Elicitation of 2-additive bi-capacity parameters

Antoine Rolland, Julien Ah-Pine & Brice Mayag

EURO Journal on Decision Processes

ISSN 2193-9438

EURO J Decis Process
DOI 10.1007/s40070-015-0043-3

EURO J Decis Process
DOI 10.1007/s40070-015-0043-3

ORIGINAL ARTICLE

Elicitation of 2-additive bi-capacity parameters

Antoine Rolland[1] · Julien Ah-Pine[1] · Brice Mayag[2]

Received: 15 December 2013 / Accepted: 15 April 2015

Abstract In some multi-criteria decision-making problems, it is more convenient to express the decision maker (DM) preferences in bipolar scales. In such cases, the bipolar Choquet integral with respect to bi-capacities was introduced as a versatile tool to model these kind of preferences. However, this aggregation function is useful in practice only if its parameters can be set up easily. To this end, elicitation techniques aim at finding the parameters values that best fit some given examples. In this paper, we address the problem of eliciting a bipolar Choquet integral with respect to a 2-additive bi-capacity. We present several techniques based on solving an optimization problem, taking into account the possible interaction, or not, with the DM. We deal with possible inconsistencies in the observed preferences and we also discuss the parsimonious character of the different models to favor simple models when several solutions exist.

Keywords 2-Additive bi-capacity identification · Bipolar Choquet integral · Preference elicitation

✉ Antoine Rolland
antoine.rolland@univ-lyon2.fr

Julien Ah-Pine
julien.ah-pine@univ-lyon2.fr

Brice Mayag
brice.mayag@dauphine.fr

1 ERIC Lab, University of Lyon 2, Bron, France

2 LAMSADE, University Paris-Dauphine, Paris, France

Published online: 07 May 2015

1 Introduction

Multi-criteria decision making (MCDM) aims at representing the preferences of a decision maker (DM) over a set of options (or alternatives) $X = X_1 \times \cdots \times X_n$, in regard to a finite set of criteria $N = \{1, \ldots, n\}$ or attributes $X_1, \ldots, X_n$. It then seeks to formalize the DM's decision process through mathematical tools in order to help her to make decisions over X. The DM's decision process is assumed to be guided by the importance and the relationships she wants to take into account regarding the criteria. Concerning the preferences representations, one possible model is the Multi-Attribute Utility Theory (MAUT) which assumes that each attribute X_i (or criterion $i \in N$) provides a utility value (or score) over the set of alternatives $u_i : X_i \rightarrow \mathbb{R},\ i \in N$. Then, an aggregation function is used to combine, for each option, its scores distribution (or profile) in an overall score. The latter global utility values are then employed to make decisions. There are many types of aggregation functions to model a decision process. The Choquet integral has been proved to be a versatile tool to construct overall scores (see for example Choquet 1954; Grabisch 1997; Grabisch and Labreuche 2005c). This aggregation function is intimately based on the concept of a capacity (or fuzzy measure). In particular, it assumes that partial utilities belong to non-negative or unipolar scales.

Unipolar scales are not always appropriate to represent the DM's preferences (see the motivating example in Grabisch and Labreuche 2008). In some problems, bipolar scales are more convenient. These type of scales are typically composed of a negative, a positive and a neutral part which respectively allow representing a negative, a positive and a neutral affect towards an option. The bipolar Choquet integral (BCI) was introduced in Grabisch and Labreuche (2002) and Greco et al. (2002) to apply the Choquet integral in the case of bipolar scales. In this paper, we particularly focus on BCI which uses the concept of a bi-capacity (BC) introduced in Grabisch and Labreuche (2002) and which was further studied in Grabisch and Labreuche (2005a, b), Fujimoto (2004) and Fujimoto and Murofushi (2005). The BCI typically requires the DM to set $3^n - 3$ values where n is the number of attributes. This cardinal corresponds to 3^n possible disjoint subsets minus 3 normalized parameters. When n exceeds some units, it is impossible for the DM to set all parameters of her decision model. To better cope with this combinatorial burden, the BCI with respect to (w.r.t.) a 2-additive bi-capacity (2A-BC) was introduced Fujimoto (2004) and Grabisch and Labreuche (2003). The 2-additivity property implies that only subsets of at most two elements are considered. This property enables to reduce the number of parameters from $3^n - 3$ to $2n^2$. Hence the bipolar Choquet integral w.r.t. a 2-additive bi-capacity appears more useful in practice.

Even though there have been many papers studying bi-capacities (Fujimoto et al. 2007; Grabisch and Labreuche 2005a, b, c; Greco et al. 2002; Greco and Rindone 2013, 2014), most of them have focused on theoretical aspects. In this contribution, we study the practical problem of identifying a 2-additive bi-capacity on the basis of information provided by the DM. This problem is also known as preference elicitation, which is an important step of a decision process. The DM gives her preferences represented by a binary relation $\succsim$ over X, where $\succ$ (respectively $\sim$) is the asymmetric (respectively the symmetric) part of $\succsim$. There are different contexts in which we can proceed to the elicitation of the preference model of a DM. In our case, we assume that the DM provides the bipolar scores for a subset of (real or fictitious) options w.r.t. all criteria of the decision problem. In addition, she provides the overall bipolar scores of the same set of alternatives. These evaluated examples constitute the only information we have at our disposal. Then, the elicitation model consists in inferring from this information, the parameters of a 2A-BC such that the associated BCI is consistent with the given preferences of the DM. We propose optimization models to tackle this kind of preference elicitation problem.

Eliciting preference models is a research topic that has been studied by many researchers (see for example Grabisch et al. 2008; Jacquet-Lagreze and Siskos 1982). However, the BCI w.r.t. 2A-BC has not been studied very much so far. The elicitation process we deal with has many relationships with the problems addressed in statistical machine learning. The interconnections between preference elicitation on the one hand and machine learning on the other hand were highlighted in Waegeman et al. (2009). There has been a growing interest for the last years about cross-fertilizing these two domains by studying how the concepts developed in one field can be applied in the other one. For example, the research works described in Fallah Tehrani et al. (2012), Hullermeier and Fallah Tehrani (2012) and Hullermeier et al. (2012) investigated the use of Choquet integral in classification tasks.

Elicitation processes for bi-capacities have been introduced in Mayag et al. (2012) with ternary alternatives and in Ah-Pine et al. (2013) with evaluated examples provided by the DM. This paper aims at extending these two approaches into a more global point of view. Note that the use of a BCI for the PROMETHEE method was already proposed in Greco and Figueira (2003) and an elicitation process was framed in Corrente et al. (2014) as well. However, the latter research works are specific to the PROMETHEE method as it includes specific features like bi-capacity symmetry. We are not concerned with such aspects so our proposal is different from these papers. In the next section, after recalling the basic definitions, we motivate, by some examples, the need for an elicitation process based on bi-capacities. In Sect. 3, we present a formal process to elicit the parameters of a 2A-BC that generalizes the ones introduced in Mayag et al. (2012) and Ah-Pine et al. (2013). In the last section, we study the parsimonious aspect of the proposed elicitation process and we present a real-life example.

2 Bi-capacities and bipolar Choquet integral

2.1 Choquet integral and its limitations

The Choquet integral (Grabisch and Labreuche 2005c) is an aggregation function known as the generalization of the arithmetic mean. One of its most notable properties is to take into account interactions between criteria. This function is based on the notion of capacity μ defined as a set function from the power set of criteria 2^N to $[0, 1]$ such that:

1. $\mu(\emptyset) = 0$
2. $\mu(N) = 1$
3. $\forall A, B \in 2^N, [A \subseteq B \Rightarrow \mu(A) \leq \mu(B)]$ (monotonicity).

The Choquet integral of a vector $(y_1, \ldots, y_n) \in \mathbb{R}_+^n$ w.r.t. a capacity μ is defined by:

$$C_\mu(y_1, \ldots, y_n) := \sum_{i=1}^{n} \left(y_{\tau(i)} - y_{\tau(i-1)}\right) \mu(\{\tau(i), \ldots, \tau(n)\}) \tag{1}$$

where τ is a permutation on N such that

$$y_{\tau(1)} \leq y_{\tau(2)} \leq \cdots \leq y_{\tau(n-1)} \leq y_{\tau(n)}, \quad \text{and} \quad y_{\tau(0)} := 0.$$

The preferences of a DM cannot always be represented by a Choquet integral. The following examples allow us to illustrate this point:

Example 1 The limitation of Choquet integral: a classical example when utility functions are fixed.

A classical example that shows the limitation of the Choquet integral model is given in Grabisch and Labreuche (2005c): the students of a faculty are evaluated on three subjects Mathematics (*M*), Statistics (*S*) and Language skills (*L*). Each course can be represented, in a natural way, on a bipolar scale where the neutral level is the aspiration level of the dean of the faculty. Hence we assume that all marks are taken from the same scale from 0 to 20 with the mark 10 corresponding to the neutral level. The evaluations of eight students are given by the table below:

	1: Mathematics (*M*)	2: Statistics (*S*)	3: Language (*L*)
A	16	13	7
B	16	11	9
C	6	13	7
D	6	11	9
E	14	16	7
F	14	15	8
G	9	16	7
H	9	15	8

To select the best students, the dean expresses his preferences:

– For a student good in Mathematics, Language is more important than Statistics

$$\Longrightarrow A \prec B \quad \text{and} \quad \mathrm{E} \prec \mathrm{F},$$

– For a student bad in Mathematics, Statistics is more important than Language

$$\Longrightarrow D \prec C \quad \text{and} \quad \mathrm{H} \prec \mathrm{G}.$$

The two preferences $A \prec B$ and $D \prec C$ lead to a contradiction with the arithmetic mean model because

$$\begin{cases} A \prec B \Rightarrow 16\,w_M + 13\,w_S + 7\,w_L < 16\,w_M + 11\,w_S + 9\,w_L \\ \\ D \prec C \Rightarrow 6\,w_M + 11\,w_S + 9\,w_L < 6\,w_M + 13\,w_S + 7\,w_L. \end{cases}$$

Furthermore it is not difficult to see that the other two preferences, $E \prec F$ and $H \prec G$, cannot be modeled by a Choquet integral C_μ since

$$\begin{cases} E \prec F \Rightarrow 7 + 7\mu(\{M,S\}) + 2\mu(\{S\}) < 8 + 6\mu(\{M,S\}) + \mu(\{S\}) \\ \\ H \prec G \Rightarrow 8 + \mu(\{M,S\}) + 6\mu(\{S\}) < 7 + 2\mu(\{M,S\}) + 7\mu(\{S\}) \end{cases}$$

$$\text{i.e.} \begin{cases} E \prec F \Rightarrow \mu(\{M,S\}) + \mu(\{S\}) < 1 \\ \\ H \prec G \Rightarrow \mu(\{M,S\}) + \mu(\{S\}) > 1 \end{cases}$$

Example 1 shows a decision situation where the strategies of the DM depend on some evaluations on criteria that are judged as good or bad. Therefore, the DM's decision behavior can be interpreted as a bipolar behavior (Grabisch et al. 2008).

Example 2 The limitation of Choquet integral: an example when utility functions are not fixed *a priori*.

Let $X_1 = \{a_1, b_1, c_1, d_1, e_1, f_1\}$ and $X_2 = \{a_2, b_2, c_2, d_2, e_2, f_2\}$. Suppose that the relation $\succsim$ is such that:

$$\begin{aligned} (a_1, e_2) &\sim (b_1, d_2) \\ (c_1, d_2) &\sim (a_1, f_2) \\ (c_1, e_2) &\not\sim (b_1, f_2) \end{aligned} \tag{2}$$

and

$$\begin{aligned} (d_1, b_2) &\sim (e_1, a_2) \\ (f_1, a_2) &\sim (d_1, c_2) \\ (f_1, b_2) &\not\sim (e_1, c_2) \end{aligned} \tag{3}$$

If we assume that

$$\begin{aligned} u_1(a_1) \le u_1(b_1) \le u_1(c_1) \le u_1(d_1) \le u_1(e_1) \le u_1(f_1) \\ u_2(a_2) \le u_2(b_2) \le u_2(c_2) \le u_2(d_2) \le u_2(e_2) \le u_2(f_2) \end{aligned} \tag{4}$$

then it is proved in Bouyssou et al. (2012) that $\succsim$ cannot be represented by a Choquet integral.

Compared to Example 1 where utility functions are fixed, this example shows that it is possible to find some preferences which cannot be modeled by a Choquet integral whatever the utility functions.

2.2 Bi-capacities and bipolar Choquet Integral

In the previous examples, the Choquet integral based on capacities fails at modeling preferences expressed in bipolar scales. A solution to this problem is to use a model related to this type of scale, namely the bipolar Choquet integral. This aggregation function is based on the concept of bi-capacities which is a set function defined over $3^N := \{(A, B) \in 2^N \times 2^N | A \cap B = \emptyset\}$, the set of couples of subsets of N with an empty intersection.

Definition 1 (*Bi-capacity* (*BC*) Grabisch and Labreuche 2005b, 2008) A function $v : 3^N \to \mathbb{R}$ is a BC on 3^N if it satisfies the following two conditions:

$$v(\emptyset, \emptyset) = 0 \tag{5}$$

$$\forall (A_1, A_2), (B_1, B_2) \in 3^N : [(A_1, A_2) \sqsubseteq (B_1, B_2) \Rightarrow v(A_1, A_2) \leq v(B_1, B_2)] \tag{6}$$

Where $(A_1, A_2) \sqsubseteq (B_1, B_2) \Leftrightarrow [A_1 \subseteq B_1 \quad \text{and} \quad \mathrm{B}_2 \subseteq \mathrm{A}_2]$.

The condition (6) is a monotonicity condition. In addition, a BC is said to be normalized if it satisfies :

$$v(N, \emptyset) = 1 \quad \text{and} \quad v(\emptyset, \mathrm{N}) = -1 \tag{7}$$

Besides, a BC is said to be additive if it holds:

$$\forall (A_1, A_2) \in 3^N : v(A_1, A_2) = \sum_{i \in A_1} v(\{i\}, \emptyset) + \sum_{\{j\} \in A_2} v(\emptyset, \{j\}) \tag{8}$$

An additive BC assumes that the attributes are independent from each other and this situation boils down to linear decision models.

To better formalize some of the properties of BC, the following definition of a (bipolar) Möbius (Fujimoto 2004; Fujimoto et al. 2007) transform[1] of a BC was proposed.

Definition 2 (*Bipolar Möbius transform of a bi-capacity* Fujimoto 2004; Fujimoto et al. 2007) Let v be a BC on 3^N. The bipolar Möbius transform of v is a set function $b : 3^N \to \mathbb{R}$ defined for any $(A_1, A_2) \in 3^N$ by:

[1] Grabisch and Labreuche (Grabisch and Labreuche 2003, 2005a, b) proposed a definition of the Möbius transform of a BC different to the one given in Fujimoto (2004). However, there is a one-to-one correspondence between the two Möbius transform definitions. This equivalence was established in Fujimoto and Murofushi (2005).

$$\begin{aligned} b(A_1,A_2) &:= \sum_{\substack{B_1\subseteq A_1\\ B_2\subseteq A_2}} (-1)^{|A_1\setminus B_1|+|A_2\setminus B_2|} v(B_1,B_2) \\ &= \sum_{(\emptyset,A_2)\sqsubseteq(B_1,B_2)\sqsubseteq(A_1,\emptyset)} (-1)^{|A_1\setminus B_1|+|A_2\setminus B_2|} v(B_1,B_2) \end{aligned} \tag{9}$$

Conversely, for any $(A_1,A_2) \in 3^N$, it holds that :

$$v(A_1,A_2) := \sum_{\substack{B_1\subseteq A_1\\ B_2\subseteq A_2}} b(B_1,B_2). \tag{10}$$

Note that using b, condition (5) is equivalent to:

$$b(\emptyset,\emptyset) = 0 \tag{11}$$

BC on 3^N generally require $3^n - 3$ parameters. To reduce this number, Grabisch and Labreuche (2005a, b, 2008) proposed the concept of k-additivity of a BC. This concept translates as follows in terms of the bipolar Möbius transform.

Definition 3 (Fujimoto et al. 2007) Given a positive integer $k<n$, a BC v is k-additive if and only if the two following conditions are satisfied:

$$\forall (A_1,A_2) \in 3^N : |A_1 \cup A_2| > k \Rightarrow b(A_1,A_2) = 0 \tag{12}$$

$$\exists (A_1,A_2) \in 3^N : |A_1 \cup A_2| = k \wedge b(A_1,A_2) \neq 0 \tag{13}$$

Note that if condition (13) is omitted, then the BC is said to be "at most k"-additive. For the sake of simplicity, we will suppose in the elicitation process that we are seeking for "at most 2"-additive BC and not "2-additive" BC strictly speaking, and we will use abusively the notation 2A-BC for "at most 2"-additive bi-capacities.

To avoid a heavy notation, we use the following shorthands for all $i,j \in N$, $i \neq j$:

- $v_{i|} := v(\{i\},\emptyset)$, $v_{|j} := v(\emptyset,\{j\})$, $v_{i|j} := v(\{i\},\{j\})$, $v_{ij|} := v(\{i,j\},\emptyset)$, $v_{|ij} := v(\emptyset,\{i,j\})$,
- $b_{i|} := b(\{i\},\emptyset)$, $b_{|j} := b(\emptyset,\{j\})$, $b_{i|j} := b(\{i\},\{j\})$, $b_{ij|} := b(\{i,j\},\emptyset)$, $b_{|ij} := b(\emptyset,\{i,j\})$.

Whenever we use i and j together, it always means that they are different.

Using the above definitions, we propose the following properties of a 2A-BC v and its bipolar Möbius transform b:

Proposition 1

1. Let v be a 2A-BC and b its bipolar Möbius transform. For any $(A_1, A_2) \in 3^N$ we have:

$$v(A_1, A_2) = \sum_{i \in A_1} b_{i|} + \sum_{j \in A_2} b_{|j} + \sum_{\substack{i \in A_1 \\ j \in A_2}} b_{i|j} + \sum_{\{i,j\} \subseteq A_1} b_{ij|} + \sum_{\{i,j\} \subseteq A_2} b_{|ij} \tag{14}$$

2. If the coefficients $b_{i|}$, $b_{|j}$, $b_{i|j}$, $b_{ij|}$, $b_{|ij}$ are given for all $i, j \in N$, then the necessary and sufficient conditions to get a 2A-BC generated by (14) are:

$$\forall (A, B) \in 3^N, \forall k \in A : b_{k|} + \sum_{j \in B} b_{k|j} + \sum_{i \in A \setminus k} b_{ik|} \geq 0 \tag{15}$$

$$\forall (A, B) \in 3^N, \forall k \in A : b_{|k} + \sum_{j \in B} b_{j|k} + \sum_{i \in A \setminus k} b_{|ik} \leq 0 \tag{16}$$

3. The inequalities (15) and (16) can be respectively reformulated in terms of the BC v as follows:

$$\forall (A, B) \in 3^N, \forall k \in A : \sum_{j \in B} v_{k|j} + \sum_{i \in A \setminus k} v_{ik|} \geq (|B| + |A| - 2) v_{k|} + \sum_{j \in B} v_{|j} + \sum_{i \in A \setminus k} v_{i|}$$

$$\forall (A, B) \in 3^N, \forall k \in A : \sum_{j \in B} v_{j|k} + \sum_{i \in A \setminus k} v_{|ik} \leq (|B| + |A| - 2) v_{|k} + \sum_{j \in B} v_{j|} + \sum_{i \in A \setminus k} v_{|i}$$

Proof (Sketch of)

1. Because v is 2-additive, the proof of (14) is given by using the relation (10) between v and b.
2. The proof of the second point is based on the expression of $v(A_1, A_2)$ given in (14) and on these equivalent monotonicity properties (which are easy to check): $\forall (A, B) \in 3^N$ and $\forall A \subseteq A'$,
   (a) $v(A, B) \leq v(A', B) \Leftrightarrow \{\forall k \in A : v(A \setminus k, B) \leq v(A, B)\}$;
   (b) $v(B, A') \leq v(B, A) \Leftrightarrow \{\forall k \in A : v(B, A) \leq v(B, A \setminus k)\}$.
3. These inequalities are obtained departing from (15) and (16) and by using the relation (10) between v and b.

□

Hence, according to Proposition (1) and (14), the computation of a 2A-BC v only requires the values of b on the elements $(i, \emptyset)$, $(\emptyset, i)$, (i, j), $(ij, \emptyset)$, $(\emptyset, ij)$, $\forall i, j \in N$. However, to satisfy the monotonicity condition given in (6), a 2A-BC should also satisfy the inequalities (15) and (16). In addition, we can add the following normalized conditions:

$$v_{N|} = \sum_{i \in N} b_{i|} + \sum_{\{i,j\} \subseteq N} b_{ij|} = 1 \text{ and } v_{|N} = \sum_{i \in N} b_{|i} + \sum_{\{i,j\} \subseteq N} b_{|ij} = -1 \tag{17}$$

Example 3 The sign $\checkmark$ in the following matrix corresponds to an element of 3^N which is required to compute a 2A-BC for $N = \{1, 2, 3\}$.

$$\begin{array}{cc}
(A,B) & \begin{array}{cccccccc} \emptyset & 1 & 2 & 3 & 12 & 13 & 23 & 123 \end{array} \\
\begin{array}{c} \emptyset \\ 1 \\ 2 \\ 3 \\ 12 \\ 13 \\ 23 \\ 123 \end{array} &
\begin{pmatrix}
\checkmark & \checkmark & \checkmark & \checkmark & \checkmark & \checkmark & \checkmark & . \\
\checkmark & . & \checkmark & \checkmark & . & . & . & . \\
\checkmark & \checkmark & . & \checkmark & . & . & . & . \\
\checkmark & \checkmark & \checkmark & . & . & . & . & . \\
\checkmark & . & . & . & . & . & . & . \\
\checkmark & . & . & . & . & . & . & . \\
\checkmark & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & .
\end{pmatrix}
\end{array}$$

We now define the Choquet integral based on bi-capacities as follows:

Definition 4 (*Bipolar Choquet integral* (*BCI*) (*w.r.t. a BC*) Grabisch and Labreuche 2005b) Let v be a BC on 3^N and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. The expression of the BCI of x w.r.t. v is given by

$$\mathcal{C}_v(x) := \sum_{i=1}^{n} |x_{\sigma(i)}| \Big[v\big(N_{\sigma(i)} \cap N^+, N_{\sigma(i)} \cap N^-\big) - v\big(N_{\sigma(i+1)} \cap N^+, N_{\sigma(i+1)} \cap N^-\big) \Big] \tag{18}$$

where $N^+ = \{i \in N | x_i \geq 0\}$, $N^- = N \setminus N^+$, $N_{\sigma(i)} := \{\sigma(i), \ldots, \sigma(n)\}$ and σ is a permutation on N such that $|x_{\sigma(1)}| \leq |x_{\sigma(2)}| \leq \cdots \leq |x_{\sigma(n)}|$.

We also have the following equivalent expression of the BCI w.r.t. b, given by Fujimoto and Murofushi (2005):

$$\mathcal{C}_b(x) = \sum_{(A_1, A_2) \in 3^N} b(A_1, A_2) \left(\bigwedge_{i \in A_1} x_i^+ \wedge \bigwedge_{j \in A_2} x_j^- \right) \tag{19}$$

where $\begin{cases} x_i^+ = x_i & \text{if} \quad \mathrm{x_i} > 0 \\ x_i^+ = 0 & \text{if} \quad \mathrm{x_i} \leq 0 \end{cases}$ and $\begin{cases} x_i^- = -x_i & \text{if} \quad \mathrm{x_i} < 0 \\ x_i^- = 0 & \text{if} \quad \mathrm{x_i} \geq 0 \end{cases}$.

Note that $\mathcal{C}_v(x) = \mathcal{C}_b(x)$ and the subscript is meant to clarify whether it is v or b which is used in the computation. Besides, the BCI of x w.r.t. a 2A-BC represented by b reduces to:

$$\begin{aligned}\mathcal{C}_b(x) = &\sum_{i=1}^{n} b_{i|}\, x_i^+ + \sum_{i=1}^{n} b_{|i}\, x_i^- + \sum_{i,j=1}^{n} b_{i|j}\left(x_i^+ \wedge x_j^-\right) \\ &+ \sum_{\{i,j\}\subseteq N} b_{ij|}\left(x_i^+ \wedge x_j^+\right) + \sum_{\{i,j\}\subseteq N} b_{|ij}\left(x_i^- \wedge x_j^-\right)\end{aligned} \tag{20}$$

Example 4 In Example 2, we saw that preferences (2) and (3) cannot be recovered by the Choquet integral w.r.t. a capacity. In contrast, these relations can now be modeled by a Choquet integral w.r.t. a bi-capacity v given by its Möbius transform b:

$$\begin{aligned} b_{1|} &= 2 & b_{|1} &= -2 \\ b_{2|} &= 3 & b_{|2} &= -3 \\ b_{1|2} &= -1 & b_{2|1} &= 1 \\ b_{12|} &= 5 & b_{|12} &= -5. \end{aligned}$$

It is sufficient to choose the utility functions u_1 and u_2 as follows:

$$\begin{aligned} u_1(a_1) &= -1 & u_2(a_2) &= 0.125 \\ u_1(b_1) &= 0 & u_2(b_2) &= 0.2 \\ u_1(c_1) &= 0.8/7 & u_2(c_2) &= 0.3 \\ u_1(d_1) &= 1/7 & u_2(d_2) &= 0.4 \\ u_1(e_1) &= 0.3 & u_2(e_2) &= 0.8 \\ u_1(f_1) &= 0.45 & u_2(f_2) &= 1. \end{aligned}$$

and then we obtain

$$\begin{aligned} \mathcal{C}_v(a_1, e_2) &= 1.2 & \mathcal{C}_v(b_1, d_2) &= 1.2 \\ \mathcal{C}_v(c_1, d_2) &= 2 & \mathcal{C}_v(a_1, f_2) &= 2 \\ \mathcal{C}_v(c_1, e_2) &= 3.2 & \mathcal{C}_v(b_1, f_2) &= 3 \\ \mathcal{C}_v(d_1, b_2) &= 1.6 & \mathcal{C}_v(e_1, a_2) &= 1.6 \\ \mathcal{C}_v(f_1, a_2) &= 1.9 & \mathcal{C}_v(d_1, c_2) &= 1.9 \\ \mathcal{C}_v(f_1, b_2) &= 2.5 & \mathcal{C}_v(e_1, c_2) &= 3. \end{aligned}$$

3 Identifying a 2-additive bi-capacity

The use of a BCI based on 2A-BC as a decision-aiding tool requires to determine the parameters of the model before computing the score of each considered solution. This step is called the "elicitation process". In MCDM one can generally consider two types of paradigms for elicitation processes: direct methods and indirect methods. In a direct method, the DM is supposed to be able to directly provide the value of each parameter. On one hand, it means that the parameters are not so numerous. On another hand, it also implicitly means that the DM has understood the

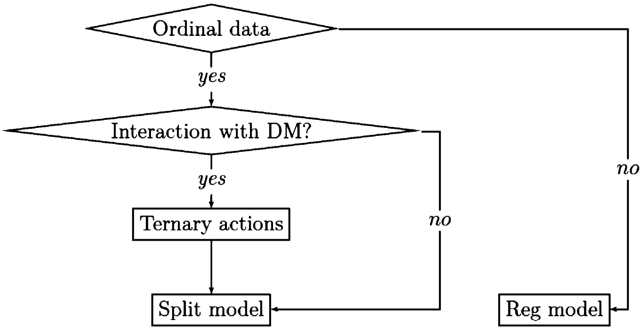

Fig. 1 Choice of adequate elicitation method

model totally so that she knows how to control the effect of each parameter on the model outcomes to recover her preferences. In the case of a BCI, it is obvious that these two conditions fail since a BC typically requires $3^n - 3$ values to be set. To reduce this complexity, 2A-BC was introduced. Nevertheless, this latter case cannot be applied in practice either. Even if the number of parameters reduces to $2n^2$, it remains very high even when n is low. Moreover, a BCI is a too complex aggregation operator to ensure that a DM will understand the influence of each parameter on the final result. Even with simpler aggregation rules, it has been shown that there is no clear link between the parameter values provided by the DM and the way these values are used in the decision model (see Bouyssou et al. 2006).

In the indirect methods, the DM does not give information about her decision model but she provides information on the outputs of her decision strategy. Then, given these judged examples, we have to infer the parameter values of the DM decision model which we assume to be based on a BCI w.r.t. a 2A-BC. The estimated BCI should predict overall scores, $\mathcal{C}_b(x)$, that are consistent with the preference relations provided by the DM. In other words, if $x \succsim x'$, which means that x is preferred or equivalent to x', then the inferred decision model should satisfy $\mathcal{C}_b(x) \geq \mathcal{C}_b(x')$.

Elicitation process

The key point is the type of information provided by the DM: does she give a precise global score to each alternative, or does she only provide a (complete or not) ordinal preference relation over the set of alternatives? As presented in Fig. 1, the available pieces of information will determine the method to be used. As we shall see, we can make the distinction between two situations:

– If only a preference relation over the alternatives is known, then a linear program can be used to determine whether there exists a bi-capacity compatible with the given preference relation or not. Note that if it is possible to interact with the DM to know her preferences on specific alternatives, then the simpler approach is to use some specific fictitious alternatives named ternary alternatives.

Table 1 Number of constraints due to the use of a BC or a 2A-BC depending of the number of criteria

# Criteria	3	4	5	6	8	10
# Constraints (2A-BC)	48	208	800	2904	34976	393260
# Constraints (BC)	117	609	3093	15561	390369	9764601

– If a global score $S(x)$ is provided for each alternative x, then we frame other optimization problems that we can solve by means of quadratic programming. Note that in this case, the alternatives could be real or fictitious, but no interaction with the DM is needed.

We detail the elicitation process with ternary actions in Sect. 3.1, with only ordinal information in Sect. 3.2 and with cardinal information in Sect. 3.3

Constraints due to the 2-additivity

We argued that the use of a 2A-BC instead of a general BC is preferable to alleviate the complexity of the model. In the same time, it reduces the complexity for the user as the number of parameters to set falls from $3^n - 3$ to $2n^2$. Furthermore, the number of constraints decrease as well, thanks to the monotonicity property: in the case of a 2A-BC this number equals

$$2\left(\sum_{i=1}^{n} i \times \binom{n}{i} \times \left(\sum_{j=0}^{n-i}\binom{n-i}{j}\right) - n\right)$$

whereas in the case of a BC it is

$$\sum_{i=0}^{n}\sum_{j=0}^{n-i}\left[\binom{n}{i}\binom{n-i}{j}\left(\sum_{k=0}^{i}\binom{i}{k} \times \sum_{k=0}^{j}\binom{j}{k}\right)\right]$$

However, in both cases, the number of constraints exponentially increases with respect to the number of criteria as shown in Table 1.

3.1 Elicitation process with ternary alternatives

The elicitation process of a 2A-BC with ternary alternatives was introduced in Mayag et al. (2012). We present below the basic foundations of the elicitation process. It consists of four steps:

1. Determine with the help of the DM three reference levels:
   (a) A reference level $\mathbf{1_i}$ in X_i which she considers as good and completely satisfying if she could reach 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).
   (b) A reference level $\mathbf{0_i}$ in X_i which she considers neutral on i. The neutral level is the absence of attractiveness and repulsiveness. The existence of

this neutral level has roots in psychology (Slovic et al. 2002), and is used in bipolar models (Tversky and Kahneman 1992).

(c) A reference level $-\mathbf{1_i}$ in X_i which she considers completely unsatisfying.

2. Build a set of ternary alternatives, i.e. options where all criteria are set to the neutral level except at most two criteria which are set to the satisfactory or unsatisfactory level.
3. Ask the DM to give her preferences over pairs of ternary alternatives.
4. Determine the different parameters of the model through a linear program.

The aim of this linear program is to find a 2A-BC that is able to represent the preferences given by the DM, the objective function being $\mathrm{MinC}_v(\mathrm{x}_0)$, where x_0 is an arbitrary ternary alternative. The variables are the values $v_{i|}, v_{|j}, v_{i|j}, v_{ij|}, v_{|ij} \forall i,j \in N$, i.e. $1 + 2n^2$ variables. There are two types of constraints:

– Constraints due to the monotonicity and the 2-additivity of the 2A-BC :

$$\forall (A,B) \in 3^N, \forall k \in A \mathrm{such that}(|\mathrm{A}| + |\mathrm{B}| - 2) \geq 0$$
$$\sum_{j \in B} v_{k|j} + \sum_{i \in A \setminus k} v_{ik|} \geq (|B| + |A| - 2) v_{k|} + \sum_{j \in B} v_{|j} + \sum_{i \in A \setminus k} v_{i|} \tag{21}$$

$$\sum_{j \in B} v_{j|k} + \sum_{i \in A \setminus k} v_{|ik} \leq (|B| + |A| - 2) v_{|k} + \sum_{j \in B} v_{j|} + \sum_{i \in A \setminus k} v_{|i} \tag{22}$$

– Constraints due to the representation of the preference relation:

$$C_v(x) = C_v(y), \quad \forall (x,y), \quad x \sim y \tag{23}$$

$$d_{\min} \leq C_v(x) - C_v(y), \quad \forall (x,y), \quad x \succ y \tag{24}$$

where $d_{\min}$ is an arbitrary strictly positive constant.

The optimization problem considered above has $2n^2 + n + 1$ variables. The number of constraints due to the representation of the preference relation is equal to the number of observed preference and indifference relations. Note that, strictly speaking, the constraints due to the monotonicity just guarantee that the bi-capacity is at most 2-additive, i.e. it can be 2-additive or even only additive. We present in the following example a model using 2A-BC that represents the preferences presented in Example 1. This model is obtained using an elicitation process with ternary alternatives.

Example 5 In Example 1, the Choquet integral fails at representing the given preferences. Let us show how these preferences could be represented by the use of a bi-capacity model. Suppose here that we set the neutral level at 10, the satisfactory level at 20 and the unsatisfactory level at 0. The assessment rules are the following ones:

– For a student good in Mathematics, Language is more important than Statistics.
– For a student bad in Mathematics, Statistics is more important than Language.

They can be expressed as follows:

– $(1,0,1) \succ (1,1,0)$
– $(-1,1,0) \succ (-1,0,1)$

Then the constraints due to the representation of the preference relation are:

– $v_{13|} > v_{12|}$
– $v_{2|1} > v_{3|1}$

For example the following bi-capacity can express the above preferences: $v_{\emptyset} = 0$, $v_{1|} = 0$, $v_{2|} = 0$, $v_{3|} = 0$, $v_{|\mathbf{1}} = -\mathbf{0.5}$, $v_{|2} = 0$, $v_{|3} = 0$, $v_{1|2} = 0$, $v_{1|3} = 0$, $v_{\mathbf{2}|\mathbf{1}} = -\mathbf{0.2}$, $v_{2|3} = 0$, $v_{\mathbf{3}|\mathbf{1}} = -\mathbf{0.4}$, $v_{3|2} = 0$, $v_{\mathbf{12}|} = \mathbf{0.6}$, $v_{\mathbf{13}|} = \mathbf{0.8}$, $v_{23|} = 0$, $v_{|\mathbf{12}} = -\mathbf{0.6}$, $v_{|\mathbf{13}} = -\mathbf{0.6}$, $v_{|23} = 0$, $v_{123|} = 1$, , $v_{|123} = -1$

3.2 Elicitation process with ordinal information

The elicitation process of a 2A-BC with real or fictitious alternatives has been introduced in Ah-Pine et al. (2013), inspired by methods detailed in Grabisch et al. (2008) in the case of the unipolar Choquet integral. In this section, we suppose that the DM gives for some examples $x \in X' \subseteq X$ their partial utilities for all criteria and a (complete or not) preference relation on the set of examples. Then, we assume that there is no further interaction with the DM. The aim of the elicitation process is to find a 2A-BC which correctly represents the observed preference relation. We use a linear program to reach this objective by extending the maximum split method introduced in Marichal and Roubens (2000). The objective function consists in maximizing the difference (split) $\mathcal{C}_b(x) - \mathcal{C}_b(x')$ for any $x \neq x' \in X'$ such that $x \succ x'$.

The variables are the values of the Möbius transform of the bi-capacity $b_{i|}$, $b_{|j}$, $b_{i|j}$, $b_{ij|}$, $b_{|ij} \forall i,j \in N$, i.e. $1 + 2n^2$ variables. We deal with the Möbius transform of the bi-capacity as it is easier to represent the 2-additivity property using the latter set function than with the corresponding bi-capacity.

There are three types of constraints:

– Constraints due to the monotonicity and the 2-additivity of the 2A-BC. To have a normalized 2A-BC in terms of b, we need to integrate the following relations in our optimization problems: (11), (12) with $k = 2$, (15), (16) and (17).
– Constraints due to the representation of the preference relation on the subset of examples X'. If $x \succsim x'$ then the BCI should be in concordance with this preference. Accordingly, we have the following second set of constraints:

$$\forall x, x' \in X', x \neq x' : x \succsim x' \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq \delta_c \tag{25}$$

where δ_c is a non-negative indifference threshold which is a parameter of the model.

– Constraints related to the computation of the BCI. Indeed, in (25) or (26), we need to calculate $\mathcal{C}_b(x)$ for each $x \in X'$. As a consequence, we need to add the constraints provided by (20) in our models. Note that despite the fact that the latter equations involve the minimum function, we can pre-compute the terms $(x_i^{\pm} \wedge x_j^{\pm})$ since they are parameters of the models. Consequently, the constraints (20) are linear equations.

Therefore, the linear program considered above has $2n^2 + n + 1$ variables, $2\left(\sum_{i=1}^{n} i \times \binom{n}{i} \times \left(\sum_{j=0}^{n-i} \binom{n-i}{j}\right) - n\right)$ constraints due to the monotonicity property and the 2-additivity of the 2A-BC and $|X'| \times (|X'| - 1)/2$ constraints due to the representation of the preference relation, where X' is the subset of observed alternatives.

3.3 Elicitation process with cardinal information

We suppose that the DM gives, for some examples $x \in X' \subseteq X$, not only their partial utilities for all criteria but also their overall scores $S(x)$. We assume that there is no further interaction with the DM. Then, the aim of the elicitation process is to find the 2A-BC that best fits the observed values.

Following a regression approach, the objective function seeks to minimize the sum of square errors between S and $\mathcal{C}_b$ which results in the following objective function: $\min \sum_{x,x' \in X'} (S(x) - \mathcal{C}_b(x))^2$. It leads to a quadratic program.

The variables are still the values of the Möbius transform of the bi-capacity $b_{i|}$, $b_{|j}$, $b_{i|j}$, $b_{ij|}$, $b_{|ij} \forall i,j \in N$, i.e. $1 + 2n^2$ variables. The constraints are the same as in Sect. 3.2.

3.4 Inconsistencies

Consistency means that the inferred parameters of the Choquet integral are such that $\mathcal{C}_b(x) \geq \mathcal{C}_b(x')$ whenever $x \succsim x'$. It might happen that this condition is not fulfilled for some pairs (x, x'). There are two main reasons for that: either the judgements provided by the DM herself are not consistent or the restriction of the decision model to 2A-BC does not allow one to fit the DM preferences correctly. In MCDM, inconsistencies are usually treated in an interaction loop with the DM (see Mayag et al. 2010 for example). It is assumed that the DM preferences can change to fix these inconsistencies when they are encountered. In our setting, the interaction loop is not permitted. Note that in the linear programs presented above, the set of constraints does not allow inconsistencies. Indeed, the inferred 2A-BC b could not be flexible enough to satisfy $\mathcal{C}_b(x) - \mathcal{C}_b(x') \geq \delta_c$ for some pairs (x, x'). In that case the optimization problem is infeasible.

Consequently, to cope with this issue, we propose a second version of our models which allows inconsistencies and attempts to infer a model that minimizes the errors due to such situations as much as possible. We transform the previous constraints as follows:

$$\forall x, x' \in X', x \neq x' : x \succsim x' \geq 0 \Rightarrow \mathcal{C}_b(x) - \mathcal{C}_b(x') \geq \delta_c - \xi_{xx'} \tag{26}$$

where $\xi_{xx'}$ are non-negative slack variables which allow inconsistencies. However, we want $\xi_{xx'}$ to be as low as possible and thus there should be a term in the objective function seeking to minimize $\sum_{x,x':x \succsim x'} \xi_{xx'}$ as well. Note that when the latter term is null, it means that the inferred model does not produce any inconsistency. On the contrary, if for some pairs (x, x'), $\xi_{xx'} > \delta_c$ then the optimal solution of the problem has not been able to satisfy the preference relations on these pairs.

3.5 Example

Let us consider again the situation introduced in Example 1. Suppose that the DM is able to give a global utility to each student such that these overall scores are consistent with the expressed preferences (see Table 2).

Let us consider the neutral level to be 10. Then we center all the scores with respect to 10 to obtain a bipolar scale (see Table 3).

Afterwards, we solve the optimization problem using the regression approach to obtain the Möbius coefficients (rounded to 10^{-2}), presented in Table 4.

Table 2 Data given by the DM

	1(*M*)	2(*S*)	3(*L*)	Score
A	16	13	7	13
B	16	11	9	14
C	6	13	7	8
D	6	11	9	7
E	14	16	7	11
F	14	15	8	12
G	9	16	7	9
H	9	15	8	10

Table 3 Data transformed to bipolar scale

	1(*M*)	2(*S*)	3(*L*)	Score
A	6	3	−3	3
B	6	1	−1	4
C	−4	3	−3	−2
D	−4	1	−1	−3
E	4	6	−3	1
F	4	5	−2	2
G	−1	6	−3	−1
H	−1	5	−2	0

Table 4 Möbius coefficients related to example 1

$b[*,*]$	$\emptyset$	M	S	MS	L	ML	SL	MSL
$\emptyset$	0	−0.88	0.2	0.08	−0.45	0.33	0.12	0
M	0.78	–	−0.57	–	−0.2	–	0	–
S	0.1	0.55	–	–	0	0	–	–
MS	-0.1	–	–	–	0	–	–	–
L	0.2	−0.2	0	0	–	–	–	–
ML	−0.2	–	0	–	.–	–	–	–
SL	0.22	0	–	–	–	–	–	–
MSL	0	–	–	–	–	–	–	–

Table 5 Result scores obtained by regression approach

	sc	*choq_val*
A	3	2.69
B	4	4
C	−2	−1.95
D	−3	−3
E	1	1.33
F	2	1.89
G	−1	−0.77
H	0	−0.41

Table 6 Result scores obtained by regression approach

	sc	*choq_val*
A	3	2.79
B	4	4
C	−2	−2.31
D	−3	−2.73
E	1	1.26
F	2	1.87
G	**−2.5**	**−1.58**
H	0	−1.05

The obtained scores are described in Table 5. We can see that, even if the obtained values are not exactly equal to the observed scores, the preference relation is preserved since $\forall x, x' \in X,\ x \succsim x' \Longleftrightarrow \mathcal{C}_b(x) \geq \mathcal{C}_b(x')$.

Suppose now that an inconsistency occurs in the observed scores given by the DM. It happens for example., if the score of alternative G is 7.5 (instead of s9). There is an inconsistency as G has a better value than C on each criterion, but G's global score is worse than C's global score. Then, we should use the regression approach with the constraint (26). The obtained values are detailed in Table 6. We can see that these scores are consistent with a dominance relation, but they fit the observed scores worse.

4 Parsimony and example

4.1 Parsimony

Parsimony is an interesting property when trying to elicit parameters of a given model: does the solution proposed by the elicitation program is as simple as possible? In our case, it means that a simple additive bi-capacity should be preferred to a 2A-BC, and that a 2-additive capacity should be preferred to a bi-capacity. Then the questions are:

1. Does the elicitation program give the parameters of an additive bi-capacity if a 2-additive bi-capacity is not needed to obtain the observed results?
2. Does the elicitation program give the parameters of a capacity if no bi-capacity is needed to obtain the observed results?

In its basic form, our elicitation process does not guarantee the parsimony of the model. The chosen constraints due to the monotonicity just guarantee that the bi-capacity is at most 2-additive. Thus, we can obtain an additive bi-capacity, or even a capacity with the proposed optimization programs. But if the aim is to obtain the simplest possible model, then we have to modify our program to cope with parsimony. We propose to add a penalty term in the objective function which favors simpler solutions. In that case, we would prefer a 1A-BC (1 additive bi-capacities) in comparison with a 2A-BC. To this end, we would like the terms of the bipolar Möbius transform for elements (A_1, A_2) such that $|A_1 \cup A_2| = 2$, to be as low as possible. We thus propose to minimize $\sum_{(A_1,A_2)\in 3^N:|A_1\cup A_2|=2} |b(A_1, A_2)|$ as well. It is well-known in machine learning that penalizing a vector w.r.t. its L_1 norm leads to sparse solutions. However, since we have several criteria to take into account, one way to solve our general problem is to consider a linear combination of the different measures we want to optimize. Accordingly, the general objective function we propose is the following one:

$$\begin{aligned} \min \quad & \alpha \sum_{x,x'\in X'} (S(x) - \mathcal{C}_b(x))^2 \\ & + \beta \sum_{x,x':S(x)\geq S(x')} \xi_{xx'} \\ & + \gamma \sum_{(A_1,A_2)\in 3^N:|A_1\cup A_2|=2} |b(A_1, A_2)| \end{aligned} \tag{27}$$

where α, β, γ are non-negative real numbers that allow balancing the importance of each criteria.

Let us take again Example 1, with the new scores presented in Table 7.

Table 7 Data given by the DM

	1(*M*)	2(*S*)	3(*L*)	Score
A	16	13	7	11.2
B	16	11	9	12
C	6	13	7	8.6
D	6	11	9	9
E	14	16	7	11.1
F	14	15	8	11.3
G	9	16	7	10.1
H	9	15	8	10.3

Table 8 Möbius coefficients without penalty

$b[*,*]$	$\emptyset$	*M*	*S*	MS	*L*	ML	SL	MSL
$\emptyset$	0	−0.2	−0.2	−0.2	−0.4	−0.2	−0.2	0
M	0.4	–	−0.2	–	0	–	0	–
S	0.3	0.2	–	–	−0.1	0	–	–
MS	−0.2	–	–	–	0	–	–	–
L	0.2	0	0	0	–	–	–	–
ML	−0.2	–	0	–	–	–	–	–
SL	0.5	0	–	–	–	–	–	–
MSL	0	–	–	–	–	–	–	–

We first decide to not penalize the complexity of the BC, by setting $\alpha = \beta = 1$ and $\gamma = 0$. This weighting scheme assumes that it is as important to recover as much as possible the original scores provided by the DM as satisfying his underlying preference relations among the alternatives. The obtained preference model is able to exactly reproduce the original scores given by the DM. The bipolar Möbius transform b is given in Table 8 and we can observe that it is a 2A-BC.

Suppose now, that we are interested in finding a simpler preference model than this 2A-BC. In that perspective we set $\alpha = 1$, $\beta = 2$ and $\gamma = 4$. For simplicity, if we assume that the sub-criteria are commensurable, then this weighting scheme states that finding a simple model is twice as important as the satisfaction of the preference relations which is twice as important as recovering the same scores given by the DM. Solving this problem on the example given in Table 7, we obtain the inferred overall scores given in Table 9. With the penalty term, the obtained model does make errors in regard to the original overall scores. However, the preference relations are all satisfied, meaning that the learned model does not generate any inconsistency regarding to the DM preferences. Then, the most interesting fact is that the obtained preference model is actually a 1A-BC as shown in Table 10.

It is also interesting to give the related BC v which is shown in Table 11. In this table we can easily see that $\forall (A_1, A_2) \in 3^N : v(A_1, A_2) = v_1(A_1) - v_2(A_2)$ with v_1 and v_2 being two additive capacities such that $v_1(\{M\}) = 0.33$, $v_1(\{S\}) = 0.18$,

Table 9 Result scores obtained by regression approach

	sc	*choq_val*
A	1.2	1.35
B	2	1.75
C	−1.4	−1.32
D	−1	−0.92
E	1.1	1.25
F	1.3	1.45
G	0.1	−0.23
H	0.3	−0.03

Table 10 Möbius coefficients with penalty

b[*,*]	∅	*M*	*S*	MS	*L*	ML	SL	MSL
∅	0	−0.18	−0.44	0	−0.38	0	0	0
M	0.33	–	0	–	0	–	0	–
S	0.18	0	–	–	0	0	–	–
MS	0	–	–	–	0	–	–	–
L	0.49	0	0	0	–	–	–	–
ML	0	–	0	–	–	–	–	–
SL	0	0	–	–	–	–	–	–
MSL	0	–	–	–	–	–	–	–

Table 11 Bi-Capacity coefficients with penalty

*b*b*,*]	∅	*M*	*S*	MS	*L*	ML	SL	MSL
∅	0	−0.18	−0.44	−0.62	−0.38	−0.56	−0.82	−1
M	0.33	–	−0.11	–	−0.05	–	−0.49	–
S	0.18	0.0	–	–	−0.2	−0.38	–	–
MS	0.51	–	–	–	0.13	–	–	–
L	0.49	0.31	0.05	−0.13	–	–	–	–.
ML	0.82	–	0.38	–	–	–	–	–
SL	0.67	0.49	–	–	–	–	–	–
MSL	1	–	–	–	–	–	–	–

$v_1(\{L\}) = 0.49$ and $v_2(\{M\}) = 0.18$, $v_2(\{S\}) = 0.44$, $v_2(\{L\}) = 0.38$. This case reduces to a preference model of CPT (Cumulative Prospect Theory) type as explained in Grabisch and Labreuche (2005a).

However, note that it is not possible to favor a model based on a simple capacity instead of a bi-capacity without adding new penalty terms. As shown in Grabisch and Labreuche (2005b), a bi-capacity can be reduced to a capacity if it can be expressed in a CPT way:

$$v(A, B) = \mu^+(A) - \mu^-(B)$$

with the adding property that μ should be a symmetric ($\mu^+ = \mu^-$) or asymmetric ($\mu^+ = -\mu^-$) capacity. Including these constraints into the mathematical program should lead to a very complex model with a great number of penalty terms. As an answer to the second question, we decide not to take into account these penalty terms, and so we cannot guarantee that our approach outputs the simplest possible preference model.

4.2 Real-world example

Preference relations based on the use of bipolar scales are not only interesting in ad hoc examples. In real-life situations, one can also face many situations where preferences are expressed using a bipolar scale. The following example allows us to illustrate this comment.

Example 6 A restaurant measures the satisfaction of its consumers concerning its services using a sample survey. The survey contains 198 exploitable consumers. The consumers' satisfaction is assessed through scores obtained on four criteria: quality/price ratio (C_1); cleanliness (C_2); team service (C_3); food quality (C_4). A global satisfaction score is also given by the customer. All these scores are between 0 (worst score) and 10 (best score). The aim is to find a model which can explain the global score by a combination of the criteria scores.

Since there are interactions between criteria, a simple weighted mean cannot represent the global satisfaction score. For example in Table 12, lines 5, 6 show that a global score of 8 can be obtained with combination (8,9,8,7) or (9,8,8,7). But we can see on line 7 that combination (9,9,8,7) does not give a better global score : there is a negative interaction between C_1 and C_2 which cannot be modeled with a weighted mean.

The satisfaction score appears also to be bipolar: the customer uses different scales to express her satisfaction or her unsatisfaction. The restaurant specified that a score of 7 should be considered as neutral. For our study we have then to center all the scores w.r.t. 7. Scores higher than 7 should be considered as positive scores,

Table 12 Satisfaction scores—example data

	*C*1	*C*2	*C*3	*C*4	Sglob
1	3	6	8	2	3
2	6	6	8	5	5
3	4	6	9	9	7
4	10	8	8	10	10
5	8	9	8	7	8
6	9	8	8	7	8
7	9	9	8	7	8
8	8	7	7	7	7
9	6	7	7	7	6
…	…	…	…	…	…

whereas scores under 7 should be considered as negative. Then the 2A-BC seems to be an adequate model to represent such a satisfaction function.

Following the recommendations developed in Sect. 3, we used a split linear method to determine the parameters of the BCI. The results are presented in Table 13. The interpretation of such parameters is not easy. We can however notice that, among others:

- The bi-capacity is not symmetric since $b(\{C_1\}, \emptyset) \neq b(\emptyset, \{C_1\})$.
- The inequality $|b(\{C_3\}, \emptyset)| > |b(\emptyset, \{C_3\})|$, shows that the satisfaction on criterion $C3$ has a greater impact on the global satisfaction than the unsatisfaction on criterion C_3 has on the global unsatisfaction. Inspired by a terminology introduced in the field of social psychology in Helzberg et al. (1959) or in marketing in Llosa (1999), we can say that criterion C_3 appears to be a *bonus criterion*: a good score on C_3 contributes to increase the global satisfaction, but a bad score on C_3 has only a small effect to the global unsatisfaction.
- On another hand, criterion $C4$ appears to be important both in satisfaction and unsatisfaction cases, as $|b(\{C_4\}, \emptyset)| = |b(\emptyset, \{C_4\})|$. We can then see criterion C_4 as a *key criterion*: the score on criterion C_4 has an influence both on global satisfaction and unsatisfaction.

With the split method, we are able to represent 91.4 % of the 39204 preference relations deduced from the scores comparisons. Note that a weighted mean, with parameters obtained by a simple multivariate linear regression, can represent 89.1 % of the preference relations. A simple Choquet integral obtained by a mean

Table 13 Möbius coefficients related to restaurant satisfaction

b[$*, *$]	$\emptyset$	C_1	C_2	$C_1\,C_2$	C_3	$C_1\,C_3$	$C_2\,C_3$	C_4
$\emptyset$	0	−0.22	−0.22	0.04	−0.06	0	0	−0.70
C_1	0.16	.	−0.08	–	−0.01	.	.	0.22
C_2	0.18	0.03	–	–	0	–	.	−0.010
$C_1\,C_2$	−0.08	–	–	–	.	–	–	–
C_3	0.15	0.10	0	–	–	–	.	−0.15
$C_1\,C_3$	0.02	–	–	–	–	–	–	–
$C_2\,C_3$	0.07	–	–	–	–	–	–	.
C_4	0.70	−0.15	−0.52	.	−0.04	–	–	–
$C_1\,C_4$	−0.07	–	–	–	–	–	–	–
$C_2\,C_4$	−0.10	–	–	–	–	–	–	–
$C_3\,C_4$	−0.04	–	–	–	–	–	–	–

b[*,*]	$C_1\,C_4$	$C_2\,C_4$	$C_3\,C_4$
$\emptyset$	−0.08	0.17	0.06
...	–	–	–

square method described in Grabisch et al. (2008) represents 87.2 % of the preference relations.

If we want to focus on the ability of the model to fit the real scores (and not the preference relation), we can notice that the squared sum distance between the real score and the estimate one is 23.22 for the BCI (with a regression method) and 30.04 for the weighted mean, which shows that a BCI model better fits the data.

It is not surprising that a bi-capacity leads to a small improvement of the performances over a weighted mean. In most cases, the use of a weighted mean is sufficient to explain the global satisfactory score and a bi-capacity is used only to refine the weighted mean.

5 Conclusion

A general process to determine the parameters of a 2A-BC from the data has been proposed. The decision model is inferred from examples that the DM evaluated both regarding their partial utilities and their global scores. Several methods have been proposed whether an interaction with the DM is possible or not. Whatever the context, our techniques all lead to the resolution of optimization problems. We investigated both the traditional preference elicitation setting where no inconsistency is allowed and the more flexible case where we have to cope with such inconsistencies. We also proposed a parsimonious BC model through the use of penalty terms.

The main next step we intend to undertake is to study the variation of computing time as a function of the number of criteria and the number of alternatives in the learning set.

References

Ah-Pine J, Mayag B, Rolland A (2013) Identification of a 2-additive bi-capacity using mathematical programming. In: Proceeding of ADT '13

Bouyssou D, Couceiro M, Labreuche C, Marichal J-L, Mayag B (2012) Using Choquet integral in machine learning: What can MCDA bring? In: DA2PL, 2012 Workshop: From Multiple Criteria Decision Aid to Preference Learning. Mons, Belgique

Bouyssou D, Marchant T, Pirlot M, Tsoukias A, Vincke P (2006) Evaluation and Decision models with multiple criteria: stepping stones for the analyst, volume 86 of International Series in Operations Research & Management Science. Springer, New York

Choquet G (1954) Theory of capacities. Annales de l'institut Fourier 5:131–295

Corrente S, Figueira JR, Greco S (2014) Dealing with interaction between bipolar multiple criteria preferences in PROMETHEE methods. Ann OR 217(1):137–164

Fallah Tehrani A, Cheng W, Dembczynski K, Hullermeier E (2012) Learning monotone nonlinear models using the Choquet integral. Mach Learn 89(1):183–211

Fujimoto K (2004) New characterizations of *k*-additivity and *k*-monotonicity of bi-capacities. In: Joint 2nd Int. Conf. on Soft Computing and Intelligent Systems and 5th International Symposium on Advanced Intelligent Systems

Fujimoto K, Murofushi T (2005) Some characterizations of k-monotonicity through the bipolar möbius transform in bi-capacities. JACIII 9(5):484–495

Fujimoto K, Murofushi T, Sugeno M (2007) k-additivity and c-decomposability of bi-capacities and its integral. Fuzzy Sets Syst 158(15):1698–1712

Grabisch M (1997) k-order additive discrete fuzzy measures and their representation. Fuzzy Sets Syst 92(2):167–189

Grabisch M, Kojadinovic I, Meyer P (2008) A review of methods for capacity identification in Choquet integral based multi-attribute utility theory: Applications of the kappalab R package. EJOR 186(2):766–785

Grabisch M, Labreuche Ch (2002) Bi-capacities for decision making on bipolar scales. In: Proc. of the EUROFUSE 02 Workshop on Information Systems. pp 185–190

Grabisch M, Labreuche Ch (2003) The Choquet integral for 2-additive bi-capacities. In: EUSFLAT Conf. pp 300–303

Grabisch M, Labreuche Ch (2005a) Bi-capacities-I: definition, Möbius transform and interaction. Fuzzy Sets Syst 151(2):211–236

Grabisch M, Labreuche Ch (2005b) Bi-capacities-II: the Choquet integral. Fuzzy Sets Syst 151(2):237–259

Grabisch M, Labreuche Ch (2005c) Fuzzy measures and integrals in MCDA. In: Multiple Criteria Decision Analysis: State of the Art Surveys, volume 78 of Int. Series in Op. Res. & Manag. Sci. Springer, New York, pp 563–604

Grabisch M, Labreuche Ch (2008) A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. 4OR 6(1):1–44

Grabisch M, Greco S, Pirlot M (2008) Bipolar and bivariate models in multicriteria decision analysis: Descriptive and constructive approaches. Int J Intell Syst 23(9):930–969

Greco S, Figueira JR (2003) Dealing with interaction between bipolar multiple criteria preferences in outranking methods. In: Research Report 11–2003. INESC-Coimbra, Portugal

Greco S, Matarazzo B, Slowinski R (2002) Bipolar Sugeno and Choquet integrals. In: Proceedings of the EUROFUSE 02 Workshop on Information Systems

Greco S, Rindone S (2013) Bipolar fuzzy integrals. Fuzzy Sets Syst 220:21–33

Greco S, Rindone F (2014) The bipolar Choquet integral representation. Theory Decis 77:1:1–29

Helzberg F, Mausner B, Snyderman BB (1959) The Motivation to Work. Wiley, New York

Hullermeier E, Fallah Tehrani A (2012) Computational intelligence in intelligent data analysis. studies in computational intelligence. In: Moewes C, Nurnberger A (eds) Computational Intelligence in Intelligent Data Analysis. Studies in Computational Intelligence. Springer, New York, pp 17–30

Hullermeier E, Fallah Tehrani A (2012) On the vc dimension of the Choquet integral. In: International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems

Jacquet-Lagreze E, Siskos J (1982) Assessing a set of additive utility functions for multicriteria decision-making, the UTA method. Eur J Oper Res 10(2):151–164

Llosa S (1999) Four ways to contribute to satisfaction in services: Basic, Plus, Key and secondary elements. In: AMA SERVSIG Service Research Conference. New Orleans, pp 121–123

Marichal J-L, Roubens M (2000) Determination of weights of interacting criteria from a reference set. EJOR 124(3):641–650

Mayag B, Grabisch M, Labreuche Ch (2010) An interactive algorithm to deal with inconsistencies in the representation of cardinal information. In: IPMU 2010, pp 148–157

Mayag B, Rolland A, Ah-Pine J (2012) Elicitation of a 2-additive bi-capacity through cardinal information on trinary actions. In: IPMU (4), pp 238–247

Simon H (1956) Rational choice and the structure of the environment. Psychol Rev 63(2):129–138

Slovic P, Finucane M, Peters E, MacGregor DG (2002) The affect heuristic. In: Heuristics and biases: the psychology of intuitive judgment. Cambridge University Press, pp 397–420

Tversky A, Kahneman D (1992) Advances in prospect theory: cumulative representation of uncertainty. J. Risk Uncertain 5:297–323

Waegeman W, De Baets B, Boullart L (2009) Kernel-based learning methods for preference aggregation. 4OR 7(2):169–189