

Using multiple reference levels in Multi-Criteria Decision Aid: the Generalized-Additive Independence model and the Choquet integral approaches

Christophe Labreuche, Michel Grabisch

▶ **To cite this version:**

Christophe Labreuche, Michel Grabisch. Using multiple reference levels in Multi-Criteria Decision Aid: the Generalized-Additive Independence model and the Choquet integral approaches. 2018. halshs-01815028

HAL Id: halshs-01815028
https://shs.hal.science/halshs-01815028

Submitted on 13 Jun 2018

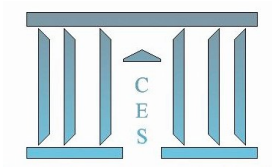
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS 1
PANTHÉON SORBONNE

Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers



Using multiple reference levels in Multi-Criteria Decision Aid: the Generalized-Additive Independence model and the Choquet integral approaches

Christophe LABREUCHE, Michel GRABISCH

2018.09



Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
https://centredeconomiesorbonne.univ-paris1.fr/
ISSN : 1955-611X

Using multiple reference levels in Multi-Criteria Decision Aid: the Generalized-Additive Independence model and the Choquet integral approaches

Christophe Labreuche*

Thales Research & Technology, 1 avenue Augustin Fresnel, 91767 Palaiseau Cedex, France
christophe.labreuche@thalesgroup.com

Michel Grabisch

Paris School of Economics, Université Paris I - Panthéon-Sorbonne, Paris, France
michel.grabisch@univ-paris1.fr

Abstract

In many Multi-Criteria Decision problems, one can construct with the decision maker several reference levels on the attributes such that some decision strategies are conditional on the comparison with these reference levels. The classical models (such as the Choquet integral) cannot represent these preferences. We are then interested in two models. The first one is the Choquet with respect to a p-ary capacity combined with utility functions, where the p-ary capacity is obtained from the reference levels. The second one is a specialization of the Generalized-Additive Independence (GAI) model, which is discretized to fit with the presence of reference levels. These two models share common properties (monotonicity, continuity, properly weighted,...), but differ on the interpolation means (Lovász extension for the Choquet integral, and multi-linear extension for the GAI model). A drawback of the use of the Choquet integral with respect to a p-ary capacity is that it cannot satisfy decision strategies in each domain bounded by two successive reference levels that are completely independent of one another. We show that this is not the case with the GAI model.

Keywords: Multiple criteria analysis; Generalized Additive Independence;

*Corresponding author
Email address: christophe.labreuche@thalesgroup.com (Christophe Labreuche)

Choquet integral; Reference Levels; Interpolation.

1. Introduction

Many models have been defined and studied in Multi-Criteria Decision Analysis (MCDA) with various levels of preference representation power and of elicitation complexity. While many of them are sufficient when the Decision Maker (DM) has only *standard* preferences, there are however situations in which the DM wishes to express much richer and specific decision strategies. We consider in this paper the case where there is an inherent complexity in the DM preferences coming from the presence of reference levels in the decision problem.

Reference elements on the attributes play an important role in MCDA. In the MACBETH approach, two reference elements are defined on each attribute, and they serve to give a semantics to the concept of weight of criteria[1] [2]. In MR-Sort, an alternative is assigned to the upper class if the set of criteria above a given threshold forms a winning coalition [4, 5]. A similar idea operates in reference point methods [41]. In psychology, reference levels such as the *neutral level* – demarcating between positive and negative stimulis [28, 45]– and the *satisficing level* that is considered as good and completely satisfactory if the DM could obtain it, even if more attractive elements could exist [44], are well-developed. The existence of a neutral level characterizes bipolar scales.

Decision strategies are often affected by reference elements. Let us cite a few examples. *Conditional Relative Importance* depicts a decision strategy that is dependent on the value of a bipolar attribute being above or below the neutral level: the relative importance between two criteria is conditional on a third (bipolar) criterion being good (above the neutral level) or bad (below the neutral level). This strategy arises in many situations – see [35] for an example.

[1] The importance of a criterion is the difference of utility going from the lower reference level to the upper one on an attribute, the other attributes being fixed to any value.

The following example also illustrates decision strategies that are dependent on the values of some attributes.

Example 1. *The chief technical officer (CTO) of a company has to decide on which products to develop and then sell to customers. Five attributes are considered to evaluate potential products:*

- *Expected Revenue (ER), which is a number in M Euro;*
- *Risk Level (RL), taking values in {low, medium, high};*
- *Reputation of the Company (RC) on the market targeted by the product, taking qualitative values between 1 (very low) to 5 (very high);*
- *Position with respect to Competitors (PC), taking values among DT(disruptive technology), SOTA(at the state of the art), BSOTA(below state of the art);*
- *Risk Mitigation (RM), taking values among T(collaborate with another partner to develop the technology – called "Team" strategy, to support the development of the technology), II(internal investment), N(do nothing).*

Usually, the risk and revenue attributes can be split into two domains – low/high risk and low/high revenue respectively. The decision strategies depend on which domain the potential product belongs to. The ideal situation is clearly for low risk and high revenue products, whereas the worst situation is for high risk and low revenue products. Most products are represented by the other two situations. Low risk and low revenue products can correspond to niche products. Here the reputation (RC) is a very important criterion. The product does not necessarily need to be sophisticated, because of the "niche" situation. The last two criteria PC and RM are thus less important. Lastly, high risk and high revenue products can secure the future of the company. But the competition is high on these markets, so that criteria PC and RM are now very important.

The previous example implicitly exhibits reference levels on attributes ER and RL, and the decision strategies depend on these levels.

The question that naturally arises is how to represent such decision strategies with MCDA models. As we will see more precisely, it is not possible to represent these preferences with standard MCDA models. The additive utility model is not suitable as the relative importance between some criteria depends on the values on some other criteria, and more generally, criteria interact each other. Reference-based models cannot help to compare alternatives in the same domain (e.g. high risk and high revenue products in Ex. 1) [41]. Despite its ability to model interacting criteria, the Choquet integral fails to represent these preferences as it cannot model different decision strategies depending on whether attributes are judged good or bad [35]. Actually the Choquet integral is based on only two levels and it cannot model various changes of preferences depending on the ranges of the attributes: The decision strategy at intermediate values is simply an interpolation of those for extreme values.

We show that two models can represent decision strategies that are conditional on the comparison with the reference levels. The first one is the Choquet integral w.r.t. a p-ary capacity [17, 18], combined with utility functions. Here the p-ary capacity represents the overall value at the different reference levels. The second one is a specialization of the Generalized Additive Independence (GAI) model that is discretized to fit with the presence of reference levels.

The concept of p-ary capacity and the associated Choquet integral have been defined as a mathematical tool that naturally extends capacities. However, their complexity makes it very hard to use them in real applications. Up to our knowledge, the only attempt in this direction was for the representation of restaurant customer satisfaction [42], although limited to bi-capacities. The aim of this paper is to show the interest of the GAI model and the Choquet integral w.r.t. p-ary capacities in applications, and especially in those exhibiting reference levels.

The main contributions of this paper can be summarized as follows:

- We present a motivating example taken from a real application in the engineering domain, to illustrate situations where decision strategies are

dependent on several reference levels (Section 3). In engineering, the different levels are successive levels of requirements from the simplest to satisfy to the most difficult one. There are different decision strategies associated to these levels. For instance, the smallest level is generally considered as a hard constraint, whereas the largest level is seen as a bonus (the customer is expected to have a crush in such a situation). This example is running throughout the document, and we show that it cannot be represented by the classical MCDA models.

- Which model can represent the motivating example? (Section 4). We provide two models able to represent the reference levels-dependent multi-criteria decision problems: the Choquet integral w.r.t. p-ary capacities and the GAI models.

  Regarding the GAI model, we provide a specialization of this model to the presence of reference levels, which is new. The GAI model is then seen as a generalization of the UTA approach, where continuous attributes are discretized. This is important to do so in practice in order to reduce the computational complexity [20].

- How to choose between the two aggregation models? (Section 5). These two models share common properties (monotonicity, continuity, properly weighted,...), but differ on the interpolation method (Lovász extension for the Choquet integral, and multi-linear extension for the GAI model). A drawback of the use of the Choquet integral w.r.t. a p-ary capacity is that the possible decision strategies satisfied in each domain bounded by two successive reference levels are not completely independent of one another. We call this phenomenon *contamination*. This was already noted with the *conditional* relative importance strategy [35]. We show that this is not the case with the GAI model.

  We also present the pros and cons of choosing each of these two models.

- How to elicit the two models? (Section 6). The question of the elicitation

of MCDA models is central in applications. This problem might be seen as intractable for the GAI model or the Choquet integral w.r.t. a p-ary capacity due to the large number of parameters. We propose a simple elicitation procedure that focuses on alternatives with balanced profiles (i.e., alternatives in-between two identical levels on all attributes) and fill the model on the other cases by using a cautious principle.

2. Preliminaries

2.1. General model

We are given a set of n attributes indexed by $N = \{1, \ldots, n\}$. Each attribute $i \in N$ is represented by a set X_i. The alternatives are characterized by a value on each attribute, and are thus represented by an element in $X = X_1 \times \cdots \times X_n$. We assume that we are given a preference relation $\succsim$ over X. It is supposed to be represented by an overall utility function

$$U : X \to \mathbb{R}, \tag{1}$$

i.e., such that $x \succsim y$ iff $U(x) \geq U(y)$.

For $x, y \in X$ and $A \subseteq N$, we denote by X_A the set $\prod_{i \in A} X_i$, by x_A the restriction of x on attributes A, and by $(x_A, y_{N \setminus A}) \in X$ the compound alternative taking value x_i for attribute i in A, and value y_i otherwise. For $x, y, z \in X$ and $A, B \subseteq N$ with $A \cap B = \emptyset$, we also denote by $(x_A, y_B, z_{N \setminus (A \cup B)})$ the alternative taking value x_i for attribute i in A, value y_i for i in B, and value z_i otherwise.

Preference relation $\succsim$ is said to satisfy *weak separability* [31] if: for all $i \in N$, all $x_i, y_i \in X_i$, and all $a_{N \setminus \{i\}}, b_{N \setminus \{i\}} \in X_{N \setminus \{i\}}$

$$(x_i, a_{N \setminus \{i\}}) \succsim (y_i, a_{N \setminus \{i\}}) \iff (x_i, b_{N \setminus \{i\}}) \succsim (y_i, b_{N \setminus \{i\}}). \tag{2}$$

Under this assumption, one can derive, for every $i \in N$, a preference relation $\succsim_i$ over attribute i from $\succsim$: for all $x_i, y_i \in X_i$

$$x_i \succsim_i y_i \quad \Leftrightarrow \quad (x_i, a_{N \setminus \{i\}}) \succsim (y_i, a_{N \setminus \{i\}}) \tag{3}$$

for some $a_{N\setminus\{i\}} \in X_{N\setminus\{i\}}$. Weak separability simply says that the preference over the values of an attribute, ceteris paribus, should be independent of the values of the other attributes. If we consider the simple example of choosing a car on the basis of its price, its comfort and its maximal speed, weak separability is clearly satisfied: all other criteria being fixed, the smaller the price the better, and similarly for all criteria: the higher the comfort, the better, and the higher the maximal speed, the better. The latter statements are depicted by relations $\succsim_i$'s.

We denote by $\succ_i$ and $\sim_i$ the asymmetric and symmetric parts of $\succsim_i$ respectively. In order to avoid unnecessary heavy notation and details, we assume that the partial preferences $\succsim_i$ are monotone: the larger the value of the attributes, the better ($x_i \succsim_i y_i \iff x_i \geq y_i$).

Under weak separability, utility U shall fulfill the following monotonicity conditions, which states that it should be consistent with each relation $\succsim_i$:

$$\forall x, y \in X \text{ with } y_i \succsim_i x_i \text{ for every } i \in N, \qquad U(y) \geq U(x) \tag{4}$$

There exist many different utility models of the form (1). In the rest of this section, we focus on two models: (k-ary) capacities and the Choquet integral (section 2.2), and the GAI model (section 2.3).

2.2. Capacities, the multi-linear model and the Choquet integral

When $\succsim$ satisfies weak separability and other properties, function U takes the decomposition form [29]

$$U(x) = F(u_1(x_1), \ldots, u_n(x_n)), \tag{5}$$

where $u_i : X_i \to \mathbb{R}$ is the *utility function* (also called *value function*) on X_i and $F : \mathbb{R}^n \to \mathbb{R}$ is an *aggregation function*. By aggregation function, we simply mean a non-decreasing function [21]. Utility function u_i shall be consistent with $\succsim_i$ (i.e. $u_i(x_i) \geq u_i(y_i)$ whenever $x_i \succsim_i y_i$). Hence u_i is non-decreasing.

We now introduce the concept of a capacity to describe a class of aggregation functions F.

Definition 1. *A* fuzzy measure *[47] or* capacity *[8] on N is a set function $\mu : 2^N \to \mathbb{R}$ satisfying (1) the monotonicity conditions: $\mu(A) \leq \mu(B)$ for every $A \subseteq B$, and (2) the normalization conditions: $\mu(\emptyset) = 0$, $\mu(N) = 1$.*

Capacities are related to the concept of pseudo-Boolean function. A *pseudo-Boolean function* is any function $f : \{0,1\}^N \to \mathbb{R}$ [24]. Writing $2^N \equiv \{0,1\}^N$, there is a one-to-one correspondence between set functions and pseudo-Boolean functions: $f(1_A) = \mu(A)$ for all $A \subseteq N$. From this correspondence, the problem of defining an aggregation function from a capacity is similar to the one of extending a pseudo-Boolean function on $[0,1]^N$.

In order to define extensions of pseudo-Boolean functions to $[0,1]^N$, it is convenient to use the Möbius transform [43] of μ, defined by

$$m^\mu(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} \mu(B). \tag{6}$$

Conversely, μ can be derived from m^μ by the formula $\mu(A) = \sum_{B \subseteq A} m^\mu(B)$. We can now give two standard extensions of a pseudo-Boolean function f (or equivalently μ) to $[0,1]^N$ [25]

$$f^\Pi(t) = \sum_{A \subseteq N} m^\mu(A) \cdot \prod_{i \in A} t_i \tag{7}$$

$$f^\wedge(t) = \sum_{A \subseteq N} m^\mu(A) \cdot \bigwedge_{i \in A} t_i. \tag{8}$$

The first expression is the multi-linear extension of f or μ, and the second one is the *Lovász* extension of f or μ. Yet $f^\wedge$ corresponds to the Choquet integral w.r.t. μ [8].

For a permutation τ on N, we define $\Omega_\tau = \{t \in \mathbb{R}^N : t_{\tau(1)} \leq t_{\tau(2)} \leq \cdots \leq t_{\tau(n)}\}$. The Choquet integral of $t \in [0,1]^N$ can be written in terms of the capacity μ [8]

$$C_\mu(t) = \sum_{i=1}^{n} \left(t_{\tau(i)} - t_{\tau(i-1)}\right) \mu\left(\{\tau(i), \cdots, \tau(n)\}\right), \tag{9}$$

where $t_{\tau(0)} := 0$ and τ is a permutation on N such that $t \in \Omega_\tau$. The Choquet integral is clearly a simple weighted sum in each domain Ω_τ. In Figure 1, the

shaded area represents Ω_τ with $n = 2$, $\tau(1) = 2$ and $\tau(2) = 1$. The Choquet integral is one of the most versatile aggregation function as it is able to capture various decision strategies representing interaction among criteria [8, 15, 19].

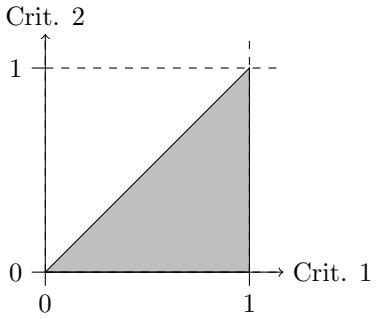


Figure 1: Set of vectors $t = (t_1, t_2)$ such that $t_1 \geq t_2$.

The use of the decomposable model (5) with a Choquet integral implies that the scales of the marginal utility functions u_i are commensurate. A scale $L_i = \{u_i(x_i) : x_i \in X_i\}$ ordered by $\geq$ is constructed on each criterion i. The marginal utility functions u_i are then said to be *commensurate* if all scales L_i's can be represented by the same scale L_u [10]. This means that one shall be able to compare the values of any two scales L_i and L_j, with a preference relation $\succsim_{L_i \times L_j}$ defined on $L_i \times L_j$. A simple way to fulfill this assumption is to ensure that all scales are identical: $L_1 = \cdots = L_n = L_u$, and $a \sim_{L_i \times L_j} a$ (value a on scale L_i is indifferent to the same value a on scale L_j) for every $a \in L_u$ and $i, j \in N$. In other words, if $u_i(x_i) = u_j(x_j)$, then value x_i on attribute X_i has the same satisfaction/attractiveness as value x_j on attribute X_j. This is the approach adopted in MACBETH for the weighted sum [2]. The basic principle is to construct interval scales given up to an affine transformation over each attribute X_i, and to fix the two degrees of freedom of each affine transformation by assuming the existence of two reference levels on the associated attribute. It has been extended to the Choquet integral [33, 32].

When commensurability is fulfilled, models of the form (5) are assumed to satisfy the idempotency property [38], which states that $F(\alpha, \ldots, \alpha) = \alpha$ for every $\alpha \in \mathbb{R}$. In other words, an alternative which is scored α on all criteria,

should have α as overall evaluation, because value α has the same meaning throughout all criteria.

2.3. GAI model

There are alternative utility models which do not need the commensurability assumption. The best-known model of this class is the *additive utility* model [29]

$$U(x) = \sum_{i \in N} u_i(x_i) \tag{10}$$

where $u_i : X_i \to \mathbb{R}$. Note that there is an induced relation between the attributes in this model. For instance, increasing attribute i from value x_i to y_i is equivalent to increasing attribute j from value x_j to y_j, which corresponds to even-swaps [26]. But this relation is not sufficient to compare directly the values on two attributes, so that the commensurability assumption is not satisfied.

The additive utility model has been generalized to allow some interaction among criteria – under the name of the Generalized Additive Independence (GAI) model [12, 13, 1]. The GAI model takes the form of the sum of utilities over subsets of attributes:

$$U(x) = \sum_{S \in \mathcal{S}} u_S(x_S) \tag{11}$$

where $\mathcal{S}$ is a collection of subsets of N, $x_S \in X_S$ is the restriction of x over attributes in coalition S, and

$$u_S : X_S \to \mathbb{R}.$$

Note that [13, Section 11.3] considers (in an expected utility context) GAI models with no assumption on the attributes (they can be discrete or continuous). The set $\mathcal{S}$ contains all subsets of attributes that interact one another. Hence the additive model (10) is a particular case of the GAI model where $\mathcal{S}$ is composed only of singletons. One may consider for instance that $\mathcal{S}$ is the collection of all singletons and pairs of criteria, as in the utility model underlying $UTA^{GMS}-INT$ [23], which have strong connections with the GAI model.

It is important to note that the GAI does not necessarily satisfy the weak separability condition [7]. A well-known example of such a violation is the following [3]: the problem is to choose menus in a restaurant on the basis of two attributes X_1 (main course: 'meat' or 'fish') and X_2 (wine: 'white' or 'red'). Then 'red wine' is preferred to 'white wine' if the main course is 'meat', but 'white wine' is preferred to 'red wine' if the main course is 'fish'.

Hence according to the assumption made in section 2.1, we will consider in this paper only GAI models that satisfy weak separability. It shall be noted that examples like the previous one, violating weak separability, are quite unusual in MCDA. The large majority of MCDA applications satisfy this condition. Moreover, most MCDA methods in OR require weak separability: Pareto front computation in multi-objective optimization, the rough set approach, outranking methods, the additive utility, the Choquet integral,... It is thus not a restriction to assume it.

3. A motivating example

This section introduces and discusses a motivating example in which decision strategies are conditional on where the alternatives are located relatively to some reference levels.

3.1. Description of the example

Consider the following example.

Example 2. *The design of a system, like an embedded system, is a complex activity [37, 40]. The aim is to help the system architect to select the best architecture. System engineers classically define (soft) requirements on key attributes. Such attributes could be the latency of a functional chain, or the weight of the embedded system. Requirements can be specified with the help of several levels*[2] *such as "the system* 'Must have' *at least/most value* x_i^{M}*", "the system*

[2]See MoSCoW method for setting priorities among requirements, among "Must have"

Documents de travail du Centre d'Economie de la Sorbonne - 2018.09

'Should have' *at least/most value* x_i^{S}*", and "it is* 'Nice to have' *for the system at least/most value* x_i^{N}*".*

The system architect expresses the following preferences. If any of the attributes is below the 'Must have' *level, then the system is unacceptable. If the two attributes are between the* 'Must have' *and* 'Should have' *levels, then the experts are intolerant and look at the worst satisfied criterion. Architectures which values on the two attributes are between the* 'Should have' *and* 'Nice to have' *levels are completely satisfactory so that the two attributes are fully compensatory. The situation where the two attributes are above the* 'Nice to have' *levels is ideal, and the DM basically looks at bonus represented by the best satisfied criterion. And so on. The preferences are depicted in Figure 2.*

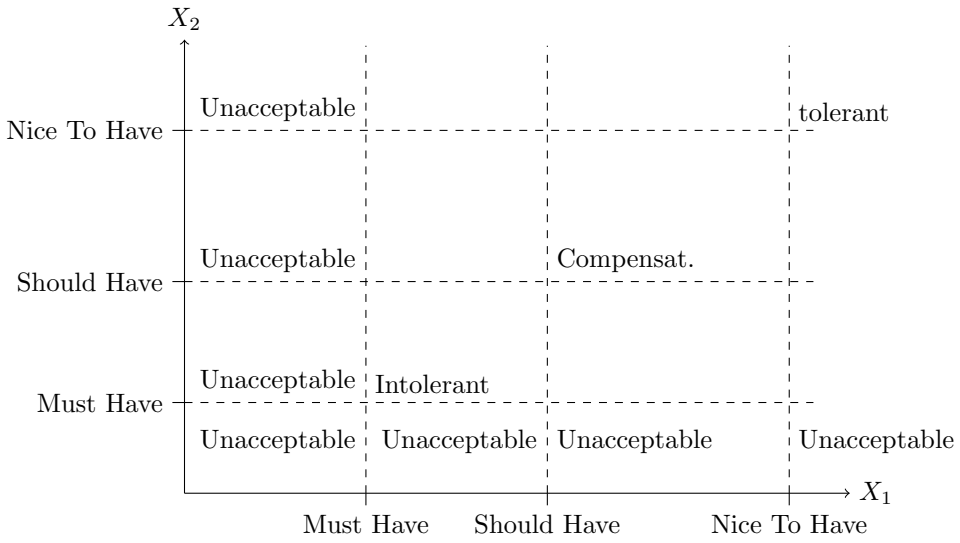


Figure 2: Decision strategies in Example 2.

(critical to the delivery), "Should have" (important but not necessary for delivery), "Could have" (desirable but not necessary), "Won't have/Would like" (not included in the current delivery.

3.2. Attempt to represent these preferences

Let us try to model the previous example with a Choquet integral. It is convenient, for the construction of the model, that each decision strategy occurs in a domain that is bounded by two reference levels. This is not entirely the case in Example 2, as there is a decision strategy below the lowest level 'Must have' and one above the highest level 'Nice to have'. As a result, we add two other reference levels: 'Worst value' (below level 'Must have') and 'Best value' (above level 'Nice to have'). Hence, at the end, we have five reference levels ($p_i = 5$).

We need first to define utility functions. We can choose u_i(Worst value) $= 0$, u_i(Must have) $= \alpha$, u_i(Should have) $= \beta$, u_i(Nice to have) $= \gamma$ and u_i(Best value) $= 1$, where $0 < \alpha < \beta < \gamma < 1$ are first identified (using for instance MACBETH approach [2]).

The architect expresses separate decision strategies for each range of values of the attributes. They can be modeled by the following capacities:

- $v(\emptyset) = v(\{1\}) = v(\{2\}) = 0$ and $v(\{1,2\}) = 1$, below the 'Must have' levels, as the two criteria are veto;
- $v(\emptyset) = 0$, $v(\{1\}) = v(\{2\}) = 0.2$, $v(\{1,2\}) = 1$, between the 'Must have' and 'Should have' levels, which is indeed an intolerant model;
- $v(\emptyset) = 0$, $v(\{1\}) = v(\{2\}) = 0.5$, $v(\{1,2\}) = 1$, between the 'Should have' and 'Nice to have' levels, which is a simple weighted sum and is thus indeed a completely compensatory model;
- $v(\emptyset) = 0$, $v(\{1\}) = v(\{2\}) = 0.8$, $v(\{1,2\}) = 1$, above the 'Nice to have' level, which indeed represents a tolerant model.

The problem with the Choquet integral is that enforcing a veto below the 'Must have' levels, implies that there is a veto for all values of the attributes. A similar reasoning applies for the other decision strategies. Hence we cannot model the four decision strategies at the same time with a simple Choquet integral.

Reference point methods are extensions of the outranking methods based on the presence of reference levels. Non-Compensatory Sorting assigns at least a category to an alternative if the set of attributes that beats the reference profile of the category forms a winning coalition [4, 5]. In ranking with multiple profiles, alternatives are compared only on the basis of the set of attributes for which each alternative beats each reference level [41, 6]. In particular, there is no possible strict preference between two alternatives which belong to the same domain (same comparison with the reference levels). Hence it is not possible to represent the preferences of Ex. 2 with reference point methods.

4. Two aggregation models based on several reference levels

This section describes two models able to distinguish decision strategies depending on the position with respect to reference elements. The first one is an extension of the Choquet integral to cope with multiple levels (see Section 4.1). The second one is a specialization of the GAI model that allows, by construction, separate decision strategies in each domain (see Section 4.2). We will show that these models are able to capture the preferences expressed in the motivating example.

The set of reference levels on attribute X_i is denoted by $D_i = \{a_i^0, a_i^1, \ldots, a_i^{p_i}\}$, with $a_i^0 < \cdots < a_i^{p_i}$. We set $p = (p_1, \ldots, p_n)$.

4.1. p-ary capacity and the Choquet integral

The concept of a capacity is based on reference levels $\{0, 1\}$ on each criterion. It has been generalized to an arbitrary number of reference levels over each criterion. For $p = (p_1, \ldots, p_n)$, we define

$$\mathcal{Q}_p(N) = \{0, \ldots, p_1\} \times \cdots \times \{0, \ldots, p_n\}, \tag{12}$$

and $\leq$ on $\mathcal{Q}_p(N)$ by, for every $q = (q_1, \ldots, q_n), q' = (q'_1, \ldots, q'_n) \in \mathcal{Q}_p(N)$

$$q \leq q' \quad \text{iff} \quad a_i^{q_i} \leq a_i^{q'_i} \quad \forall i \in N. \tag{13}$$

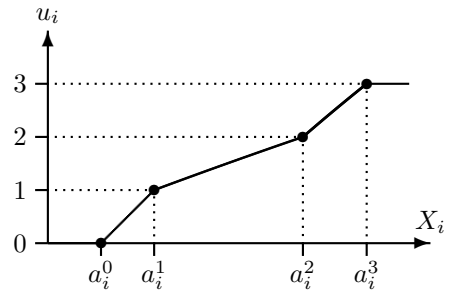


Figure 3: Piecewise affine utility function u_i with $p_i = 3$.

We may call $\mathcal{Q}_p(N)$ the *grid* of reference levels. A standard capacity is characterized by two reference elements a_i^0 and a_i^1 that are mapped onto 0 and 1 respectively, thanks to the utility function u_i. Likewise, the $p_i + 1$ reference elements $a_i^0, a_i^1, \ldots, a_i^{p_i}$ are mapped onto values $\{0, 1, 2, \ldots, p_i\}$ respectively, by u_i:

$$\forall i \in N \quad \forall l \in \{0, 1, 2, \ldots, p_i\} \qquad u_i(a_i^l) = l. \tag{14}$$

We consider here the simplest utility function fulfilling this condition, that is making linear interpolations (see Figure 3):

$$u_i(x_i) = \begin{cases} 0 \quad \text{if } x_i \leq a_i^0 \\ \frac{x_i - a_i^l}{a_i^{l+1} - a_i^l}\,(l+1) + \frac{a_i^{l+1} - x_i}{a_i^{l+1} - a_i^l}\,l = l + \frac{x_i - a_i^l}{a_i^{l+1} - a_i^l} \quad \text{if } a_i^l \leq x_i < a_i^{l+1} \\ p_i \quad \text{if } x_i \geq a_i^{p_i} \end{cases} \tag{15}$$

A more general form of u_i satisfying (14) can be envisaged.

We can now define p-ary capacities, where a usual capacity is a $(1, \ldots, 1)$-ary capacity.

Definition 2 ([16, 18]). *A* p-ary capacity *on* N *is a function* $v : \mathcal{Q}_p(N) \to \mathbb{R}$ *satisfying the monotonicity conditions:*

$$\forall q, q' \in \mathcal{Q}_p(N) \text{ s.t. } q \leq q' \quad , \quad v(q) \leq v(q'), \tag{16}$$

and the normalization conditions: $v(0, \ldots, 0) = 0$, $v(p_1, \ldots, p_n) = 1$.

This definition is a slight generalization of that presented in [16, 18] in the sense that $p_1 = p_2 = \cdots = p_n =: k$ in these references.

A p-ary capacity is said to be *non-normalized* if we relax the normalization condition $v(p_1, \ldots, p_n) = 1$. In the context of cooperative game theory, this concept is similar to *multichoice games* [27].

Let us define $\overline{\mathcal{Q}}_p(N) = \{q \in \mathcal{Q}_p(N) \ : \ q_i < p_i \ \ \forall i \in N\}$. For $q \in \overline{\mathcal{Q}}_p(N)$, let $V_q = [a_1^{q_1}, a_1^{q_1+1}] \times [a_2^{q_2}, a_2^{q_2+1}] \times \cdots \times [a_n^{q_n}, a_n^{q_n+1}]$.

Let $x \in X$. A p-ary capacity can be seen as a superposition of $p_1 \times p_2 \times \cdots \times p_n$ capacities (over each domain V_q for $q \in \overline{\mathcal{Q}}_p(N)$) so that the Choquet integral with respect to v at x is simply the Choquet integral with respect to one of these capacities. One needs thus to identify the capacity associated to x. The latter is characterized by two successive reference levels on each attribute. We define $q(x) \in \mathcal{Q}_p(N)$ by

$$q_i(x) = \begin{cases} 0 & \text{if } x_i < a_i^1 \\ l & \text{if } a_i^l \leq x_i < a_i^{l+1} \ (l \in \{1, \ldots, p_i - 2\}) \\ p_i - 1 & \text{if } x_i \geq a_i^{p_i - 1} \end{cases} \tag{17}$$

for all $i \in N$. Then x is attached to domain $V_{q(x)}$. The (non-normalized) capacity attached to x is $\mu_{q(x)}$ associated to domain $V_{q(x)}$, where, for $q \in \overline{\mathcal{Q}}_p(N)$ and $S \subseteq N$, $\mu_q(S)$ is defined by

$$\mu_q(S) = v((q+1)_S, q_{N \setminus S}) - v(q), \tag{18}$$

where $((q+1)_S, q_{N \setminus S})$ takes value $q_i + 1$ if $i \in S$ and value q_i otherwise. We clearly have that $u_i(x_i) \in [q_i(x), q_i(x) + 1]$. An illustration of $q(x)$ can be found in Figures 4 and 5.

We define

$$\forall i \in N \qquad \phi_i(x) = u_i(x_i) - q_i(x) \in [0, 1]. \tag{19}$$

Figure 5 illustrates the concepts of utilities and ϕ_i in the criteria space (output of the utility functions).

Remark 1. *There is a clear connection between the previous definitions and some notions used in signal compression [9, 14]. In this theory,* nearest neighbour quantizer *of a point is the element of the lattice that is closest to this*

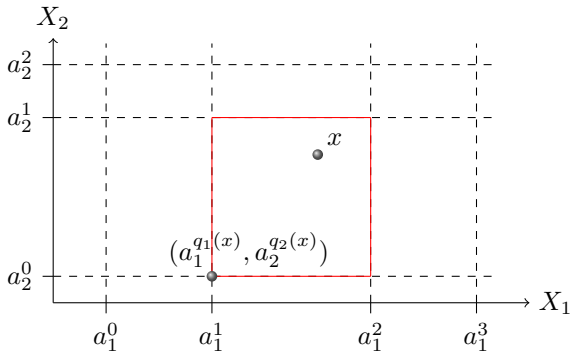


Figure 4: Point x belongs to the square $[a_1^1, a_1^2] \times [a_2^0, a_2^1]$. Hence $q_1(x) = 1$ and $q_2(x) = 0$.

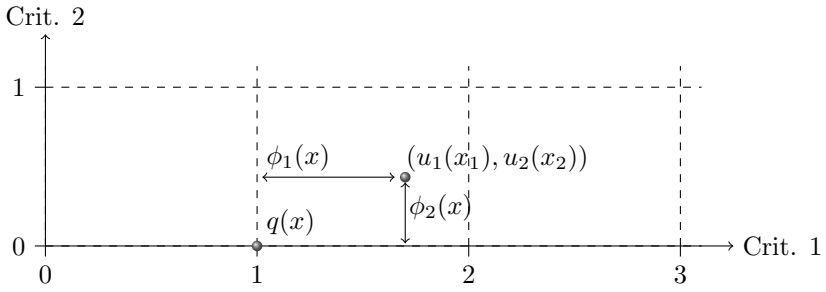


Figure 5: Illustration of $q(x)$, $\phi_1(x)$ and $\phi_2(x)$.

point. This is similar to $q(x)$, which can be defined as the largest element of lattice $\mathcal{Q}_p(N)$ that is smaller than vector $(u_1(x_1), \ldots, u_n(x_n))$. The Voronoi cell of the lattice associated with $q \in \mathcal{Q}_p(N)$ is the set of points for which $q(x) = q$. In our case, the set of point x for which $q(x) = q$ is the hypercube $\times_{i \in N}[a_i^{q_i}, a_i^{q_i+1}]$. Lastly, the modulo-$\mathcal{Q}_p(N)$ operation w.r.t. the lattice *is defined as the difference between the point and the associated nearest neighbour quantizer. We note that $\phi_i(x)$ is the modulo-$\mathcal{Q}_p(N)$ operation w.r.t. the lattice of $u_i(x_i)$ (and could be denoted by $u_i(x_i)$ mod $\mathcal{Q}_p(N)$). The following table compares the concepts used in this paper and in signal compression.*

Our definitions	*Counterpart in signal compression*
$\mathcal{Q}_p(N)$	*lattice*
$q(x)$	*nearest neighbour quantizer*
$V_{(0,\dots,0)}$	*basic Voronoi cell of the lattice*
V_q	*Voronoi cell associated to* q
$\phi_i(x)$	*modulo-*$\mathcal{Q}_p(N)$ *operation w.r.t. the lattice*

Then the Choquet integral w.r.t. v for alternative x is given by

$$U_v^{\wedge}(x) = C_{\mu_{q(x)}}(\phi(x)) + v(q(x)). \tag{20}$$

Note that notation $U_v^{\wedge}$ corresponds to the Lovász extension of v. The Choquet integral w.r.t. a p-ary capacity will be called later *p-ary Choquet integral* by abuse of language.

Let $\Omega_{q,\tau} = \{t \in [q, q+1]^N\,,\ t_{\tau(1)} - q_{\tau(1)} \leq \cdots \leq t_{\tau(n)} - q_{\tau(n)}\}$. The Choquet integral w.r.t. v is clearly a simple weighted sum in each domain $\Omega_{q,\tau}$. In Figure 6, the shaded areas are the sets $\Omega_{q,\tau}$ for all values of q, with $\tau(2) = 1$ and $\tau(1) = 2$.

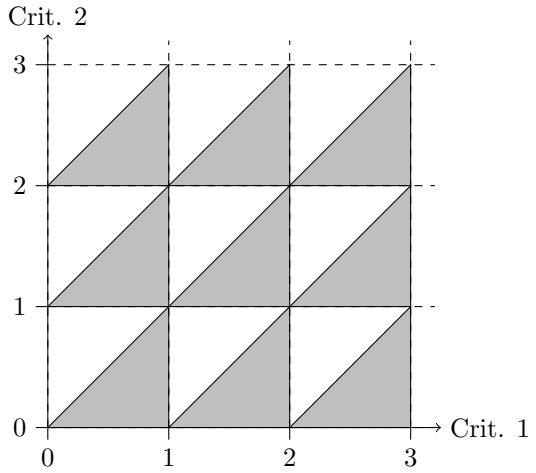


Figure 6: Example with $k = 3$. The shaded triangles represent the areas where $\phi_1(x) \geq \phi_2(x)$.

Example 3. *Let us illustrate the previous construction on Ex. 2. Consider*

the following p-ary capacity – only filled around the diagonal. For instance, $v(0,0) = v(1,0) = v(0,1) = 0$, $v(2,3) = 0.6$ *and* $v(4,4) = 1$.

	'Worst value' (a_1^0)	*'Must have'* (a_1^1)	*'Should have'* (a_1^2)	*'Nice to have'* (a_1^3)	*'Best value'* (a_1^4)
'Best value' (a_2^4)				*0.95*	*1*
'Nice to have' (a_2^3)			*0.6*	*0.8*	*0.95*
'Should have' (a_2^2)		*0.2*	*0.5*	*0.7*	
'Must have' (a_2^1)	*0*	*0.1*	*0.2*		
'Worst value' (a_2^0)	*0*	*0*			

For x in $V_{(0,0)}$*: We have* $q_1(x) = q_2(x) = 0$ *and* $\phi_i(x) = \dfrac{x_i - a_i^0}{a_i^1 - a_i^0}$. *Then* $\mu_{(0,0)}(\emptyset) = \mu_{(0,0)}(\{1\}) = \mu_{(0,0)}(\{2\}) = 0$ *and* $\mu_{(0,0)}(\{1,2\}) = 0.1$. *Hence* $U_v^{\wedge}(x) = 0.1\,\big(\phi_1(x) \wedge \phi_2(x)\big)$. *This indeed corresponds to a veto (see Ex. 2).*

For x in $V_{(1,1)}$*: We have* $q_1(x) = q_2(x) = 1$ *and* $\phi_i(x) = \dfrac{x_i - a_i^1}{a_i^2 - a_i^1}$. *Then* $\mu_{(1,1)}(\emptyset) = 0$, $\mu_{(1,1)}(\{1\}) = \mu_{(1,1)}(\{2\}) = 0.1$ *and* $\mu_{(1,1)}(\{1,2\}) = 0.4$. *Hence* $U_v^{\wedge}(x) = 0.1 + 0.1\,\phi_1(x) + 0.1\,\phi_2(x) + 0.2\,\big(\phi_1(x) \wedge \phi_2(x)\big)$. *This corresponds to an intolerant behaviour (see Ex. 2).*

For x in $V_{(2,2)}$*: We have* $q_1(x) = q_2(x) = 2$ *and* $\phi_i(x) = \dfrac{x_i - a_i^2}{a_i^3 - a_i^2}$. *Then* $\mu_{(2,2)}(\emptyset) = 0$, $\mu_{(2,2)}(\{1\}) = 0.2$, $\mu_{(2,2)}(\{2\}) = 0.1$ *and* $\mu_{(2,2)}(\{1,2\}) = 0.3$. *Hence* $U_v^{\wedge}(x) = 0.5 + 0.2\,\phi_1(x) + 0.1\,\phi_2(x)$. *This corresponds to a complete compensation between the two attributes (see Ex. 2).*

For x in $V_{(3,3)}$*: We have* $q_1(x) = q_2(x) = 3$ *and* $\phi_i(x) = \dfrac{x_i - a_i^3}{a_i^4 - a_i^3}$. *Then* $\mu_{(3,3)}(\emptyset) = 0$, $\mu_{(3,3)}(\{1\}) = \mu_{(3,3)}(\{2\}) = 0.15$ *and* $\mu_{(3,3)}(\{1,2\}) = 0.2$. *Hence* $U_v^{\wedge}(x) = 0.8 + 0.05\,\phi_1(x) + 0.05\,\phi_2(x) + 0.1\,\big(\phi_1(x) \vee \phi_2(x)\big)$. *This corresponds to a tolerant behaviour (see Ex. 2).*

Let us note that there is another extension of capacities – known as interval dependent (or more generally level dependent) capacities [22]. This concept is different from p-ary capacities [22, section 7]. An interval dependent capacity is composed of a partition of [0,1], in which each interval in the partition is assigned a capacity. The Choquet integral w.r.t. interval dependent capacities of a vector $t \in [0,1]^n$ uses all capacities associated to levels between the smallest and the largest values of t. The Choquet integral w.r.t. interval dependent capacities satisfies *cardinal tail independence* [Greco, Matarazzo, Giove., 2011 - Theorem

1]. According to this property, the contribution of the l (for any l) best criteria of an alternative t is independent on the precise values of t on the other criteria (provided that they keep lower scores). In view of Example 2, consider an alternative t below the *"must have"* level on criterion 1, and above the *"nice to have"* levels on the other criteria. Hence, all the different capacities provided in Section 3.2 are used in the computation of the score of t, and thus criteria $N \setminus \{1\}$ are aggregated using the capacities related to the *"must have"*, *"should have"* and *"nice to have"* levels. As these last two capacities are tolerant, t gets a relatively large score, which is not what the DM wants. Hence the *cardinal tail independence* property prevents the level dependent capacities to satisfy the preferences in Ex. 2.

By contrast, the different capacities that are used in a p-ary capacities are assigned to a specific area – namely some domain V_q – so that the contribution of each capacity in the p-ary capacity v is clear.

4.2. GAI model

We have seen that the Choquet integral w.r.t. p-ary capacities can represent the motivating example. Let us see if the GAI model can also solve it. We need to specify, given the collection $\mathcal{S}$, each utility u_S. We treat the utility function u_S for every subset $S \in \mathcal{S}$ separately.

Utility u_S is obtained from a learning process and thus through learning instances. The first idea is to make no assumption on u_S and thus introduce an unknown $u_S(x_S)$ for every learning instance x. This is for instance the approach adopted in $UTA^{GMS}-INT$ [23]. The drawback of this method is that the number of unknowns depends on the number of values x_S for all learning instances. This number can be quite large in some applications. Moreover, the problem is also how to make a recommendation on new alternatives y. Indeed, if y_S is not equal to a value x_S for a learning instance, then how to compute $u_S(y_S)$? This is why a robust preference relation is proposed in $UTA^{GMS}-INT$ [23].

If we consider the additive utility restriction, the most well-known method for

learning it, is the UTA method. It consists in discretizing X_i in $a_i^0, \ldots, a_i^{p_i}$. The unknowns are the utility at these points, and are denoted by $u_i^D(a_i^0), \ldots, u_i^D(a_i^{p_i})$. We set $\mathcal{U}^\Pi := \{u_S^D(y_S) \ : \ S \in \mathcal{S} \ , \ y_S \in D_S\}$. The number of unknowns is equal to $\sum_{S\in\mathcal{S}} \prod_{i\in S}(p_i + 1)$. In UTA, the value of u_i for any x_i is performed by affine interpolation, as for (15). Instead of a piecewise affine model, one could use a parametric model, for instance based on polynomial model [46][chapter 7] or multiple sigmoid functions [48]. The major difficulty is to extend these one-dimensional parametric models to subsets S of larger cardinality. For this reason, we prefer piecewise affine utility functions.

When $S \in \mathcal{S}$ contains more than one attribute, the idea is to perform a multi-linear interpolation (see (7)). Let us recall that $D_i = \{a_i^0, a_i^1, \ldots, a_i^{p_i}\}$ is the set of levels, and $D_S = \times_{i\in S} D_i$. The following set

$$I = \{i \in N \ : \ x_i \in [a_i^0, a_i^{p_i}] \text{ and } x_i \notin D_i\}$$

contains all attributes for which an interpolation is required. Given x_i, we define $\underline{x}_i$ and $\overline{x}_i$ by

$$\begin{aligned}
&\underline{x}_i = a_i^{q_i} \ , \ \overline{x}_i = a_i^{q_i+1} \quad \text{if } x_i \in (a_i^{q_i}, a_i^{q_i+1}) \\
&\underline{x}_i = \overline{x}_i = x_i^0 \quad \text{if } x_i \leq a_i^0 \\
&\underline{x}_i = \overline{x}_i = x_i^{p_i} \quad \text{if } x_i \geq a_i^{p_i} \\
&\underline{x}_i = \overline{x}_i = x_i^{q_i} \quad \text{if } x_i = a_i^{q_i}
\end{aligned}$$

Note that $\overline{x}_i = \underline{x}_i$ iff $i \in N \setminus I$. We wish to generalize the UTA approach and the multi-linear extension (7). For this reason, u_S will be denoted by u_S^Π.

Function u_S^Π shall interpolate u_S^D:

$$u_S^\Pi(z_S) = u_S^D(z_S) \quad \forall z_S \in D_S. \tag{21}$$

Using multi-linear interpolation, we obtain for every $x_S \in X_S$

$$u_S^\Pi(x_S) = \sum_{A\subseteq I\cap S} \left[\prod_{i\in A} \frac{\overline{x}_i - x_i}{\overline{x}_i - \underline{x}_i} \times \prod_{i\in (I\cap S)\setminus A} \frac{x_i - \underline{x}_i}{\overline{x}_i - \underline{x}_i} \times u_S^D(\underline{x}_A, \overline{x}_{(I\cap S)\setminus A}, x_{S\setminus I}) \right] \tag{22}$$

where $(\underline{x}_A, \overline{x}_{(I\cap S)\setminus A}, x_{S\setminus I})$ is an alternative that is equal to $\underline{x}_k$ if $k \in A$, to $\overline{x}_k$ if $k \in (I \cap S) \setminus A$, and to x_k if $k \in S \setminus I$. The GAI model becomes

$$U^{\Pi}(x) = \sum_{S\in\mathcal{S}} u_S^{\Pi}(x_S). \tag{23}$$

Note that the choice of the multi-linear interpolation is motivated by the fact that it is a barycentric interpolation among all extreme points $\prod_{i\in I\cap S}\{\underline{x}_i, \overline{x}_i\}$. Moreover, this is the usual interpolation. Finally it satisfies the *weak difference independence* [11], which is a well-known property in MAUT [29]. This property implies that one can construct interval scales over each attribute independently of the remaining attributes. The scales over each attribute need not be commensurate as expression (22) uses only sums of products (there is no comparison of the values of attributes).

Example 4 (Ex. 2 cont.). *Consider the p-ary capacity given in Example 4.*

For x in $V_{(0,0)}$: We have $\phi_i(x) = \dfrac{x_i - a_i^0}{a_i^1 - a_i^0}$ *and* $U_v^{\Pi}(x) = 0.1\ \phi_1(x) \times \phi_2(x)$. *This indeed corresponds to a veto.*

For x in $V_{(1,1)}$: We have $\phi_i(x) = \dfrac{x_i - a_i^1}{a_i^2 - a_i^1}$ *and* $U_v^{\Pi}(x) = 0.1 + 0.1\ \phi_1(x) + 0.1\ \phi_2(x) + 0.2\ \phi_1(x) \times \phi_2(x)$. *This corresponds to an intolerant behaviour.*

For x in $V_{(2,2)}$: We have $\phi_i(x) = \dfrac{x_i - a_i^2}{a_i^3 - a_i^2}$ *and* $U_v^{\Pi}(x) = 0.5 + 0.2\ \phi_1(x) + 0.1\ \phi_2(x)$. *This corresponds to a complete compensation between the two attributes.*

For x in $V_{(3,3)}$: We have $\phi_i(x) = \dfrac{x_i - a_i^3}{a_i^4 - a_i^3}$ *and* $U_v^{\Pi}(x) = 0.8 + 0.15\ \phi_1(x) + 0.15\ \phi_2(x) - 0.1\ \phi_1(x) \times \phi_2(x)$. *This corresponds to a tolerant behavior.*

One can easily compare these expressions with those given in Example 3. The only difference is a matter of extension – namely Lovász extension for $U^{\wedge}$ and multi-linear for U^{Π} (see Section 2.2). This is also true between a_i^3 and a_i^4, as $U_v^{\wedge}$ can be rewritten as $U_v^{\wedge}(x) = 0.8 + 0.15\,\phi_1(x) + 0.15\,\phi_2(x) - 0.1\,\phi_1(x) \wedge \phi_2(x)$.

5. How to choose between the two aggregation models?

We have shown in Section 4 that the GAI model U^{Π} and the Choquet integral $U^{\wedge}$ w.r.t. p-ary capacities are in fact very similar models, despite the apparent

difference between their mathematical formulation. Both are based on a p-ary capacity and only differ from the interpolation used. As the Lovász extension requires appropriate normalization of the attributes, which is not the case of the multi-linear extension, the Choquet integral model requires utility functions. This explains the differences of their formulation.

The question that arises is which model shall one choose in applications. We first analyse the properties that these two models have in common (Section 5.1). Then we discuss on a specific property satisfied by the Choquet integral that might appear as a drawback in some situations (Section 5.2). Lastly we provide a comparison of the pros and cons of each of these two models (Section 5.3).

5.1. Properties of $U^\wedge$ and U^Π

$U^\wedge$ and U^Π have in common to provide the ability to define a different model in each domain V_q. This allows to have different decision strategies in each area of X.

$U^\wedge$ and U^Π have as unknown v and u^D respectively. Formally these unknowns are linked by relation $u^D(x) = v(r(x))$ for all $x \in D$, where $r(x) = (r_1(x_1), \ldots, r_n(x_n)$ with $r_i(x_i^k) = k$. Conversely, $v(q) = u^D(a_1^{q_1}, \ldots, a_n^{q_n})$ for all $q \in \mathcal{Q}_p(N)$. It is thus equivalent to consider an interpolation of v or u^D. For convenience, we will focus on u^D. Let us show that $U^\wedge$ and U^Π share similar properties as interpolation functions of v and u^D.

5.1.1. General properties

This section describes important properties that the interpolation function shall satisfy. In order to comply with the GAI decomposability property, we are interested in interpolation operators $\mathcal{I}^S$ for every non-empty $S \subseteq N$. Operator $\mathcal{I}^S$ maps a function $f_S^D : D_S \to \mathbb{R}$ to its interpolation $f_S : X_S \to \mathbb{R}$. Let us give some wished properties.

The following property states that $\mathcal{I}^S$ shall be an extension.

Interpolation. For all $x \in D_S$, $\mathcal{I}^S\left(f_S^D\right)(x) = f_S^D(x)$.

We note that this is a generalization of the *Properly Weighted* property of the Choquet integral [39].

Continuity is an essential property of the interpolation.

Continuity. Function $x \mapsto \mathcal{I}^S\left(f_S^D\right)(x)$ is continuous in X_S.

We are interested in interpolating a utility model satisfying the GAI form (11). We must thus perform two operations: interpolation and addition over the subsets in $\mathcal{S}$. The next axiom says that whatever the order in which these two operations are done, the result shall be the same.

Stability of interpolation regarding additivity. For all $x \in X$,

$$\sum_{S\in\mathcal{S}} \mathcal{I}^S\left(f_S^D\right)(x_S) = \mathcal{I}^N\left(\sum_{S\in\mathcal{S}} f_S^D\right)(x).$$

Finally, we impose monotonicity.

Monotonicity. For any $i \in S$, $\mathcal{I}^S\left(f_S^D\right)$ is monotone in $\prod_{i\in N}[a_i^0, a_i^{p_i}]$.

These four properties are simple and appealing in applications. We now show that they are satisfied by the two models.

5.1.2. Satisfaction of these properties by the GAI model

The next lemma shows that the four properties are fulfilled by the multi-linear expression (22).

Lemma 1. *The multi-linear extension satisfies* **Interpolation**, **Continuity**, **Stability of interpolation regarding additivity** *and* **monotonicity**.

Proof : Property **Interpolation** is clearly fulfilled by (21).

On the other hand, we have

$$\begin{aligned}
&\mathcal{I}^N\left(\sum_{S\in\mathcal{S}} u_S^D\right)(x)\\
&= \sum_{A\subseteq I}\left[\prod_{i\in A}\frac{\overline{x}_i - x_i}{\overline{x}_i - \underline{x}_i} \times \prod_{i\in I\setminus A}\frac{x_i - \underline{x}_i}{\overline{x}_i - \underline{x}_i} \times \left(\sum_{S\in\mathcal{S}} u_S^D(\underline{x}_{A\cap S}, \overline{x}_{(S\cap I)\setminus A}, x_{S\setminus I})\right)\right]\\
&= \sum_{S\in\mathcal{S}}\sum_{A\subseteq I\cap S}\left[\prod_{i\in A}\frac{\overline{x}_i - x_i}{\overline{x}_i - \underline{x}_i} \times \prod_{i\in (S\cap I)\setminus A}\frac{x_i - \underline{x}_i}{\overline{x}_i - \underline{x}_i}\ T_S\ u_S^D(\underline{x}_{A\cap S}, \overline{x}_{(S\cap I)\setminus A}, x_{S\setminus I})\right]
\end{aligned}$$

where

$$T_S = \sum_{B\subseteq I\setminus S}\prod_{i\in B}\frac{\overline{x}_i - x_i}{\overline{x}_i - \underline{x}_i}\times\prod_{i\in I\setminus(S\cup B)}\frac{x_i - \underline{x}_i}{\overline{x}_i - \underline{x}_i} = \prod_{i\in I\setminus S}\left[\frac{\overline{x}_i - x_i}{\overline{x}_i - \underline{x}_i} + \frac{x_i - \underline{x}_i}{\overline{x}_i - \underline{x}_i}\right] = 1$$

Hence

$$\mathcal{I}^N\left(\sum_{S\in\mathcal{S}} u_S^D\right)(x) = \sum_{S\in\mathcal{S}}\mathcal{I}^S\left(u_S^D\right)(x_S)$$

Hence **Stability of interpolation regarding additivity** is fulfilled.

To show the continuity, we only have to prove continuity w.r.t. an attribute, say $k \in N$. Let then $x, x', x'' \in X$ such that

$$x_i = x'_i = x''_i \quad \forall i \in N\setminus\{k\}$$
$$x'_k < \overline{x}'_k = x_k = \underline{x}''_k < x''_k \quad \text{and} \quad x_k \in D_k$$

Let $S \ni k$ and $I = \{i \in N \ : \ x_i \in [a_i^0, a_i^{p_i}] \text{ and } x_i \notin D_i\}$. We have $k \notin I$ (as $x_k \in D_k$) and

$$u_S^\Pi(x'_S) = \sum_{A\subseteq I\cap S}\left[\prod_{i\in A}\frac{\overline{x}_i - x_i}{\overline{x}_i - \underline{x}_i}\times\prod_{i\in(I\cap S)\setminus A}\frac{x_i - \underline{x}_i}{\overline{x}_i - \underline{x}_i}\times\frac{x'_k - \underline{x}'_k}{\overline{x}'_k - \underline{x}'_k}\times u_S^D(\underline{x}_A, \overline{x}_{(I\cap S)\setminus A}, \overline{x}'_k, x_{S\setminus(I\cup k)})\right]$$
$$+\sum_{A\subseteq I\cap S}\left[\prod_{i\in A}\frac{\overline{x}_i - x_i}{\overline{x}_i - \underline{x}_i}\times\prod_{i\in(I\cap S)\setminus A}\frac{x_i - \underline{x}_i}{\overline{x}_i - \underline{x}_i}\times\frac{\overline{x}'_k - x'_k}{\overline{x}'_k - \underline{x}'_k}\times u_S^D(\underline{x}_A, \underline{x}'_k, \overline{x}_{(I\cap S)\setminus A}, x_{S\setminus(I\cup k)})\right]$$
$$\xrightarrow[x'_k\to x_k]{}\sum_{A\subseteq I\cap S}\left[\prod_{i\in A}\frac{\overline{x}_i - x_i}{\overline{x}_i - \underline{x}_i}\times\prod_{i\in(I\cap S)\setminus A}\frac{x_i - \underline{x}_i}{\overline{x}_i - \underline{x}_i}\times u_S^D(\underline{x}_A, \overline{x}_{(I\cap S)\setminus A}, x_{S\setminus I})\right] = u_S^\Pi(x_S)$$

and

$$u_S^\Pi(x''_S) = \sum_{A\subseteq I\cap S}\left[\prod_{i\in A}\frac{\overline{x}_i - x_i}{\overline{x}_i - \underline{x}_i}\times\prod_{i\in(I\cap S)\setminus A}\frac{x_i - \underline{x}_i}{\overline{x}_i - \underline{x}_i}\times\frac{x''_k - \underline{x}''_k}{\overline{x}''_k - \underline{x}''_k}\times u_S^D(\underline{x}_A, \overline{x}_{(I\cap S)\setminus A}, \overline{x}''_k, x_{S\setminus(I\cup k)})\right]$$
$$+\sum_{A\subseteq I\cap S}\left[\prod_{i\in A}\frac{\overline{x}_i - x_i}{\overline{x}_i - \underline{x}_i}\times\prod_{i\in(I\cap S)\setminus A}\frac{x_i - \underline{x}_i}{\overline{x}_i - \underline{x}_i}\times\frac{\overline{x}''_k - x''_k}{\overline{x}''_k - \underline{x}''_k}\times u_S^D(\underline{x}_A, \underline{x}''_k, \overline{x}_{(I\cap S)\setminus A}, x_{S\setminus(I\cup k)})\right]$$
$$\xrightarrow[x''_k\to x_k]{}\sum_{A\subseteq I\cap S}\left[\prod_{i\in A}\frac{\overline{x}_i - x_i}{\overline{x}_i - \underline{x}_i}\times\prod_{i\in(I\cap S)\setminus A}\frac{x_i - \underline{x}_i}{\overline{x}_i - \underline{x}_i}\times u_S^D(\underline{x}_A, \overline{x}_{(I\cap S)\setminus A}, x_{S\setminus I})\right] = u_S^\Pi(x_S)$$

Hence **continuity** is proved.

Finally, it is well-known that the multi-linear extension is monotone in $[0,1]^N$. Hence U^{Π} is monotone is monotone in each domain V_q and thus in $\prod_{i\in N}[a_i^0, a_i^{p_i}]$. ■

5.1.3. Satisfaction of these properties by the Choquet integral

We show some properties for the Choquet integral in this section.

Lemma 2. *$U_v^\wedge$ satisfies* **Interpolation**, **Continuity**, **Monotonicity** *and* **Stability of interpolation regarding additivity** *(If $u^D(x) = \sum_{S\in\mathcal{S}} u_S^D(x_S)$, and $u^D(x) = v(r(x))$, $u_S^D(x_S) = v_S(r(x_S))$, then $U_v^\wedge(x) = \sum_{S\in\mathcal{S}} U_{v_S}^\wedge(x_S)$).*

Proof : Let us first show that $U_v^\wedge$ satisfies **Continuity**. Let $x \in X$ such that $J := \left\{i \in N \ : x_i \in D_i \setminus \{a_i^0, a_i^{p_i}\}\right\} \neq \emptyset$. The Choquet integral is continuous. Hence $U_v^\wedge$ is continuous in each domain V_q. We need to show that it is continuous across these domains.

Let $q \in \mathcal{Q}_p(N)$ such that $x \in V_q$. For $i \in N \setminus J$, x_i belongs to $[a_i^{q_i}, a_i^{q_i+1}]$ and not to another interval. For $i \in J$, x_i can belong to $[a_i^{q_i-1}, a_i^{q_i}]$ or $[a_i^{q_i}, a_i^{q_i+1}]$. To show the continuity, we need to show that $U_v^\wedge(x)$ does not depend on which interval x_i is associated to. Let $S \subseteq J$ such that x_i is associated to $[a_i^{q_i-1}, a_i^{q_i}]$ (resp. $[a_i^{q_i}, a_i^{q_i+1}]$) for $i \in S$ (resp. $i \in J \setminus S$). Then $\phi_i(x_i) = 1$ if $i \in S$, and $\phi_i(x_i) = 0$ if $i \in J \setminus S$. The elements of $N \setminus J$ are denoted by $\{\pi(1), \ldots, \pi(f)\}$, where $f = |N \setminus J|$ and $0 \leq \phi_{\pi(1)}(x_{\pi(1)}) \leq \cdots \leq \phi_{\pi(f)}(x_{\pi(f)}) \leq 1$. We set $q' \in \mathcal{Q}_p(N)$ by $q'_i = q_i$ if $i \in N \setminus S$ and $q'_i = q_i - 1$ if $i \in S$. Then $x \in V_{q'}$. Moreover

$$
\begin{aligned}
&U_v^\wedge(x) = v(q') + C_{\mu_{q'}}(\phi(x)) \\
&= v(q') + \sum_{i=1}^{f} \phi_{\pi(i)}(x_{\pi(i)}) \left[\mu_{q'}(\{\pi(i), \ldots, \pi(q)\} \cup S) - \mu_{q'}(\{\pi(i+1), \ldots, \pi(q)\} \cup S)\right] + \mu_{q'}(S)
\end{aligned}
$$

For $A \subseteq N \setminus J$ and $B \subseteq J$, $\mu_{q'}(A \cup B) = v((q'+1)_{A\cup B}, q'_{N\setminus(A\cup B)}) - v(q')$. Then

$$
\begin{aligned}
\mu_{q'}(A \cup S) &= v((q'+1)_A, q'_{N\setminus(J\cup A)}, ((q-1)+1)_S, q_{J\setminus S}) - v(q') \\
&= v((q+1)_A, q_{N\setminus(J\cup A)}, q_J) - v(q')
\end{aligned}
$$

We note that $\mu_{q'}(A \cup S)$ is independent of S. Hence $U_v^\wedge(x)$ is continuous.

Taking the previous expressions with $x \in D$, we obtain that $U_v^\wedge(a_1^{q_1}, \ldots, a_n^{q_n}) = v(q) = U^D(a_1^{q_1}, \ldots, a_n^{q_n})$ for every $q \in \mathcal{Q}_p(N)$. Hence $U_v^\wedge$ satisfies **Interpolation**.

The **Monotonicity** property is satisfied. Indeed, the Choquet integral is known to be monotone, so that $U^\wedge$ is monotone in each domain V_q. Moreover, monotonicity is ensured across the boundaries, as $U^\wedge$ is continuous (in particular around the boundaries of domains V_q).

Consider finally axiom **Stability of interpolation regarding additivity**. If $u^D(x) = \sum_{S \in \mathcal{S}} u_S^D(x_S)$, and $u^D(x) = v(r(x))$, $u_S^D(x_S) = v_S(r(x_S))$, then $v(A) = \sum_{S \in \mathcal{S}} v_S(A \cap S)$. The result follows from the linearity of the Choquet integral with respect to the capacity. ■

5.2. Contamination effect of the Choquet integral

When the Choquet integral satisfies some properties in a given area of the space, then this property is usually also satisfied in a neighbourhood. This is what we call *contamination effect*. We show that this is not the case with the GAI model. We first describe this effect on the condition relative importance (section 5.2.1). Then we study more general decision strategies (section 5.2.2).

The contamination effect of the Choquet integral is important in practice, as it might forbid the DM to express very different decision strategies in two neighbour cells.

5.2.1. Case of conditional relative importance

Let us consider *conditional relative importance* [35] that is the baseline of most of motivations of the Choquet integral [19, 34]. Given three criteria, with the first criterion being bipolar, the decision strategy is as follows: criterion 2 is more important than criterion 3 when criterion 1 is repulsive (below 0), and

criterion 2 is less important than criterion 3 when criterion 1 is attractive (above the neutral level 0). Hence with $X_1 = [-1, 1]$ and $X_2 = X_3 = [0, 1]$

$$\forall x_1 \in [-1, 0)\, \forall x_2, x_3 \in [0, 1] \qquad \frac{\partial U}{\partial x_2}(x_1, x_2, x_3) > \frac{\partial U}{\partial x_3}(x_1, x_2, x_3) \tag{24}$$

$$\forall x_1 \in (0, 1]\, \forall x_2, x_3 \in [0, 1] \qquad \frac{\partial U}{\partial x_2}(x_1, x_2, x_3) < \frac{\partial U}{\partial x_3}(x_1, x_2, x_3) \tag{25}$$

where the importance of criterion k at point x is measured by $\frac{\partial U}{\partial x_k}(x)$ [32]. It has been shown that this cannot be strictly represented by the Choquet integral or its extensions (and even any piecewise affine functions) [35]. More precisely, if (24) is satisfied, then necessarily, we also have $\frac{\partial U}{\partial x_2}(x_1, x_2, x_3) > \frac{\partial U}{\partial x_3}(x_1, x_2, x_3)$ whenever $0 \leq x_1 \leq \min(x_2, x_3)$ [35].

The next lemma shows that this decision strategy can easily be modelled by the GAI model with a multi-linear interpolation.

Lemma 3. *We write $U(x_1, x_2, x_3) = U_{1,2}(x_1, x_2) + U_{1,3}(x_1, x_3)$. We assume 3 reference levels $-1, 0, +1$ on attribute 1, and two reference levels $0, 1$ on the other attributes. Then the GAI model with multi-linear interpolation can simultaneously fulfil (24) and (25).*

Proof : We consider the unknowns $u^D_{1,2} : \{-1, 0, 1\} \times \{0, 1\} \to \mathbb{R}$ of $u_{1,2}$. If $x_1 < 0$, then, by (22)

$$\begin{aligned}
&u_{1,2}(x_1, x_2) = u^D_{1,2}(-1, 0) \times (-x_1) \times (1 - x_2) + u^D_{1,2}(0, 0) \times (x_1 + 1) \times (1 - x_2) \\
&+ u^D_{1,2}(-1, 1) \times (-x_1) \times x_2 + u^D_{1,2}(0, 1) \times (x_1 + 1) \times x_2 \\
&= u^D_{1,2}(0, 0) + x_1 \left[u^D_{1,2}(0, 0) - u^D_{1,2}(-1, 0)\right] + x_2 \left[u^D_{1,2}(0, 1) - u^D_{1,2}(0, 0)\right] \\
&+ x_1\, x_2 \left[u^D_{1,2}(-1, 0) - u^D_{1,2}(0, 0) - u^D_{1,2}(-1, 1) + u^D_{1,2}(0, 1)\right]
\end{aligned}$$

and thus

$$\frac{\partial u_{1,2}}{\partial x_2}(x_1, x_2) = u^D_{1,2}(0, 1) - u^D_{1,2}(0, 0) + x_1 \left[u^D_{1,2}(0, 1) - u^D_{1,2}(0, 0) + u^D_{1,2}(-1, 0) - u^D_{1,2}(-1, 1)\right].$$

Now if $x_1 > 0$, then

$$\begin{aligned}
&u_{1,2}(x_1,x_2) = u^D_{1,2}(0,0)\times(1-x_1)\times(1-x_2) + u^D_{1,2}(1,0)\times x_1\times(1-x_2)\\
&+ u^D_{1,2}(0,1)\times(1-x_1)\times x_2 + u^D_{1,2}(1,1)\times x_1\times x_2\\
&= u^D_{1,2}(0,0) + x_1\left[u^D_{1,2}(1,0) - u^D_{1,2}(0,0)\right] + x_2\left[u^D_{1,2}(0,1) - u^D_{1,2}(0,0)\right]\\
&+ x_1\,x_2\left[u^D_{1,2}(0,0) - u^D_{1,2}(1,0) - u^D_{1,2}(0,1) + u^D_{1,2}(1,1)\right]
\end{aligned}$$

and thus

$$\frac{\partial u_{1,2}}{\partial x_2}(x_1,x_2) = u^D_{1,2}(0,1) - u^D_{1,2}(0,0) + x_1\left[u^D_{1,2}(0,0) - u^D_{1,2}(1,0) - u^D_{1,2}(0,1) + u^D_{1,2}(1,1)\right].$$

We consider the unknowns $u^D_{1,3} : \{-1,0,1\}\times\{0,1\} \to \mathbb{R}$ of $u_{1,3}$. If $x_1 < 0$, then

$$\frac{\partial u_{1,3}}{\partial x_3}(x_1,x_3) = u^D_{1,3}(0,1) - u^D_{1,3}(0,0) + x_1\left[u^D_{1,3}(0,1) - u^D_{1,3}(0,0) + u^D_{1,3}(-1,0) - u^D_{1,3}(-1,1)\right].$$

Now if $x_1 > 0$, we obtain

$$\frac{\partial u_{1,3}}{\partial x_3}(x_1,x_3) = u^D_{1,3}(0,1) - u^D_{1,3}(0,0) + x_1\left[u^D_{1,3}(0,0) - u^D_{1,3}(1,0) - u^D_{1,3}(0,1) + u^D_{1,3}(1,1)\right].$$

Condition (24) gives

$$\begin{aligned}
&\forall x_1\in[-1,0)\,\forall x_2,x_3\in[0,1] \qquad \frac{\partial u_{1,2}}{\partial x_2}(x_1,x_2) > \frac{\partial u_{1,3}}{\partial x_3}(x_1,x_3)\\
&\iff \forall x_1\in[-1,0) \qquad (-x_1)\left[u^D_{1,2}(-1,1) - u^D_{1,2}(-1,0)\right] + (1+x_1)\left[u^D_{1,2}(0,1) - u^D_{1,2}(0,0)\right]\\
&\qquad\qquad > (-x_1)\left[u^D_{1,3}(-1,1) - u^D_{1,3}(-1,0)\right] + (1+x_1)\left[u^D_{1,3}(0,1) - u^D_{1,3}(0,0)\right]\\
&\iff \qquad u^D_{1,2}(-1,1) - u^D_{1,2}(-1,0) > u^D_{1,3}(-1,1) - u^D_{1,3}(-1,0)\\
&\qquad\qquad \text{and } u^D_{1,2}(0,1) - u^D_{1,2}(0,0) \geq u^D_{1,3}(0,1) - u^D_{1,3}(0,0)
\end{aligned}$$

Condition (25) gives

$$\begin{aligned}
&\forall x_1\in(0,1]\,\forall x_2,x_3\in[0,1] \qquad \frac{\partial u_{1,2}}{\partial x_2}(x_1,x_2) < \frac{\partial u_{1,3}}{\partial x_3}(x_1,x_3)\\
&\iff \forall x_1\in(0,1] \qquad (1-x_1)\left[u^D_{1,2}(0,1) - u^D_{1,2}(0,0)\right] + x_1\left[u^D_{1,2}(1,1) - u^D_{1,2}(1,0)\right]\\
&\qquad\qquad < (1-x_1)\left[u^D_{1,3}(0,1) - u^D_{1,3}(0,0)\right] + x_1\left[u^D_{1,3}(1,1) - u^D_{1,3}(1,0)\right]\\
&\iff \qquad u^D_{1,2}(0,1) - u^D_{1,2}(0,0) \leq u^D_{1,3}(0,1) - u^D_{1,3}(0,0)\\
&\qquad\qquad \text{and } u^D_{1,2}(1,1) - u^D_{1,2}(1,0) < u^D_{1,3}(1,1) - u^D_{1,3}(1,0)
\end{aligned}$$

Hence preferences (24) and (25) are equivalent to the following conditions

$$u^D_{1,2}(-1,1) - u^D_{1,2}(-1,0) > u^D_{1,3}(-1,1) - u^D_{1,3}(-1,0) \tag{26}$$

$$u^D_{1,2}(0,1) - u^D_{1,2}(0,0) = u^D_{1,3}(0,1) - u^D_{1,3}(0,0) \tag{27}$$

$$u^D_{1,2}(1,1) - u^D_{1,2}(1,0) < u^D_{1,3}(1,1) - u^D_{1,3}(1,0) \tag{28}$$

These constraints are not conflicting so that the GAI model with multi-linear interpolation can fulfill (24) and (25). ■

5.2.2. A more general situation

For the sake of simplicity, we assume that $a^k_i = k$ for every level k.

Lemma 4. *For a given $q \in \mathcal{Q}_p(N)$ with $q_i < p_i$ for every $i \in N$, and $j \in N$, we have*

$$\frac{\partial U^\wedge}{\partial x_j}(x) = v(q_j + 1, q_{N\setminus j}) - v(q) \tag{29}$$

for every $x \in \mathbb{R}^N$ such that $q(x)$ (see (17)) has the form $(q_S - 1, q_{N\setminus S})$ for some $S \subseteq N \setminus j$, and $\phi_k(x_k) \leq \phi_j(x_j) \leq \phi_i(x_i)$ (see (19)) for all $i \in S$ and $k \in N \setminus (S \cup j)$.

This lemma shows that the same weight $v(q_j + 1, q_{N\setminus j}) - v(q)$ of criterion j can be obtained from all cells

$$\mathcal{C} = \Big\{[r, r+1] \text{ with } r = (q_S - 1, q_{N\setminus S}) \text{ and } S \subseteq N \setminus j\Big\}.$$

This is contamination as if the relative importance of criterion j is fixed in one of these cells, then the relative importance of criterion j necessarily takes the same value for all other cells in $\mathcal{C}$. This contamination phenomenon is very important in practice. It can be expected in some applications as it ensures some kind of smoothness of the decision strategies across the cells. In other applications, it can be undesirable as it forbids representing very different decision strategies in neighbour cells.

Proof : We have

$$\frac{\partial U^{\wedge}}{\partial x_j}(x) = \frac{\partial u_j(x_j)}{\partial x_j} \times \left[\mu_{q(x)}(\{j, \pi(k+1), \ldots, \pi(n)\}) - \mu_{q(x)}(\{\pi(k+1), \ldots, \pi(n)\})\right]$$

where $\phi_{\pi(1)}(x_{\pi(1)}) \leq \cdots \leq \phi_{\pi(n)}(x_{\pi(n)})$ and $\pi(k) = j$. As $a_i^k = k$ for all k, then $\frac{\partial u_j(x_j)}{\partial x_j} = 1$.

Let $S \subseteq N \setminus j$. Assume that $q(x) = (q_S - 1, q_{N\setminus S})$ and $\phi_k(x_k) \leq \phi_j(x_j) \leq \phi_i(x_i)$ for all $i \in S$ and $k \in N \setminus (S \cup j)$. Then

$$\begin{aligned}
\frac{\partial U^{\wedge}}{\partial x_j}(x) &= \left[\mu_{q(x)}(S \cup \{j\}) - \mu_{q(x)}(S)\right] \\
&= v(q_S, q_j + 1, q_{N\setminus(S\cup\{j\})}) - v(q_S, q_{N\setminus(S)}) \\
&= v(q_j + 1, q_{N\setminus j}) - v(q)
\end{aligned}$$

■

The previous contamination phenomenon does not occur the GAI model as the multi-linear extension interpolates with non-zero coefficients among all vertices of the cell $[q(x), q(x) + 1]$.

5.3. Synthesis: pros and cons of the two models

We have seen that the two models satisfied the four basic properties (**Interpolation**, **Continuity**, **Monotonicity** and **Stability of interpolation regarding additivity**). Moreover, unlike the GAI model, the Choquet integral satisfies a contamination property: some decision strategies fulfilled by the Choquet integral in a cell are necessarily fulfilled also in some neighbour cells. This property comes from the parsimonious interpolation used in the Lovász extension.

To continue the comparison between these two models, we propose the following table presenting the assets and disadvantages of these models.

Advantages of U^{Π}	**Advantages of $U^{\wedge}$**
• No separation between utilities and weights. Hence the model is linear in its parameters and is thus easier to learn (same as UTA). • no "contamination" for GAI (see Section 5.2) • no commensurability requirement among the partial utilities u_S	• interpretability of model: the separation between utilities and weights allows us to better interpret the model and thus explain a decision • the Choquet integral is the most parsimonious interpolation method in the sense that it uses only $n+1$ point in n-dimension [21]
Disadvantages of U^{Π}	**Disadvantages of $U^{\wedge}$**
• Needs to have bounded attributes (otherwise, multilinear extension is non-monotone)	• Separation between utilities and weights. Hence it is complex to learn all parameters of the model at the same time. • The "contamination" phenomenon is unavoidable (see Section 5.2) • Commensurability across the criteria is required, which can be complex to obtain

6. How to elicit the two models?

This section addresses the elicitation of the previous models, which is very important in practice. We have seen in Section 5.1 that the parameters of a GAI model can be put in the form of a p-ary capacity. Hence in order to address the elicitation of the two models that are considered in this paper, it is sufficient to consider the elicitation of a p-ary capacity.

6.1. Elicitation by linear programming

A capacity contains 2^n parameters, whose elicitation requires a significant amount of preference information. In order to reduce the elicitation burden, some sub-models of capacities have been defined. One can mention the 2-additive capacities using $\frac{n(n+1)}{2}$ parameters. These models are identified using Linear Programming (LP) in which the variables are the parameters of the (2-additive) capacity, the constraints are the monotonicity conditions and the fulfilment of the preference information of the DM (e.g. some comparisons of options), and the objective function aims for instance at maximizing the entropy of the capacity [30].

One could think of directly extending this approach to p-ary capacities, which are defined by $\prod_{i\in N}(p_i+1)$ parameters. However, the DM cannot reasonably provide enough preference information in order to identify this model, and one has to find suitable submodels. A two-additive p-ary capacity is a p-ary capacity of the form:

$$v(q) = \sum_{\{i,j\}\in\mathcal{S}_2} v_{i,j}(q_i, q_j) + \sum_{i\in N} v_i(q_i), \tag{30}$$

where $\mathcal{S}_2$ is a subset of the set of pairs in N, $v_{i,j}$ depends only on criteria i and j, and v_i depends only on criterion i. It contains thus $\sum_{\{i,j\}\in\mathcal{S}_2}(p_i+1)\times(p_j+1)+\sum_{i\in N}(p_i+1)$ parameters. It has been shown in this case that the number of monotonicity conditions on v can be reduced from an exponential number to a quadratic number, as it is sufficient to enforce monotonicity conditions on terms v_i and $v_{i,j}$ with no restriction on modelling power of v [20].

The elicitation of a two-additive p-ary capacity using LP has been described in [23] for a utility model which has strong connections with the GAI model, and also in [36]. In this LP, the unknowns are the parameters of the p-ary capacity, and the constraints are the monotonicity conditions and the fulfilment of the preference information given by the DM. In [23, 36], a Mixed Integer LP is used in order to find the p-ary capacity of the form (30), satisfying the previous constraints, and minimizing the number of terms in $\mathcal{S}_2$.

6.2. Elicitation by expert knowledge

We propose here a more parsimonious approach, where some domain expertise brought by the DM allows the analyst to save some elicitation time. This expertise concerns the decision strategies related to the various reference levels and their relationship.

We assume for simplicity that we have the same number of levels for all attributes: $p_1 = p_2 = \cdots = p_n =: k$. Moreover, levels $a_1^l, a_2^l, \ldots, a_n^l$ (for any $l \in \{0, 1, \ldots, k\}$) are supposed to have the same semantics and thus to be commensurate, as in Example 2. We denote by $a^0, a^2, \ldots, a^k$ the $k+1$ semantic levels. They are ordered from the lowest one a^0 to the best one a^k.

We suppose that the $k+1$ reference levels delimit k different decision strategies, namely between a^0 and a^1, between a^1 and a^2, etc.

In order to elicit the p-ary capacity v, we proceed in three steps:

- **Elicitation of the p-ary capacity on the diagonal :** We construct a utility function w on the $k+1$ semantic levels $\{a^0, a^2, \ldots, a^k\}$, with the normalization $w(a^0) = 0$ and $w(a^k) = 1$. This can be done for instance with the help of the MACBETH approach [2]. Due to the commensurability assumption among levels $a_1^l, a_2^l, \ldots, a_n^l$, and the idempotency property, alternative $(a_1^l, a_2^l, \ldots, a_n^l)$ has the same level $w(a^l)$ on all attributes and thus $w(a^l)$ should be its overall evaluation. Hence for every $l \in \{0, \ldots, k\}$

$$v(l, \ldots, l) = U(a_1^l, a_2^l, \ldots, a_n^l) = w(a^l). \tag{31}$$

- **Elicitation of the p-ary capacity around the diagonal :** For each $l \in \{0, 1, \ldots, k-1\}$, we construct the decision strategy of the DM when all criteria take value between a^l and a^{l+1}, i.e., inside domain $V_{(l,\ldots,l)}$. It is represented by a non-normalized capacity $\nu_l : 2^n \to \mathbb{R}$ (which can be 2-additive) such that $\nu_l(\emptyset) = 0$ and $\nu_l(N) = w(a^{l+1}) - w(a^l)$, the latter condition ensuring that (31) holds. This capacity can be constructed by standard capacity elicitation techniques, and permits to determine the p-ary capacity on $V_{(l,\ldots,l)}$:

$$\forall S \subseteq N \qquad v((l+1)_S, l_{N\setminus S}) = \nu_l(S) + w(a^l). \tag{32}$$

  Observe that (31) holds.

- **Implicit determination of the p-ary capacity outside the diagonal:** Formula (32) specifies v at points $\big\{((l+1)_S, l_{N\setminus S}), l \in \{0, 1, \ldots, k-1\}, S \subseteq N\big\}$ (i.e., around the diagonal). It remains to fix v for the other points. We propose to do so without asking any further information from the DM, in order to reduce the elicitation burden.

  Let us fix $q = (q_1, \ldots, q_n) \in \mathcal{Q}_q(N)$. The basic idea is that the decision strategy in domain V_q is inherited from the decision strategy in domain on the diagonal associated to the worst level of q, namely $V_{\underline{q}}$, where $\underline{q} = (h, \ldots, h)$, and $h = \min(q_1, \ldots, q_n)$. This induces a decision behaviour focused on the worst levels, which seems to be natural in most applications. In Example 2, when assessing an alternative x, the architect looks firstly at whether x is above all *"must have"* thresholds, secondly at whether x is above the *"should have"* thresholds, and lastly at whether some *"nice to have"* requirements are met. For instance, fulfilling a *"nice to have"* requirement cannot compensate being below *"must have"* on another attribute.

  In order to define the p-ary capacity for the points in V_q, we first give the expression of $U_v^{\wedge}(x)$ for $x \in V_{\underline{q}}$, and then extend this formula to points in

V_q. By (20), the overall score of an alternative $x \in V_{\underline{q}}$ is given by

$$U_v^\wedge(x) = C_{\mu_{\underline{q}}}(\phi(x)) + v(\underline{q}) = C_{\nu_h}(u_1(x_1) - h, \ldots, u_n(x_n) - h) + w(a^h), \quad (33)$$

as $\mu_{\underline{q}} = \nu_h$. In order to extend the right hand side of (33) to $x \in V_q$, one may directly use this formula with $x \in V_q$. However, $u_i(x_i) - h$ is equal to 0 at $x_i = a_i^h$, to 1 at a_i^{h+1}, to 2 at a_i^{h+2}, ..., to $q_i - h$ at $a_i^{q_i}$, ... These values are somewhat arbitrary. As a matter of fact, utility values $0, 1, \ldots, k$ in u_i (see (15)) have no precise numerical meaning, and shall be seen only as successive labels. Instead, we replace u_i by a utility function derived from w, as w represents the DM preferences regarding the difference of satisfaction between the different levels. It is defined by

$$w_i(x_i) = \begin{cases} w(a^0) & \text{if } x_i \leq a_i^0 \\ w(a^l) + (w(a^{l+1}) - w(a^l)) \frac{x_i - a_i^l}{a_i^{l+1} - a_i^l} & \text{if } a_i^l \leq x_i \leq a_i^{l+1} \\ w(a^k) & \text{if } x_i \geq a_i^k \end{cases}$$

so that $w_i(a_i^l) = w(a^l)$ for every $l \in \{0, \ldots, k\}$. As $u_i(x_i) - h = \frac{w_i(x_i) - w(a^h)}{w(a^{h+1}) - w(a^h)}$ for $x_i \in [a_i^h, a_i^{h+1}]$, we have

$$U_v^\wedge(x) = C_{\nu_h}\left(\frac{w_1(x_1) - w(a^h)}{w(a^{h+1}) - w(a^h)}, \ldots, \frac{w_n(x_n) - w(a^h)}{w(a^{h+1}) - w(a^h)}\right) + w(a^h) \quad (34)$$

for every $x \in V_{\underline{q}}$. This formula will be used for $x \in V_q$. Hence the value of the p-capacity at q is given by:

$$v(q) = C_{\nu_h}\left(\frac{w_1(q_1) - w(a^h)}{w(a^{h+1}) - w(a^h)}, \ldots, \frac{w_n(q_n) - w(a^h)}{w(a^{h+1}) - w(a^h)}\right) + w(a^h). \quad (35)$$

Example 5 (Ex. 2 continued). *Figure 7 gives the expression of $U_v^\wedge$ for the four decision strategies. Between levels a^0 and a^1, the overall utility is given by $z_1 \wedge z_2$ (for $z_1, z_2 \in [0, 0.1]$). This expression is used to derive $v(l, 0) = 0$ and $v(0, l) = 0$ (for $l \in \{2, 3, 4\}$). Between levels a^1 and a^2, the overall utility is given by $\frac{z_1 + z_2}{4} + \frac{z_1 \wedge z_2}{2}$ (for $z_1, z_2 \in [0.1, 0.5]$). This expression is used to derive $v(l, 1)$ and $v(1, l)$ (for $l \in \{3, 4\}$). For instance $v(1, 3) = 0.3$ is obtained by applying the previous expression with $z_1 = 0.1$ and $z_2 = 0.9$. Between levels a^2 and a^3, the overall utility is given by $0.375\, z_1 + 0.625\, z_2$ (for $z_1, z_2 \in [0.5, 0.9]$). This*

expression is used to derive $v(4,2)$ *and* $v(2,4)$*. For instance* $v(4,2) = 0.8125$ *is obtained by applying the previous expression with* $z_1 = 0.5$ *and* $z_2 = 1$*. All values of* v *are given in Figure 8.*

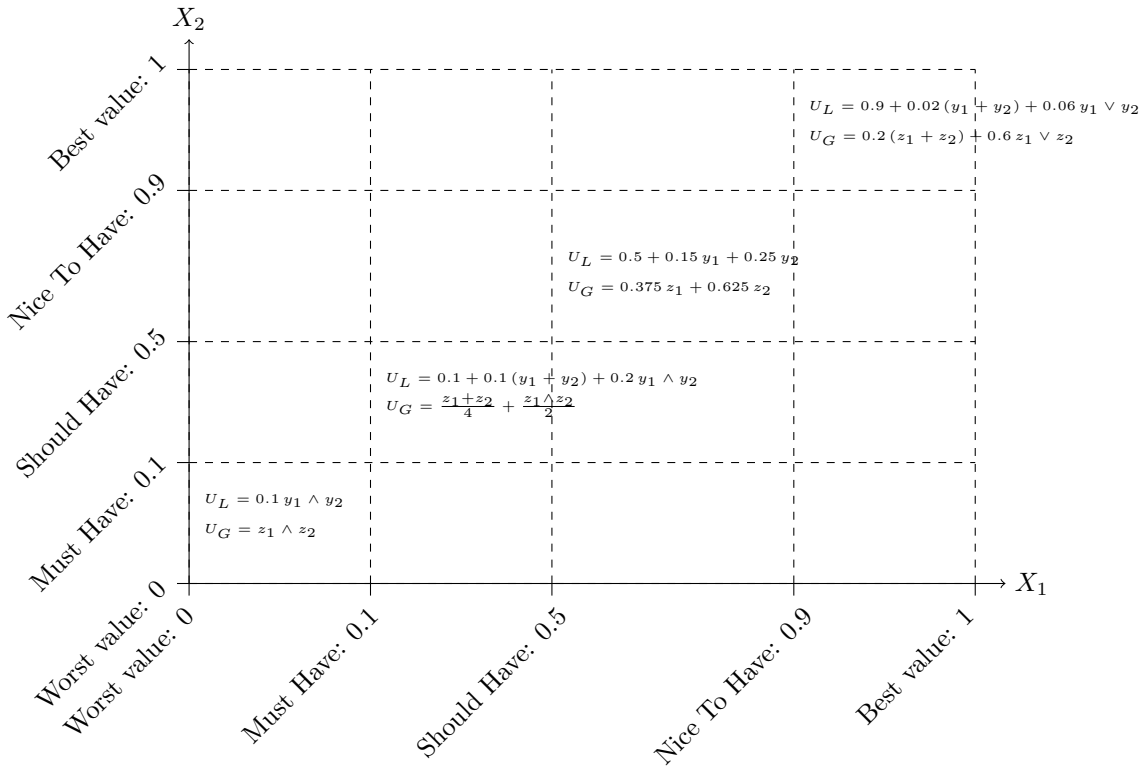


Figure 7: Expression of the overall utility in Example 2. Formula $U_v^{\wedge}(x) = C_{\nu_h}(u_1(x_1) - h, \ldots, u_n(x_n) - h) + w(a^h)$ is written in two different ways. In the first one, we use new variables $y_i := u_i(x_i) - h$. Then $U_v^{\wedge}(x)$ is equal to $U_L(y) := C_{\nu_h}(y) + w(a^h)$ (where L refers to the fact that variables y are Local to domain $V_{\underline{q}}$). The second expression is based on (34). With the new variables $z_i = w_i(x_i)$, (34) is equal to $U_G(z) := C_{\nu_h}\left(\frac{z_1 - w(a^h)}{w(a^{h+1}) - w(a^h)}, \ldots, \frac{z_n - w(a^h)}{w(a^{h+1}) - w(a^h)}\right) + w(a^h)$ (where G stands for Global). The figure presents the two expressions $U_L(y)$ and $U_G(z)$ for the four capacities. They are presented at the four diagonal blocs.

7. Conclusion

This paper considers MCDA situations in which the decision strategies are conditional on the position of the values of the alternative on the attributes

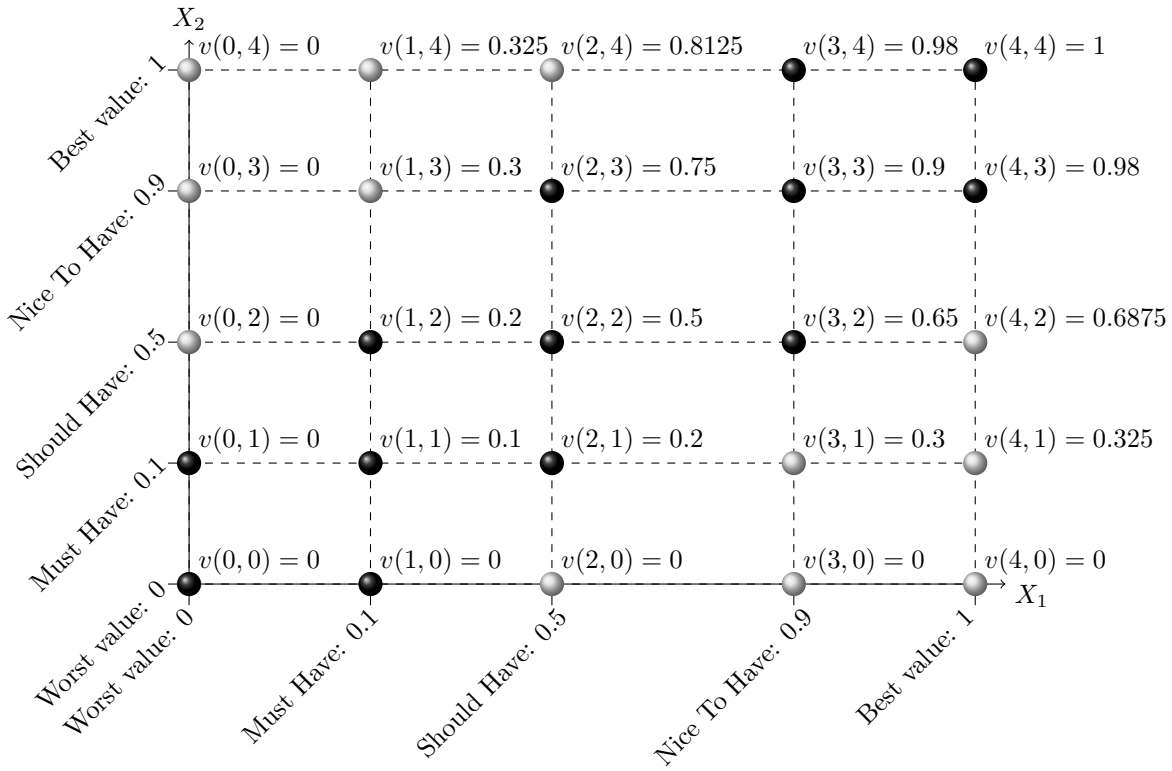


Figure 8: Value of the p-ary capacity in Example 2. The black balls correspond to the values of v that are directly given by $\nu_0, \nu_1, \nu_2, \nu_3$. The other values of v (gray balls) are obtained by (35).

comparatively to several reference levels. Two examples of such situations have been given. Standard MCDA approaches cannot handle these cases. We have seen that the Choquet w.r.t. a p-ary capacity and a specialization of the GAI model, which is discretized to fit with the presence of reference levels, can fulfil these examples. The latter model can be seen as an extension of the UTA approach. A convenient representation of the unknowns of these two models is given by a function u^D defined on the Cartesian product D of the reference levels over all attributes. The overall model is then obtained as an interpolation of this function u^D. Four basic properties have been defined on this interpolation: **Interpolation** (the overall utility shall be identical to u^D on D), **Continuity**, **Monotonicity** and **Stability of interpolation regarding additivity** (that is stability of the interpolation with respect to an additive decomposability property of the overall utility). These properties are satisfied by the two models. A major difference between these models is related to the contamination property: If the Choquet integral satisfies some decision strategies in a cell, then this strongly sets constraints on some neighbour cells. This is not the case with the GAI model.

Finally, we have described two approaches to elicit the models at stake. The first one relies on Linear Programming, which is the standard approach in OR. The second one consists first in identifying the decision strategies between any two successive levels, and then extending these strategies for profiles having diverse values on the attributes using a pessimistic principle: a profile inherits from the decision strategy corresponding to the worst level in the profile. The latter elicitation approach can be used in real applications as the elicitation workload is kept to a reasonable amount.

References

[1] F. Bacchus and A. Grove. Graphical models for preference and utility. In *Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 3–10, Montreal, Canada, July 1995.

[2] C. A. Bana e Costa, J.M. De Corte, and J.-C. Vansnick. MACBETH. *International Journal of Information Technology and Decision Making*, 11:359–387, 2012.

[3] C. Boutilier, F. Bacchus, and R. Brafman. UCP-networks: A directed graphical representation of conditional utilities. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence (UAI-01)*, pages 56–64, Seattle, 2001.

[4] D. Bouyssou and T. Marchant. An axiomatic approach to noncompensatory sorting methods in MCDM I: The case of two categories. *Eur. J. of Operational Research*, 178(1):217–245, 2007.

[5] 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, 2007.

[6] D. Bouyssou and Th. Marchant. Multiattribute preference models with reference points. *Eur. J. of Operational Research*, 229:470–481, 2013.

[7] D. Braziunas. *Decision-theoretic elicitation of generalized additive utilities*. PhD thesis, University of Toronto, 2012.

[8] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1953.

[9] J H. Conway and N.J.A. Sloane. *Sphere Packings, Lattices and Groups*. Springer-Verlag, New York, 1988.

[10] D. Dubois, H. Fargier, and P. Perny. Qualitative decision theory with preference relations and comparison uncertainty: an axiomatic approach. *Artificial Intelligence*, 148:219–260, 2003.

[11] J.S. Dyer and R.K. Sarin. Measurable multiattribute value functions. *Operations Research*, 27:810–822, 1979.

[12] P. Fishburn. Interdependence and additivity in multivariate, unidimensional expected utility theory. *International Economic Review*, 8:335–342, 1967.

[13] P. Fishburn. *Utility Theory for Decision Making*. J. Wiley & Sons, 1970.

[14] A. Gersho and R.M. Gray. *Vector Quantization and Signal Compression*. Kluwer Academic Pub, Boston, 1992.

[15] M. Grabisch. The application of fuzzy integrals in multicriteria decision making. *European J. of Operational Research*, 89:445–456, 1996.

[16] M. Grabisch and Ch. Labreuche. Capacities on lattices and k-ary capacities. In *Int. Conf. Of the Euro Society for Fuzzy Logic and Technology (EUSFLAT)*, Zittau, Germany, September 10-12 2003.

[17] M. Grabisch and Ch. Labreuche. Bi-capacities - II : the Choquet integral. *Fuzzy Sets and Systems*, 151:237–259, 2005.

[18] M. Grabisch and Ch. Labreuche. Bipolarization of posets and natural interpolation. *Journal of Mathematical Analysis and Application*, 343:1080–1097, 2008.

[19] M. Grabisch and Ch. Labreuche. A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. *Annals of Operation Research*, 175:247–286, 2010.

[20] M. Grabisch and Ch. Labreuche. Monotone decomposition of 2-additive generalized additive independence models. *Mathematical Social Sciences*, DOI: https://doi.org/10.1016/j.mathsocsci.2017.09.007, 2017.

[21] M. Grabisch, J.L. Marichal, R. Mesiar, and E. Pap. *Aggregation functions*. Cambridge University Press, 2009.

[22] S. Greco, B. Matarazzo, and S. Giove. The Choquet integral with respect to a level dependent capacity. *Fuzzy Sets and Systems*, 175(1):1–35, 2011.

[23] S. Greco, V. Mousseau, and R. Słowinski. Robust ordinal regression for value functions handling interacting criteria. *European Journal of Operational Research*, 239(3):711–730, 2014.

[24] P. L. Hammer and R. Holzman. On approximations of pseudo-Boolean functions. *ZOR - Methods and Models of Operations Research*, 36:3–21, 1992.

[25] P. L. Hammer and S. Rudeanu. *Boolean Methods in Operations Research and Related Areas*. Springer, 1968.

[26] J. Hammond, R. Keeney, and H. Raiffa. Even Swaps: a rational method for making trade-offs. *Harvard Business Review*, March-June:137–149, 1998.

[27] C. R. Hsiao and T. E. S. Raghavan. Shapley value for multichoice cooperative games, I. *Games and Economic Behavior*, 5:240–256, 1993.

[28] D. Kahneman, P. Slovic, and A. Tversky. *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press, 2001.

[29] R. L. Keeney and H. Raiffa. *Decision with Multiple Objectives*. Wiley, New York, 1976.

[30] I. Kojadinovic. Estimation of the weights of interacting criteria from the set of profiles by means of information-theoretic functionals. *European Journal of Operational Research*, 155:741–751, 2004.

[31] D.H. Krantz, R.D. Luce, P. Suppes, and A. Tversky. *Foundations of measurement*, volume 1: Additive and Polynomial Representations. Academic Press, 1971.

[32] Ch. Labreuche. An axiomatization of the Choquet integral and its utility functions without any commensurability assumption. *submitted*.

[33] 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.

[34] Ch. Labreuche and M. Grabisch. Generalized Choquet-like aggregation functions for handling ratio scales. *European Journal of Operational Research*, 172:931–955, 2006.

[35] Ch. Labreuche and M. Grabisch. The representation of conditional relative importance between criteria. *Annals of Operation Research*, 154:93–122, 2007.

[36] Ch. Labreuche and M. Grabisch. Use of the GAI model in multi-criteria decision making: inconsistency handling, interpretation. In *Int. Conf. Of the Euro Society for Fuzzy Logic and Technology (EUSFLAT)*, Milano, Italy, 2013.

[37] Ch. Labreuche, B. Mayag, and B. Duqueroie. Extension of the MACBETH approach to elicit an OWA operator. *EURO J. on Decision Processes*, 3:65–105, 2015.

[38] J.-L. Marichal. *Aggregation operators for multicriteria decision aid.* PhD thesis, University of Liège, 1998.

[39] J.-L. Marichal. An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria. *IEEE Tr. on Fuzzy Systems*, 8(6):800–807, 2000.

[40] J.P. Pignon, Ch. Labreuche, and P. Ponthoreau. Combining experimentation and multi-criteria decision aid: A way to improve NEC systems of systems architecting. In *NATO SCI panel symposium on agility, resilience and control in NEC*, Amsterdam, The Netherland, May 20-25 2008.

[41] A. Rolland. Reference-based preferences aggregation procedures in multi-criteria decision making. *European J. of Operational Research*, 225:479–486, 2013.

[42] A. Rolland, J. Ah-Pine, and B. Mayag. Elicitation of 2-additive bi-capacity parameters. *EURO Journal on Decision Processes*, 3:5–28, 2015.

[43] G. C. Rota. On the foundations of combinatorial theory I. Theory of Möbius functions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 2:340–368, 1964.

[44] H. Simon. Rational choice and the structure of the environment. *Psychological Review*, 69:129–138, 1956.

[45] P. Slovic, M. Finucane, E. Peters, and D.G. MacGregor. The affect heuristic. In *T. Gilvitvh, D. Griffin, D. Kahneman (eds) Heuristics and biases: the psychology of intuitive judgement*, pages 397–420, Cambridge University Press, Cambridge, 2002.

[46] O. Sobrie. *Learning preferences with multiple-criteria models*. PhD thesis, University of Mons, CentraleSupelec, 2016.

[47] M. Sugeno. *Theory of fuzzy integrals and its applications*. PhD thesis, Tokyo Institute of Technology, 1974.

[48] A. Fallah Tehrani, Ch. Labreuche, and E. Hüllermeier. Choquistic utilitaristic regression. In *Decision Aid to Preference Learning (DA2PL) workshop*, Chatenay-Malabry, France, November 2014.