 0, \quad v_i(\mathbf{1}_i, \mathbf{0}_{-i}) = 1$$

where  $\mathbf{0} := (\mathbf{0}_1, \dots, \mathbf{0}_n)$ . Due to this normalization, we define  $u_i(x_i) := v_i(x_i, \mathbf{0}_{-i})$ , for all  $i \in N$ . Hence all value functions are built, with the property they all coincide for the  $\mathbf{0}_i$  levels and for the  $\mathbf{1}_i$  levels. For this reason, they are called *commensurate*<sup>2</sup>.

---

<sup>2</sup> A more precise definition would be the following. Two scales  $u_i, u_j$  on criteria  $i$  and  $j$  are said to be *commensurate* if for every  $x_i, x_j$  such that  $u_i(x_i) = u_j(x_j)$ , the degrees of satisfaction felt by the DM on criteria  $i$  and  $j$  are equal. One convenient way to achieve

*Construction of the Choquet integral.* We consider another subset of  $X$ :

$$X]_{\{0,1\}} := \{(\mathbf{1}_A, \mathbf{0}_{-A}) \mid A \subseteq N\},$$

Again, we use the MACBETH approach, under the same assumptions, to build an interval scale  $u_{\{0,1\}}$  on  $X]_{\{0,1\}}$ , encoding the attractiveness of its elements. Among all possible interval scales, we choose the one satisfying

$$u_{\{0,1\}}(\mathbf{0}) = 0, \quad u_{\{0,1\}}(\mathbf{1}) = 1$$

where  $\mathbf{1} = (\mathbf{1}_1, \dots, \mathbf{1}_n)$ . Note that since  $(\mathbf{1}_i, \mathbf{0}_{-i}) \in X_i] \cap X]_{\{0,1\}}$ ,  $v_i(\mathbf{1}_i, \mathbf{0}_{-i}) = 1$ , and both scales  $v_i$  and  $u_{\{0,1\}}$  have the same 0, it is necessary to have  $u_{\{0,1\}}(\mathbf{1}_i, \mathbf{0}_{-i}) > 0$ , otherwise the DM is inconsistent (since we assumed  $(\mathbf{1}_i, \mathbf{0}_{-i}) \succ \mathbf{0}$ ).

We assume that the DM satisfies a dominance property for alternatives in  $X]_{\{0,1\}}$ , i.e., if  $A \subseteq B$ , we have  $u_{\{0,1\}}(\mathbf{1}_A, \mathbf{0}_{-A}) \leq u_{\{0,1\}}(\mathbf{1}_B, \mathbf{0}_{-B})$ . Let us define  $\mu(A) := u_{\{0,1\}}(\mathbf{1}_A, \mathbf{0}_{-A})$  for all  $A \subseteq N$ . Then clearly  $\mu$  is a normalized capacity on  $N$ .

From the above assumptions, it can be proved that  $u_{\{0,1\}}(\mathbf{1}_A, \mathbf{0}_{-A}) = F(\mathbf{1}_A, \mathbf{0}_{-A})$  for all  $A \subseteq N$ , i.e.,  $F$  is an extension on  $S^n$  of the capacity  $\mu$ .

The fact that  $F$  can be taken as the Choquet integral  $\mathcal{C}_\mu$  is now justified by equation 2.2. Since  $\mu$  is determined by the above procedure, and the value functions are built, the construction is complete.

## 2.5 The Choquet integral as a parsimonious linear interpolator

In the previous section, we have taken for granted that  $F$  was the Choquet integral. In this section, we show that it is in some sense the best possible choice.

Coming back to the previous section, we have determined numbers  $u_{\{0,1\}}(\mathbf{1}_A, \mathbf{0}_{-A})$  for all  $A \subseteq N$ , which are the values of  $F$  on the vertices of the hypercube  $[0, 1]^n$ . Determining  $F$  on the whole hypercube then becomes a problem of interpolation. As many types of interpolation exist, we are looking here for a linear interpolation using the fewest possible points (parsimonious linear interpolation). For a given  $x := (x_1, \dots, x_n)$  in  $[0, 1]^n$ , let us denote by  $\mathcal{V}(x)$  the set of vertices used for the linear interpolation, which is written as

$$F(x) = \sum_{A \subseteq N \mid (\mathbf{1}_A, \mathbf{0}_{-A}) \in \mathcal{V}(x)} \left[ \alpha_0(A) + \sum_{i=1}^n \alpha_i(A) x_i \right] F(\mathbf{1}_A, \mathbf{0}_{-A}), \quad (2.4)$$

---

this is the use of two absolute levels existing on each scale as above, provided the scales are interval scales. The commensurateness issue is crucial for the method we present here. Note however that this assumption is not necessary in (2.1) and other models of conjoint measurement and multiattribute utility theory.

where  $\alpha_i(A) \in \mathbb{R}$ ,  $i = 0, \dots, n$ ,  $\forall A \subseteq N$ . To keep the meaning of interpolation, we impose that the convex hull  $\text{conv}(\mathcal{V}(x))$  contains  $x$ , and any  $x \in [0, 1]^n$  should belong to a unique polyhedron  $\text{conv}(\mathcal{V}(x))$  (except for common facets), and that continuity should hold. To ensure a minimal number of vertices for these polyhedra, they should be  $(n + 1)$ -dimensional simplices.

Still many different triangulations using simplices are possible, but there is one which is of particular interest, since it leads to an interpolation where all constant terms  $\alpha_0(A)$  are null. This triangulation uses the  $n!$  *canonical simplices* of  $[0, 1]^n$ :

$\text{conv}(\mathcal{V}_\sigma) = \{x \in [0, 1]^n \mid x_{\sigma(1)} \leq \dots \leq x_{\sigma(n)}\}$ , for some permutation  $\sigma$  on  $N$ .

**Proposition 2.6.** *The linear interpolation (2.4) using the canonical simplices is written as*

$$F(x) = \sum_{i=1}^n [x_{\sigma(i)} - x_{\sigma(i-1)}] \mu(\{\sigma(i), \dots, \sigma(n)\}), \quad \forall x \in \text{conv}(\mathcal{V}_\sigma), \quad (2.5)$$

where  $\mu(A) = F(1_A, 0_{-A})$ . Moreover,  $F$  is continuous on  $[0, 1]^n$ .

We recognize in (2.5) the Choquet integral. Lovász (1983) discovered this formula by considering the problem of extending the domain of pseudo-Boolean functions to  $\mathbb{R}^n$ . Later, Singer (1984) proved the above result (uniqueness of the interpolation). The fact that the so-called *Lovász extension* is the Choquet integral was observed by Marichal (1998, 2002).

## 2.6 Construction of the capacity by optimization methods

The construction presented in Section 2.4 allows us to construct both the value functions and the capacity. It is complete, well founded, and can be done in practice. However, although the construction of value functions has a complexity in  $O(n)$  and thus is always possible, the construction of the capacity has an exponential complexity, hence is tractable only for small values of  $n$ . Moreover, since the options of  $X \upharpoonright_{\{0,1\}}$  do not correspond to real alternatives, the DM may not feel comfortable when comparing these options. He may prefer to use more realistic options. Hence, the interest of the method described in Section 2.4 to construct the capacity is more theoretical than practical. An alternative is to make best use of the information provided by the DM on some set of alternatives, or any other kind of information (importance of criteria, interaction, etc.). In other words, once the value functions have been determined, the construction of the capacity reduces to an optimization problem under constraints.

There are various choices for the objective function and the constraints, which we summarize below (see Grabisch et al. (2008b) for a more detailed survey). The general form is:

$$\begin{array}{l} \text{Minimize } g(\mu, \dots) \\ \text{Subject to } \left\{ \begin{array}{l} \mu(A \cup i) - \mu(A) \geq 0, \quad \forall i \in N, \forall A \subseteq N \setminus i \\ \mu(N) = 1 \\ \mathcal{C}_\mu(u(x)) - \mathcal{C}_\mu(u(x')) \geq \delta_C \\ \vdots \\ \phi_i(\mu) - \phi_j(\mu) \geq \delta_S \\ \vdots \\ \text{constraints on } I^\mu \\ \vdots \end{array} \right. \end{array}$$

The variables of the optimization problem are the  $2^n - 1$  unknown values of  $\mu$ . Since the transforms presented in Definition 2.5 are linear and invertible, the values of the Möbius transform  $m^\mu$  or the interaction transform  $I^\mu$  can be used for variables as well, without altering the linear character of the constraints. In particular, if one uses  $k$ -additive capacities instead of general capacities (see Section 2.7.1), then the Möbius transform is the best suited representation since it allows to considerably reduce the number of variables. In some cases, one may have additional variables, such as thresholds, etc.

The first set of constraints simply depicts the monotonicity of the capacity, thus it should always be present. The second constraint is the normalization of the capacity. It is necessary only if idempotency is required, i.e.,  $\mathcal{C}_\mu(a, a, \dots, a) = a$ , for any  $a \in \mathbb{R}_+$ . The third set of constraints depicts the preference of the decision maker. Here we put for convenience  $u(x) := (u_1(x_1), \dots, u_n(x_n))$ . It translates equation (2.1) with a threshold  $\delta_C$ , which has to be fixed arbitrarily. The fourth set of constraints translates information concerning the importance of criteria. As it will be explained in Section 5.1, the Shapley value represents the overall importance of each criterion, and this set of constraints permits to translate statements like “criterion  $i$  is more important than criterion  $j$ ”. The last set of constraints concerns interaction among criteria (see again Section 5.1). Most of the time, it only concerns interaction between two criteria, and could take various forms (sign of the interaction, difference above a given threshold, etc.). Note that only the first set of constraints is mandatory, while all the others may be absent.

An important remark is that in some cases the problem may be infeasible due to contradictory constraints provided by the decision maker (e.g., violation of dominance), or because the model is not powerful enough to

take into account the preference of the decision maker (e.g., bi-capacities may be required, see Section 3). It is possible to get rid of infeasibility by reconsidering the information provided by the decision maker (Labreuche and Le Huédé, 2006).

We give some examples of objective functions.

- **Minimizing the total squared error.** Suppose that for a subset  $X_0$  of alternatives, we know the overall score  $y(x)$  that should be assigned to an alternative  $x \in X_0$ . Then a natural criterion is to minimize the total squared error between the desired output  $y(x)$  and the output of the model  $\mathcal{C}_\mu(u(x))$ :

$$g(\mu) = \sum_{x \in X_0} [\mathcal{C}_\mu(u(x)) - y(x)]^2.$$

This is a quadratic program, not always strictly convex, hence it does not have in general a unique solution (see an investigation on this aspect in Miranda and Grabisch (1999)). Usually, the set of constraints reduces to monotonicity and normalization of  $\mu$ . This approach has been first proposed by Tanaka and Murofushi (1989), and extensively used by Grabisch for classification problems (Grabisch and Nicolas, 1994), and subjective evaluation (Grabisch et al., 2002).

Although the use of such a criterion is somewhat odd in decision theory (usually  $y(x)$  is not known) and much more related to the field of estimation theory, it has been widely applied. In order to avoid the use of quadratic solvers which need a huge amount of memory (in  $O(2^{2n})$ , if general capacities are used) and sometimes give strange results, there exist several other approaches which are less greedy (generally in  $O(2^n)$  for memory allocation), although suboptimal:

- HLMS (Heuristic Least Mean Squares) is a gradient descent type algorithm tailored for the Choquet integral, and proposed by Grabisch (1995). It has the advantage over the quadratic approach to be able to work with very few learning data, and to provide less extremal solutions (i.e., closer to the uniformly distributed capacity). However, the algorithm uses the values of  $\mu$  as variables, and only monotonicity constraints can be taken into account, hence it is not possible to handle  $k$ -additive models. A statistical comparison of HLMS and the quadratic approach is described in Grabisch and Raufaste (to appear).
- There exist numerous algorithms based on genetic algorithms, most of them being restricted to  $\lambda$ -measures (see Section 2.7). A good representative of such a family, not restricted to  $\lambda$ -measures, is given by Wang et al. (1999). Grabisch (2003b) proposed a version handling  $k$ -additive capacities, while Combarro and Miranda (2003) proposed an original approach exploiting the convexity of the set of capacities.

- Lastly we mention Kwon and Sugeno (2000) and Sugeno and Kwon (1995) who propose, in order to avoid the exponential complexity, to replace the Choquet integral w.r.t. a single capacity  $\mu$  by a sum of  $p$  Choquet integrals w.r.t.  $\mu_1 \dots, \mu_p$  defined on subsets  $C_1, \dots, C_p$  of  $N$ , such that  $C_1, \dots, C_p$  form a covering of  $N$ . The idea stems from the work of Fujimoto about inclusion-exclusion covering (Fujimoto and Murofushi, 2000).
- **Maximum separation of alternatives.** This method proposed by Marichal and Roubens (2000) consists in maximizing the difference in overall scores among alternatives: if the decision maker prefers (in the strict sense)  $x$  to  $x'$ , then this should be reflected in the model by two sufficiently different outputs. The objective function is simply  $g = -\epsilon$ , and the third set of constraints becomes:

$$C_\mu(u(x)) - C_\mu(u(x')) \geq \delta_C + \epsilon.$$

This gives a linear program, hence very easy to solve with standard algorithms. However, as the least squares approach, this method does not necessarily give a unique solution. Moreover, the solution can be sometimes considered as too extreme.

Based on this method, Meyer and Roubens (2005), Marichal et al. (2005) have built a MCDA method called TOMASO. It starts from ordinal information, which is transformed into cardinal information (see Section 4) by computing the number of times a given alternative is better than other ones for a given criterion.

- **Minimum variance method.** The idea of the minimum variance method (Kojadinovic, 2007a) is to produce a “least specific” capacity (i.e., closest to the uniformly distributed capacity) compatible with the preference of the decision maker. The objective function is similar to a variance:

$$g(\mu) = \frac{1}{n} \sum_{i \in N} \sum_{S \subseteq N \setminus i} \frac{(n-s-1)!s!}{n!} \left( \sum_{T \subseteq S} m^\mu(T \cup i) - \frac{1}{n} \right)^2,$$

and the constraints are the three first sets of constraints. Minimizing  $g$  amounts to maximizing the extended Havrda and Charvat entropy of order 2 (Havrda and Charvat, 1967). It is a strictly convex quadratic program, with a unique solution. It has similar features as HLMS, since it can work with very few learning data (here, these are preferences over alternatives), and does not produce extreme solutions.

## 2.7 Particular submodels

There are two main drawbacks of the Choquet integral, which are interrelated: its exponential complexity ( $2^n - 2$  real values are needed to define

a normalized capacity), and the difficulty to interpret these values, and consequently to analyze the behaviour of the Choquet integral. Several particular families of capacities, hence inducing submodels, have been proposed to solve this issue, the most important ones being the  $k$ -additive capacities (Grabisch, 1997a), the  $p$ -symmetric capacities (Miranda et al., 2002), and the  $k$ -tolerant and  $k$ -intolerant capacities (Marichal, 2004), which are presented below. Note that a very commonly used family is that of  $\lambda$ -measures, introduced by Sugeno (Sugeno, 1973, 1977). Although convenient, they are merely distorted probabilities, and thus are too restrictive for MCDA.

### 2.7.1 $k$ -additive capacities

**Definition 2.7.** (Grabisch, 1997a) A capacity  $\mu$  is  $k$ -additive if its Möbius transform satisfies  $m^\mu(A) = 0$  for all  $A \subseteq N$  such that  $|A| > k$ , and there exists  $A \subseteq N$ ,  $|A| = k$ , such that  $m^\mu(A) \neq 0$ .

An important property is that  $\mu$  is  $k$ -additive if and only if (i) for all  $A \subseteq N$ ,  $|A| = k$ , we have  $I^\mu(A) = m^\mu(A)$ , and (ii) for all  $A \subseteq N$ ,  $|A| > k$ , we have  $I^\mu(A) = 0$ .

1-additive capacities are ordinary additive capacities. The  $k$ -additivity property fixes the degree of interaction between criteria: 1-additivity does not permit interaction, 2-additivity allows interaction up to 2 criteria, etc.

A  $k$ -additive capacity needs only  $\sum_{i=1}^k \binom{n}{i}$  coefficients to be defined, instead of  $2^n - 2$ . In practice, 2-additivity is probably the best compromise between low complexity and richness of the model.

As we already know, the Choquet integral w.r.t 1-additive capacities is a weighted arithmetic mean.

The expression of the Choquet integral w.r.t. 2-additive capacities is of particular interest. For any 2-additive capacity  $\mu$ , and any real-valued function  $f$  on  $N$ , we obtain (see Grabisch (1997b))

$$\mathcal{C}_\mu(f) = \sum_{i,j \in N | I_{ij} > 0} (f_i \wedge f_j) I_{ij} + \sum_{i,j \in N | I_{ij} < 0} (f_i \vee f_j) |I_{ij}| + \sum_{i \in N} f_i \left[ \phi_i - \frac{1}{2} \sum_{j \neq i} |I_{ij}| \right] \quad (2.6)$$

where  $\phi_i$  is the Shapley value of  $\mu$ , and  $I_{ij} := I^\mu(\{i, j\})$  is the interaction index between criteria  $i$  and  $j$ . The formula is remarkable for two reasons:

- It explains well the meaning of the interaction index and Shapley value (see Section 5.1): a positive interaction induces a conjunctive aggregation of scores (*necessarily both* scores have to be high to produce a high overall score), while a negative interaction induces a disjunctive aggregation (it is *sufficient that one* score is high). Clearly, the Shapley value is the linear part of the model, while interaction is the nonlinear part.
- Coefficients are nonnegative, and moreover, if the capacity is normalized, they sum up to 1. In other words, this means that the Choquet integral

is a convex combination of the scores  $f_i$  on all criteria, and of all disjunctive and conjunctive combinations of scores on pairs of criteria. Hence, the coefficient of a given term can be interpreted as the percentage of contribution of such term to the overall score. This feature is highly appreciated in practice (see Sections 6.1 and 7.3).

There is an alternative expression of the Choquet integral w.r.t. 2-additive capacities:

$$\mathcal{C}_\mu(f) = \sum_{i=1}^n \phi_i f_i - \sum_{\{i,j\} \subseteq N} \frac{I_{ij}}{2} |f_i - f_j|.$$

Remarkably the weights of the linear part are the coefficients of the Shapley value. One sees that if  $I_{ij} > 0$ , the more  $f_i$  is different from  $f_j$ , the more the interaction phenomenon penalizes the overall assessment  $\mathcal{C}_\mu(f)$ . More precisely, if  $f_i > f_j$ , the good evaluation of  $f$  on criterion  $i$  is *penalized* to a degree  $I_{ij}/2$  by the worse evaluation on criterion  $j$ . If  $I_{ij} < 0$ , the more  $f_i$  is different from  $f_j$ , the more the interaction phenomenon increases the overall assessment. More precisely, if  $f_i < f_j$ , the bad evaluation of  $f$  on criterion  $i$  is *saved* to a degree  $|I_{ij}|/2$  by the better evaluation on criterion  $j$ .

*2.7.2 p-symmetric capacities*  $k$ -additive capacities generalize the notion of additivity. Similarly,  $p$ -symmetric capacities, introduced by Miranda et al. (2002), generalize symmetric capacities, and also offer a hierarchy of more and more complex models.

A subset  $A$  is a *subset of indifference* for  $\mu$  if for all  $B_1, B_2 \subseteq A$  such that  $|B_1| = |B_2|$ , we have  $\mu(C \cup B_1) = \mu(C \cup B_2)$ , for all  $C \subseteq N \setminus A$ . Observe that any subset of a subset of indifference is also a subset of indifference, and that any singleton is a subset of indifference.

**Definition 2.8.** A capacity  $\mu$  on  $N$  is  $p$ -symmetric if the (unique) coarsest partition of  $N$  into subsets of indifference contains exactly  $p$  subsets  $A_1, \dots, A_p$ . The partition  $\{A_1, \dots, A_p\}$  is called the *basis* of  $\mu$ .

In the above definition, a partition  $\pi$  is coarser than another partition  $\pi'$  if all subsets of  $\pi$  are union of some subsets of  $\pi'$ .

Clearly, a 1-symmetric capacity is a symmetric capacity. Considering a basis  $\{A_1, \dots, A_p\}$ , any subset  $B \subseteq N$  can be identified with a  $p$ -dimensional vector  $(b_1, \dots, b_p)$ , with  $b_i := |B \cap A_i|$ . Hence, a  $p$ -symmetric capacity needs  $\prod_{i=1}^p (|A_i| + 1)$  coefficients to be defined.

The Choquet integral for 1-symmetric capacities is just an OWA (see Section 2.3). For  $p$ -symmetric capacities with basis  $\{A_1, \dots, A_p\}$ , the formula becomes:

$$\mathcal{C}_\mu(f) = \sum_{i=1}^p \mathcal{C}_{\mu|_{A_i}}(f|_{A_i}) + \sum_{B|B \not\subseteq A_j, \forall j} m^\mu(B) \bigwedge_{i \in B} f_i,$$

where  $m^\mu$  is the Möbius transform of the  $p$ -symmetric capacity,  $\mu|_{A_i}$  is the restriction of  $\mu$  to  $A_i$ , i.e.,

$$\mu|_{A_i}(C) := \mu(C), \forall C \subseteq A_i,$$

and  $x|_{A_i}$  is the restriction of  $x$  to  $A_i$ .

The above expression is slightly simpler than the one given in Miranda et al. (2002). An important observation is that for  $i = 1, \dots, p$ ,  $\mu|_{A_i}$  is a symmetric nonnormalized capacity on  $A_i$ , therefore  $\mathcal{C}_{\mu|_{A_i}}$  is a classical OWA operator on  $A_i$ , with nonnegative weights  $w_{i_1}, \dots, w_{i_{|A_i|}}$  satisfying  $\sum_{j=1}^{|A_i|} w_{i_j} = \mu(A_i)$ .

**2.7.3  $k$ -intolerant capacities** Suppose a Choquet integral  $\mathcal{C}_\mu$  is used to aggregate scores on criteria, and suppose the output value  $\mathcal{C}_\mu(f)$  of  $f$  is always bounded above by the  $k$ th lowest coordinate  $f_{\sigma(k)}$  of  $f$ . Then, clearly, this Choquet integral has a somehow intolerant behavior. The lower the value of  $k$ , the more intolerant the behavior. This suggests the following definition (Marichal, 2004), where as before  $\sigma$  is a permutation on  $N$  such that  $f_{\sigma(1)} \leq \dots \leq f_{\sigma(n)}$ .

**Definition 2.9.** A Choquet integral  $\mathcal{C}_\mu$  (or equivalently its underlying capacity  $\mu$ ) is at most  $k$ -intolerant if  $\mathcal{C}_\mu(f) \leq f_{\sigma(k)}$ . It is  $k$ -intolerant if, in addition,  $\mathcal{C}_\mu(f) \not\leq f_{\sigma(k-1)}$ , where  $f_{\sigma(0)} := 0$  by convention.

It can be shown that  $\mathcal{C}_\mu$  is at most  $k$ -intolerant if and only if  $\mu(A) = 0$ ,  $\forall A \subseteq N$  such that  $|A| \leq n - k$ .

The dual notion of  $k$ -tolerant capacities can be introduced as well: then  $\mathcal{C}_\mu \geq f_{\sigma(k)}$ , which is equivalent to  $\mu(A) = 1$ ,  $\forall A \subseteq N$  such that  $|A| \geq k$ .

Another form of intolerance can be expressed through the concept of *veto criterion* (Grabisch, 1997b).

**Definition 2.10.** A criterion  $i \in N$  is a veto for a Choquet integral  $\mathcal{C}_\mu$  (or equivalently its underlying capacity  $\mu$ ) if  $\mathcal{C}_\mu(f) \leq f_i$ , for all  $f \in \mathbb{R}_+^n$ .

It can be shown that  $i$  is a veto for  $\mu$  if and only if  $\mu(A) = 0$  whenever  $A \not\ni i$ . More generally, a coalition  $A$  of criteria is a veto if  $\mathcal{C}_\mu(f) \leq \bigwedge_{i \in A} f_i$ , for all  $f \in \mathbb{R}_+^n$ , which is equivalent to  $\mu(B) = 0$  whenever  $B \not\supseteq A$ .

The dual notion of veto is called *favor*. A coalition  $A$  of criteria  $i$  is a favor for  $\mu$  if  $\mathcal{C}_\mu(f) \geq \bigvee_{i \in A} f(i)$ , for all  $f \in \mathbb{R}_+^n$ , which is equivalent to  $\mu(B) = 1$  whenever  $A \cap B \neq \emptyset$ .

### 3 Bipolar models

We have seen in Section 2 the case of the Choquet integral on unipolar scales. These scales are not always appropriate. In particular, the DM may

not feel comfortable with the reference elements  $\mathbf{U}_i$  since they correspond to very extreme preference states (Grabisch and Labreuche, 2005b; Grabisch, 2005). Bipolar scales are alternative measurement scales, for which the extreme preference states are not explicitly considered. The Choquet integral defined in Section 2.3 aggregates nonnegative scores measuring the preferences of the DM on unipolar scales. As we will see in this section, there are many ways to extend the Choquet integral to bipolar scales, i.e., to  $\mathbb{R}^n$ .

### 3.1 Bipolar scale

A scale on  $X_i$  is said to be *bipolar* if there exists in  $X_i$  a particular element or level  $\mathbf{0}_i$ , called *neutral level*<sup>3</sup>, such that the elements of  $X_i$  preferred to  $\mathbf{0}_i$  are considered as “good”, while the elements of  $X_i$  less preferred than  $\mathbf{0}_i$  are considered as “bad” for the DM. A unipolar scale (see Section 2.1) has no neutral level. Bipolar scales depict attractiveness or desirability. The notion opposite to the attractiveness exists and corresponds to repulsiveness. The neutral level is the absence of attractiveness and repulsiveness. A bipolar scale is encoded in  $\mathbb{R}$ , where the zero value corresponds to the neutral element, the positive values to the attractive elements of  $X_i$ , and the negative values to the repulsive elements. One has  $u_i(\mathbf{0}_i) = 0$ . The satisficing level  $\mathbf{S}_i$  defined in Section 2.1, is the second reference level used on bipolar scales. For convenience, we will denote it again by  $\mathbf{1}_i$ . The existence of these absolute neutral and satisficing levels is debatable, however it has roots in psychology (Slovic et al., 2002) and in the theory of bounded rationality of Simon (1956) (see Grabisch and Labreuche (2005b) and Grabisch (2005) for more details), and it has been supported by many applications (Bana e Costa and Vansnick, 1997). We will also see an example in Section 7.1. For a detailed presentation of bipolarity, see Grabisch et al. (2008a).

### 3.2 Motivating example

The examples presented in the literature to show the flaws of the weighted sum and to motivate the Choquet integral are usually based on *conditional relative importance* among criteria. A classical example is the assessment of students with the help of three criteria: mathematics, statistics and language skills. Each course can be naturally represented on a bipolar scale where the neutral level is the aspiration level of the director of the school. Assuming an evaluation scale from  $-10$  to  $10$ , consider four students  $a, b, c, d$  with the

---

<sup>3</sup> This should not be confused with our notation of the unsatisfactory level of Section 2.

following marks:

$$\begin{aligned} u_1(a) &= 8 & , & & u_1(b) &= 8 & , & & u_1(c) &= -5 & , & & u_1(d) &= -5 \\ u_2(a) &= 6 & , & & u_2(b) &= 5 & , & & u_2(c) &= 6 & , & & u_2(d) &= 5 \\ u_3(a) &= -3 & , & & u_3(b) &= -2 & , & & u_3(c) &= -3 & , & & u_3(d) &= -2 \end{aligned}$$

The director expresses the following preferences

$$b \succ a \succ c \succ d. \quad (3.1)$$

The comparison  $b \succ a$  results from the fact that, since  $a$  and  $b$  are both good in mathematics, the director prefers the student that is better in language. The comparison  $a \succ c$  is trivial. Finally  $c \succ d$  since for two students bad in mathematics, the director prefers the student that is better in statistics.

Clearly, the weighted sum fails to represent 3.1. In order to check whether the Choquet integral succeeds, let us first define the asymmetric Choquet integral in  $\mathbb{R}^n$ :

$$\mathcal{C}_\mu(f) := \mathcal{C}_\mu(f^+) - \mathcal{C}_{\bar{\mu}}(f^-)$$

where  $f^+ := f \vee 0$  (componentwise), and  $f^- = (-f)^+$ . Observe that the formula in Definition 2.3 is still valid for the asymmetric integral with real-valued integrands. It is easy to see that any normalized capacity fulfilling  $\mu(\{1, 2\}) = \mu(\{1\}) = 0.7$  and  $\mu(\{2, 3\}) = \mu(\{2\}) = 0.3$  satisfies to Relation (3.1).

Since the justification given by the director on relations (3.1) is not restricted to the students  $a, b, c, d$ , the same preferences are obtained for other students  $a', b', c', d'$

$$\begin{aligned} u_1(a') &= 4 & , & & u_1(b') &= 4 & , & & u_1(c') &= -1 & , & & u_1(d') &= -1 \\ u_2(a') &= 6 & , & & u_2(b') &= 5 & , & & u_2(c') &= 6 & , & & u_2(d') &= 5 \\ u_3(a') &= -3 & , & & u_3(b') &= -2 & , & & u_3(c') &= -3 & , & & u_3(d') &= -2 \end{aligned}$$

Hence

$$b' \succ a' \succ c' \succ d'. \quad (3.2)$$

Strangely enough, these preferences cannot be represented by an asymmetric Choquet integral since  $b' \succ a'$  is equivalent to  $\mu(\{1, 2\}) + \mu(\{2\}) < 1$ , and  $c' \succ d'$  is equivalent to  $\mu(\{1, 2\}) + \mu(\{2\}) > 1$ . We conclude that, by changing a little bit the above example presented as a motivating example to the Choquet integral, a limitation of this model is obtained.

### 3.3 Notion of bi-capacity and bipolar Choquet integral

The previous example shows a situation where the decision strategies of the DM are conditional on some criteria being good or bad. The decision behavior of the DM is thus bipolar. Due to equation (2.2), the capacity focuses only on the positive part of the bipolar scale in the asymmetric Choquet integral. Hence the latter cannot represent bipolar decision strategies. The limitation of this model thus comes from the notion of capacity. The idea is to define a concept that gathers all combinations of positive and negative values on the criteria. Instead of focusing the attention on all binary acts  $(1_A, 0_{-A})$ , for all  $A \subseteq N$ , it seems more appropriate to look at all ternary acts  $(1_A, -1_B, 0_{-(A \cup B)})$ , for all disjoint subsets  $A, B$  of  $N$ .

Let

$$\mathcal{Q}(N) := \{(A, B) \in 2^N \times 2^N \mid A \cap B = \emptyset\}.$$

Many independent works have led to the definition of functions on  $\mathcal{Q}(N)$ . A *ternary voting game* is a function  $v : \mathcal{Q}(N) \rightarrow \{-1, 1\}$ , and is used to model abstention in voting games (Felsenthal and Machover, 1997). *Bi-cooperative games* defined as functions  $v : \mathcal{Q}(N) \rightarrow \mathbb{R}$  satisfying  $v(\emptyset, \emptyset) = 0$  are extension of ternary voting games (Bilbao et al., 2000). The generalization has recently been independently rediscovered in the context of MCDA:

**Definition 3.1 (Labreuche and Grabisch (2006a)).** *A bi-capacity is a function  $v : \mathcal{Q}(N) \rightarrow \mathbb{R}$  satisfying  $v(\emptyset, \emptyset) = 0$ ,  $v(A, B) \leq v(A', B)$  whenever  $A \subseteq A'$  (monotonicity w.r.t. the first argument), and  $v(A, B) \geq v(A, B')$  whenever  $B \subseteq B'$  (monotonicity w.r.t. the second argument). A bi-capacity  $v$  is said to be normalized if  $v(N, \emptyset) = 1$ ,  $v(\emptyset, N) = -1$ .*

Note that the concept of *bipolar capacity*, which is close to that of bi-capacity, has been independently introduced (Greco et al., 2002). Bipolar capacities and bi-capacities turn out to be similar in the context of MCDA (see Grabisch and Labreuche (2005c)).

The dual of a bi-capacity  $v$  is a bi-capacity  $\bar{v}$  defined by

$$\bar{v}(A, B) := -v(B, A) \quad \forall (A, B) \in \mathcal{Q}(N).$$

Let  $\Sigma_A := \{f \in \mathbb{R}^n, f_A \geq 0, f_{-A} < 0\}$ .

**Definition 3.2 (Labreuche and Grabisch (2006a)).** *For any  $A \subseteq N$ ,  $f \in \Sigma_A$ , the bipolar Choquet integral of  $f$  w.r.t. a bi-capacity  $v$  is given by*

$$\mathcal{BC}_v(f) := \mathcal{C}_\mu(|f|)$$

where  $\mu(C) := v(C \cap A, C \setminus A)$ .

Note that  $\mu$  is not in general a capacity but a game, since it is not necessarily monotonic. The fundamental equation (2.2) is generalized as follows:

$$\mathcal{BC}_v(1_A, -1_B, 0_{-(A \cup B)}) = v(A, B), \quad \forall (A, B) \in \mathcal{Q}(N). \quad (3.3)$$

### 3.4 Representation of the motivating example

Relation (3.2) is equivalent to the following two inequalities (Labreuche and Grabisch, 2006a)

$$\begin{aligned} v(\{1, 2\}, \emptyset) - v(\{1, 2\}, \{3\}) &> v(\{2\}, \emptyset), \\ v(\{2\}, \{3\}) &> 0. \end{aligned}$$

Since there is no contradiction between these two inequalities, the preferential information given in Relation (3.2) can be represented by a Choquet integral w.r.t. a bi-capacity. An essential question is whether there exists again a slight modification of the same decision strategies such that the resulting preferences  $b'' \succ a'' \succ c'' \succ d''$ , for some alternatives  $a'', b'', c'', d''$ , cannot be represented by the bi-capacity model. To see this, one can generalize the idea behind preference orderings (3.1) and (3.2) in terms of two general rules

**(R1)**: For a student good at mathematics (criterion  $i$ ), language (criterion  $j^+$ ) is more important than statistics (criterion  $j^-$ ).

**(R2)**: For a student bad in mathematics (criterion  $i$ ), statistics (criterion  $j^-$ ) is more important than language (criterion  $j^+$ ).

In these preferences, the relative importance of language compare to statistics is conditional on the student being good or bad in mathematics. Throughout this subsection, we will denote by  $i, j^+, j^-$  the indices in rules **(R1)** and **(R2)**.

It has been shown in Labreuche and Grabisch (2007) that the Choquet integral w.r.t. a bi-capacity fails to represent **(R1)** and **(R2)** in general. Bi-capacities have been introduced to represent complex preferences that cannot be modeled with a capacity. Likewise, one may seek for a more general model than bi-capacity able to represent rules **(R1)** and **(R2)**. Such a model, that can be described by an aggregation function  $F : \mathbb{R}^n \rightarrow \mathbb{R}$ , must be continuous and non-decreasing. It should also be piecewise linear as a natural generalization of the Choquet integral. As a matter of fact, there does not exist any aggregation function satisfying both the previous three conditions and rules **(R1)** and **(R2)** (Labreuche and Grabisch, 2007). More precisely, the following result holds.

**Proposition 3.3 (Labreuche and Grabisch (2007))**. *Assume that  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  is continuous, non-decreasing and piecewise linear. Let  $\Phi^+ \subseteq \{f \in \mathbb{R}^n, f_i \geq 0\}$ ,  $\Phi^- \subseteq \{f \in \mathbb{R}^n, f_i \leq 0\}$  such that  $F$  is linear in  $\Phi^+$  and in  $\Phi^-$ . If there exists a nonempty open set  $B \subseteq \mathbb{R}^2$  and  $f_{-\{i, j^+, j^-\}} \in \mathbb{R}^{n-3}$  such that*

$$\Phi^+ \cap \Phi^- \supseteq \{(0_i, g_{j^+}, g_{j^-}, f_{-\{i, j^+, j^-\}}), \forall (g_{j^+}, g_{j^-}) \in B\}$$

then rules **(R1)** and **(R2)** cannot be represented by  $F$  in the two domains  $\Phi^+$  and in  $\Phi^-$  (i.e., criterion  $j^+$  is more important than criterion  $j^-$  in  $\Phi^+$ , and criterion  $j^+$  is less important than criterion  $j^-$  in  $\Phi^-$ ).

This proposition proves that if, for two neighbor domains  $\Phi^+$  and  $\Phi^-$  such that the value of criterion  $i$  can be arbitrarily small independently of criteria  $j^+$  and  $j^-$  in both  $\Phi^+$  and  $\Phi^-$ , then rules **(R1)** and **(R2)** cannot be satisfied in both  $\Phi^+$  and  $\Phi^-$ . This explains why bi-capacities cannot represent Relation (3.2). In short, rules **(R1)** and **(R2)** cannot be satisfied if criterion  $i$  is the one closest to the neutral level among criteria  $i, j^+, j^-$ .

One cannot gain a lot by extending bi-capacities to more complex models. Actually, bi-capacities enable to represent the following rules.

**(R1')**: If the value w.r.t. criterion  $i$  is attractive ( $> 0$ ), and  $i$  is not the closest to the neutral level among criteria  $i, j^+, j^-$ , then criterion  $j^+$  is more important than criterion  $j^-$ .

**(R2')**: If the value w.r.t. criterion  $i$  is repulsive ( $< 0$ ), and  $i$  is not the closest to the neutral level among criteria  $i, j^+, j^-$ , then criterion  $j^+$  is less important than criterion  $j^-$ .

One can interpret this restriction in the following way. When criterion  $i$  has the closest value to the neutral level among criteria  $i, j^+$  and  $j^-$ , the distinction between  $i$  being attractive or repulsive is not so meaningful to the DM and shall be removed from rules **(R1)** and **(R2)**.

### 3.5 Particular models

The number of terms in a bi-capacity, which is  $|\mathcal{Q}(N)| = 3^n$ , is much larger than that for a capacity, which is  $2^n$ . So, the submodels of the bi-capacities containing much less terms than bi-capacities are of particular interest.

First of all, if a bi-capacity  $v$  satisfies  $v(A, B) = v(N \setminus B, N \setminus A)$  for all  $(A, B) \in \mathcal{Q}(N)$ , then  $\mathcal{BC}_v$  is the asymmetric Choquet integral of the capacity  $\mu$  given by  $\mu(C) = v(C, \emptyset)$  for all  $C \subseteq N$  (Labreuche and Grabisch, 2006a).

The *Šipoš integral* (Šipoš, 1979) – also called *symmetric Choquet integral* – is defined by

$$\check{C}_\mu(f) := \mathcal{C}_\mu(f^+) - \mathcal{C}_\mu(f^-),$$

for all  $f \in \mathbb{R}^n$ . If a bi-capacity  $v$  satisfies  $v(A, B) = -v(B, A)$  for all  $(A, B) \in \mathcal{Q}(N)$ , then  $\mathcal{BC}_v$  is the Šipoš integral of the capacity  $\mu$  given by  $\mu(C) = v(C, \emptyset)$  for all  $C \subseteq N$ .

Tversky and Kahneman have proposed a model that encompasses both the symmetric and asymmetric Choquet integrals, known as the *Cumulative Prospect Theory (CPT)* model (Tversky and Kahneman, 1992):

$$CPT(f) := \mathcal{C}_{\mu_1}(f^+) - \mathcal{C}_{\mu_2}(f^-),$$

where  $\mu_1$  and  $\mu_2$  are two capacities. If a bi-capacity  $v$  satisfies  $v(A, B) - v(A, B') = v(A', B) - v(A', B')$  for any  $(A, B), (A, B'), (A', B), (A', B') \in \mathcal{Q}(N)$ , then  $\mathcal{BC}_v$  becomes the CPT model with the two capacities  $\mu_1$  and  $\mu_2$  given by  $\mu_1(C) = v(C, \emptyset)$  and  $\mu_2(C) = -v(\emptyset, C)$  for all  $C \subseteq N$ .

The CPT model fails to represent the preferences given in Section 3.2 and all decision strategies that are conditional on some criteria being good or bad (i.e., that depend on the sign of some criteria). It is thus necessary to define other submodels of bi-capacities. It can be noticed that in most MCDA problems with sign-dependent decision strategies, the bipolar nature is not generally compulsory on all criteria. Let us denote by  $P \subseteq N$  the set of criteria for which the DM's behavior is clearly of bipolar nature. In the example given in Section 3.2,  $P$  is reduced to criterion Mathematics. The approach proposed in Labreuche and Grabisch (2004) is to allow more degrees of freedom on the criteria  $P$  compared to the remaining criteria  $N \setminus P$  that do not need bipolarity. This is done by enforcing some symmetry conditions on the criteria  $N \setminus P$ , which state that the interaction between positive and negative values vanishes for these criteria. More precisely, the interaction between two criteria  $i$  and  $j$ , where  $i$  is attractive and  $j$  is repulsive, in the presence of bi-coalition  $(A, B) \in \mathcal{Q}(N \setminus \{i, j\})$  is defined by:

$$\delta_{\{i, \{j\}\}}^{A, B}(v) = v(A \cup \{i\}, B \cup \{j\}) - v(A \cup \{i\}, B) - v(A, B \cup \{i\}) + v(A, B)$$

**Definition 3.4.** A bi-capacity  $v$  is said to be symmetric outside  $P$  (called  $P$ -nonsymmetric) if  $\delta_{\{i, \{j\}\}}^{A, B}(v) = 0$  for all  $(A, B) \in \mathcal{Q}(N \setminus \{i, j\})$  and all  $\{i, j\} \subseteq N \setminus P$ .

This is derived from a property satisfied by the asymmetric Choquet integral and the CPT model. Set

$$\begin{aligned} \mathcal{Q}^P(N) := \{ & (A \cup A', B \cup B'), (A, B) \in \mathcal{Q}(P), (A', B') \in \mathcal{Q}(N \setminus P) \\ & \text{with } A' = \emptyset \text{ or } B' = \emptyset \}. \end{aligned}$$

Let  $v_P$  be the restriction of  $v$  on  $\mathcal{Q}^P(N)$ .  $v_P$  contains  $3^{n-p} \times (2^{p+1} - 1)$  terms and is thus bipolar on  $P$  and unipolar on  $N \setminus P$ . The next lemma shows that  $v$  is determined only from the knowledge of  $v_P$ .

**Lemma 3.5.** (Labreuche and Grabisch, 2004) If  $v$  is  $P$ -nonsymmetric, then for all  $(A, B) \in \mathcal{Q}(P)$  and all  $(A', B') \in \mathcal{Q}(N \setminus P)$

$$v(A \cup A', B \cup B') = v_P(A \cup A', B) + v_P(A, B \cup B') - v_P(A, B).$$

Usual bi-capacities are recovered when  $P = N$ . Moreover,  $\emptyset$ -nonsymmetric bi-capacities correspond to the CPT model.

### 3.6 Construction of the value functions based on the MACBETH approach

*3.6.1 Case of the Šipoš integral* This section is based on Grabisch et al. (2003). The construction of the value function for the Šipoš integral is quite similar to that presented in Section 2.4. The only difference is the use of the two reference levels  $\mathbf{0}_i$  and  $\mathbf{1}_i$ . To construct the value function on  $X_i$ , the following set is introduced:

$$X_i] := \{(\mathbf{0}_1, \dots, \mathbf{0}_{i-1}, x_i, \mathbf{0}_{i+1}, \dots, \mathbf{0}_n \mid x_i \in X_i\}.$$

As in Section 2.4, a unique scale  $v_i$  is constructed on  $X_i]$  by enforcing  $v_i(\mathbf{0}) = 0$  and  $v_i(\mathbf{1}_i, \mathbf{0}_{-i}) = 1$ . It remains to define the value function on attribute  $i$  as follows:

$$\forall x_i \in X_i, \quad u_i(x_i) := v_i(x_i, \mathbf{0}_{-i}). \quad (3.4)$$

We need to justify the previous construction and in particular Relation (3.4). The two interval scales  $v_i$  and  $\check{C} \circ u$ , where  $u = (u_1, \dots, u_n)$ , on  $X_i]$  are equivalent. Hence there exists  $\alpha > 0$  and  $\beta$  such that for all  $x_i \in X_i$ ,

$$\check{C}_\mu(u_1(\mathbf{0}_1), \dots, u_{i-1}(\mathbf{0}_{i-1}), u_i(x_i), u_{i+1}(\mathbf{0}_{i+1}), \dots, u_n(\mathbf{0}_n)) = \alpha v_i(x_i, \mathbf{0}_{-i}) + \beta.$$

Hence  $\check{C}_\mu(u_i(x_i), \mathbf{0}_{-i}) = \alpha u_i(x_i) + \beta$  for all  $x_i \in X_i$ . Since the value functions take positive and negative values on bipolar scales, we obtain

$$\forall a \in \mathbb{R}, \quad \check{C}_\mu(a_i, \mathbf{0}_{-i}) = \alpha a + \beta. \quad (3.5)$$

This relation holds for the Šipoš integral with  $\alpha = \mu(\{i\})$  and  $\beta = 0$  so that the previous construction (3.4) is valid.

*3.6.2 Case of the general bipolar Choquet integral* The construction made for the Šipoš integral cannot be used for the bipolar Choquet integral since Relation (3.5) is not true for  $\mathcal{BC}_v$ . The reason is that the coefficients  $\alpha$  and  $\beta$  for  $\mathcal{BC}_v(a_i, \mathbf{0}_{-i})$  depend on the sign of  $a$ .

This leads to constructing the positive and negative part of  $u_i$  separately (Labreuche and Grabisch, 2006a). Let us denote by  $X_i^+$  the elements of  $X_i$  that are more attractive than  $\mathbf{0}_i$ , and by  $X_i^-$  the elements of  $X_i$  that are more repulsive than  $\mathbf{0}_i$ . As a consequence, we introduce the two subsets  $X_i^+]$  and  $X_i^-]$  of  $X$  defined by

$$X_i^+] := \{(x_i, \mathbf{0}_{-i}), x_i \in X_i^+\}, \quad X_i^-] := \{(x_i, \mathbf{0}_{-i}), x_i \in X_i^-\}.$$

A unique scale  $v_i^+$  is constructed on  $X_i^+]$  by enforcing  $v_i^+(\mathbf{0}) = 0$  and  $v_i^+(\mathbf{1}_i, \mathbf{0}_{-i}) = 1$ . We obtain  $u_i^+(x_i) := v_i^+(x_i, \mathbf{0}_{-i})$  for all  $x_i \in X_i^+$ .

Concerning the negative part of the scale, we need to assume the existence in  $X_i^-]$  of an element denoted by  $-\mathbf{1}_i$  that is considered as bad and

unsatisfactory. It is symmetric to the level  $\mathbf{1}_i$  and corresponds to a reference level. A unique scale  $v_i^-$  is constructed on  $X_i^-$  by enforcing  $v_i^-(\mathbf{0}) = 0$  and  $v_i^-(-\mathbf{1}_i, \mathbf{0}_{-i}) = -1$ . We obtain  $u_i(x_i) := v_i^-(x_i, \mathbf{0}_{-i})$  for all  $x_i \in X_i^-$ .

Hence the value function  $u_i$  is thoroughly determined on  $X_i$  and the positive and negative parts are identical at  $\mathbf{0}_i$  since  $v_i^-(\mathbf{0}) = v_i^+(\mathbf{0})$ .

### 3.7 The bipolar Choquet integral as a parsimonious linear interpolator

As for the classical Choquet integral, the bipolar Choquet integral is also a parsimonious linear interpolator on the hypercube  $[-1, 1]^n$ , between all points of the form  $(1_A, -1_B, 0_{-(A \cup B)})$ , for all  $(A, B) \in \mathcal{Q}(N)$  (Grabisch, 2004b). Indeed, the bipolar Choquet integral w.r.t. a bi-capacity  $v$  is a symmetrization of the classical Choquet integral on  $[0, 1]^n$  over  $[-1, 1]^n$ . More precisely, any  $f \in [-1, 1]^n$  is mapped to  $[0, 1]^n$  taking its absolute value  $|f|$ , and a game  $\mu$  is defined from  $v$  and  $f$  by  $\mu(C) := v(C \cap A, C \setminus A)$ , where  $A := \{i \in N \mid f_i \geq 0\}$ .

### 4 The ordinal case: the Sugeno integral

In many applications, scores on criteria are expressed on a finite ordinal scale or a qualitative scale. Most of the time, this ordinal information is turned in an arbitrary way into cardinal information, or is treated as such (e.g., A, B, C, D, E are coded by numbers 1, 2, 3, 4, 5, and then these numbers are treated as real numbers in the calculations). From a measurement theoretical point of view (see, e.g., Roberts (1979)), numbers on an ordinal scale cannot be handled by standard arithmetic operators, like sum, product, difference, etc. Only comparisons can be done, which considerably reduces our possibility of building models. Moreover, the finiteness of the scale adds further technical intricacies.

There are basically two attitudes for solving this problem.

- The first one consists in trying to turn the ordinal problem into a cardinal one. The naive (and most often used) way of converting levels of an ordinal scale into numbers illustrates this approach. A sounder way to proceed is to get cardinal information from the ordinal one, for example by counting the number of times an alternative is better or worse than the other ones on a given criterion. This number can then be used as a (cardinal) score on this criterion, and the Choquet integral can be used with these new scores. This is basically what was proposed by Roubens (2001), and in the TOMASO method (Meyer and Roubens, 2005; Marichal et al., 2005). Note however that the scores obtained in this way have a very different meaning from the original ones, in particular, they are *relative* to the set of alternatives.

- The second tries to directly deal with the ordinal scores and the poor algebra underlying them, with limitations due to finiteness.

In this survey paper, we develop the latter attitude, so as to highlight the limits of this approach.

#### 4.1 Why the Sugeno integral

A fundamental result obtained by Marichal (to appear) shows that if one restricts to the use of max and min as operators, then the Sugeno integral is the only solution for aggregating scores. Specifically, let us call *lattice weighted polynomial*  $P(f_1, \dots, f_n)$  any expression formed with  $\vee, \wedge$ , parentheses, constants, and variables  $f_1, \dots, f_n$  defined on some lattice  $L$ . Obviously, a Sugeno integral  $\mathcal{S}_\mu(f)$  on  $N$ , where  $f$  is the vector of scores expressed on some ordinal scale, is an example of weighted lattice polynomial (see Definition 2.4). What is remarkable is that conversely, any lattice polynomial satisfying  $P(0, \dots, 0) = 0$  and  $P(1, \dots, 1) = 1$  is necessarily a Sugeno integral.

We refer the reader to survey papers (Dubois et al., 2001a; Murofushi and Sugeno, 2000) and to (Marichal, 2001, 2000) for properties of the Sugeno integral, especially in a decision making perspective. We mention that in the context of decision under uncertainty, an axiomatic construction similar to the one of Savage has been done by Dubois et al. (2000b, 2001b).

#### 4.2 The symmetric Sugeno integral

We have defined for the Choquet integral its symmetric and asymmetric versions for real-valued integrands, i.e., which may take negative values. The same should be done for the Sugeno integral, but a first question is: what is a negative number on an ordinal scale? This amounts to first define a zero level, then to perform some order-reversing symmetry around this zero level. The second question is: how to extend min and max operators on this symmetrized scale so as to keep good properties? Curiously, this second question happens to be much more difficult than one might think (see details in Grabisch (2004a, 2003a)).

Let us call  $L^+$  some ordinal scale, with least element denoted by 0, and define  $L := L^+ \cup L^-$ , where  $L^-$  is a reversed copy of  $L^+$ , i.e. for any  $a, b \in L^+$ , we have  $a \leq b$  iff  $-b \leq -a$ , where  $-a, -b$  are the copies of  $a, b$  in  $L^-$ . Moreover, the two elements 0,  $-0$  are merged into a single one denoted by 0. Hence, the zero level of  $L$  is 0, and levels above or below it are the positive or negative values of the scale.

We want to endow  $L$  with operations  $\oplus, \otimes$  satisfying (among possible other conditions):

- (C1)  $\oplus, \otimes$  coincide with  $\vee, \wedge$  respectively on  $L^+$   
(C2)  $-a$  is the symmetric of  $a$ , i.e.  $a \otimes (-a) = 0$ .

Hence we may extend to  $L$  what exists on  $L^+$  (e.g. the Sugeno integral), and a difference operation could be defined. The problem is that conditions (C1) and (C2) imply that  $\oplus$  would be non-associative in general. Indeed, take  $0 < a < b$  and consider the expression  $(-b) \oplus b \oplus a$ . Depending on the place of parentheses, the result differs since  $((-b) \oplus b) \oplus a = 0 \oplus a = a$ , but  $(-b) \oplus (b \oplus a) = (-b) \oplus b = 0$ .

It can be shown that the best solution (i.e., associative on the largest domain) for  $\oplus$  is given by:

$$a \oplus b := \begin{cases} -(|a| \vee |b|) & \text{if } b \neq -a \text{ and } |a| \vee |b| = -a \text{ or } = -b \\ 0 & \text{if } b = -a \\ |a| \vee |b| & \text{otherwise.} \end{cases} \quad (4.1)$$

Except for the case  $b = -a$ ,  $a \oplus b$  equals the absolutely larger one of the two elements  $a$  and  $b$ .

The extension of  $\wedge$ , viewed as the counterpart of multiplication, is simply done on the principle that the rule of sign should hold:  $-(a \otimes b) = (-a) \otimes b$ ,  $\forall a, b \in L$ . It leads to an associative operator, defined by:

$$a \otimes b := \begin{cases} -(|a| \wedge |b|) & \text{if } \text{sign } a \neq \text{sign } b \\ |a| \wedge |b| & \text{otherwise.} \end{cases} \quad (4.2)$$

Mimicking the definition of the symmetric Choquet integral, the symmetric Sugeno integral is defined as follows:

$$\tilde{\mathcal{S}}_\mu(f) := \mathcal{S}_\mu(f^+) \otimes (-\mathcal{S}_\mu(f^-)),$$

for any  $L$ -valued  $f$ , with same notation as in Section 2.3.

There is no suitable definition of asymmetric integral, since the conjugate of a capacity is difficult to define in a proper way on an ordinal scale.

Lastly, we mention Denneberg and Grabisch (2004), who have proposed a general formulation of the Sugeno integral on arbitrary symmetric scales.

### 4.3 Construction of the model

As for the Choquet integral, we need a procedure to build our model, hence the value functions and the capacity.

It is possible to follow an approach similar to the one presented in Section 2, by the use of neutral and satisfactory levels and particular sets of alternatives  $X_i$ ,  $i \in N$ , and  $X \uparrow_{\{0,1\}}$  (see a full description in Grabisch and Labreuche (2005b)). However, a special difficulty arises here since for any alternative  $(x_i, \mathbf{0}_{-i})$  in  $X_i$ , its evaluation by the Sugeno integral is:

$$\mathcal{S}_\mu(u(x_i, \mathbf{0}_{-i})) = u_i(x_i) \wedge \mu(\{i\}).$$

Then, if  $u_i(x_i) \geq \mu(\{i\})$ , the value of  $u_i(x_i)$  cannot be observed since it is “hidden” by  $\mu(\{i\})$ , which acts like a threshold (compare the situation with the Choquet integral, where  $\mu(\{i\})$  is simply a multiplicative factor).

Greco et al. (2004) have proposed another approach based on decision rules, which does not use neutral nor satisfactory levels, because commensurate scales are no more necessary in their framework. Moreover, this result gives a characterization of the Sugeno integral solely based on the preference relation. This result having been stated without proof, Bouyssou et al. have proposed a proof of this deep result (Bouyssou et al., 2006). Specifically, the preference relation  $\succ$  is said to be *strongly 2-graded* if for all  $x, y, z, w \in X$ , all  $a_i \in X_i$ , and all  $i \in N$

$$\left. \begin{array}{l} x \succ z \\ \text{and} \\ y \succ w \\ \text{and} \\ z \succ w \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} (a_i, x_{-i}) \succ z \\ \text{or} \\ (x_i, y_{-i}) \succ w. \end{array} \right.$$

The characterization is the following.

**Proposition 4.1.** *Let  $\succeq$  be a binary relation on  $X$ . This relation can be represented by the Sugeno integral (see equation (2.1)) if and only if  $\succeq$  is transitive and complete, it satisfies the order-denseness condition (i.e., it exists a countable subset  $Y \subseteq X$  that is dense in  $X$  for  $\succeq$ ), and it is strongly 2-graded.*

#### 4.4 Identification of capacities

In situations where value functions are known, the problem of the identification of capacities when the model is a Sugeno integral in an ordinal context, or even when  $L = [0, 1]$  or  $[-1, 1]$ , appears to be rather different from the case of the Choquet integral. The main reason is that we are not able to write the identification problem as a minimization problem *stricto sensu*, since the notion of difference between values, hence of error, is not defined in a way which is suitable on an ordinal scale, to say nothing about “squared errors” or “average values”.

Even if we take  $L$  as a real interval, which permits to define a squared error criterion as for the Choquet integral, the minimization problem obtained is not easy to solve, since it involves non-linear, non-differentiable operations  $\vee, \wedge, \otimes, \oplus$ . In such cases, only meta-heuristic methods can be used, as genetic algorithms, simulated annealing, etc. There exist some works in this direction, although most of the time used for the Choquet integral, which is questionable (Wang et al., 1999; Grabisch, 2003b).

An alternative to this option is to find the set of capacities (possibly empty) which enable the representation of the preference of the decision maker over a set of alternatives of interest by the Sugeno integral. A detailed study of this problem has been done by Rico et al. (2005) for the Sugeno integral. We also mention the work of Słowiński et al. (2002) based on decision rules.

#### 4.5 Drawbacks, and how to get rid of them

Making decision with the Sugeno integral has some drawbacks, which are clearly put into light by the following result (Murofushi, 2001). Let us consider w.l.o.g.  $L := [0, 1]$ , and  $\succeq$  be a weak order (complete, reflexive, transitive) on  $[0, 1]^n$ , and for  $a, b \in [0, 1]^n$ , denote  $a \geq b$  if  $a_i \geq b_i$  for all  $i \in N$ , and  $a > b$  if  $a \geq b$  and  $a_i > b_i$  for some  $i \in N$ , and  $a \gg b$  if  $a_i > b_i$  for all  $i \in N$ . We say that  $\succeq$  satisfies *monotonicity* if  $a \geq b$  implies  $a \succeq b$ , the *strong Pareto condition* if  $a > b$  implies  $a \succ b$ , and the *weak Pareto condition* if  $a \gg b$  implies  $a \succ b$ . Then the following holds.

**Proposition 4.2.** *Let  $\mu$  be a capacity on  $N$ , and  $\succeq_\mu$  the weak order induced on  $[0, 1]^n$  by the Sugeno integral  $\mathcal{S}_\mu$  (i.e.,  $a \succeq_\mu b$  if  $\mathcal{S}_\mu(a) \geq \mathcal{S}_\mu(b)$ ).*

- (i)  $\succeq_\mu$  always satisfies monotonicity.
- (ii)  $\succeq_\mu$  satisfies the weak Pareto condition if and only if  $\mu$  is 0-1 valued.
- (iii)  $\succeq_\mu$  never satisfies the strong Pareto condition.

Note that the Choquet integral always satisfies the weak Pareto condition, and the strong one if and only if  $\mu$  is strictly monotone.

The main reason of these poor properties is that the Sugeno integral may remain constant over large domains, so it is not discriminative for decision since many alternatives will receive the same overall score. A natural solution to make it more discriminative is to use lexicographic approaches. For example, the lexicographic approach applied on the min or max leads to the well known leximin and leximax (Moulin, 1988), which are far more discriminative than min and max:

$$x \prec_{\text{leximin}} y \Leftrightarrow (x_{\sigma(1)}, \dots, x_{\sigma(n)}) \leq^l (y_{\sigma(1)}, \dots, y_{\sigma(n)}),$$

where  $\leq^l$  denotes the lexicographic order, and similarly for the leximax  $\prec_{\text{leximax}}$ . Dubois and Fargier (2005), and Fargier and Sabbadin (2005), have proposed a clever way of defining the lexicographic Sugeno integral (see also Grabisch (2006) and Murofushi (2001) for other approaches and a survey). Basically, considering  $m \times n$  matrices  $u, v$ , we can compare them by the complete preorder  $\prec_{\text{maxmin}}$  defined by

$$u \preceq_{\text{maxmin}} v \Leftrightarrow \bigvee_{i=1}^n \bigwedge_{j=1}^m u_{ij} \leq \bigvee_{i=1}^n \bigwedge_{j=1}^m v_{ij}.$$

Observe that the Sugeno integral has exactly this max-min form, with  $m = 2$ . Now, the order  $\preceq_{\text{lmax}(\prec_{\text{lmin}})}$ , obtained by ordering the columns of the matrix by  $\prec_{\text{lmin}}$ , and then by using the  $\preceq_{\text{lmax}}$  order on them, is a refinement of  $\prec_{\text{maxmin}}$ , and thus defines a lexicographic Sugeno integral, much more discriminating than the original Sugeno integral.

If the ordinal scale is finite, by using a suitable encoding of the scale on the real line, it is shown in Dubois and Fargier (2005) that the  $\preceq_{\text{lmax}(\prec_{\text{lmin}})}$  ordering amounts to compare alternatives by a Choquet integral. This result generalizes the one of Moulin, saying that under the same conditions, the leximin and leximax are equivalent to a sum, and shows that, even on the ordinal case, we are eventually back to the Choquet integral.

## 5 Intrinsic analysis of models

The aim of this section is to provide tools for analysing the obtained model, that is, for interpreting the capacity in terms of importance of criteria, interaction, and typical decision behaviours. We restrict here to the Choquet integral, the main reason being that similar attempts done for the Sugeno integral did not provide, up to this time, sufficiently convincing results.

### 5.1 Importance and interaction indices

*5.1.1 Case of capacities* The complexity of the notion of capacity comes from the fact that it is defined by  $2^n$  values. The behaviour of the DM does not appear clearly when looking at the values taken by the capacity.

One is interested in particular in knowing what is the importance of a given criterion in the decision. We may say that a criterion  $i$  is important if whenever added to some coalition  $A$  of criteria, the score of  $(\mathbf{1}_{A \cup \{i\}}, \mathbf{0}_{-A \cup \{i\}})$  is significantly larger than the score of  $(\mathbf{1}_A, \mathbf{0}_{-A})$ . Hence, an importance index should compute an average value of the quantity  $\delta_i^A(\mu) := \mu(A \cup \{i\}) - \mu(A)$  for all  $A \subseteq N \setminus \{i\}$ . Another requirement is that the sum of importance indices should be a constant, say 1. Lastly, the importance index should not depend on the numbering of the criteria. These three requirements plus a linearity assumption, which states that the importance index should be a weighted arithmetic mean of the  $\delta_i^A(\mu)$  coefficients, uniquely determines the importance index, known as the *Shapley importance index* (Shapley, 1953)

$$\phi_i(\mu) = \sum_{A \subseteq N \setminus \{i\}} \frac{|A|!(n - |A| - 1)!}{n!} (\mu(A \cup \{i\}) - \mu(A)).$$

This value turns out to be exactly equal to the average weight of criterion  $i$  in the Choquet integral over all possible profiles in  $[0, 1]^n$  (Marichal, 1998;

Kojadinovic, 2007b). Note that the Shapley value is a particular case of the interaction index (2.3):  $\phi_i(\mu) = I^\mu(\{i\})$ .

One would also like to quantify the way two criteria  $i$  and  $j$  interact together. Recall that  $\mu(A)$  is the overall score of an option that is perfectly satisfactory (with score 1) on criteria  $A$  and completely unacceptable (with score 0) on the remaining criteria. Let  $A \subseteq N \setminus \{i, j\}$ . Consider an option that is very good on criteria  $A$  and unacceptable on criteria  $N \setminus (A \cup \{i, j\})$ . One wonders whether it is really beneficial for this option to be good at both criteria  $i$  and  $j$ . Values  $\delta_i^A(\mu) = \mu(A \cup \{i\}) - \mu(A)$ ,  $\delta_j^A(\mu) = \mu(A \cup \{j\}) - \mu(A)$  and  $\delta_{ij}^A(\mu) = \mu(A \cup \{i, j\}) - \mu(A)$  correspond to the added value for this option to be good at criterion  $i$  but not at criterion  $j$ , to be good at criterion  $j$  but not at criterion  $i$  and to be good at both criteria  $i$  and  $j$  respectively. When  $\delta_{ij}^A(\mu) > \delta_i^A(\mu) + \delta_j^A(\mu)$ , improving both  $i$  and  $j$  gives strictly more than improving  $i$  and  $j$  separately. In this case, criteria  $i$  and  $j$  deserve to be well-satisfied together. We say in this case that there is *complementarity* among criteria  $i$  and  $j$ . When  $\delta_{ij}^A(\mu) < \delta_i^A(\mu) + \delta_j^A(\mu)$ , it is not interesting to improve both criteria  $i$  and  $j$  together. We say in this case that there is *substitutability* among criteria  $i$  and  $j$ . We observe that  $\delta_{\{i,j\}}^A(\mu) := \delta_{ij}^A(\mu) - \delta_i^A(\mu) - \delta_j^A(\mu) = \mu(A \cup \{i, j\}) - \mu(A \cup \{i\}) - \mu(A \cup \{j\}) + \mu(A)$ . The interaction index is a weighted arithmetic mean of these differences over all  $A \subseteq N \setminus \{i, j\}$  (Murofushi and Soneda, 1993)

$$I_{ij}(\mu) = \sum_{A \subseteq N \setminus \{i, j\}} \frac{|A|!(n - |A| - 2)!}{(n - 1)!} \delta_{\{i, j\}}^A(\mu).$$

A positive interaction describes complementarity among criteria, and a negative interaction depicts substitutability among criteria (Grabisch, 1996, 1997b). The interaction index  $I_{ij}(\mu)$  can also be interpreted as the variation of the mean weight of criterion  $i$  in the Choquet integral when criterion  $j$  switches from the least satisfied criterion to the best satisfied criterion (Kojadinovic, 2007b). There are two axiomatizations of the interaction index (Grabisch and Roubens, 1999; Fujimoto et al., 2006). Note that the interaction index  $I_{ij}(\mu)$  is a particular case of the interaction index (2.3):  $I_{ij}(\mu) = I^\mu(\{i, j\})$ .

**5.1.2 Case of bi-capacities** As for capacities, due to the complexity of the bi-capacity model, involving  $3^n$  coefficients, it is important in practice to be able to analyze a bi-capacity in terms of decision behaviours, namely importance of criteria and interaction among them.

Let us first define the importance  $\phi_i(v)$  for a bi-capacity  $v$ . Unlike capacities where we have previously seen that we come up quite easily to a unique definition, many definitions seem suitable for bi-capacities (see Felsenthal

and Machover (1997); Grabisch and Labreuche (2005a); Bilbao et al. (to appear); Labreuche and Grabisch (2006b); Kojadinovic (2007b)). Note that the last two proposals are identical. Cooperative Game Theory is a good approach to select the most appropriate definition. In this setting, the bi-capacity  $v$  is interpreted as a bi-cooperative game. More precisely, in the context of cost sharing problems,  $v(A, B)$  is the stand alone price of serving agents in  $A \cup B$  when  $A$  have decided to contribute positively to the game and  $B$  have decided to contribute negatively to the game (Labreuche and Grabisch, submitted). Unlike usual games, where at the end all players join the grand coalition, it is not assumed here that all players have decided to be positive contributors. We denote by  $S$  the set of players that have decided to be positive contributors, and by  $T$  the set of players that have decided to be negative contributors. The remaining players  $N \setminus (S \cup T)$  have chosen not to participate to the game. As a result, the share of the total cost among the players depends on the bi-coalition  $(S, T)$ . We denote by  $\varphi_i^{S,T}(v)$  the payoff allotted to player  $i$ . This share is uniquely obtained by extending the requirements characterizing the Shapley value, and by adding a monotonicity requirement (Labreuche and Grabisch, submitted)

$$\begin{aligned} \varphi_i^{S,T}(v) &= \sum_{K \subseteq (S \cup T) \setminus \{i\}} \frac{k!(s+t-k-1)!}{(s+t)!} \\ &\quad \times [v(S \cap (K \cup \{i\}), T \cap (K \cup \{i\})) - v(S \cap K, T \cap K)]. \end{aligned}$$

From this expression, the payoff for positive contributors (i.e., players in  $S$ ) is non-negative, the payoff for negative contributors (i.e., players in  $T$ ) is non-positive, and the payoff for the remaining players is zero. The idea is thus to define the importance of criterion  $i$  relatively to bi-coalition  $(S, T)$  as  $\phi_i^{S,T}(v) = \left| \varphi_i^{S,T}(v) \right|$  in order to obtain non-negative values. One can then define the mean importance of criterion  $i$  as the average value of  $\phi_i^{S,T}(v)$  over all bi-coalitions  $(S, T)$  such that  $S \cup T = N$  (Labreuche and Grabisch, 2006b):

$$\begin{aligned} \phi_i(v) &= \frac{1}{2^{n-1}} \sum_{S \subseteq N, i \in S} \phi_i^{S, N \setminus S}(v) \\ &= \sum_{(A,B) \in \mathcal{Q}(N \setminus \{i\})} \frac{(a+b)!(n-a-b-1)!}{2^{a+b} n!} \\ &\quad \times (v(A \cup \{i\}, B) - v(A, B \cup \{i\})) \end{aligned}$$

This value turns out to be exactly the average weight of criterion  $i$  in the bipolar Choquet integral (Labreuche and Grabisch, 2006b; Kojadinovic, 2007b).

The interaction index  $I_{ij}(v)$  can be obtained from the importance indices by using the recursive axiom of Grabisch and Roubens (1999):

$$I_{ij}(v) = \sum_{(A,B) \in \mathcal{Q}(N \setminus \{i\})} \frac{(a+b)!(n-a-b-2)!}{2^{a+b}(n-1)!} \times \left( \delta_{\{i,j\}, \emptyset}^{A,B}(v) - \delta_{\emptyset, \{i,j\}}^{A,B}(v) \right)$$

where  $\delta_{\{i,j\}, \emptyset}^{A,B}(v) = v(A \cup \{i, j\}, B) - v(A \cup \{i\}, B) - v(A \cup \{j\}, B) + v(A, B)$  and  $\delta_{\emptyset, \{i,j\}}^{A,B}(v) = v(A, B \cup \{i, j\}) - v(A, B \cup \{i\}) - v(A, B \cup \{j\}) + v(A, B)$ . The interaction index  $I_{ij}(v)$  can be interpreted in terms of the variation of the mean weight of criterion  $i$  in the bipolar Choquet integral when criterion  $j$  varies (Kojadinovic, 2007b).

## 5.2 The index of average improvement

This section is based on Grabisch and Labreuche (2001). The main concern here is to know on which criteria acts should be improved on average so as to reach the highest possible global score. As an example of application, acts could be employees in a company. Their overall level is evaluated from their skills or know-how in different fields. The company is interested in helping these people to reach the highest possible level so that they will work more efficiently. So, we wish to construct an index of importance  $W_A$  (where  $A \subseteq N$  is a coalition of criteria) that will be large if improving the score of acts in criteria in  $A$  yields in general a large improvement of the overall evaluation. Such an index  $W_A$  should be defined and characterized with the help of axioms that are very intuitive for a DM. Basically, we wish to construct such an index for the Choquet integral. However the Choquet integral has a complicated expression and is not directly understandable by a DM. Hence,  $W_A$  must be defined for general aggregation functions  $F$  in  $L^2([0, 1]^n)$ , and characterized by axioms based on special evaluation functions that are much more intuitive than the Choquet integral. The main axiom (called *Step Evaluation*) considers a special family of aggregators that can only take the values 0 and 1. For these  $\{0, 1\}$ -evaluation functions, a natural expression for  $W_A$  comes up. The expression of  $W_A(F)$  is also natural when  $F$  is the weighted sum, which provides the second axiom. From these two axioms, one can deduce the expression of  $W_A(F)$  for any  $F$ , assuming that  $W_A(F)$  is linear and continuous w.r.t.  $F$ :

$$W_A(F) = 3 \cdot 2^{|A|} \int_{f \in [0, 1]^n} \int_{g_A \in [f_A, 1]} [F(g_A, f_{-A}) - F(f)] df dg_A ,$$

where  $g_A \in [f_A, 1]$  means that for any  $i \in A$ ,  $g_i \in [f_i, 1]$ .

This expression is then computed for the Choquet integral:

$$W_i(\mathcal{C}_\mu) = 6 \sum_{K \subset N \setminus \{i\}} \frac{(|K|+1)!(n-|K|)!}{(n+2)!} [\mu(K \cup \{i\}) - \mu(K)] .$$

This final expression is close to the Shapley index. Interestingly, (Marichal and Mathonet, 2007, Corollary 19) have shown that the  $W_i(\mathcal{C}_\mu)$  coefficients are also exactly the dominant coefficients in the best approximation of the Choquet integral by a linear function.

### 5.3 Andness and orness

The andness and orness degrees have been introduced by Dujmović (1974) for root-mean powers, and express the relative location of a given aggregation function with respect to minimum and maximum, respectively. More precisely, for any aggregation function  $F$  on  $[0, 1]^n$ ,

$$\begin{aligned} \text{andness}(F) &:= \frac{\overline{\max} - \overline{F}}{\overline{\max} - \overline{\min}}, \\ \text{orness}(F) &:= \frac{\overline{F} - \overline{\min}}{\overline{\max} - \overline{\min}}, \end{aligned}$$

where  $\overline{F}$  indicates the expected value of  $F$  on  $[0, 1]^n$ , assuming that the inputs are independent and uniformly distributed (similarly for min and max). Clearly, these degrees are in  $[0, 1]$ , and  $\text{andness}(F) + \text{orness}(F) = 1$  always holds. From the relation  $\overline{\min} = \frac{1}{n+1}$  and the fact that min and max are dual aggregation functions, one obtains

$$\begin{aligned} \text{andness}(F) &:= \frac{n}{n-1} - \frac{n+1}{n-1} \overline{F}, \\ \text{orness}(F) &:= -\frac{1}{n-1} + \frac{n+1}{n-1} \overline{F}. \end{aligned}$$

Marichal (2004) has shown that the orness degree for the Choquet integral is

$$\begin{aligned} \text{orness}(\mathcal{C}_\mu) &= \sum_{\substack{T \subseteq N \\ 0 < t < n}} \frac{1}{(n-1) \binom{n}{t}} \mu(T) \\ &= \sum_{T \subseteq N} \frac{n-t}{(n-1)(t+1)} m^\mu(T) \end{aligned}$$

where  $m^\mu$  is the Möbius transform of  $\mu$ .

## 5.4 Tolerance and intolerance

In Section 2.7.3, we have presented the notions of  $k$ -(in)tolerant capacities, and of veto or favor criteria. We give in this section complementary notions, leading to useful indices, introduced by Marichal (2007).

The notion of veto and favor being rather extreme, they rarely occur in practice, and it is more useful to define a degree of veto and favor. This can be achieved by considering the fact that, for every  $x \in [0, 1]^n$ , if criterion  $i$  is a veto, then  $\mathcal{C}_\mu(0_i, x_{-i}) = 0 = \min(0_i, x_{-i})$ , and if  $i$  is a favor then  $\mathcal{C}_\mu(1_i, x_{-i}) = 1 = \max(1_i, x_{-i})$ . Then the degrees of veto and favor for criterion  $i$  are respectively defined by

$$\begin{aligned} \text{veto}(\mathcal{C}_\mu, i) &:= \frac{E(\max(0_i, Z_{-i})) - E(\mathcal{C}_\mu(0_i, Z_{-i}))}{E(\max(0_i, Z_{-i})) - E(\min(0_i, Z_{-i}))} \\ \text{favor}(\mathcal{C}_\mu, i) &:= \frac{E(\mathcal{C}_\mu(1_i, Z_{-i})) - E(\min(1_i, Z_{-i}))}{E(\max(1_i, Z_{-i})) - E(\min(1_i, Z_{-i}))}, \end{aligned}$$

where  $E$  denotes expectation, assuming that the random inputs  $Z_1, \dots, Z_n$  are independent and uniformly distributed. This gives, after computation (Marichal, 2007),

$$\begin{aligned} \text{veto}(\mathcal{C}_\mu, i) &= 1 - \sum_{T \subseteq N \setminus i} \frac{1}{(n-1) \binom{n-1}{t}} \mu(T) \\ &= 1 - \sum_{T \subseteq N \setminus i} \frac{n}{(n-1)(t+1)} m^\mu(T) \\ \text{favor}(\mathcal{C}_\mu, i) &= \sum_{T \subseteq N \setminus i} \frac{1}{(n-1) \binom{n-1}{t}} \mu(T \cup i) - \frac{1}{n-1} \\ &= \sum_{T \subseteq N \setminus i} \frac{n}{(n-1)(t+1)} [m^\mu(T \cup i) + m^\mu(T)] - \frac{1}{n-1}. \end{aligned}$$

Just as for veto and favor indices, it seems interesting to introduce indices measuring the degree to which a Choquet integral is  $k$ -tolerant or  $k$ -intolerant. The basic property used for defining a degree of being  $k$ -intolerant is  $\mathcal{C}_\mu(x) = 0$  for all  $x \in [0, 1]^n$  such that  $x_{\sigma(k)} = 0$ , while for  $k$ -tolerance the equivalent property is  $\mathcal{C}_\mu(x) = 1$  for all  $x \in [0, 1]^n$  such that  $x_{\sigma(n-k+1)} = 1$ . This leads to the following definitions, for any  $0 \leq k < n$ :

$$\begin{aligned} \text{intol}_k(\mathcal{C}_\mu) &= \frac{n-k+1}{(n-k) \binom{n}{k}} \sum_{\substack{K \subseteq N \\ |K|=k}} E(\mathcal{C}_\mu(0_K, Z_{-K})), \\ \text{tol}_k(\mathcal{C}_\mu) &= \frac{n-k+1}{(n-k) \binom{n}{k}} \sum_{\substack{K \subseteq N \\ |K|=k}} E(\mathcal{C}_\mu(1_K, Z_{-k})) - \frac{1}{n-k}, \end{aligned}$$

which gives, after computation

$$\begin{aligned} \text{intol}_k(\mathcal{C}_\mu) &= 1 - \frac{1}{n-k} \sum_{t=0}^{n-k} \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \mu(T), \\ \text{tol}_k(\mathcal{C}_\mu) &= \frac{1}{n-k} \sum_{t=k}^n \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \mu(T) - \frac{1}{n-k}. \end{aligned}$$

## 6 Analysis of the models based on alternatives

This section is devoted to the latest phase of a decision aiding process where the multi-criteria model has already been constructed from interview with the DM. In this last phase, the model is evaluated against several options that are brought by the DM. Yet, it is often the case that the DM needs more than just an evaluation of his options. He wants to have an explanation of these evaluations and in particular of the ordering of the options resulting from these evaluations. It often happens that, looking at the evaluations, the DM wants to work on an option so as to improve it. He needs thus to have recommendations on the criteria on which the improvement is the most beneficial.

### 6.1 Explanation of the result of the Choquet integral

This section is based on Labreuche (2005, 2006). The DM wants to have an explanation of the result of the evaluation made by the model. Consider two profiles  $g, h \in \mathbb{R}^n$ . Assuming without loss of generality that  $F(g) > F(h)$ , the recommendation of the model  $F$  is that option  $g$  is preferred to  $h$ . The DM wants to have an argumentation of the reason of this preference.

The central ingredient here is the notion of *argument*. The arguments that can be used are the elementary decision behaviours  $P$  represented in the aggregation function. Consider a family of aggregation functions parametrized by coefficients  $\alpha = \{\alpha_i\}_{i \in P}$ :

$$F_\alpha(f) = \sum_{i \in P} \alpha_i F_i(f) \quad (6.1)$$

where for all  $i \in P$ ,  $\alpha_i \geq 0$  and  $\sum_{i \in P} \alpha_i = 1$ . By equation (2.6), the 2-additive model can be put into this form. It is also the case of the general Choquet integral, even if this representation is not unique and not always simple to obtain. The  $F_i$  functions are min/max combinations of the criteria for the Choquet integral. Each elementary decision behavior  $F_i$  is very

simple to understand and corresponds to an argument. Hence  $P$  is identified to the set of arguments that can be given to the DM. An argument  $i \in P$  is said to be positive (resp. negative) regarding the comparison  $F_\alpha(g) > F_\alpha(h)$  if  $F_i(g) > F_i(h)$  (resp.  $F_i(g) < F_i(h)$ ).

The complexity of the argumentation to be presented to the DM depends basically on how tight the comparison between  $g$  and  $h$  is. For instance, if  $g_i > h_i$  for all  $i \in N$  (Situation I), then relation  $F_\alpha(g) > F_\alpha(h)$  is trivial and it is not necessary to use the specificities of the values  $\alpha$  in the generated argumentation. Now, if  $F_{\hat{\alpha}}(g) \leq F_{\hat{\alpha}}(h)$  where  $\hat{\alpha}_i = 1/|P|$  for all  $i \in P$  (Situation II), there are more negative arguments than positive ones. Thus relation  $F_\alpha(g) > F_\alpha(h)$  implies that the positive arguments are generally speaking stronger than the negative arguments, i.e., the weights  $\alpha_i$  of the positive arguments are on average larger than that of the negative arguments. More generally, depending on  $\alpha$ ,  $f$  and  $g$ , there are several patterns of argumentation that can be generated to the DM. These types are called *anchors* (Labreuche, 2006). Situations I and II are two instances of anchors. We denote by  $\Psi(\alpha, g, h)$  the set of anchors that can be used in the comparison  $F_\alpha(g) > F_\alpha(h)$ . An anchor  $\psi$  can be basically used to specify the reason for discarding some arguments of  $P$ . This reason might be the existence of reference parameters  $\alpha_\psi : P \rightarrow \mathbb{R}$  assigned to the anchor  $\psi$ . This is the case of Situation II where  $\alpha_\psi := \hat{\alpha}$  for this anchor.

The set of arguments that can be used by an anchor  $\psi \in \Psi$  is denoted by  $P(\psi) \subseteq P$ . One has  $P(\psi) = \emptyset$  in the anchor of Situation I, and  $P(\psi) = P$  for Situation II. When presenting to the DM the argumentation associated to an anchor  $\psi$ , it is usually not necessary to mention all arguments  $P(\psi)$ . In order to determine the very arguments that are at the root of the result of the comparison of  $g$  and  $h$ , the elementary decision behaviours are “removed” from the explanation one at a time while the decision remains the same. This means that the parameters  $\alpha$  are replaced by the reference ones for some elements of  $P(\psi)$ . We define the weights  $\alpha \otimes_A \alpha_\psi$  by  $(\alpha \otimes_A \alpha_\psi)_d = \alpha_d$  if  $d \in A$ , and  $= (\alpha_\psi)_d$  otherwise. Then one looks for  $A \subseteq P(\psi)$  with the lowest cardinality for which  $F_{\alpha \otimes_A \alpha_\psi}(g) > F_{\alpha \otimes_A \alpha_\psi}(h)$  (Labreuche, 2006). Such subset  $A$  is displayed to the DM for the anchor  $\psi$ .

## 6.2 Recommendation of the criteria on which improvement is the most beneficial

It is quite usual in MCDA that the options that are evaluated are not fixed and that the DM wishes to obtain recommendations on how to improve an option of interest. Section 7.1 gives an example of this situation. The DM wants to know on which criteria an improvement of the option should be done in order to get the maximal possible improvement of the overall score. The option is described by a profile  $f = (f_1, \dots, f_n) \in \mathbb{R}^n$ . Most of the

time the DM wants to know how to improve option  $f$  into a new profile  $f'$  such that the overall evaluation  $F(f')$  reaches a given expectation level (Montmain, 2007):

$$\min c(f, f') \quad \text{under} \quad \begin{cases} f' \in \mathbb{R}^n \\ \forall i \in N, f'_i \geq f_i \\ F_\mu(f') = e \end{cases}$$

where  $F$  is the aggregation function,  $e$  is the expectation level, and  $c(f, f')$  quantifies the cost to improve option  $f$  into a new profile  $f'$ . The previous optimization problem provides the new profile  $f'$  that should be reached. The drawback of this approach is that the DM is often not able to easily construct a new option corresponding to the profile  $f'$  (see Section 7.1). He will thus proceed iteratively by transforming  $f$  into  $f_1$ , then  $f_1$  into  $f_2$ , and so on, until the expectation level  $e$  is reached. The recommendation the DM wants to have is an indication of the criteria of  $f$  that should be improved in priority. As we have seen, we have no idea of the intensity of improvement that the DM will obtain.

To solve this problem, an index denoted by  $\omega_A(F)(f)$  quantifying the worth for the profile  $f$  to be improved in criteria among  $A \subseteq N$ , subject to the evaluation function  $F$ , has been proposed in Labreuche (2004). As shown by the following example, one shall not restrict the subsets  $A$  to singletons. Consider the case of an expert that is very intolerant, described by the min aggregation function  $F(f) = \min_{i \in N} f_i$ . If all criteria are equally satisfied, then improving only one criterion will not change the overall evaluation, so that it is useless to work on a single criterion, whereas it is worth improving all of them at the same time.

Let  $V$  be the set of piecewise continuous functions defined on  $[0, 1]^n$ . This space is endowed with the norm  $\|u\|_V = \sup_{x \in [0, 1]^n} |u(x)|$ . The index  $\omega_A$  is seen as an operator from  $V$  onto itself.

The index  $\omega_A$  is defined axiomatically for any  $F \in V$ . First of all, if  $F$  is constant over criteria  $A$ , then  $\omega_A(F)(f) = 0$ . Moreover, if  $F$  does not depend on criterion  $i$ , then  $\omega_{A \cup \{i\}}(F)(f) = \omega_A(F)(f)$ . When  $F$  can be decomposed into  $n$  functions  $F_i$  of each criterion, another requirement describes an optimistic decomposability of  $\omega_A(F)$  from the  $\omega_i(F_i)$ . Lastly, an invariance property  $\omega_A(F)$  for  $\{0, 1\}$ -valued functions  $F$  is described. Previous requirements combined with linearity, symmetry and continuity (i.e.  $\sup_{F \in V, F \neq 0} \frac{\|\omega_A(F)\|_V}{\|F\|_V} < \infty$ ) of  $\omega_A$  uniquely defines  $\omega$  (Labreuche, 2004):

$$\omega_A^\wedge(F)(f) := \int_0^1 [F((1 - \tau)f_A + \tau, f_{N \setminus A}) - F(f)] d\tau. \quad (6.2)$$

This expression gives the mean value of the difference  $F(g_A, f_{N \setminus A}) - F(f)$  only for  $g_A$  on the diagonal from  $f_A$  to  $1_A$ . Therefore, equation 6.2 gives the mean impact of uniformly improving all criteria of  $A$  at the same time, where one assumes that all possible levels of improvement (from sticking to  $f_A$  up to reaching the ideal profile  $1_A$ ) have the same probability to occur.

By relaxing the decomposability axiom to its most general form, one obtains infinitely many operators  $\omega$  satisfying the axioms. All operators are identical when  $A$  reduces to a singleton. However, the index  $\omega^\wedge$  gives the largest values among all previous operators  $\omega$  for any  $F \in V$ ,  $f \in [0, 1]^n$  and  $A \subseteq N$  (Labreuche, 2004). Hence equation (6.2) favors the coalitions with large cardinality. Finally note that (6.2) can be extended so as to take into account the improvement cost  $c$ :

$$\omega_A^\wedge(F)(f) := \int_0^1 \frac{F((1-\tau)f_A + \tau, f_{N \setminus A}) - F(f)}{c(f, ((1-\tau)f_A + \tau, f_{N \setminus A}))} d\tau.$$

## 7 Applications and softwares

The Choquet integral is more and more used by researchers around the world as a versatile tool that models interaction among criteria. Let us cite a few (among many others) new applications of the Choquet integral that have been carried out during the last ten years: policy capturing in strategic decision making (Liginlal and Ow, 2005), analysis of root dispersal where interactions model the competition among wood species in forests (Näther and Wälder, 2007), computation of the number of citations (Torra and Narukawa, to appear), clinical diagnosis (Saito et al., 2007), detection of line arrows in technical drawings (Wendling and Tabbone, 2003), monitoring of the improvement of an overall industrial performance (Berrah et al., 2008), selection of groups of genes with high classifying power in gene expression data analysis (Fagnelli and Moretti, to appear), evaluation of discomfort in sitting position when driving a car (Grabisch et al., 2002), to cite a few.

All details about these applications can be found in the referenced papers. For space limitations, in Section 7.1 below we only develop one specific industrial application (Pignon and Labreuche, 2007).

There exist softwares providing a toolbox for the Choquet integral and related notions, which can also be used in applications since they contain identification methods. In Sections 7.2 and 7.3, we develop two of them, namely Kappalab and Myriad. Another recently available toolbox has been developed by Beliakov. It is called `fntools`, and is available at <http://www.deakin.edu.au/~gleb/fntools.html>.

## 7.1 A class of industrial applications: design of complex systems

An interesting class of applications is the design of an industrial product or a complex system. One can think of the body of a car, a chemical process, or a military information architecture (Pignon and Labreuche, 2007). Such a complex system is characterized by input parameters  $z_1, \dots, z_p$ , e.g., the size of the components of a car body, the parameters controlling a chemical process, or the precise definition of all entities in the military force and their links. The set of all possible values of the vector of variables  $(z_1, \dots, z_p)$  is denoted by  $\Gamma$ . A system is thus defined by an element  $\gamma \in \Gamma$ . Not all elements of  $\Gamma$  lead to admissible systems for the customer since some requirements of the customer must usually be fulfilled. The set of elements  $\Gamma_{\mathcal{F}} \subseteq \Gamma$  for which the associated system satisfies these requirements are the feasible values of the input parameters.

Yet, all elements of  $\Gamma_{\mathcal{F}}$  are not indifferent to the customer. The company needs to construct a model of the preferences of the customer based on his decision criteria. These criteria are often a refinement of the requirements. The set of attributes is again denoted by  $X_1, \dots, X_n$ , and the set of alternatives is  $X = X_1 \times \dots \times X_n$ . For the body of a car, these attributes describe how the body buckles during a crash test, where the crash test is often simulated with the help of a numerical code. For a chemical process, they quantify the quality (purity, concentration,...) of the result of the process. For the military architecture, they quantify the fulfilment of the operational mission and are obtained thanks to large simulations on architecture-labs (Pignon and Labreuche, 2007). Let  $T : \Gamma_{\mathcal{F}} \rightarrow X$  be the transformation that provides the values on the attributes of the system obtained from a vector  $\gamma \in \Gamma_{\mathcal{F}}$  of the input parameters. As we have just seen, the determination of  $T(f)$  for  $f \in \Gamma_{\mathcal{F}}$  is not easy for complex systems. It requires complex simulations or experiments, and is thus costly and time consuming.

The overall evaluation of a system characterized by  $\gamma \in \Gamma_{\mathcal{F}}$  is

$$F(u(T(\gamma))) = F(u_1(T_1(\gamma)), \dots, u_n(T_n(\gamma)))$$

where  $T_1, \dots, T_n$  are the  $n$  components of  $T$ ,  $u_1, \dots, u_n$  are the value functions and  $F$  is the aggregation function.

The preferences of the customer are usually complex and require an elaborate multi-criteria model.

First of all, bipolar scales are most of the time well-suited for this kind of problems. The three reference levels are meaningful for complex systems. Indeed, the neutral value  $\mathbf{0}_i$  corresponds to the performance of today's generation of systems. It is indeed neither good nor bad to perform as today's systems. The satisficing value  $\mathbf{1}_i$  is the targeted aspiration level to be reached. One is satisfied if this level is reached by the new product, even if it would be possible to do better than this level. Finally, the anti-satisficing

value  $-1_i$  corresponds to the level of performance of the previous generation of systems. Reaching this level would be considered as a decline.

Concerning the aggregation part of the evaluation model, a weighted sum is generally not sufficient. Among the criteria of the customer, one usually have operational and monetary ones. For a customer that aims to possess a complex system, the performance of the product is compulsory. A low cost cannot compensate for bad operational performance. As a result, the operational criteria act as veto. Many other interactions such as conditional relative importance of criteria are most often encountered.

Once the model has been constructed from interviews with the customer, this model will help the company in designing the best system for the customer at the best price. It is important that the recommendations resulting from the multi-criteria model are explained to the customer. This means that for  $\gamma, \gamma' \in \Gamma_{\mathcal{F}}$ , one needs to understand the comparison of  $T(\gamma)$  with  $T(\gamma')$  by  $\succeq$ , e.g., why  $T(\gamma) \succ T(\gamma')$  (Labreuche, 2005, 2006).

The company would like at the end to determine

$$\text{Argmax}_{\gamma \in \Gamma_{\mathcal{F}}} F(u_1(T_1(\gamma)), \dots, u_n(T_n(\gamma))).$$

This is a very complex operation since we have seen that  $T$  is not known explicitly, and it is very complex to perform one computation of  $T$ . There are two possible solutions. The first one consists in approximating  $T$  by a model, e.g., a second order model. Statistics provide powerful tools for such approximations, such as experiment design. Experiment design is well-suited for our problem since it tries to minimize the number of computations of  $T$ . An alternative approach is to use optimization algorithms, such as steepest descent to iteratively converge to the optimal solution. In the steepest descent method, one needs to know the direction where it will be more rewarding to change the current vector  $\gamma \in \Gamma_{\mathcal{F}}$ . The  $\omega_i$  indicator described in Section 6.2 has been designed to help the company in determining the right direction for the modification of the current solution.

## 7.2 The KAPPALAB tool

Kappalab, which is an acronym for “laboratory for capacities”, is a free package designed for the GNU R statistical system. It can be downloaded from the *Comprehensive R Archive Network* (<http://cran.r-project.org>) or from <http://www.stat.auckland.ac.nz/~ivan/kappalab>, and simply needs the R system to be installed. It provides a toolbox under a Matlab-like environment, that is, the user can do various operations by using an on-line high level language. Among other possibilities, one can:

- enter or construct a capacity, compute its Möbius or interaction transform and back, its Shapley value, perform some tests on a given capacity (e.g., monotonicity);
- compute the Choquet or Sugeno integral for general or  $k$ -additive capacities, even if they are expressed by their Möbius or interaction transform;
- construct a model for data fitting, using any of the techniques described in Section 2.6, including TOMASO.

Kappalab allows to work up to  $n = 32$ . However, due to memory limitation, for such high values of  $n$ , only 2 or 3-additive capacities can be considered. For general capacities, one can work comfortably with up to  $n = 10$  criteria. A detailed example of its utilization can be found in (Grabisch et al., 2008b).

### 7.3 The MYRIAD tool

MYRIAD is a multi-criteria decision aiding software developed by THALES<sup>4</sup> based on the Choquet integral. It aims at first helping the DM to construct the model, and then to analyse a set of options on the constructed model (Labreuche and Le Huédé, 2005). MYRIAD has a user-friendly graphical interface. Even if this tool is available for the moment only for the THALES customers, we briefly present it since it covers most aspects of decision aid and it includes the major theoretical advances on the Choquet integral. Here are the main features of this tool.

- For the construction of the model:
  - enter any hierarchy of criteria (see Figure 7.1);
  - choice of the model among the 2-additive capacities and general capacities;
  - analysis of the inconsistencies of the preferential information (Labreuche and Le Huédé, 2006);
  - sensitivity analysis of the values of the capacity (see Figure 7.2);
  - possibility to model context-dependent value functions (Labreuche, 2007).
- For the analysis of several options on the multi-criteria model obtained in the previous step:
  - display of the assessment scores of two options on the criteria hierarchy;
  - graphical representation of the results of the aggregation computations (see Figure 7.3). One can “plot” the result of the aggregation  $F_\alpha$  on an option  $f$  in a pie-chart in which each segment represents an elementary behavior  $F_k$  in equation 6.1. The aperture of the segment related to  $F_k$  is  $2\pi\alpha_k$ , and this segment is covered at rate  $F_k(f)$ .

---

<sup>4</sup> THALES is a world leader in Mission-critical information systems for the Aerospace, Defence and Security markets.

Hence, the surface covered by this segment is  $\alpha_k F_k(f)$  so that the overall covering of the disk is precisely  $F_\alpha(f)$ . This graphical representation makes it easy to understand why result  $F_\alpha(f)$  is rather high (the disk is pretty filled up) or low (the disk is almost empty). It is displayed in the graphical interface (see Figure 7.3);

- generation of a textual argumentation of the result of the aggregation computations (See Section 6.1);
- determination of the criteria to be improved first (See Section 6.2).

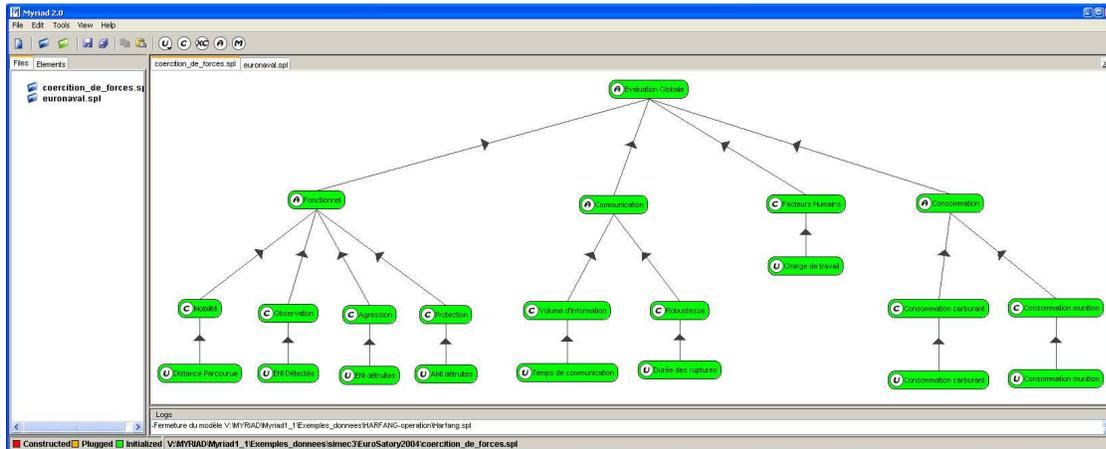


Fig. 7.1. Hierarchy of criteria.

## 8 Some directions for future research

As described above, many developments around the Choquet and Sugeno integrals have occurred in the last decade. Nevertheless, these developments leave open many questions. In particular, the discovery of the new bipolar models opens many new questions on their better understanding and management. Here is a non-exhaustive list of some future works still to be done.

- Axiomatization of the Choquet integral in the context of MCDA based on  $\succeq$ , in the same spirit as Greco et al. (2004) and Bouyssou et al. (Bouyssou et al., 2006; Bouyssou and Marchant, 2007a,b), is still missing.
- Define ordinal models that do not have the usual drawback of lacking of discriminatory power (see, e.g., the *drowning effect* (Dubois et al., 2001a)).

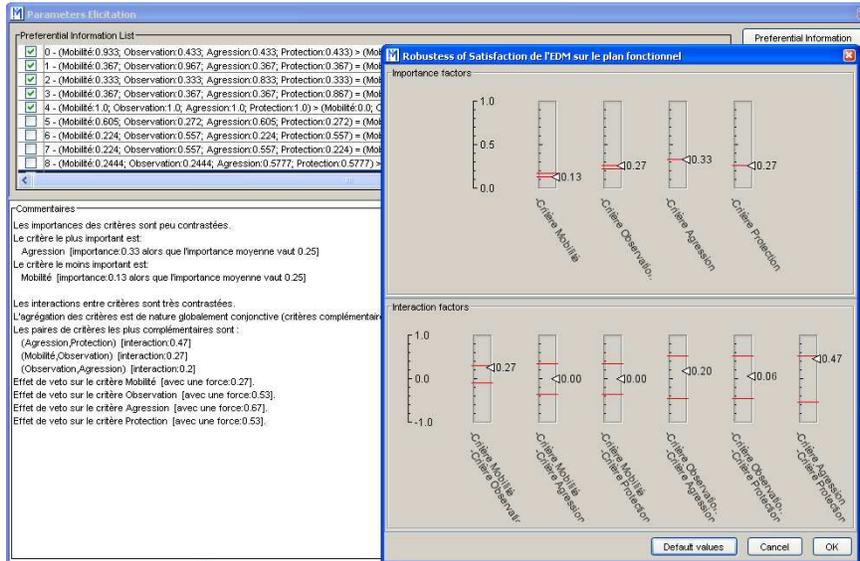


Fig. 7.2. Sensitivity analysis on the parameters of the capacity.

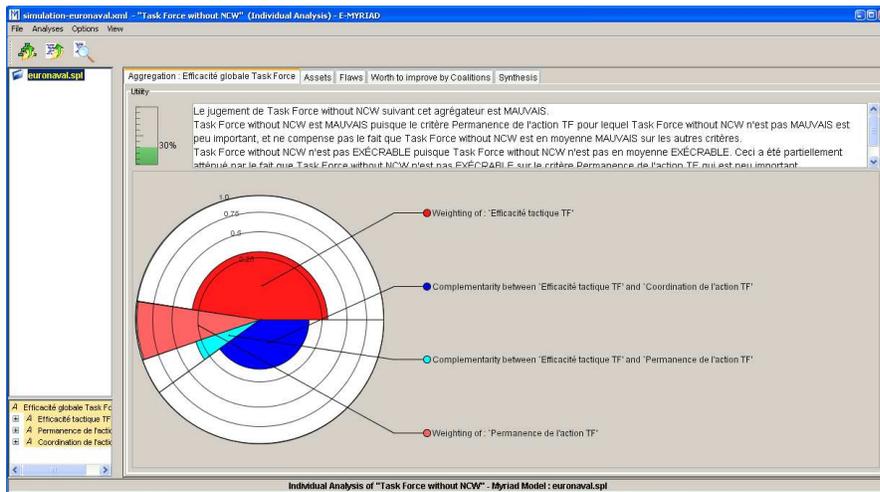


Fig. 7.3. Pie chart representing a 2-additive Choquet integral.

- Bi-capacities are very attractive in practice due to their ability to represent bipolar decision behaviour. Unfortunately, it is not easy to use them in real applications since one would need much more preferential information than what the DM can bear. Simplified sub-models and smart algorithms are still needed.

We hope that these questions will receive satisfactory answers in that next decade.

## References

- C. A. Bana e Costa and J.-C. Vansnick. A theoretical framework for Measuring Attractiveness by a Categorical Based Evaluation Technique (MACBETH). In *Proc. XIth Int. Conf. on MultiCriteria Decision Making*, pages 15–24, Coimbra, Portugal, August 1994.
- C. A. Bana e Costa and J.-C. Vansnick. Applications of the MACBETH approach in the framework of an additive aggregation model. *J. of Multicriteria Decision Analysis*, 6:107–114, 1997.
- C. A. Bana e Costa and J.-C. Vansnick. The MACBETH approach: basic ideas, software and an application. In N. Meskens and M. Roubens, editors, *Advances in Decision Analysis*, pages 131–157. Kluwer Academic Publishers, 1999.
- L. Berrah, G. Mauris, and J. Montmain. Monitoring the improvement of an overall industrial performance based on a Choquet integral aggregation. *Omega*, 36(3):340 – 351, 2008.
- J. M. Bilbao, J. R. Fernandez, A. Jiménez Losada, and E. Lebrón. Bicooperative games. In J. M. Bilbao, editor, *Cooperative games on combinatorial structures*. Kluwer Acad. Publ., 2000.
- J. M. Bilbao, J. R. Fernández, N. Jiménez, and J. J. López. The value for bi-cooperative games. *Annals of Operations Research*, to appear.
- D. Bouyssou and Th. Marchant. An axiomatic approach to noncompensatory scoring methods in MCDM, I: The case of two categories. *Eur. J. of Operational Research*, 178:217–245, 2007a.
- D. Bouyssou and Th. Marchant. An axiomatic approach to noncompensatory scoring methods in MCDM, II: More than two categories. *Eur. J. of Operational Research*, 178:246–276, 2007b.
- D. Bouyssou, Th. Marchant, and M. Pirlot. A conjoint measurement approach to the discrete sugeno integral. a note on a result of greco, matarazzo and slowiński. *Annales du LAMSADE*, 6:39–60, October 2006.
- G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1953.
- E. Combarro and P. Miranda. A genetic algorithm for the identification of fuzzy measures from sample data. In *Int. Fuzzy Systems Association World Congress (IFSA 2003)*, pages 163–166, Istanbul, Turkey, June 2003.
- D. Denneberg and M. Grabisch. Measure and integral with purely ordinal scales. *J. of Mathematical Psychology*, 48:15–27, 2004.
- D. Denneberg and M. Grabisch. Interaction transform of set functions over a finite set. *Information Sciences*, 121:149–170, 1999.
- D. Dubois and H. Fargier. Bridging the gap between discrete Sugeno and Choquet integrals. Presented at RUD'2005 (workshop on Risk, Utility and Decision), Heidelberg, Germany, 22-24 june 2005, 2005.
- D. Dubois, W. Ostasiewicz, and H. Prade. Fuzzy sets: history and basic notions. In D. Dubois and H. Prade, editors, *Fundamentals of Fuzzy Sets*, volume 1 of *The Handbooks of Fuzzy Sets*, pages 21–124. Kluwer Academic Publishers, 2000a.
- D. Dubois, H. Prade, and R. Sabbadin. Qualitative decision theory with Sugeno integrals. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integrals — Theory and Applications*, pages 314–332. Physica Verlag, Heidelberg, 2000b.

- D. Dubois, J.-L. Marichal, H. Prade, M. Roubens, and R. Sabbadin. The use of the discrete Sugeno integral in decision making: a survey. *Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9(5):539–561, 2001a.
- D. Dubois, H. Prade, and R. Sabbadin. Decision-theoretic foundations of qualitative possibility theory. *Eur. J. of Operational Research*, 128:459–478, 2001b.
- J. J. Dujmović. Weighted conjunctive and disjunctive means and their application in system evaluation. *Univ. Beograd. Publ. Elektrotechn. Fak.*, pages 147–158, 1974.
- H. Fargier and R. Sabbadin. Qualitative decision under uncertainty: back to expected utility. *Artificial Intelligence*, 164:245–280, 2005.
- D. Felsenthal and M. Machover. Ternary voting games. *Int. J. of Game Theory*, 26:335–351, 1997.
- V. Fragnelli and S. Moretti. A game theoretical approach to the classification problem in gene expression data analysis. *Computers and Mathematics with Applications*, to appear.
- K. Fujimoto and T. Murofushi. Hierarchical decomposition of the Choquet integral. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integrals — Theory and Applications*, pages 94–104. Physica Verlag, 2000.
- K. Fujimoto, I. Kojadinovic, and J.-L. Marichal. Axiomatic characterizations of probabilistic and cardinal-probabilistic interaction indices. *Games and Economic Behavior*, 55:72–99, 2006.
- M. Grabisch. Set function over finite sets: transformations and integrals. In E. Pap, editor, *Handbook of Measure Theory*, pages 1381–1401. Elsevier Science Publ., 2002.
- M. Grabisch. The Möbius function on symmetric ordered structures and its application to capacities on finite sets. *Discrete Mathematics*, 287(1-3):17–34, 2004a.
- M. Grabisch. The symmetric Sugeno integral. *Fuzzy Sets and Systems*, 139:473–490, 2003a.
- M. Grabisch. Modelling data by the Choquet integral. In V. Torra, editor, *Information fusion in data mining*, pages 135–148. Physica Verlag, 2003b.
- M. Grabisch. Évaluation subjective. In D. Bouyssou, D. Dubois, M. Pirlot, and H. Prade, editors, *Concepts et Méthodes pour l’Aide à la Décision. Vol. 3: Analyse multicritère*, IC2, pages 175–232. Hermès, 2005.
- M. Grabisch. The Choquet integral as a linear interpolator. In *10th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2004)*, pages 373–378, Perugia, Italy, July 2004b.
- M. Grabisch. Some lexicographic approaches to the Sugeno integral. In *Proc. Int. Conf. on Information Processing and Management of Uncertainty (IPMU’06)*, pages 572–579, Paris, France, July 2006.
- M. Grabisch. A new algorithm for identifying fuzzy measures and its application to pattern recognition. In *Int. Joint Conf. of the 4th IEEE Int. Conf. on Fuzzy Systems and the 2nd Int. Fuzzy Engineering Symposium*, pages 145–150, Yokohama, Japan, March 1995.
- M. Grabisch. The application of fuzzy integrals in multicriteria decision making. *European J. of Operational Research*, 89:445–456, 1996.
- M. Grabisch.  $k$ -order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, 92:167–189, 1997a.
- M. Grabisch. Alternative representations of discrete fuzzy measures for decision making. *Int. J. of Uncertainty, Fuzziness, and Knowledge Based Systems*, 5:587–607, 1997b.
- M. Grabisch. The interaction and Möbius representations of fuzzy measures on finite spaces,  $k$ -additive measures: a survey. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integrals — Theory and Applications*, pages 70–93. Physica Verlag, 2000.
- M. Grabisch and Ch. Labreuche. How to improve acts: an alternative representation of the importance of criteria in mcdm. *Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9:145–157, 2001.
- M. Grabisch and Ch. Labreuche. Bi-capacities. Part I: definition, Möbius transform and interaction. *Fuzzy Sets and Systems*, 151:211–236, 2005a.

- M. Grabisch and Ch. Labreuche. Fuzzy measures and integrals in MCDA. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis*, pages 563–608. Kluwer Academic Publishers, 2005b.
- M. Grabisch and Ch. Labreuche. Bi-capacities. Part II: the Choquet integral. *Fuzzy Sets and Systems*, 151:237–259, 2005c.
- M. Grabisch and Ch. Labreuche. A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. *JOR*, 6:1–44, 2008. doi 10.1007/s10288-007-0064-2.
- M. Grabisch and J.M. Nicolas. Classification by fuzzy integral — performance and tests. *Fuzzy Sets & Systems, Special Issue on Pattern Recognition*, 65:255–271, 1994.
- M. Grabisch and E. Raufaste. An empirical study of statistical properties of Choquet and Sugeno integrals. *IEEE Tr. on Fuzzy Systems*, to appear.
- M. Grabisch and M. Roubens. An axiomatic approach to the concept of interaction among players in cooperative games. *Int. Journal of Game Theory*, 28:547–565, 1999.
- M. Grabisch, J.-L. Marichal, and M. Roubens. Equivalent representations of set functions. *Mathematics of Operations Research*, 25(2):157–178, 2000.
- M. Grabisch, J. Duchêne, F. Lino, and P. Perny. Subjective evaluation of discomfort in sitting position. *Fuzzy Optimization and Decision Making*, 1(3):287–312, 2002.
- M. Grabisch, Ch. Labreuche, and J.-C. Vansnick. On the extension of pseudo-Boolean functions for the aggregation of interacting bipolar criteria. *Eur. J. of Operational Research*, 148:28–47, 2003.
- M. Grabisch, S. Greco, and M. Pirlot. Bipolar and bivariate models in multicriteria decision analysis: descriptive and constructive approaches. *Int. J. of Intelligent Systems*, 23(9):930–969, 2008a.
- M. Grabisch, I. Kojadinovic, and P. Meyer. A review of capacity identification methods for Choquet integral based multi-attribute utility theory — applications of the Kappalab R package. *Eur. J. of Operational Research*, 186:766–785, 2008b.
- S. Greco, B. Matarazzo, and R. Słowiński. Bipolar Sugeno and Choquet integrals. In *EUROFUSE Workshop on Informations Systems*, pages 191–196, Varenna, Italy, September 2002.
- S. Greco, B. Matarazzo, and R. Słowiński. Axiomatic characterization of a general utility function and its particular cases in terms of conjoint measurement and rough-set decision rules. *Eur. J. of Operational Research*, 158(2):271–292, 2004.
- J. Havrda and F. Charvat. Quantification method of classification processes. concept of structural  $\alpha$ -entropy. *Kybernetika*, 3:30–95, 1967.
- I. Kojadinovic. Minimum variance capacity identification. *Eur. J. of Operational Research*, 177:498–514, 2007a.
- I. Kojadinovic. A weight-based approach to the measurement of the interaction among criteria in the framework of aggregation by the bipolar choquet integral. *European Journal of Operational Research*, 179:498–517, 2007b.
- D. H. Krantz, R. D. Luce, P. Suppes, and A. Tversky. *Foundations of measurement*, volume 1: Additive and Polynomial Representations. Academic Press, 1971.
- S. H. Kwon and M. Sugeno. A hierarchical subjective evaluation model using non-monotonic fuzzy measures and the Choquet integral. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integrals — Theory and Applications*, pages 375–391. Physica Verlag, 2000.
- Ch. Labreuche. Determination of the criteria to be improved first in order to improve as much as possible the overall evaluation. In *Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU)*, pages 609–616, Perugia, Italy, July 4-9 2004.
- Ch. Labreuche. Argumentation of the results of a multi-criteria evaluation model in individual and group decision aiding. In *Int. Conf. Of the Euro Society for Fuzzy Logic and Technology (EUSFLAT)*, pages 482–487, Barcelona, Spain, September 7-9 2005.
- Ch. Labreuche. Argumentation of the decision made by several aggregation operators based on weights. In *Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU)*, pages 683–691, Paris, France, July 2-7 2006.

- Ch. Labreuche. Modeling the context in multi-criteria decision analysis. In *Int. Conf. Of the Euro Society for Fuzzy Logic and Technology (EUSFLAT)*, Ostrava, Czech Republic, September 11-14 2007.
- Ch. Labreuche and M. Grabisch. The Choquet integral for the aggregation of interval scales in multicriteria decision making. *Fuzzy Sets & Systems*, 137:11–26, 2003.
- Ch. Labreuche and M. Grabisch. Partially unipolar bi-capacities in MCDM. In *SCIS-ISIS*, Yokohama, Japan, September 21-24 2004.
- Ch. Labreuche and M. Grabisch. Generalized Choquet-like aggregation functions for handling ratio scales. *European Journal of Operational Research*, 172:931–955, 2006a.
- Ch. Labreuche and M. Grabisch. The representation of conditional relative importance between criteria. *Annals of Operations Research*, 154:93–122, 2007.
- Ch. Labreuche and M. Grabisch. Definition of importance indices for bicapacities in MCDA. In *Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU)*, pages 292–299, Paris, France, July 2-7 2006b.
- Ch. Labreuche and M. Grabisch. A value for bi-cooperative games, submitted.
- Ch. Labreuche and F. Le Huédé. Myriad: a tool suite for MCDA. In *Int. Conf. of the Euro Society for Fuzzy Logic and Technology (EUSFLAT)*, pages 204–209, Barcelona, Spain, September 7-9 2005.
- Ch. Labreuche and F. Le Huédé. Inconsistencies in the determination of a capacity. In *European Conference on Artificial Intelligence (ECAI) - Workshop on Advances in Preference Handling*, pages 73–80, Riva del Garda, Italy, August 28 - September 1 2006.
- D. Liginlal and T.T. Ow. On policy capturing with fuzzy measures. *European Journal of Operational Research*, 167:461–474, 2005.
- L. Lovász. Submodular functions and convexity. In A. Bachem, M. Grötschel, and B. Korte, editors, *Mathematical programming. The state of the art*, pages 235–257. Springer Verlag, 1983.
- J.-L. Marichal. An axiomatic approach of the discrete Sugeno integral as a tool to aggregate interacting criteria in a qualitative framework. *IEEE Tr. on Fuzzy Systems*, 9(1):164–172, 2001.
- J.-L. Marichal. On Sugeno integral as an aggregation function. *Fuzzy Sets and Systems*, 114:347–365, 2000.
- J.-L. Marichal. Aggregation of interacting criteria by means of the discrete Choquet integral. In T. Calvo, G. Mayor, and R. Mesiar, editors, *Aggregation operators: new trends and applications*, volume 97 of *Studies in Fuzziness and Soft Computing*, pages 224–244. Physica Verlag, 2002.
- J.-L. Marichal. Tolerant or intolerant character of interacting criteria in aggregation by the Choquet integral. *Eur. J. of Operational Research*, 155(3):771–791, 2004.
- J.-L. Marichal. Weighted lattice polynomials. *Discrete Mathematics*, to appear.
- J.-L. Marichal.  $k$ -intolerant capacities and Choquet integrals. *Eur. J. of Operational Research*, 177(3):1453–1468, 2007.
- J.-L. Marichal. *Aggregation operators for multicriteria decision aid*. PhD thesis, University of Liège, 1998.
- J.-L. Marichal and P. Mathonet. Approximations of lovász extensions and their induced interaction index. *Discrete Applied Mathematics*, 2007. doi:10.1016/j.dam.2007.07.011.
- J.-L. Marichal and M. Roubens. Determination of weights of interacting criteria from a reference set. *Eur. J. of Operational Research*, 124:641–650, 2000.
- J.-L. Marichal, P. Meyer, and M. Roubens. Sorting multi-attribute alternatives: the TOMASO method. *Computers & Operations Research*, 32:861–877, 2005.
- P. Meyer and M. Roubens. Choice, ranking and sorting in fuzzy multiple criteria decision aid. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis*, pages 471–506. Kluwer Academic Publishers, 2005.
- P. Miranda and M. Grabisch. Optimization issues for fuzzy measures. *Int. J. of Uncertainty, Fuzziness, and Knowledge-Based Systems*, 7(6):545–560, 1999.
- P. Miranda, M. Grabisch, and P. Gil.  $p$ -symmetric fuzzy measures. *Int. J. of Uncertainty, Fuzziness, and Knowledge-Based Systems*, 10 (Suppl.):105–123, 2002.

- J. Montmain. Interprétation de la contribution d'un critère dans le contexte d'une amélioration multicritère basée sur une intégrale de Choquet. In *Logique Floue et Applications (LFA)*, Nîmes, France, 2007.
- H. Moulin. *Axioms of cooperative decision making*. Wiley, 1988.
- T. Murofushi. Lexicographic use of Sugeno integrals and monotonicity conditions. *IEEE Tr. on Fuzzy Systems*, 9(6):783–794, 2001.
- T. Murofushi. A technique for reading fuzzy measures (I): the Shapley value with respect to a fuzzy measure. In *2nd Fuzzy Workshop*, pages 39–48, Nagaoka, Japan, October 1992. In Japanese.
- T. Murofushi and S. Soneda. Techniques for reading fuzzy measures (III): interaction index. In *9th Fuzzy System Symposium*, pages 693–696, Sapporo, Japan, May 1993. In Japanese.
- T. Murofushi and M. Sugeno. Fuzzy measures and fuzzy integrals. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integrals — Theory and Applications*, pages 3–41. Physica Verlag, 2000.
- W. Näther and K. Wälder. Applying fuzzy measures for considering interaction effects in root dispersal models. *Fuzzy Sets & Systems*, 158:572–582, 2007.
- J.P. Pignon and Ch. Labreuche. A methodological approach for operational and technical experimentation based evaluation of systems of systems architectures. In *Int. Conference on Software & Systems Engineering and their Applications (ICSSEA)*, Paris, France, December 4–6 2007.
- A. Rico, M. Grabisch, Ch. Labreuche, and A. Chateaufneuf. Preference modelling on totally ordered sets by the Sugeno integral. *Discrete Applied Mathematics*, 147:113–124, 2005.
- F. S. Roberts. *Measurement Theory*. Addison-Wesley, 1979.
- M. Roubens. Ordinal multiattribute sorting and ordering in the presence of interacting points of view. In D. Bouyssou, E. Jacquet-Lagrèze, P. Perny, R. Slowinsky, D. Vanderpooten, and Ph. Vincke, editors, *Aiding Decisions with Multiple Criteria: Essays in Honor of Bernard Roy*, pages 229–246. Kluwer Academic Publishers, 2001.
- K. Saito, Y. Watanabe, H. Hashimoto, and A. Uchiyama. An application of fuzzy integral model for the clinical diagnosis. *J. of Biomedical Fuzzy Systems Association*, 1(1): 17–24, 2007.
- D. Schmeidler. Integral representation without additivity. *Proc. of the Amer. Math. Soc.*, 97(2):255–261, 1986.
- D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989.
- L. S. Shapley. A value for  $n$ -person games. In H. W. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games, Vol. II*, number 28 in Annals of Mathematics Studies, pages 307–317. Princeton University Press, 1953.
- H. Simon. Rational choice and the structure of the environment. *Psychological Review*, 63(2):129–138, 1956.
- I. Singer. Extensions of functions of 0-1 variables and applications to combinatorial optimization. *Numerical Functional Analysis and Optimization*, 7(1):23–62, 1984.
- P. Slovic, M. Finucane, E. Peters, and D. G. MacGregor. The affect heuristic. In T. Gilovitch, D. Griffin, and D. Kahneman, editors, *Heuristics and biases: the psychology of intuitive judgment*, pages 397–420. Cambridge University Press, 2002.
- R. Słowiński, S. Greco, and B. Matarazzo. Axiomatization of utility, outranking and decision-rule preference models for multiple-criteria classification problems under partial inconsistency with the dominance principle. *Control and Cybernetics*, 31(4):1005–1035, 2002.
- M. Sugeno. Construction of fuzzy measures and evaluation of similarity patterns by fuzzy integral. *Trans. S.I.C.E. (Keisoku Jidōseigyō Gakkai)*, 9:361–368, 1973. In Japanese.
- M. Sugeno. *Theory of fuzzy integrals and its applications*. PhD thesis, Tokyo Institute of Technology, 1974.
- M. Sugeno. Fuzzy measures and fuzzy integrals — a survey. In Gupta, Saridis, and Gaines, editors, *Fuzzy Automata and Decision Processes*, pages 89–102. North Holland, 1977.
- M. Sugeno and S. H. Kwon. A clusterwise regression-type model for subjective evaluation. *J. of Japan Society for Fuzzy Theory and Systems*, 7(2):291–310, 1995.

- A. Tanaka and T. Murofushi. A learning model using fuzzy measures and the Choquet integral. In *5th Fuzzy System Symposium*, pages 213–218, Kobe, Japan, June 1989. In Japanese.
- V. Torra and Y. Narukawa. The h-index and the number of citations: two fuzzy integrals. *IEEE Trans. on Fuzzy Systems*, to appear.
- A. Tversky and D. Kahneman. Advances in prospect theory: cumulative representation of uncertainty. *J. of Risk and Uncertainty*, 5:297–323, 1992.
- J. Šipoš. Integral with respect to a pre-measure. *Math. Slovaca*, 29:141–155, 1979.
- Z. Wang, K. S. Leung, and J. Wang. A genetic algorithm for determining nonadditive set functions in information fusion. *Fuzzy Sets and Systems*, 102:462–469, 1999.
- L. Wendling and S. Tabbone. Recognition of arrows in line drawings based on the aggregation of geometric criteria using the choquet integral. In *Seventh International Conference on Document Analysis and Recognition - ICDAR2003*, pages 299–303, Edinburgh, UK, August 2003.
- R. R. Yager. On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Trans. Systems, Man & Cybern.*, 18:183–190, 1988.