HAL open science

# New directions in ordinal evaluation: Sugeno integrals and beyond

Miguel Couceiro, Didier Dubois, Hélène Fargier, Michel Grabisch, Henri Prade, Agnès Rico

## To cite this version:

Miguel Couceiro, Didier Dubois, Hélène Fargier, Michel Grabisch, Henri Prade, et al.. New directions in ordinal evaluation: Sugeno integrals and beyond. Michalis Doumpos; José Rui Figueira; Salvatore Greco; Constantin Zopounidis. New Perspectives in Multiple Criteria Decision Making: Innovative Applications and Case Studies, Springer, Cham, pp.177-228, 2019, Multiple Criteria Decision Making book series (MCDM), 978-3-030-11481-7. 10.1007/978-3-030-11482-4_7 . hal-01941776

**HAL Id: hal-01941776**

**https://inria.hal.science/hal-01941776v1**

Submitted on 2 Dec 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.

# New directions in ordinal evaluation: Sugeno integrals and beyond

Miguel Couceiro[1], Didier Dubois[2], Hélène Fargier[2], Michel Grabisch[3], Henri Prade[2], Agnès Rico[4]

(1) Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
(2) IRIT, CNRS, Université Paul Sabatier,118 route de Narbonne F-31062 Toulouse, France
(3) Paris School of Economics, Université Paris I Panthéon-Sorbonne, 106-112, Boulevard de l'Hôpital, 75013 Paris, France
(4) ERIC, Université Claude Bernard Lyon 1, 43 bld du 11-11, F-69100 Villeurbanne, France
{quentin.brabant, miguel.couceiro}@loria.fr, {dubois, prade}@irit.fr, agnes.rico@univ-lyon1.fr,
michel.grabisch@univ-paris1.fr

**Abstract.** This chapter provides a state-of-the-art account of the use of Sugeno integrals in decision evaluation, when it is difficult to use meaningful figures of merit when assessing the worth of a decision and when only a finite scale of, e.g., linguistic categories, can be used. Here, Sugeno integrals are thought of as idempotent lattice polynomial functions on a finite bounded chain, which makes it possible to assign importance weights to groups of criteria or states. Algebraic and behavioral characterizations of the Sugeno integral are presented and discussed, including in the special cases of weighted minima and maxima. Extensions of this framework are also surveyed, namely: the lexicographic refinements that increase the discrimination power of this approach; the use of local utility functions in order to cope with criteria having distinct rating scales; and the generalization of the criteria weighting scheme at work in Sugeno integrals. Another kind of extension considered is when ratings belong to a bipolar scale where good and bad figures are explicitly present, thus giving rise to the symmetric Sugeno integral or to the separate evaluation of pros and cons. Moreover, it is pointed out that Sugeno integrals encode decision rules and that this bridge leads to methods for extracting knowledge from qualitative data. The results of empirical studies of the latter are also presented and discussed, accordingly.
**Keywords** Sugeno integral, lattice polynomial, bipolarity, qualitative decision theory, decision rule.

## 1 Introduction: motivation for qualitative evaluation methods

In the setting of Artificial Intelligence (for instance, recommender systems, cognitive robotics, but other fields as well), the use of decision rules based on numerical aggregation functions is not always natural. In particular, probabilities, utilities, importance weights cannot always be easily elicited from the user, by lack of time or lack of precision. Information systems advising persons cannot ask too many questions to users for modeling their preferences, nor collect from them meaningful numbers representing probabilities or criteria importance levels, or yet utility values. Even if they get them, making numerical operations on them needs justification, for instance, does the scale used authorize such calculations? To illustrate, suppose that a referee fills a form to assess the merits of a paper for a journal, and numerical ratings are required by the system. What is the precise meaning of these ratings? Does it make sense to compute averages from them?

In such situations it is more natural to resort to a qualitative approach to multicriteria evaluation. The rationale is to refrain from using numbers that look arbitrary or hard to collect, namely address decision problems in the ordinal setting. Gigerenzer and Todd [70] have argued that human decisions are often made on the basis of an ordinal ranking of the strength of criteria rather than on numerical evaluations, and hence the qualitative nature of the decision process. In daily life, people seldom resort to explicit numerical computations of figures of merit. This idea has also been exploited in Artificial Intelligence for a long time in qualitative decision theory [33]. For instance, so-called conditional preference networks (CP-nets) [9] allow for an easier representation of ordinal preference relations on multidimensional sets of alternatives, using local conditional preference statements interpreted ceteris paribus. See [40] for a survey of qualitative decision rules under uncertainty.

There are two advantages to using a qualitative approach: (i) a gain in robustness and the need for less data; (ii) qualitative methods lend themselves to a logical representation (which makes proposed

choices more easily explainable). There are two possible choices of qualitative settings for representing notions such as utility ratings stemming from several agents, importance levels and likelihood degrees:

- use distinct non-commensurate scales, which makes the framework very restrictive as impossibility theorems regarding rational aggregation processes are often obtained (e.g., in voting theory).
- use finite commensurate scales (taking advantage of notions facilitating commensurateness such as certainty equivalents), which leads to a finite ordered set of value classes.

In multi-criteria decision making, Sugeno integrals [101, 102] are commonly used as qualitative aggregation functions [75] using finite scales under the commensurability assumption between them. The definition of these integrals is based on a monotonic set-function named capacity or fuzzy measure that aims to qualitatively represent the likelihood of sets of possible states of nature, the importance of sets of criteria, etc. These set functions are currently used in many areas such as uncertainty modeling [59, 60], multiple criteria aggregation [10, 71, 78] or in game theory [97]. See also a recent book devoted to capacities in such areas [74]. Moreover, Sugeno integrals naturally lend themselves to a representation in terms of if-then rules involving thresholds [**?**,53, **?**], which makes them easy to interpret.

This chapter surveys several results around Sugeno integral as well as some of its extensions in the problem of evaluating decisions under uncertainty or objects according to several criteria. We shall speak of *alternatives* to stand for acts, decisions or objects to be evaluated. In section 2 we recall basic definitions and various mathematical properties of Sugeno integrals, including the links with conjunctive and disjunctive normal forms of lattice polynomials, and the close connection between Sugeno integrals and medians. Axiomatic characterisations of Sugeno integrals are also surveyed, as well as the expression of Sugeno integrals in possibilistic logic. Section 3 reviews generalisations of Sugeno integrals, namely their lexicographic refinements, the use of local utility functions that cope with the use of several local scales, and the use of more general conjunction and implication connectives for combining weights of criteria and local ratings. Section 4 discusses the notion of bipolar qualitative evaluation by means of special cases or variants of Sugeno integrals, which use bipolar scales with explicit positive and negative values, or yet a positive and a negative scale, in order to balance the pros and the cons. The last section deals with qualitative data analysis, namely how to represent qualitative data by means of Sugeno integrals or by a set of if-then rules.

## 2 Sugeno integrals

We consider a finite set of criteria $\mathcal{C} = \{1, \cdots, n\}$, also denoted by $[n]$. The alternatives considered are evaluated using these criteria. Here, the evaluation scale $L$ is supposed to be common to all criteria and is assumed to be a finite totally ordered set, for instance a subset of the interval $[0, 1]$. In any case, the bottom of $L$ is denoted by $0$ and the top is denoted by $1$. Using a single evaluation scale means that the ratings according to different criteria are commensurate, which is a strong assumption and that will be lifted later on in the paper.

The maximum (resp. minimum) will be denoted by $\vee$ (resp. $\wedge$). An alternative is represented by a function $f : \mathcal{C} \to L$, or equivalently, by a tuple of ratings on the different criteria, i.e., by $f = (f_1, \cdots, f_n) \in L^n$ where $f_i$ is the rating of $f$ according to criterion $i$. We assume that the rating scale $L$ is equipped with an involutive operation, denoted by $1 - \cdot$, such that whenever $\lambda \in L, 1 - \lambda \in L$ as well.

### 2.1 Basic definitions and preliminaries

In usual multicriteria evaluation based on weighted average, importance weights are assigned to criteria. In this paper, importance levels can be assigned to sets of criteria (instead of single ones) by means of a capacity which is a mapping $\mu : 2^{\mathcal{C}} \to L$ such that $\mu(\emptyset) = 0$, $\mu(\mathcal{C}) = 1$, and if $A \subseteq B$ then

$\mu(A) \leq \mu(B)$. This generalized importance assignment enables dependencies between criteria to be accounted for; namely, redundant criteria in a set $A$ are such that $\mu(A) = \max_{i \in A} \mu(\{i\})$ (since the weight of group $A$ is the one of the most important criterion in it), while a synergy between them is expressed when $\mu(A) > \max_{i \in A} \mu(\{i\})$. Capacities qualify to represent uncertainty in decision theory as well as importance weights in multicriteria decision-making.

The discrete Sugeno integral, widely used to aggregate qualitative local evaluations in multiple attribute evaluation [101], is defined by:

$$S_\mu(f) = \bigvee_{A \subseteq \mathcal{C}} (\mu(A) \wedge \bigwedge_{i \in A} f_i) \tag{1}$$

The name "integral" for such an expression may sound surprising. However, it was proposed first by Sugeno [101] under the name "fuzzy integral" in analogy with Lebesgue integral under the following equivalent form:

$$S_\mu(f) = \bigvee_{\lambda \in L} \lambda \wedge \mu(\{i : f_i \geq \lambda\})$$

The idea was to replace integral (sum) and product in Lebesgue integral by fuzzy set union ($\max$) and intersection ($\min$). For further background see, e.g., [76, 101, 102].

There are alternative expressions of Sugeno integral as follows [84, 101, 102]:

$$S_\mu(f) = \bigvee_{A \subseteq \mathcal{C}} (\mu(A) \wedge \bigwedge_{i \in A} f_i) = \bigwedge_{A \subseteq \mathcal{C}} (\mu(\overline{A}) \vee \bigvee_{i \in A} f_i)) \tag{2}$$

$$= \bigvee_{i=1}^{n} f_{(i)} \wedge \mu(\{(i), \ldots, (n)\}) = \bigwedge_{i=1}^{n} f_{(i)} \vee \mu(\{(i+1), \ldots, (n)\}). \tag{3}$$

$$= \bigvee_{\lambda \in L} \lambda \wedge \mu(\{i : f_i \geq \lambda\}) = \bigwedge_{\lambda \in L} a \vee \mu(\{i : f_i > \lambda\}). \tag{4}$$

where we have supposed $f_{(1)} \leq \cdots \leq f_{(n)}$ and $\overline{A}$ is the complement of $A$ in $\mathcal{C}$. It turns out [81] that $S_\mu(f)$ is the median of the set

$$\{f_1, \ldots, f_n\} \cup \{\mu(\{(n)\}), \ldots, \mu(\{(2), \ldots, (n)\})\}.$$

For instance, if $f_i = \lambda$ for $i \in A$ and $\theta < \lambda$ otherwise, it is easily seen that $S_\mu(f)$ is the median of $\lambda, \theta, \mu(A)$.

Note that Sugeno integral has exponential complexity in terms of the number of criteria, but its expression can be reduced to one of linear size by ranking the values $f_i$. Besides not all values $\mu(A)$ are always useful for defining $S_\mu$. The qualitative Mœbius transform $\mu_\#$ of the capacity $\mu$ [73, 86] is a mapping from $2^C$ to $L$ defined by

$$\mu_\#(E) = \begin{cases} \mu(E) \text{ if } \mu(E) > \bigvee_{B \subsetneq E} \mu(B) \\ 0 \text{ otherwise.} \end{cases}$$

It contains the minimal information needed to reconstruct the capacity $\mu$. Due to monotonicity of $\mu$ we can replace $\bigvee_{B \subsetneq E} \mu(B)$ in the above equation by $\bigvee_{i \in E} \mu(E \setminus \{i\})$. The function $\mu_\#$ is also the qualitative counterpart of a basic probability assignment in evidence theory [46], since it holds $\mu(A) = \bigvee_{E \subseteq A} \mu_\#(E)$. The set $\mathcal{F}(\mu) = \{E : \mu_\#(E) > 0\}$ is called the set of focal subsets of $\mu$. Then, Sugeno integral then can be expressed in a simplified form as

$$S_\mu(f) = \bigvee_{A \in \mathcal{F}(\mu)} (\mu(A) \wedge \bigwedge_{i \in A} f_i) \tag{5}$$

which contains no mathematically redundant min-terms.

A special case of capacity is a possibility measure [49, 52, 105] which is a maxitive capacity, i.e., a capacity $\Pi$ such that $\Pi(A \cup B) = \Pi(A) \vee \Pi(B)$. Since the set of criteria is finite, the possibility distribution $\pi : \pi_i = \Pi(\{i\})$, here representing individual criteria weights, is enough to recover the set-function: $\forall A \subseteq \mathcal{C}$, $\Pi(A) = \bigvee_{i \in A} \pi_i$: focal sets of $\Pi$ are singletons. In this case, criteria are considered redundant with one another.

The conjugate $\mu^c(A)$ of capacity $\mu$ is a capacity defined by $\mu^c(A) = 1 - \mu(\overline{A}), \forall A \subseteq \mathcal{C}$. The conjugate of a possibility measure $\Pi$ is a necessity measure $N(A) = 1 - \Pi(\overline{A})$, and then $N$ is a minitive capacity, i.e., $N(A \cap B) = N(A) \wedge N(B)$. Moreover, $N(A) = \bigwedge_{i \notin A} \iota(i)$ where $\iota(i) = N(\mathcal{C} \setminus \{i\})$ (this is the degree of impossibility of $i$ when dealing with uncertainty), and $\iota(i) = 1 - \pi_i$, where $\pi$ defines the conjugate possibility measure $\Pi = N^c$. In a group $A$ of criteria, we may have $N(\{i\}) = 0, \forall i \in A$ but $N(A) > 0$ which suggests that necessity measures account for criteria in positive synergy. Focal sets of necessity measures are nested, they are the cuts of the possibility distribution $\pi$, i.e., we have that $\mathcal{F}(N) = \{\{i : \pi_i \geq \lambda\} : \lambda \in L \setminus \{0\}\}$.

It is easy to see (and well-known [44, 77]) that if the capacity is a possibility measure, the Sugeno integral simplifies in the form of a prioritized maximum [47]:

$$S_\Pi(f) = SLMAX_\pi(f) = \bigvee_{i \in \mathcal{C}} \pi_i \wedge f_i.$$

Likewise it can be shown that if the capacity is a necessity measure, Sugeno integral simplifies in the form of a prioritized minimum [47]:

$$S_N(f) = SLMIN_\pi(f) = \bigwedge_{i \in \mathcal{C}} (1 - \pi_i) \vee f_i.$$

The use of the optimistic criterion $SLMAX_\pi$ captures the so-called focus effect: the order of magnitude of the importance of a group of criteria is the one of the most important argument, in the group. This assumption perfectly suits the intuition of a qualitative scale as it means that weaker arguments are always negligible compared with a single stronger one.

### 2.2 Sugeno integrals as lattice polynomials under normal form

A convenient way to introduce the discrete Sugeno integral is via the concept of (lattice) polynomial functions, i.e., functions which can be expressed as combinations of variables and constants using the lattice operations $\wedge$ and $\vee$. More precisely, given a bounded chain $L$, by an $n$-ary polynomial function, we simply mean a function $\phi : L^n \to L$ defined recursively as follows:

(i) For each $i \in [n] = \{1, \ldots, n\}$ and each $\lambda \in L$, the projection $\phi(\lambda_1, \ldots \lambda_n) \mapsto \lambda_i$ and the constant function $\phi \mapsto \lambda$ are polynomial functions from $L^n$ to $L$.

(ii) If $\phi$ and $\psi$ are polynomial functions from $L^n$ to $L$, then $\phi \vee \psi$ and $\phi \wedge \psi$ are polynomial functions from $L^n$ to $L$.

(iii) Any polynomial function from $L^n$ to $L$ is obtained by finitely many applications of the rules (i) and (ii).

We refer to those polynomial functions constructed from projections by finitely many applications of (ii) as *lattice term functions* (or simply, *term functions*). A well-known example of a term function is the ternary median function, which is given by

$$\begin{aligned} \mathrm{median}(\lambda, \lambda', \lambda'') &= (\lambda \wedge \lambda') \vee (\lambda' \wedge \lambda'') \vee (\lambda'' \wedge \lambda) \\ &= (\lambda \vee \lambda') \wedge (\lambda' \vee \lambda'') \wedge (\lambda'' \vee \lambda). \end{aligned}$$

As shown by Marichal [85], the discrete Sugeno integrals are exactly those polynomial functions $\phi : L^n \to L$ which are *idempotent*, i.e., that satisfy $\phi(\lambda, \ldots, \lambda) = \lambda$.

In this subsection we revisit classical normal form representations of lattice polynomials and recall the median normal form representation that follows from *median decomposability*.

**Disjunctive and conjunctive normal forms** Goodstein [67] has shown that in the case of bounded distributive lattices, polynomial functions are exactly those which allow their representations in disjunctive and conjunctive normal forms. In this subsection we recall some related useful results.

**Proposition 1.** *Let $\phi\colon L^n \to L$ be a function. The following conditions are equivalent:*

*(i) $\phi$ is a polynomial function.*
*(ii) There exists a set function $\mu\colon 2^{[n]} \to L$ such that $\phi(f) = \bigvee_{I\subseteq[n]}(\mu(I) \wedge \bigwedge_{i\in I} f_i)$.*
*(iii) There exists a set function $\kappa\colon 2^{[n]} \to L$ such that $\phi(f) = \bigwedge_{I\subseteq[n]}(\kappa(I) \vee \bigvee_{i\in I} f_i)$.*

The expressions given in $(ii)$ and $(iii)$ of Proposition 1 are usually referred to as the *disjunctive normal form* (DNF) representation and the *conjunctive normal form* (CNF) representation, respectively, of the polynomial function $\phi$. Notice that the DNF and CNF representations of a polynomial function $\phi\colon L^n \to L$ are not necessarily unique.

For each $I \subseteq [n]$, let $\mathbf{1}_I$ be the element of $L^n$ whose $i$th component is 1, if $i \in I$, and 0, otherwise. Let $\mu_\phi\colon 2^{[n]} \to L$ be the set function given by $\mu_\phi(I) = \phi(\mathbf{1}_I)$. It is monotone with inclusion. It is easy to see that if moreover $\mu_\phi([n]) = 1$ and $\mu_\phi(\emptyset) = 0$, then letting $\mu = \mu_\phi$ in Proposition 1, $\phi$ is a Sugeno integral in DNF. This representation is not unique: we still get the same Sugeno integral $\phi$ if we change $\mu_\phi$ into any set-function $\mu$ in $\mathrm{DNF}(\phi) = \{\mu \in L^{2^{[n]}} : \phi(f) = \bigvee_{I\subseteq[n]} \mu(I) \wedge \bigwedge_{i\in I} f_i\} = \{\mu : \mu_\phi^{\#} \leq \mu \leq \mu_\phi\}$, using the qualitative Moebius transform $\mu_\phi^{\#}$.

Dually, let $\kappa_\phi\colon 2^{[n]} \to L$ be the function given by $\kappa_\phi(I) = \phi(\mathbf{1}_{[n]\setminus I}) = \mu_\phi(\overline{I})$. Function $\kappa_\phi$ is clearly antitone, and if $\mu_\phi$ is a capacity, we recognize again the Sugeno integral in CNF in Proposition 1 (iii).

*Remark 1.* Note that $\mu_\phi$ is the only monotone set function in $\mathrm{DNF}(\phi)$ and, similarly, $\kappa$ is the only anti-monotone set function in $\mathrm{CNF}(\phi) = \{\kappa \in L^{2^{[n]}} : \phi(f) = \bigwedge_{I\subseteq[n]}(\kappa(I) \vee \bigvee_{i\in I} f_i)\}$.

**Median normal form** For instance, consider act $f$ resulting in a good consequence $x$ if event $A$ occurs and a bad consequence $y$ (with $u(x) > u(y)$) otherwise. It is easily seen that $S_{\gamma,u}(f)$ is the median of $\{u(x), u(y), \gamma(A)\}$

It is not difficult to see that every lattice polynomial function $\phi\colon L^n \to L$ is *median decomposable*, i.e., for every $f \in L^n$,

$$\phi(f) = \mathrm{median}\left(\phi(f^{k\to 0}), f_k, \phi(f^{k\to 1})\right), \tag{6}$$

where $f^{k\to\lambda} = (f_1, \ldots, f_{k-1}, \lambda, f_{k+1}, \ldots, f_n)$ for $k \in [n]$ is obtained by changing $f_k$ into $\lambda \in L$ in the vector $(f_1, \ldots, f_n)$. In fact, the converse is also true, and thus we have the following characterization of lattice polynomial functions.

**Theorem 1 ([85, Theorem 17]).** *The solutions of the median decomposition equation (6) are exactly the lattice polynomial functions from $L^n$ to $L$.*

For further characterizations, we refer the reader to the survey paper [25]. This median decomposition scheme naturally leads to a recursive procedure for obtaining median representations of functions independent from the way functions are given [24]. Indeed, by setting a ranking of variables, we can

repeatedly apply Theorem 1 to the variables of any given function in order to derive a nested formula made of medians applied to variables and constants. To illustrate, consider the 5-ary median function $\mathrm{median}_5$. This naïve approach leads to a median representation of $\mathrm{median}_5$:

$$\begin{aligned}\mathrm{median}_5(f_1,f_2,f_3,f_4,f_5) &= \mathrm{median}\big(\mathrm{median}_5(0,f_2,f_3,f_4,f_5), f_1, \mathrm{median}_5(1,f_2,f_3,f_4,f_5)\big)\\ &= \mathrm{median}\big(\mathrm{median}\big(\mathrm{median}_5(0,0,f_3,f_4,f_5), f_2, \mathrm{median}_5(0,1,f_3,f_4,f_5)\big),\\ &\quad f_1, \mathrm{median}\big(\mathrm{median}_5(1,0,f_3,f_4,f_5), f_2, \mathrm{median}_5(1,1,f_3,f_4,f_5)\big)\big)\\ &= \dots\end{aligned}$$

In this way, we obtain a median normal form representation of $\mathrm{median}_5$ with $1+2+4+8+16=31$ occurences of $\mathrm{median}$, which is not optimal. Indeed, there exists a much smaller representation with only $4$ occurences of $\mathrm{median}$:

$$\mathrm{median}_5(f_1,f_2,f_3,f_4,f_5) = \mathrm{median}(\mathrm{median}(\mathrm{median}(f_2,f_3,f_4),f_4,f_5),\mathrm{median}(f_2,f_3,f_5),f_1).$$

Now, it is not difficult to extend the results in [21] and show that the median normal form produces representations that make use, up to polynomial equivalence, of the least number of monotonic connectives. However, an efficient procedure for computing the smallest median normal form is still unknown and the problem of deciding whether a median representation is minimal seems to be (mildly) untractable [28, 29].

### 2.3 Algebraic and behavioral characterizations

Sugeno integral has been characterized by a few properties, especially decomposability for comonotonic functions $f$ and $g : L^n \to L$, i.e., such that $f_i > f_j \Rightarrow g_i \geq g_j, \forall i,j \in [n]$. Namely,

**Theorem 2.** *Let $I : L^{\mathcal{C}} \to L$. There is a capacity $\mu$ such that $I(f) = S_\mu(f)$ for every $f \in L^{\mathcal{C}}$ if and only if the following properties are satisfied*

**1.** $I(f \vee g) = I(f) \vee I(g)$, *for any comonotonic* $f, g \in L^{\mathcal{C}}$.
**2.** $I(\lambda \wedge f) = \lambda \wedge I(f)$, *for every* $\lambda \in L$ *and* $f \in L^{\mathcal{C}}$.
**3.** $I(\mathbf{1}_{\mathcal{C}}) = 1$.

Equivalently, conditions (1 -3) can be replaced by conditions (1'-3') below [57]:

**1'.** $I(f \wedge g) = I(f) \wedge I(g)$, for any comonotonic $f, g \in L^{\mathcal{C}}$.
**2'.** $I(\lambda \vee f) = \lambda \vee I(f)$, for every $\lambda \in L$ and $f \in L^{\mathcal{C}}$.
**3'.** $I(\mathbf{0}_{\mathcal{C}}) = 0$.

Most older formulations of this theorem [13, 92] redundantly add an assumption of increasing monotonicity of the functional $I$ (if $f \geq g$ then $I(f) \geq I(g)$) to the three conditions (1-3). But these papers do not point out the equivalent conditions (1'-3'). The existence of these two equivalent characterisations is due to the possibility of equivalently writing Sugeno integral in conjunctive and disjunctive normal form (see Equation (2)). As a consequence, the De Morgan dual $1 - S_\mu(1-f)$ of a Sugeno integral $S_\mu(f)$ is also a Sugeno integral in the sense that (2) can be expressed as

$$S_\mu(f) = 1 - S_{\mu^c}(1-f) \tag{7}$$

for the conjugate capacity $\mu^c$. Note that a functional $I(f)$ satisfies conditions (1-3) if and only if $1 - I(1-f)$ satisfies conditions (1'-3'). Marichal [84] provides several similar characterizations, especially one assuming maxitive and minitive comonotonicity (conditions 1 and 1') along with idempotence, one assuming homogeneity conditions 2 and 2' plus increasing monotonicity. However, the proof that

conditions (1-3) are necessary and sufficient seems to first appear in a thesis [94] (and then used in [76, 74]).

Another characterization has been provided in the context of decision under uncertainty, in a setting similar to Savage's approach to expected utility functionals. The set of criteria $\mathcal{C}$ is replaced by a *finite* set $[n]$ of states, and alternative decisions are just functions $f$ from $[n]$ to a set of consequences $X$. We consider again a finite totally ordered scale $(L, \leq)$ with bottom 0 and top 1. A mapping $u : X \to L$ is named a utility function. We assume that $X$ contains an ideal consequence $x^*$ with $u(x^*) = 1$ and a worst consequence $x_*$ with $u(x_*) = 0$. Note that the use of a single scale for rating the consequences of acts is more natural than for rating alternatives in multiple-criteria evaluation problems.

The decision-maker is supposed to supply a preference relation $\succeq$ on the set $X^{\mathcal{C}}$ of alternatives (called acts), that is, a non trivial preorder: $\succeq$ is transitive and complete.

We introduce new notations that will be useful in the following:

- A constant act $\mathsf{x}$ is such that $\exists x \in X, \forall i \in \mathcal{C}, \mathsf{x}(i) = x$. In particular, the acts $\mathsf{x}^*$ and $\mathsf{x}_*$ are such that $\mathsf{x}^*(i) = x^*, \forall i \in \mathcal{C}$ and $\mathsf{x}_*(i) = x_*, \forall i \in \mathcal{C}$.
- For acts $f, g$, $fAg$ is the act defined by $fAg(i) = f(i)$ for all $i$ in $A$ and $fAg(i) = g(i)$ for all $i$ in $\overline{A}$.
- When using the mapping that assigns to each state $i$ the utility value $u(x)$ of its consequence $f(i) = x$ under act $f$, namely $u \circ f$, the values $u(f(i))$ will be simplified as $f_i$.

Note that the preference relation $\succeq$ induces a complete preordering $\geq_P$ on consequences: $x \geq_P y$ if and only if $\mathsf{x} \succeq \mathsf{y}$; this ordering can be extended to acts as follows: $f \geq_P g$ if and only if $f(i) \geq_P g(i), \forall i \in \mathcal{C}$. This is the Pareto-ordering. Then, one can define an act $f \vee g$ making the best of $f$ and $g$, such that $\forall i \in \mathcal{C}, (f \vee g)_i = f_i$ if $f_i \geq g_i$ and $g_i$ otherwise; and an act $f \wedge g$ making the worst of $f$ and $g$, such that $\forall i \in \mathcal{C}, (f \wedge g)_i = f_i$ if $g_i \geq_P f_i$ and $g_i$ otherwise. Acts are thus combined like fuzzy sets.

The axioms proposed in [58] are as follows:

**A1** *Totality*: $\succeq$ is a non-trivial total preorder, i.e., it is transitive and complete, and $f \succ g$ for some acts.
**WP3** *Weak compatibility with constant acts*: $\forall A \subseteq \mathcal{C}, \forall x, y \in X, \forall f, \mathsf{x} \succeq \mathsf{y}$ implies $\mathsf{x}Af \succeq \mathsf{y}Af$.
**RCD** *Restricted conjunctive dominance*: For any acts $g, h$ and any constant act $\mathsf{x}$, $\mathsf{x} \succ h$ and $g \succ h$ imply $\mathsf{x} \wedge g \succ h$.
**RDD** *Restricted max-dominance*: For any acts $g, h$ and any constant act $\mathsf{x}$, $h \succ \mathsf{x}$ and $h \succ g$ imply $h \succ \mathsf{x} \vee g$.

Axioms **A1** and **WP3** entail Pareto-dominance: if $f \geq_P g$ then $f \succeq g$ (see Lemma 4 in [59]). Moreover, **RCD** and **RDD** make sense for one-shot decisions, i.e., without repetition, making the compensation of bad results by good ones impossible.

We recall here the main result about this axiomatization for decision under uncertainty [58].

**Theorem 3.** *Let $(X^{\mathcal{C}}, \succeq)$ be a preference structure. The following propositions are equivalent:*

- *$(X^{\mathcal{C}}, \succeq)$ satisfies* **A1***, plus* **WP3, RCD, RDD**.
- *there exists a finite chain $L$ of preference levels, an $L$-valued monotonic set-function $\mu$, and an $L$-valued utility function $u$ on $X$, such that $f \succeq g$ if and only if $S_\mu(f) \geq S_\mu(g)$.*

The proof of this theorem as it appears in [58, 59] is incomplete. See [61] for a complete proof.

The area of significance of this qualitative decision theory and more precisely the one of axioms **RCD** and **RDD**, is restricted to the case where $X$ and $\mathcal{C}$ are finite and where the value scale is coarse. For instance, **RCD** means that limiting from above the potential utility values of an act $g$, that is better than another one $h$, to a constant value that is better than the utility of act $h$, still yields an act better than $h$. This is in contradiction with expected utility theory and debatable in the latter setting. Indeed, suppose $g$ is a lottery where you win 1000 euros against nothing with equal chances. Suppose the certainty

equivalent of this lottery is 400 euros, received for sure, and $h$ is the fact of receiving 390 euros for sure. Now, it is likely that, if $f$ represents the certainty-equivalent of $g$, $f \wedge g$ will be felt strictly less attractive than $h$, as the former means you win 400 euros against nothing with equal chances. Axiom **RCD** implies that such a lottery should ever be preferred to receiving $400-\epsilon$ euros for sure, for arbitrary small positive values of $\epsilon$. This axiom is thus strongly counterintuitive in the context of economic theory, with a continuous consequence set $X$. However the area of significance of qualitative decision theory is precisely when both $X$ and $S$ are finite.

Two presuppositions actually underlie axiom **RCD** (and similar ones for **RDD**)

(i) There is no compensation effect in the decision process: in case of equal chances, winning 1000 euros cannot compensate the possibility of not earning anything. It fits with the case of one-shot decisions where the notion of certainty equivalent can never materialize: you can only get 1000 euros or get nothing if you just play once. You cannot get 400 euros. The latter can only be obtained in the average, by playing several times.

(ii) There is a big step between one level $\lambda_i \in V$ in the qualitative value scale and the next one $\lambda_{i+1}$ with $V = \{1 = \lambda_1 > \cdots > \lambda_m = 0\}$. The preference pattern $f \succ h$ always means that $f$ is significantly preferred to $h$ so that the preference level of $f \wedge g$ can never get very close to that of $h$ when $g \succ h$. The counterexample above is obtained by precisely bringing these two preference levels very close to each other so that $f \wedge g$ can become less attractive than the sure gain $h$. Level $\lambda_{i+1}$ is in some sense considered negligible in front of $\lambda_i$.

Axioms **RDD** and **RCD** can be replaced in Theorem 3 by non-compensation assumptions [59]:

$$\textbf{Axiom NC}: \begin{cases} 1_L Ay \sim y \text{ or } 1_L Ay \sim 1_L A0 \\ \text{and} \\ xA0_L \sim x \text{ or } xA0_L \sim 1_L A0_L \end{cases}$$

Non-compensation formalizes the following intuition: in order to evaluate act $1_L Ay$, there is no middle term between values $u(y)$ and $\mu(1_L A0_L)$. Theorem 3 also holds if in the expression of **RCD** and **RDD** one considers any two comonotonic acts. Indeed Sugeno integrals are "linear" for operations maximum and minimum with respect to disjunctions and conjunctions of comonotonic acts as seen in condition 1 of Theorem 2 and the associated condition 1'. In this sense, Sugeno integral is a qualitative counterpart to Choquet integral. It is easy to check that these equalities hold with any two acts $f$ and $g$, for the pessimistic and the optimistic possibilistic preference functionals respectively:

$$SLMIN_\pi(f)(f \wedge g) = \min(SLMIN_\pi(f), SLMIN_\pi(g))$$

$$SLMAX_\pi(f \vee g) = \max(SLMAX_\pi(f), SLMAX_\pi(g)).$$

The criterion $SLMIN_\pi(f)$ can be axiomatized by strengthening axiom **RCD** as follows:

**Axiom CD**: $\forall f, g, h, f \succ h$ and $g \succ h$ jointly imply $f \wedge g \succ h$ (Conjunctive Dominance).

This axiom means that if two acts $f, g$ are individually better than a third one, the act $f \wedge g$ which yields the worse result of both acts still remains better than the third one. It makes sense in the scope of a one-shot-decision. Together with Sugeno integral axioms, it implies that the set-function $\mu$ is a necessity measure and so, $S_\mu(f) = SLMIN_\pi(f)$, for some possibility distribution $\pi$. In order to figure out why axiom CD leads to a pessimistic criterion, Dubois, Prade and Sabbadin [60] have noticed that it can be equivalently replaced by the following property:

**Axiom PESS** $\forall A \subseteq S, \forall f, g, fAg \succ g$ implies $g \succeq gAf$ (Pessimism).

This property can be explained as follows: if changing $g$ into $f$ when $A$ occurs results in a better act, the decision maker has enough confidence in event $A$ to consider that improving the results on $A$ is worth trying. But, in this case, there is less confidence on the complement $\overline{A}$ than in $A$, and any possible improvement of $g$ when $\overline{A}$ occurs is neglected. So, $g \succeq gAf$. For instance, $g$ means losing ($= A$) or winning ($= \overline{A}$) 10,000 euros with equal chances according to whether $A$ occurs or not, and $f$ means winning either nothing ($= A$) or 20,000 euros ($= \overline{A}$) conditioned on the same event. Then $fAg$ is clearly safer than $g$ as there is no risk of losing money. However, if axiom **PESS** holds, then the chance of winning much more money (20,000 euros) by choosing act $gAf$is neglected because there is still a good chance to lose 10,000 euros with this lottery. Such a behaviour is clearly cautious.

Similarly, the optimistic criterion $SLMAX_\pi(f)$ can be axiomatized by strengthening the axioms **RDD** as follows:

**Axiom DD**: $\forall f, g, h, h \succ f$ and $h \succ g$ jointly imply $h \succ f \vee g$ (Disjunctive Dominance.)

Together with properties appearing in Theorem 3 it implies that the set-function $\mu$ is a possibility measure and so, $S_\mu(f) = SLMAX_\pi(f)$ for some possibility distribution $\pi$. The optimistic counterpart to property axiom **PESS** that can serve as a substitute to axiom DD for the representation of criterion $SLMAX_\pi$ is:

**Axiom OPT** $\forall A \subseteq S, \forall f, g, g \succ fAg$ implies $gAf \succeq g$. (Optimism).

### 2.4 Sugeno integral and decision rules

So far, only a few works try to provide a logical reading of decision processes. One of such few attempts is given in [42] in the framework of decision under uncertainty, where uncertain knowledge and prioritized preference are respectively represented by means of two distinct possibilistic logic bases, and where the pessimistic or optimistic decision criteria that are maximized are particular cases of Sugeno integrals. Another attempt is given in [66] in the framework in a multiple criteria decision making, where a qualitative approach (in the spirit of possibilistic logic) is compared to the numerical analogue based on the Choquet integral. However, the most successful approach was initiated by Greco *et al.* in [78] where they provided a preliminary study (later completed by Bouyssou *et al.* in [10]) pointing out that the set of the elements for which a Sugeno integral is greater than a given score $\mu$ can be described by if–then rules.

**Selection rules** Consider Sugeno integral in the form given by (5), using the set $\mathcal{F}(\mu)$ of focal sets of $\mu$. It is straightforward to see that the inequality $S_\mu(f) \geq \theta$ is equivalent to $\exists T \in \mathcal{F}(\mu)$ such that $\mu(T) \geq \theta$ and $\forall i \in T,\ f_i \geq \theta$. On this basis, it can be claimed that Sugeno integral based on capacity $\mu$ is equivalent to the set of if-then rules of the form:

$$R_T^s: \quad \textbf{If } \forall i \in T,\ f_i \geq \mu(T) \textbf{ then } S_\mu(f) \geq \mu(T).$$

for $T \in \mathcal{F}(\mu)$. Note that these rules are not redundant because either focal sets are not nested or, if they are, they correspond to distinct weights, and the greater the set, the larger the weight. Moreover, they are *single-thresholded* rules, which indicates the limited expressive power of Sugeno integrals. As such rules are bounding the global evaluation from below, they are meant to select "good" alternatives, so we can call them *selection rules*.

Conversely, a set of single-thresholded selection rules of the form "**If** $\forall i \in T_j,\ f_i \geq \theta_j$ **then** $\phi(f) \geq \theta_j$ for $j = 1, \ldots, k$" can be represented by the Sugeno integral with focal sets among $\{T_j : j = 1, \ldots, k\}$ and such that $\mu(A) = \max_{j:T_j \subseteq A} \theta_j$ (the integral representation does away with redundant rules).

This set of rules can be encoded in possibilistic logic [50] as a set of weighted cubes [53]. Define for each criterion $i$ a family of Boolean predicates $P_i(\theta), \theta > 0 \in L$ such that $P_i(\theta)$ is true if $f_i \geq \theta$

and 0 otherwise (we write $f \models P_i(\theta)$). Then we consider weighted Boolean formulas of the form $[\bigwedge_{j\in T} P_j(\theta), \theta]$ and interpreted as

$$\pi_{[T,\theta]}(f) = \begin{cases} \theta \text{ if } f_i \geq \theta, \forall i \in T; \\ 0 \text{ otherwise} \end{cases}.$$

Each weighted cube $[\bigwedge_{j\in T} P_j(\mu(T)), \mu_{\#}(T)]$ for a focal set $T$ encodes a selection rule $R^s_T$ as stated above. The lower possibility distributions associated to a set of such weighted formulas is interpreted as the maximum of the lower possibility distributions associated to each weighted formula. The possibilistic base

$$G^-_\mu = \{[\bigwedge_{j\in T} P_j(\theta), \theta] \,:\, \mu(T) \geq \theta > 0,\, T \in \mathcal{F}(\mu)\}$$

with lower possibility distribution $\pi^-_\mu(f) = \max\{\pi^-_{[\phi,\theta]}(f) \;:\; \mu(T) \geq \theta > 0 \text{ and } T \in \mathcal{F}(\mu)\}$ encodes a Sugeno integral since $S_\mu(f) = \pi^-_\mu(f)$ (see Proposition 4 in [53]).

**Elimination rules** Symmetrically, we can obtain rules for the rejection of bad alternatives associated to the Sugeno integral, namely rules expressing the inequality $S_\mu(f) \leq \gamma$ [35]. The idea is to use the conjunctive min-max form of Sugeno integral in equation (2), which corresponds to possibility distributions over interpretations in standard possibilistic logic [50].

The focal sets of the conjugate of $\mu$ are sufficient to calculate the Sugeno integral, namely [35]:

$$S_\mu(f) = \min_{F\in\mathcal{F}(\mu^c)} \max(1 - \mu^c_{\#}(F), \max_{i\in F} f_i).$$

It is then clear that $S_\mu(f) \leq \theta$ if and only if $\exists F \in \mathcal{F}(\mu^c)$ with $\mu^c(F) \geq 1-\theta$ s.t. $\forall x_i \in F\ f_i \leq \theta$. This result shows that for each focal set $F$ of the conjugate $\mu^c$ we have the following single-thresholded *elimination rule*:

$$R^e_F\text{:}\quad \textbf{If } f_i \leq 1 - \mu^c_{\#}(F) \text{ for all } i \in F \textbf{ then } S_\mu(f) \leq 1 - \mu^c_{\#}(F).$$

Conversely, a set of single-thresholded elimination rules can be represented by a Sugeno integral.

The possibilistic logic encoding of elimination rules associated to Sugeno integral, is now obtained as set of weighted clauses. Define for each criterion $i$ a family of Boolean predicates $P_i(\theta), \theta > 0 \in L$ such that $P_i(\theta)$ is true for $f$ if $f_i > \theta$ and 0 otherwise.

The set of weighted clauses $\{(\bigvee_{j\in F} P_j(\theta), 1-\theta)) \,:\, \theta < 1\}$ induces an upper possibility distribution:

$$\pi^+_F(f) = \min_{\theta<1} \max(\theta, \max_{j\in F} P_j(\theta)) = \max_{j\in F} f_j.$$

Each weighted clause $(\bigvee_{j\in F} P_j(1-\mu^c(F)), \mu^c(F)))$ for a focal set $F$ of $\mu^c$ corresponds to the elimination rule $R^e_F$ stated above.

A logical rendering of the Sugeno integral in the min-max form is obtained as follows. First consider the following base of clauses $B^F_\mu = \{(\bigvee_{j\in F} P_j(\theta), 1-\theta)) : 1 - \mu^c_{\#}(F)) \leq \theta < 1\}$. It can be proved [35] that the induced upper possibility is now of the form

$$\pi^+_{B^F_\mu}(f) = \max(1 - \mu^c_{\#}(F), \max_{i\in F} f_i).$$

The possibilistic base

$$G^+_\mu = \{(\bigvee_{j\in F} P_j(\theta), 1-\theta)) \,:\, 1 - \mu^c(F)) \leq \theta < 1, F \in \mathcal{F}(\mu^c)\},$$

induced by all focal sets induces the possibility distribution with upper possibility distribution $\pi^+_\mu(f) = \min_{F\in\mathcal{F}(\mu^c)} \pi^+_{B^F_\mu}(f)$, is precisely equal to $S_\mu(f)$.

This rule-based approach to the description of Sugeno integral delivers several lessons:

- A Sugeno integral can be equivalently expressed by a set of single-thresholded selection and elimination rules, respectively squeezing it from above and from below,
- The expressive power of Sugeno integrals is limited to a class of very specific decision rules.

## 3 Extensions

The finite scale approach to qualitative decision criteria is simple (especially in the case of weighted max and min). Strictly speaking, it belongs to the class of decision problems that were coined "sorting" by B. Roy [96]. Elements of the scale $L$ correspond to a totally ordered set of classes of situations thatare more or less attractive, and computing the global evaluation of an alternative comes down to assigning it to a class. In the context of DMU, the restriction of the pessimistic approach to the most plausible states, at work in possibilistic criteria, makes them more realistic than the maximin criterion, and more flexible than purely ordinal approaches with no commensurateness assumption.

However, approaches based on an absolute qualitative value scale have their own shortcomings.

- Naturally, a complete preorder on alternatives is obtained from the Sugeno integral. But this preorder is very coarse, especially if the number of elements in $L$ is small. It cannot be large as the human mind cannot make sense of more than seven levels of absolute worth [91]. So many alternatives, some of which being intuitively better than other ones, will be put in the same class. Hence, one issue is to make Sugeno integral more discriminant via suitable refinement tools. This is the topic of the first subsection.
- Moreover, one has to accept the commensurability assumption between the dimensions of the decision problem. In decision under uncertainty this is not so problematic as there is a single value scale, and the uncertainty scale can be related to the value scale via the notion of certainty equivalent of an uncertain event. However this assumption is much more problematic in the MCDM problem, as each criterion has its own scale that may not be directly commensurate with other ones. In such a situation, Sugeno integral cannot be directly applied. Utility functions that relate the criteria scales to a single one must be introduced, which requires an extension of Sugeno integral described in the second subsection.
- Finally, the role of weights in Sugeno integral is confined to being bounds that limit the value scales from above and from below. In a weighted max, a little important criterion can only deliver poor grades, while in a weighted min it can only deliver good grades. Other ways of letting importance degrees affect the local evaluations can be envisaged, changing the minimum or the maximum operations into more general conjunction and disjunction operations. This is the topic of the last subsection.

### 3.1 Lexicographic refinements

The main reason for the lack of discrimination power of absolute qualitative criteria is the fact that they do not use all the available information to rank alternatives, since an alternative $f$ can be considered indifferent to another alternative $g$, even if $f$ is at least as good as $g$ in all criteria and strictly so for some of them (violation of the strict Pareto ordering). This is typically the case when an alternative is rated by its worst performance or its best performance across criteria. This defect is absent from the expected utility model.

**Refining min and max** The lack of discrimination of the maximin rule itself (using $\min_{i\in\mathcal{C}} f_i$ to evaluate alternative $f$) was actually addressed a long time ago by Cohen and Jaffray [18] who improved it by comparing acts on the basis of their worst consequences of *distinct* merits, i.e. one considers only the

set $D(f,g) = \{i, f_i \neq g_i)\}$ to compare alternatives $f$ and $g$. Denote the refined strict preference relation between acts by

$$f \succ_{dmin} g \iff \min_{i \in D(f,g)} f_i > \min_{s \in D(f,g)} g_i$$

and the weak preference by $f \succeq_{dmin} g \iff \neg(g \succ_{dmin} f)$. This refined rule always rates an act $f$ better than another act $g$ whenever $f$ strictly Pareto-dominates $g$. However, only a partial ordering of acts is then obtained. This last decision rule is actually no longer based on a preference functional (i.e. it cannot be encoded by a function, like expected utility). This decision rule has been independently proposed by Fargier et al. [63] and used in fuzzy constraint satisfaction problems [41] under the name *discrimin ordering*.

This ordering can be further refined by the so-called *Leximin* ordering well-known in economics [32]: The idea is to reorder vectors $f = (f_1, \dots f_n)$ by non-decreasing values as $(f_{(1)}, \dots, f_{(n)})$, where $f_{(k)}$ is the $k$th smallest component of the vector (i.e., $f_{(1)} \leq \dots \leq f_{(n)}$). Define the *Leximin* ($\succeq_{lmin}$) and *Leximax* ($\succeq_{lmax}$) rules as:

- $f \succeq_{lmin} g \Leftrightarrow$ either $\forall j, f_{(j)} = g_{(j)}$ or $\exists i, \forall j < i, f_{(j)} = g_{(j)}$ and $f_{(i)} > g_{(i)}$
- $f \succeq_{lmax} g \Leftrightarrow$ either $\forall j, f_{(j)} = g_{(j)}$ or $\exists i, \forall j > i, f_{(j)} = g_{(j)}$ and $f_{(i)} > g_{(i)}$.

Similarly, a *Leximax* preorder can be envisaged as a refinement of the one induced by the maximum. Let $f, g \in L^n$. The two possible alternatives $f$ and $g$ are indifferent if and only if the corresponding reordered vectors are the same. The *Leximin*-ordering is a refinement of the discrimin ordering, hence of both the Pareto-ordering and the maximin-ordering [39]: $f \succ_{dmin} g$ implies $f \succ_{lmin} g$. Leximin optimal alternatives are always discrimin maximal alternatives, and thus indeed min-optimal and Pareto-maximal: $\succ_{lmin}$ is the most selective among these preference relations. The *Leximin* ordering can discriminate more than any symmetric aggregation function, since when, e.g., in the numerical setting, the sum of the $f_i$'s equals the sum of the $g_i$'s, it does not mean that the reordered vectors are the same. Similar comments apply for the Leximax ordering.

**Weighted *Leximax* / *Leximin* criteria** Suppose that *Leximin* and *Leximax* orderings are defined on sets of tuples whose components belong to a totally ordered set $(V, \trianglerighteq)$, say $Leximin(\trianglerighteq)$ and $Leximax(\trianglerighteq)$. Now, suppose $(V, \trianglerighteq) = (L^p, \succeq_{lmin})$ or $(V, \trianglerighteq) = (L^p, \succeq_{lmax})$, with any positive integer $p$. Then, nested lexicographic ordering relations that enable $L$-valued matrices to be compared can be recursively obtained as $Leximin(\succeq_{lmin})$, $Leximax(\succeq_{lmin})$, $Leximin(\succeq_{lmax})$, or yet $Leximax(\succeq_{lmax})$.

Consider for instance the procedure $Leximax(\succeq_{lmin})$ defining the relation $\succeq_{lmax(\succeq lmin)}$. It applies to matrices $A$ of dimension $p \times q$ with coefficients $a_{ij}$ in $(L, \geq)$. These matrices can be totally ordered in a very refined way by this relation. Denote row $i$ of $A$ by $a_{i\cdot}$, and let $A^\star$ and $B^\star$ be rearranged matrices $A$ and $B$ such that terms in each row are reordered increasingly and rows are arranged lexicographically top-down in decreasing order. The relation $A \succ_{lmax(\succeq lmin)} B$ is defined as follows:

$$\exists k \leq p \text{ s.t. } \forall i < k, a^\star_{i\cdot} =_{lmin} b^\star_{i\cdot} \text{ and } a^\star_{k\cdot} >_{lmin} b^\star_{k\cdot}$$

Relation $\succeq_{lmax(\succeq lmin)}$ is a complete preorder. $A \simeq_{lmax(\succeq lmin)} B$ if and only if both matrices have the same coefficients up to the above described rearrangement. Moreover, $\succeq_{lmax(\succeq lmin)}$ refines the ranking obtained by the optimistic criterion:

$$\max_i \min_j a_{ij} > \max_i \min_j b_{ij} \text{ implies } A \succ_{lmax(\succeq lmin)} B.$$

and especially, if $A$ Pareto-dominates $B$ in the strict sense ($\forall i, j, a_{ij} \geq b_{ij}$ and $\exists i^*, j^*$ such that $a_{i^*j^*} > b_{i^*j^*}$), then $A \succ_{lmax(\succeq lmin)} B$.

The comparison of alternatives $f$ and $g$ using the weighted maximum $SLMAX_\pi$ can be refined using relation $\succeq_{lmax(\succeq lmin)}$ applied to $n \times 2$ matrices on $(L, \leq)$, $n$ being the number of criteria, namely comparing matrices $F_\pi$ and $G_\pi$ with coefficients $f_{i1} = \pi_i$ and $f_{i2} = f_i$, $g_{i1} = \pi_i$ and $g_{i2} = g_i$.

Likewise the comparison of alternatives $f$ and $g$ using the weighted maximum $SLMIN_\pi$ can be refined using relation $\succeq_{lmin(\succeq lmax)}$ comparing matrices $[f]_{1-\pi}$ and $[g]_{1-\pi}$ with coefficients $f_{i1} = 1-\pi_i$ and $f_{i2} = f_i$, $g_{i1} = 1 - \pi_i$ and $g_{i2} = g_i$.

**Leximaxmin criteria and weighted average.** It has been proved [64] that the above refinements of $SLMIN_\pi$ and $SLMAX_\pi$ can be represented by weighted averages (e.g., expected utility functionals) using special kinds of probability distributions and real-valued utility functions. First note that, in a finite setting, the qualitative *Leximin* and *Leximax* rules can be simulated by means of a sum of utilities provided that the levels in the qualitative (finite) utility scale $L$ are mapped to values sufficiently far away from one another on a numerical scale. Consider an increasing mapping $\phi$ from $L$ to the reals. It is possible to define this mapping in such a way as to refine the max ordering:

$$\max_{i=1,\dots n} f_i > \max_{i=1,\dots n} g_i \text{ implies } \sum_{i=1,\dots n} \phi(f_i) > \sum_{i=1,\dots n} \phi(g_i) \tag{8}$$

For instance, the transformation $\phi(\lambda_i) = N^i$ with $N > n$ achieves this goal. It is a super-increasing mapping in the sense that $\phi(\lambda_i) > \sum_{j<i} \phi(\lambda_j), \forall i = 1, \dots, m$. In order to map $L$ to [0, 1] so that $\phi(\lambda_0) = 0$ and $\phi(\lambda_n) = 1$ just let $\phi(\lambda_i) = \frac{N^i-1}{N^m-1}$. Note that it is a convex function $[n] \to \mathbb{R}^+$. It can actually be checked that the *Leximax* ordering is equivalent to applying the Bernoulli criterion with respect to such a convex utility function $\phi(.)$:

$$f >_{Leximax} g \text{ if and only if } \sum_{i=1,\dots n} \phi(f_i) > \sum_{i=1,\dots n} \phi(g_i). \tag{9}$$

A similar encoding of the *Leximin* procedure by a sum can be achieved using another super-increasing mapping (for instance, the transformation $\psi(\lambda_i) = \frac{1-N^{-i}}{1-N^{-m}}$ a concave function $L \to \mathbb{R}^+$):

$$f >_{Leximin} g \text{ if and only if } \sum_{i=1,\dots n} \psi(f_i) > \sum_{i=1,\dots n} \psi(g_i) \tag{10}$$

The *Leximin* ordering comes down to applying the Bernoulli criterion with respect to such a concave utility function $\psi(.)$. The qualitative pessimistic and optimistic criteria under total ignorance are thus refined by means of a classical criterion with respect to a risk-averse and risk-prone utility function respectively, as can be seen by plotting $L$ against numerical values in the ranges $\phi(L)$ and $\psi(L)$.

The same results apply to possibilistic criteria $SLMIN_\pi$ and $SLMAX_\pi$ [64] that can be simulated by weighted averages. Consider first the optimistic possibilistic criterion $SLMAX_\pi$ under a given possibility distribution $\pi$. We can again define an increasing mapping $\chi$ from $L$ to the reals such that $\chi(\lambda_0) = 0$ and especially:

$$\begin{array}{l} \max_i \min(\pi_i, f_i) > \max_i \min(\pi_i, g_i) \\ \text{ implies} \\ \sum_{i=1,\dots n} \chi(\pi_i) \cdot \chi(f_i) > \sum_{i=1,\dots n} \chi(\pi_i) \cdot \chi(g_i) \end{array} \tag{11}$$

A sufficient condition is that: $\forall i \in \{1, \dots, m\}, \chi(\lambda_i)^2 \geq N \cdot \chi(\lambda_{i-1}) \cdot \chi(1)$ for some $N > n$. The increasing mapping $\chi(\lambda_i) = \frac{N}{N^{2m-i}}, i = 1, \dots, m$, and $\chi(\lambda_0) = 0$, with $N = n + 1$ can be chosen, with $n = |\mathcal{C}|; m = |L|$. It is such that $\chi(\lambda_m) = 1$

Moreover, let $\{E_0, \dots, E_k\}$ be the well-ordered partition of $\mathcal{C}$ induced by $\pi$, $E_k$ containing the most important criteria, and $E_0$ the least important. Let $K = \frac{1}{\sum_{i=1,k} |E_i| \cdot \chi(\pi_i)}$. Define $\chi^*(\lambda_i) = K \odot \chi(\lambda_i)$, it holds that:

- The weights $p_i = \chi^*(\pi_i)$ define a probability assignment respectful of the possibilistic ordering of criteria. In particular, distribution $p$ is uniform on equally important criteria (the sets $E_j$). Moreover, if $i \in E_j$ then $p_i$ is greater than the sum of the probabilities of all less probable elements, that is, $p_i > P(E_{j-1} \cup \cdots \cup E_0)$. Such probabilities introduced by Snow [100], are said to be *big-stepped* in [2].
- the values $\chi(f_i)$ form a big-stepped numerical utility function (a super-increasing sequence of reals $u_l > \cdots > u_1$ such that $\forall l \geq i > 1, u_i > n \cdot u_{i-1}$) that can be encoded by a convex real mapping $[n] \to \mathbb{R}^+$.
- The preference functional
$$EU_+(f) = \sum_{i=1,\ldots n} \chi^*(\pi_i) \cdot \chi(f_i) \tag{12}$$
is an expected (big-stepped) utility criterion for a risk-seeking decision-maker, that refines the weighted maximum.

The pessimistic criterion $SLMIN_\pi$ can be similarly refined since $SLMAX_\pi(f) = 1 - SLMAX_\pi(1-f)$ using the order-reversing map of $L$. Then, choosing the same mapping $\chi^*$ as above, one may have that
$$\begin{array}{l} \min_i \max(\pi_i, f_i) > \min_i \max(\pi_i, g_i) \\ \text{implies} \\ \sum_{i=1,\ldots n} \chi^*(\pi_i) \cdot \phi(f_i) > \sum_{i=1,\ldots n} \chi^*(\pi_i) \cdot \phi(g_i) \end{array} \tag{13}$$
where $\phi(\lambda_i) = 1 - \chi(1 - \lambda_i))$ (it is equal to $1 - \frac{n+1}{(n+1)^{2^i}}$, here). Function $\phi(\cdot)$ is a super-increasing numerical utility function that can be encoded by a concave real mapping $[n] \to \mathbb{R}^+$, and the weighted average criterion
$$EU_-(f) = \sum_{i=1,\ldots n} \chi^*(\pi_i) \cdot \phi(f_i) \tag{14}$$
is a risk-averse one, that refines $SLMIN_\pi$.

The big-stepped functionals $EU_+(f)$ and $EU_-(f)$ turn out to represent the relations $\succeq_{lmax(\succeq lmin)}$ and $\succeq_{lmin(\succeq lmax)}$: it is proved in [64] that
$$EU_+(f) \geq EU_+(f) \text{ if and only if } [f]_\pi \succeq_{lmax(\succeq lmin)} [g]_\pi \tag{15}$$
$$EU_-(f) \geq EU_-(g) \text{ if and only if } [f]_{1-\pi} \succeq_{lmin(\succeq lmax)} [g]_{1-\pi}. \tag{16}$$

These results point out the deep agreement between qualitative possibilistic criteria and weighted averages. The former is just coarser than the latter, and as such cannot account for compensative effects. As a consequence, the additive preference functionals $EU_+(f)$ and $EU_-(f)$ refining the possibilistic criteria are qualitative despite their numerical encoding (numerical utility values are meaningless, but for ensuring a refined ranking over alternatives).

**Refining Sugeno integral** Applying the increasing transformation $\chi$ that changes a maxmin expression into a sum of products to the minimal disjunctive form $\bigvee_{A \in \mathcal{F}(\mu)} (\mu_\#(A) \wedge \bigwedge_{i \in A} f_i)$ of Sugeno integral $S_\mu(f)$ yields:
$$E_\#^{lsug}(f) = \sum_{A \in 2^{\mathcal{C}}} \chi(\min_{i \in A} f_i) \cdot \chi^*(\mu_\#(A)) = \sum_{A \in 2^{\mathcal{C}}} \min_{i \in A} \chi(f_i) \cdot m_\#(A), \tag{17}$$
where $\chi(\lambda_m) = 1, \chi(\lambda_0) = 0, \chi(\lambda_j) = \frac{K}{K^{2^{m-j}}}, j = 1, m-1$, and we set $K = 2^{|\mathcal{C}|}$. Function $\chi^*$ normalizes $\chi$ in such a way that $\sum_{A \in 2^{\mathcal{C}}} m_\#(A) = 1$, where the positive weights $m_\#(\cdot) = \chi^*(\mu_\#(\cdot))$ define a random set. Ranking tuples by $E_\#^{lsug}(f)$ comes down to a *Leximax($\geq_{lmin}$)* comparison of $(2^n \times 2)$ matrices with rows of the form $(\mu_\#(A), \min_{i \in A} \chi(f_i))$. It is clear that $E_\#^{lsug}(f)$ is a Choquet integral w.r.t. a belief function with basic mass assignment $m_\#$. It refines the original Sugeno integral. More details can be found in [36, 37].

### 3.2 Sugeno utility functionals

In practice, each criterion in an MCDM problem may have its own scale, which requires an extension of Sugeno integral. We will assume that there are possibly distinct scales $L_1, \ldots, L_n$, one per criterion, that are are finite chains, and, with no danger of ambiguity, we will denote the top and bottom elements of $L_i$ by 1 and 0, respectively, for all $i \in \mathcal{C}$. We say that a mapping $\varphi_i \colon L_i \to L$, $i \in [n]$, is a *local utility function* if it is order-preserving. It is a qualitative utility function since it is a mapping between finite chains. A function $\Phi : \mathbf{L} = \prod_{i=1}^{n} L_i \to L$ is said to be a *Sugeno utility functional* (SUF) if there is a Sugeno integral $S_\mu \colon L^n \to L$ and local utility functions $\varphi_i \colon L_i \to L$, $i \in [n]$ with $\varphi_i(0) = 0, \varphi_i(1) = 1$, such that

$$\Phi(f) = S_\mu(\varphi_1(f_1), \ldots, \varphi_n(f_n)). \tag{18}$$

an expression first proposed by Greco et al. [78]. We shall denote SUFs by $S_{\mu,\varphi}$. Note that Sugeno utility functionals are order-preserving. Moreover, it was shown in [30] that the set of functions obtained by composing lattice polynomials with local utility functions is the same as the set of Sugeno utility functionals.

Sugeno utility functionals can be characterized in complete analogy with polynomial functions by extending the notion of median decomposability. We say that $\Phi \colon \mathbf{L} \to L$ is *pseudo-median decomposable* if for each $k \in [n]$ there is a local utility function $\varphi_k \colon L_k \to L$ such that

$$\Phi(f) = \operatorname{med}\left(\Phi(f^{k\to 0}), \varphi_k(f_k), \Phi(f^{k\to 1})\right) \tag{19}$$

for every $f \in \mathbf{L}$.

**Theorem 4 ([31]).** *A function $\Phi \colon \mathbf{L} \to L$ a Sugeno utility functional if and only if $\Phi$ is pseudo-median decomposable.*

*Remark 2.* In [30, 31] a more general notion of pseudo-median decomposability was considered where the inner functions $\varphi_i \colon L_i \to L$, $i \in [n]$, where it is only required to satisfy boundary conditions.

Note that once the local utility functions $\varphi_i \colon L_i \to L$ ($i \in [n]$) are given, the pseudo-median decomposability formula (19) provides a disjunctive normal form of a polynomial function $p_0$ which can be used to factorize $\Phi$. To this extent, let $\widehat{\mathbf{1}}_I$ denote the characteristic vector of $I \subseteq [n]$ in $\mathbf{L}$, i.e., $\widehat{\mathbf{1}}_I \in \mathbf{L}$ is the $n$-tuple whose $i$-th component is $1_{L_i}$ if $i \in I$, and $0_{L_i}$ otherwise.

**Theorem 5 ([31]).** *If $\Phi \colon \mathbf{L} \to L$ is pseudo-median decomposable w.r.t. local utility functions $\varphi_k \colon L_k \to L$ ($k \in [n]$), then $\Phi = p_0(\varphi_1, \ldots, \varphi_n)$, where the polynomial function $p_0$ is given by*

$$p_0\left(f_1, \ldots, f_n\right) = \bigvee_{I \subseteq [n]} \left(\Phi\left(\widehat{\mathbf{1}}_I\right) \wedge \bigwedge_{i \in I} f_i\right). \tag{20}$$

In other words, this theorem characterizes SUFs, namely, $p_0\left(f_1, \ldots, f_n\right) = S_{\mu,\varphi}(f)$ if $p_0$ is idempotent (i.e., we must have $\Phi\left(\widehat{\mathbf{1}}_{[n]}\right) = 1$ and $\Phi\left(\widehat{\mathbf{1}}_{\emptyset}\right) = 0$, and $\Phi$ is order preserving).

*Remark 3.* Procedures to obtain local utility functions $\varphi_i \colon L_i \to L$ ($i \in [n]$), which can be used to factorize a given Sugeno utility functional $f \colon \mathbf{L} \to L$ into a composition (18), were presented in [30] when $L$ is an arbitrary chain, and in [31] when $L$ is a finite distributive lattice.

Another kind of axiomatization was proposed quite early by Greco et al. [78], namely conditions under which a preference relation $\succeq$ on $\mathbf{L} = \prod_{i=1}^{n} L_i \to L$ can be represented by a Sugeno utility functional. Namely, the following equivalence between the two following statements is obtained (a part of Th. 1 in [78], here recalled in the finite setting):

- the preference relation $\succeq$ on $\mathbf{L}$ is a complete preordering such that for all $f, g, h, h' \in \mathbf{L}$, $\theta, \beta \in L_i, i \in [n]$, if $f^{i\to\theta} \succeq h$, and $g^{i\to\beta} \succeq h'$, then $g^{i\to\theta} \succeq h$ or $f^{i\to\beta} \succeq h'$;

– There is a Sugeno integral and local utility functions $\varphi_k \colon L_k \to L$ $(k \in [n])$ such that

$$S_\mu(\varphi_1(f_1), \ldots, \varphi_n(f_n)) \geq S_\mu(\varphi_1(g_1), \ldots, \varphi_n(g_n)) \quad \text{if and only if} \quad f \succeq g$$

Bouyssou et al [10] reconsider this result in the scope of conjoint measurement for their non-compensatory decomposable representation model, and show that Greco et al. axiom is a strong form of their non-compensation axiom stating that for all $f, g, h, h' \in \mathbf{L}, i \in [n]$, if $f \succeq h$ and $g \succeq h'$ then $f^{i \to g_i} \succeq h$ or $g^{i \to f_i} \succeq h'$.

Finally, another, simpler axiom has been used by Couceiro et al. [20] (Th. 3.6) to characterize Sugeno utility functionals:

$$\text{for all } f, g \in \mathbf{L}, \theta \in L_k, k \in [n), \text{ if } f^{k \to \theta} \succ f^{k \to 0} \text{ and } g^{k \to 1} \succ g^{k \to \theta} \text{ then } g^{k \to \theta} \succeq f^{k \to \theta}.$$

This axiom is closely connected to the median-decomposability of Sugeno integrals. Note that when $L_i = L, \forall i \in [n]$, Sugeno utility functionals are more expressive than Sugeno integrals even in the special cases of $SLMIN_\pi$ and $SLMAX_\pi$, as shown by the counterexample given in [20]:

*Example 1.* Let $L = \{0, \lambda, 1\}$ endowed with the ordering $0 < \lambda < 1$, and consider the preference relation $\succeq$ on $\mathbf{L} = L^2$ whose linearly ordered equivalence classes are

$$\begin{aligned}
[(1,1)] &= \{(1,1), (1,\lambda), (1,0), (0,1), (\lambda,1)\}, \\
[(\lambda,\lambda)] &= \{(\lambda,\lambda), (\lambda,0)\}, \\
[(0,\lambda)] &= \{(0,\lambda), (0,0)\}.
\end{aligned}$$

This relation does not satisfy axiom **RCD**, e.g., take $f = (0, \lambda)$, $g = (0, 1)$ and $h = (\lambda, \lambda)$, where $g \succ f$ and $h \succ f$ but $f = h \wedge g$. Thus it cannot be represented by a Sugeno integral. However, letting $S_\mu(f) = f_1 \vee f_2$ (i.e. $\mu$ is the uniform possibility distribution), and utility functions $\varphi_1$ equal to the identity and $\varphi_2(1) = 1, \varphi_2(\lambda) = \varphi_2(0) = 0$, it can be checked that $\succeq$ is represented by the Sugeno utility functional $f_1 \vee \varphi_2(f_2)$.

### 3.3 Generalized Sugeno integrals

In Sugeno integrals, the role of weights is devoted to shrinking the evaluation scales of each group of criteria. For instance, in the weighted maximum, the range of evaluation for a criterion $i$ of weight $\pi_i < 1$ is restricted to $[0, \pi_i] \subset L$; in a weighted minimum, it is restricted to $[1 - \pi_i, 1] \subset L$. Any evaluation $f_i$ greater than $\pi_i$ in $SLMAX_\pi$ (resp. less than $1 - \pi_i$ in $SLMIN_\pi$) is brought back to $\pi_i$ (resp. $1 - \pi_i$). This role of weights is very peculiar. In this part we explore various alternative weighting schemes. More precisely we are going to generalize rating modification schemes by weights in Sugeno integrals. We first consider generalized forms of aggregation functions $SLMIN_\pi$ and $SLMAX_\pi$, $\bigwedge_{i=1}^n (\pi_i \rhd f_i)$ and $\bigvee_{i=1}^n (\pi_i \otimes f_i)$ respectively, where $\rhd$ and $\otimes$ are suitable rating modification operations.

We first discuss the nature of such operations $\rhd$ and $\otimes$ following intuitive requirements.

In a conjunctive aggregation $\bigwedge_{i=1}^n (\pi_i \rhd f_i)$, the following conditions are natural:

(i) The global evaluation should not be affected by a useless criterion ($\pi_i = 0$) even if the rating is maximal ($f_i = 1$), hence we assume $0 \rhd f_i = 1$.

(ii) A very poor rating on a criterion with top importance $\pi_i = 1$ should be enough to bring the global evaluation down to 0, hence we assume $1 \rhd 0 = 0$.

(iii) The better is the local rating $f_i$, the greater is the modified rating $\pi_i \rhd f_i$ ($\pi_i \rhd f_i$ should increase with $f_i$).

(iv) The less important the criterion, the more lenient should be the modified local rating ($\pi_i \rhd f_i$ should decrease with $\pi_i$).

So the operation $\rhd$ should be a generalized implication.

**Definition 1.** *A fuzzy implication is a two-place operation $\to$ on $L$ such that:*

*(i)* $0 \to 1 = 1$*;* $1 \to 0 = 0$*;* $1 \to 1 = 1$*;* $0 \to 0 = 1$*.*
*(ii)* $\alpha \to \beta$ *is increasing in the wide sense with* $\beta$ *for all* $\alpha \in L$*.*
*(iii)* $\alpha \to \beta$ *is decreasing in the wide sense with* $\alpha$ *for all* $\beta \in L$*.*

In a disjunctive aggregation $\bigvee_{i=1}^{n}(\pi_i \otimes f_i)$ similar conditions should hold:

(i) For the same reason as in the conjunctive case, we should have $0 \otimes f_i = 1$.
(ii) A top rating on a criterion with top importance should be enough to bring the global evaluation to 1, hence we assume $1 \otimes 1 = 1$.
(iii) $\pi_i \otimes f_i$ should increase with $f_i$ as in the conjunctive case.
(iv) A good rating on an important criterion should have more positive influence that one on a little important criterion ($\pi_i \otimes f_i$ should increase with $\pi_i$).

So the operation $\otimes$ should be a generalized conjunction.

**Definition 2.** *A fuzzy conjunction is a two-place operation $\otimes$ on $L$ such that:*

*(i)* $0 \otimes 1 = 0$*;* $1 \otimes 0 = 0$*;* $1 \otimes 1 = 1$*;* $0 \otimes 0 = 0$*.*
*(ii)* $\alpha \otimes \beta$ *is increasing in the wide sense with* $\alpha$*, for all* $\beta \in L$*.*
*(iii)* $\alpha \otimes \beta$ *is increasing in the wide sense with* $\beta$ *for all* $\alpha \in L$*.*

Note that to each fuzzy implication $\rhd$ in the sense of Def. 1 corresponds a fuzzy conjunction in the sense of Def. 2 $\otimes = \mathcal{S}(\rhd)$ such that $\alpha\mathcal{S}(\rhd)\beta = 1 - (\alpha \rhd (1-\beta))$. We shall say that such a pair of operations $(\rhd, \otimes)$ are *semidual*. Moreover these connectives can be exchanged, i.e., $\alpha \rhd \beta = \alpha\mathcal{S}(\otimes)\beta = 1 - (\alpha \otimes (1-\beta))$, i.e., semiduality is an involutive transformation. Note that $\alpha \rhd \beta = 1$ if and only if $\alpha\mathcal{S}(\rhd)(1-\beta) = 0$ that is, $\alpha$ and $1-\beta$, when positive, are divisors of 0 for the fuzzy conjunction $\otimes = \mathcal{S}(\rhd)$.

It is interesting to notice that the equality (2) can be expressed using the conjugate capacity $\mu^c(A) = 1 - \mu(\overline{A})$ for every $A \subseteq \mathcal{C}$ and the semi-dual of $\wedge$, namely $\alpha\mathcal{S}(\wedge)\beta = \max(1-\alpha, \beta)$, which is Kleene-Dienes implication $\to_{KD}$:

$$S_\mu(f) = \bigvee_{A \subseteq \mathcal{C}} \big(\mu(A) \wedge \bigwedge_{i \in A} f_i\big) = \bigwedge_{A \subseteq \mathcal{C}} (\mu^c(A) \to_{KD} \bigvee_{i \in A} f_i). \tag{21}$$

This equality corresponds to extending the duality relation between a capacity $\mu$ and its conjugate to Sugeno integrals [77], i.e. (7) ($S_\mu(f) = 1 - S_{\mu^c}(1-f)$).

Sugeno integral can be generalised with other inner fuzzy conjunction or implication functions linked by semiduality, using expressions we call *q-integrals* and *q-cointegrals*, denoted respectively by $\int_\mu^\otimes f$ and $\int_\mu^\to f$.

**Definition 3.** *[57] Let $\otimes$ be a fuzzy conjunction. A q-integral is the mapping $\int_\mu^\otimes : L^{\mathcal{C}} \to L$ defined by*

$$\int_\mu^\otimes f = \bigvee_{A \subseteq C} \big(\mu(A) \otimes \bigwedge_{i \in A} f_i\big), \textit{ for all } f \in L^{\mathcal{C}}.$$

**Definition 4.** *[57] Let $\to$ be a fuzzy implication, a q-cointegral is a mapping $\int_\mu^\to : L^{\mathcal{C}} \to L$ defined by*

$$\int_\mu^\to f = \bigwedge_{A \subseteq C} \big(\mu^c(A) \to \bigvee_{i \in A} f_i\big), \textit{ for all } f \in L^{\mathcal{C}}.$$

Note that, when $\otimes$ is the product and $L = [0,1]$, the q-integral is Shilkret integral [98] proposed in 1971 (and later reintroduced by A. Kaufmann [82] in 1979, under the name of "admissibility"). Grabisch et al. [77] introduce so-called Sugeno-like integrals, which are similar to q-integrals, where $\otimes$ is a triangular norm. Borzová-Molnárová et al. [8] study this type of integrals in the continuous case as well when $\otimes$ is a semicopula and $L = [0,1]$. This kind of definition is also proposed by Dvořák and Holčapek [62] assuming $(L, \otimes, 1)$ is a commutative monoid and considering the complete residuated lattice generated by this monoidal operation. In fact, what they study is an extension of Definition 3: Namely, they study fuzzy integrals of this type *extended to algebras of fuzzy sets,* that is, where $\mu$ is now a *fuzzy set* function that assigns an importance value $\mu(\tilde{A})$ to any *fuzzy* subset $\tilde{A}$ of $\mathcal{C}$.

Now if we consider the dual quantity $1 - \int_\mu^\otimes (1-f)$, it is of the form of a cointegral $\int_\mu^\rightarrow f$ with respect to the semi-dual implication $\rightarrow = \mathcal{S}(\otimes)$. But it can be checked that we no longer have the equality (2), that is now

$$\int_\mu^\otimes f \neq \int_\mu^{\mathcal{S}(\otimes)} f$$

while we do have that $S_\mu(f) = \int_\mu^\wedge f = \int_\mu^{\rightarrow KD} f$, for every capacity $\mu$ and every $f \in L^{\mathcal{C}}$, as expressed by identity (2). There are not many works considering q-cointegrals in the above sense. Grabisch et al. ([77], p. 302) notice the failure of the duality relation (7) for Sugeno-like integrals that use t-norms other than $\min$, which hints at co-integrals using implications that are semi-duals of t-norms. The study of solutions to the equality $\int_\mu^\otimes f = \int_\mu^{\mathcal{S}(\otimes)} f$ is carried out in [5]. It turns out that there exist very few operations other than Kleene-Dienes implication and Kleene conjunction for which these integrals coincide (the conjunctions must be of the form $\varphi(\alpha) \wedge \psi(\beta)$ for local utility functions $\varphi, \psi : L \to L$).

In [55], we study q-cointegrals induced by Gödel implication: $\alpha \rightarrow_G \beta = \begin{cases} 1 \text{ if } \alpha \leq \beta, \\ \beta \text{ otherwise} \end{cases}$ and the q-integral induced by its semi-dual non-commutative conjunction [45]:

$$\alpha \otimes_G \beta = 1 - (\alpha \rightarrow_G (1-\beta)) = \begin{cases} 0 \text{ if } a \leq 1-\beta, \\ \beta \text{ otherwise.} \end{cases}$$

We also study q-cointegrals induced by the contrapositive symmetric of Gödel implication defined by $\alpha \rightarrow_{GC} \beta = \begin{cases} 1 \text{ if } \alpha \leq \beta, \\ 1-\alpha \text{ otherwise} \end{cases}$ and its semidual conjunction, which is equal to $\beta \otimes_G \alpha$. Using a necessity measure for $\mu$ and implication $\rightarrow_G$, the aggregation $\int_\mu^{\rightarrow_G} f$ selects all alternatives that pass the prescribed threshold for each criterion and ranks the remaining ones according to their worst local ratings (forming a waiting list). Using $\rightarrow_{GC}$, the aggregation selects all alternatives that pass the prescribed threshold for each criterion and ranks the remaining ones according to the importance of violated criteria, putting the objects that violate the least important criteria upfront. Moreover it is shown in [55] that $\int_\mu^{\otimes_G f} > \int_\mu^{\rightarrow_G} f$ and $\int_\mu^{\otimes_{GC}} f > \int_\mu^{\rightarrow_{GC}} f$ in general. This inequality cannot even be generalised to other conjunctions. More precisely, if $\otimes$ is a triangular norm, or a copula, then $\int_\mu^\rightarrow (\theta A \lambda) > \int_\mu^\otimes (\theta A \lambda)$ in general ($\theta A \lambda$ is the alternative that takes value $\theta$ for $i \in A$ and $\lambda$ otherwise). However, if $\otimes$ is greater than the minimum (which is the case with the semi-dual of the contrapositive symmetric of Gödel implication), then $\int_\mu^\rightarrow (\alpha A \beta) < \int_\mu^\otimes (\alpha A \beta)$ in general. Other q-integrals and q-co-integrals are studied in [57].

**Elementary properties of q-integrals and co-integrals** The counterpart of the Sugeno integral expression in terms of the nested family of subsets $\{(i), \cdots, (n)\}$, induced by $f$, where $f_{(1)} \leq \cdots \leq f_{(n)}$, and the simplified form of Sugeno q-integral are still valid for q-integrals:

$$\int_\mu^\otimes f = \bigvee_{i=1}^n \mu(\{(i), \cdots, (n)\}) \otimes f_{(i)}, \qquad \int_\mu^\otimes (f) = \bigvee_{\lambda \in L} \mu(\{f \geq \lambda\}) \otimes \lambda.$$

If $\mu$ is a possibility measure $\Pi$ based on possibility distribution $\pi$ we retrieve the $\otimes$-weighted maximum: $\int_{\Pi}^{\otimes} f = MAX_{\pi}^{\otimes}(f)$. However, if $\mu$ is a necessity measure, the expression of the q-integral will not simplify when $\otimes \neq \min$. In other words we do not have that $\int_{N}^{\otimes} f = MIN_{\pi}^{\rightarrow}(f)$ for $\rightarrow = \mathcal{S}(\otimes)$ except when $\otimes = \wedge$.

Using semi-duality, q-cointegrals can be expressed in terms of q-integrals since $\int_{\mu}^{\rightarrow} f = 1 - \int_{\mu^c}^{\otimes}(1-f)$. In [57] we derive the following results from the ones on (conjunction-based) q-integrals:

$$\int_{\mu}^{\rightarrow} f \;=\; \bigwedge_{i=1}^{n} \mu^c(\{(1), \cdots, (i)\}) \rightarrow f_{(i)}, \qquad \int_{\mu}^{\rightarrow} f \;=\; \bigwedge_{\lambda \in L} \mu^c(\{f \leq \lambda\}) \rightarrow \lambda.$$

We also get that when $\mu$ is a necessity measure $N$ based on possibility measure $\pi$, the q-cointegral reduces to the $\rightarrow$-weighted minimum: $\int_{N}^{\rightarrow} f = MIN_{\pi}^{\rightarrow}(f) = \wedge_{i=1}^{n} \pi_i \rightarrow f_i$. However the q-cointegral with respect to a possibility measure $\int_{\Pi}^{\rightarrow} f$ does not reduce to a weighted maximum.

The characterization result in Theorem 2 for Sugeno integrals can be extended to q-integrals and co-integrals.

**Theorem 6.** *[57] Let $I\colon L^{\mathcal{C}} \rightarrow L$ be a mapping. There are a capacity $\mu$ and a fuzzy conjunction $\otimes$ such that $I(f) = \int_{\mu}^{\otimes} f$ for every $f \in L^{\mathcal{C}}$ if and only if*

*(i) $I(f \vee g) = I(f) \vee I(g)$, for any comonotone $f, g \in L^{\mathcal{C}}$.*
*(ii) There are a capacity $\kappa\colon 2^{\mathcal{C}} \rightarrow L$ and a binary operation $\star$ on $L$ such that $I(\lambda \wedge \mathbf{1}_A) = \kappa(A) \star \lambda$ for every $\lambda \in L$ and every $A \subseteq \mathcal{C}$.*
*(iii) $I(\mathbf{1}_{\mathcal{C}}) = 1$ and $I(\mathbf{0}_{\mathcal{C}}) = 0$.*

*In that case, we have $\mu = \kappa$ and $\otimes = \star$.*

If the fuzzy conjunction $\otimes$ satisfies $\lambda \otimes 1 = \lambda$, then under the assumptions of Theorem 6, the functional $I$ is of the form $I(f) = \int_{\mu}^{\otimes} f$ where $\mu(A) = I(\mathbf{1}_A)$. There is a specific result in [57] when the functional $I$ is fully maxitive, to characterize possibilistic q-integrals of the form $I(f) = \int_{\Pi}^{\otimes} f$.

Since the fuzzy conjunction $\otimes$ is not supposed to be commutative, there is a companion q-integral $\int_{\mu}^{\star} f$ with $\lambda \star \lambda' = \lambda' \otimes \lambda$ and a similar characterization result.

For q-cointegrals defined from fuzzy implications, similar characterization results have been derived. However they use counterparts of properties (i'), (ii') and (iii') of Sugeno integral recalled under Theorem 2. We consider a fuzzy implication $\rightarrow$, which can always be assumed to be of the form $\alpha \rightarrow \beta = 1 - \alpha \otimes (1 - \beta)$ for a fuzzy conjunction $\otimes$. This semi-duality property leads to obtain the following characterisation result.

**Theorem 7.** *Let $I\colon L^{\mathcal{C}} \rightarrow L$ be a mapping. There are a capacity $\mu$ and a fuzzy implication $\rightarrow$ such that $I(f) = \int_{\mu}^{\rightarrow} f$ for every $f \in L^{\mathcal{C}}$ if and only if the following properties are satisfied.*

*(i) $I(f \wedge g) = I(f) \wedge I(g)$, for any comonotone $f, g \in L^{\mathcal{C}}$.*
*(ii) There are a capacity $\rho\colon 2^{\mathcal{C}} \rightarrow L$ and a binary operation $\rhd$ such that $I(\lambda \vee \mathbf{1}_A) = \rho^c(\overline{A}) \rhd \lambda, \forall \lambda \in L$.*
*(iii) $I(\mathbf{1}_{\mathcal{C}}) = 1$ and $I(\mathbf{0}_{\mathcal{C}}) = 0$.*

*In that case $\rho = \mu$, and $\rhd = \rightarrow$.*

Note that the homogeneity condition $I(\lambda \vee \mathbf{1}_A) = \rho^c(\overline{A}) \rightarrow \lambda$ for q-cointegrals is better understood if we express the latter expression $(1 - \rho(\overline{A})) \rightarrow \lambda$ as $\rho(\overline{A}) \oplus \lambda$, where $\oplus$ is a disjunction built as the De Morgan dual $1 - (1 - \cdot) \otimes (1 - \cdot)$ of $\otimes = \mathcal{S}(\rightarrow)$, which means it is indeed the disjunctive counterpart of the homogeneity condition $I(a \wedge \mathbf{1}_A) = \lambda(A) \otimes a$ for q-integrals. Clearly, if the fuzzy implication function is such that $\lambda \rightarrow 0 = 1 - \lambda$, then $I(\mathbf{1}_A) = 1 - \rho^c(\overline{A}) = \rho(A)$ (for instance $\rightarrow$ is

the symmetric contrapositive of a residual fuzzy implication induced by a conjunction having two-sided identity 1, such as $\wedge$, for which we have $\lambda \rightarrow_{CG} 0 = 1 - \lambda$). There is a specific result in [57] when the functional $I$ is fully minitive, to characterize possibilistic q-integrals of the form $I(f) = \int_N^{\rightarrow} f$. In the same paper, representation results for the companion q-cointegral defined from a q-cointegral by contrapositive symmetry are proposed.

To summarize, the above results indicate that properties of Sugeno integrals remain valid for more general functionals where the weights of groups of criteria may variously act to modify local evaluations. However these set of properties is split into some for q-integrals and others for q-co-integrals:

- Comonotonic maxitivity is specific to q-integrals and comonotonic minitivity is specific to q-cointegrals.
- Q-integrals and q-cointegrals do not satisfy the same homogeneity conditions, the former being homogeneous with respect to the meet $\wedge$ and the latter with respect to the join $\vee$.
- Q-integrals simplify if computed wrt a possibility measure and q-cointegrals simplify wrt a necessity measure but not conversely.

## 4 Bipolar evaluation methods

When rating alternatives on a scale $L$, it is useful to have a clear understanding of the meaning of its end-points. The explicit handling of positive and negative judgments when selecting an alternative must be distinguished from the simple need for ranking alternatives in terms, e.g., of preference. People also need to express that some alternative is good or bad for them, a notion that simple preference relations cannot express. Using a simple preference relation, the best available choice may fail to be really suitable for the decision-maker. In other circumstances, even the worst ranked option remains somewhat acceptable. To discriminate between these two situations, one absolute landmark or reference point expressing neutrality or indifference, and explicitly separating the positive and the negative judgments, must appear in the model. [1]

Modeling this situation requires a *bipolar scale* $(L, >)$, i.e., a totally ordered set with a prescribed interior element $e$ called *neutral*, separating the positive range of evaluations $\lambda > e$ from the negative one $\lambda < e$. Mathematically, if the scale is equipped with a binary operation $\star$ (an aggregation operator), $e$ is an idempotent element for $\star$, possibly acting as an identity. Interestingly, classical utility theory does not exploit bipolarity. Utility functions are defined up to an increasing affine transformation (i.e., they rely on an interval scale), and the separation between positive and negative evaluations has no special meaning. In contrast, Cumulative Prospect Theory (CPT, for short) proposed by [103] is an attempt to explicitly account for positive and negative evaluations using the real line as a genuine bipolar scale. The simplest qualitative bipolar scale contains three elements: $\{-, 0, +\}$. In bipolar scales, the negative side of the scale (below $e$) is the mirror image of the positive one (above $e$). An object is evaluated on such a bipolar scale as being either positive or negative or neutral. It cannot be positive and negative at the same time. This is called a *univariate bipolar* framework.

However, it is known from many experiments in cognitive psychology [87, 12, 99] that humans often evaluate alternatives for the purpose of decision-making by considering positive and negative aspects *separately*. Under this bipolar view, comparing two alternatives comes down to comparing pairs of sets of arguments or features, namely, the set of pros and cons pertaining to one alternative versus the set of pros and cons pertaining to the other. This view of bipolarity requires the use of two unipolar qualitative scales $L^+$ and $L^-$ (a positive one and a negative one) instead of a unique bipolar scale. This is the *bivariate unipolar* framework. Here each scale is unipolar in the sense that the neutral level is at one end of the scale. In a *positive* scale the bottom element is neutral. In a *negative* scale the top element is

[1] Even ordinal decision methods need to inject some form of bipolarity. Note that multicriteria decision methods based on the merging of outranking relations use concordance and discordance tests between criteria [96], where the notion of veto prevents the choice of alternatives that rate too low with respect to some criteria. It can be viewed as an attempt to capture the idea of bipolar preference [88].

neutral. A bipolar scale can be viewed as the union of a positive and a negative scale $L^+ \cup L^-$ extending the ordering relations on each scale so $\forall \lambda^+ \in L^+, \lambda^- \in L^-, \lambda^+ > \lambda^-$.

There are in fact three forms of bipolarity can be found at work in the literature, we call types I, II, III in [51]:

- **Type I: Symmetric univariate bipolarity**. It relies on the use of bipolar scales.
- **Type II: Symmetric bivariate bipolarity.** It relies on the use of two unipolar scales related via duality. Positive and negative strengths are computed similarly on the basis of the same data and can be conflicting. This is the case of argumentation systems where reasons for an alternative and reasons against it are collected from the same knowledge base prior to making a decision. Psychologists have shown that the simultaneous presence of positive and negative arguments prevents decisions from being simple to make, except when their strengths have different orders of magnitude.
- **Type III: Asymmetric bipolarity.** In this form of bipolarity, the negative part of the information is not of the same nature as the positive part, while in type II bipolarity only the polarities are opposite. In decision-making, this kind of bipolarity corresponds to the opposition between soft constraints (that eliminate unwanted alternatives) and criteria (that evaluate preferred ones), as discussed in [3]. Constraints have a prominent role and first select the most tolerated alternatives; positive preferences (such as goals and desires) then act to discriminate among this set of non-rejected alternatives. Hence a positive evaluation, even if high, can never outperform a negative evaluation even if very weak. In this approach, negative features prevail over positive ones. The latter matter only when no constraint is violated. In the previous type of bipolarity, positive and negative arguments play symmetric roles.

In this section, we review two bipolar approaches to decision related to Sugeno integrals, namely the symmetric Sugeno integral, and a qualitative counterpart of CPT, where positive and negative sides of alternatives are evaluated separately in the spirit of argumentation theory.

### 4.1 The symmetric Sugeno Integral

Let $L$ be a totally ordered set with bottom element 0, and let $-L := \{-\lambda : \lambda \in L\}$ be its "symmetric" copy endowed with the reversed order. Consider the symmetric ordered structure $\tilde{L} := L \cup (-L) \setminus \{-0\}$, a bipolar scale analogous the real line where the zero that acts as a neutral element and such that $\lambda + (-\lambda) = 0$. In particular, $-(-\lambda) = \lambda$. The question is thus how to define lattice polynomial functions on such bipolar ordered structures while keeping the symmetry with respect to $0$. In particular, we seek a symmetric extension of the Sugeno integral.

As we saw, lattice polynomial functions on distributive lattices (in particular, on linearly ordered sets) are can be represented in disjunctive normal form (see, e.g., [26, 27, 67]). Thus any lattice polynomial function could be in principle defined on a symmetric structure $\tilde{L}$ if one could define the symmetric extensions of maximum and minimum on $L$ over the bipolar $\tilde{L}$.

The symmetric minimum $\oslash$ (playing the role of the product in the real line) is rather simple. Define the absolute value of $\lambda \in \tilde{L}$ as:$|\lambda| = \lambda$ if $\lambda \in L$ and $-\lambda$ otherwise. Grabisch [72, 73] proposed the following notion of *symmetric minimum*:

$$\lambda \oslash \lambda' := \begin{cases} -(|\lambda| \wedge |\lambda'|) & \text{if } \operatorname{sign}(\lambda) \neq \operatorname{sign}(\lambda') \\ |\lambda| \wedge |\lambda'| & \text{otherwise.} \end{cases} \tag{22}$$

The absolute value of $\lambda \oslash \lambda'$ equals $|\lambda| \wedge |\lambda'|$ and $\lambda \oslash \lambda' < 0$ if and only if the two elements $\lambda$ and $\lambda'$ have opposite signs. Like the usual minimum operator $\wedge$, the symmetric minimum $\oslash$ is associative.

Now, the symmetric notion of maximum playing the role of a sum is more challenging. Intuitively, the *symmetric maximum* $\obar$ should extend the maximum on $L$ with $0$ as neutral element, and should fulfill the symmetry requirement: for every $\lambda \in \tilde{L}$, $\lambda \obar (-\lambda) = 0$. However, unlike with the minimum, this symmetry requirement immediately implies that any extension $\obar$ of the maximum operator $\vee$ cannot be

associative. To illustrate this point, let $L = \mathbb{N}$ and observe that $(2 \circledvee 3) \circledvee (-3) = 3 \circledvee (-3) = 0$ where as $2 \circledvee (3 \circledvee (-3)) = 2 \circledvee 0 = 2$.

Nonetheless, Grabisch [72] showed that the "best" definition of $\circledvee$ (see Theorem 8 below) is:

$$\lambda \circledvee \lambda' = \begin{cases} -(|\lambda| \vee |\lambda'|) & \text{if } \lambda' \neq -\lambda \text{ and } |\lambda| \vee |\lambda'| = -\lambda \text{ or } = -\lambda' \\ 0 & \text{if } \lambda' = -\lambda \\ |\lambda| \vee |\lambda'| & \text{otherwise.} \end{cases} \tag{23}$$

In other words, if $\lambda' \neq -\lambda$, then $\lambda \circledvee \lambda'$ returns the element that is the larger in absolute value among the two elements $\lambda$ and $\lambda'$. Moreover, it is not difficult to see that $\circledvee$ satisfies the following properties:

(C1) $\circledvee$ coincides with the maximum on $L^2$;
(C2) $\lambda \circledvee (-\lambda) = 0$ for every $\lambda \in \tilde{C}$;
(C3) $-(\lambda \circledvee \lambda') = (-\lambda) \circledvee (-\lambda')$ for every $\lambda, \lambda' \in \tilde{C}$.

Hence, $\circledvee$ almost behaves like $+$ on the real line, but for associativity. As shown in [72], if one requires that (C1), (C2) and (C3) hold, then (23) is the best possible definition for $\circledvee$.

**Theorem 8.** *[72, Prop. 5] No binary operation satisfying (C1), (C2), (C3) is* associative *on a larger domain than* $\circledvee$.

The following result presents some further properties of $\circledvee$ and describes the sequences that reveal the nonassociativity of $\circledvee$.

**Proposition 2.** *[72, 22, Prop. 5] The symmetric maximum has the following properties:*

*(i)* $\circledvee$ *is commutative on* $\tilde{L}$.
*(ii)* $0$ *is the neutral element of* $\circledvee$.
*(iii)* $\circledvee$ *is associative on an expression involving* $\alpha_1, \ldots, \alpha_n \in \tilde{L}$, *with* $|\{i : \alpha_i \neq 0\}| > 2$, *if and only if* $\bigvee_{i=1}^n \alpha_i \neq -\bigwedge_{i=1}^n \alpha_i$.
*(iv)* $\circledvee$ *is nondecreasing in each argument on* $\tilde{L}$.

Sequences fulfilling condition (iii) were referred to as *associative* in [22].

Now the ambiguity in evaluating $\circledvee$ on nonassociative sequences makes it hard to use it for defining lattice polynomial functions on $\tilde{L}$, since the result of $\circledvee_{i=1}^k a_i$ depends on the particular way the binary $\circledvee$ is applied to the terms of $(\alpha_i)_{1 \leq i \leq k}$. Grabisch [72] suggested ways of making $\circledvee$ associative that were fully developed in [22], namely, by fixing beforehand a systematic way of putting parentheses on any sequence of $\tilde{L}^*$, procedure that was called a *rule of computation*.

**Making $\circledvee$ associative: rules of computation** We now recall the formalism of [22]. As, we will only consider countable sequences of elements of $\tilde{L}$, without loss of generality, we may assume that $\tilde{L} = \mathbb{Z}$. In this way, elements of $\tilde{L}^*$ are (finite) sequences of integers, denoted by $\sigma = (\lambda_i)_{i \in I}$ for some finite index set $I$, including the empty sequence $\varepsilon$, i.e.,

$$\tilde{L}^* = \Big( \bigcup_{n \in \mathbb{N}} (\tilde{L})^n \Big) \cup \{\varepsilon\}.$$

This convention will simplify our exposition and establish connections to the theory of integer means.

Also, as $\circledvee$ is commutative, the order of symbols in the word does not matter, and we can consider the decreasing order of the absolute values of the elements in the sequence (e.g., $5, 5, -5, -3, 2, -2, 1, 0$). Since sequences are ordered, we can consider the following convenient formalism for representing sequences. For an arbitrary sequence

$$\sigma = (\underbrace{n_1, \ldots, n_1}_{p_1 \text{ times}}, \underbrace{-n_1, \ldots, -n_1}_{m_1 \text{ times}}, \ldots, \underbrace{n_q, \ldots, n_q}_{p_q \text{ times}}, \underbrace{-n_q, \ldots, -n_q}_{m_q \text{ times}})$$

let $\theta(\sigma) = (n_1, \ldots, n_q)$ be the sequence of absolute values (magnitudes) of integers in $\sigma$, and let $\psi(\sigma) = ((p_1, m_1), \ldots, (p_q, m_q))$ be the sequence of pairs of numbers of occurrence of these integers. For instance, if $\sigma = (3, 3, -3, 2, -2, -2, 1, 1, 1, 1)$, then

$$\theta(\sigma) = (3, 2, 1); \quad \psi(\sigma) = ((2, 1), (1, 2)(4, 0)).$$

Let $\mathfrak{S}$ denote the set of all integer sequences in this formalism, including the empty sequence, and let $\mathfrak{S}_0$ be the subset of all nonassociative sequences.

This formalism facilitates the precise definition of rules of computation in terms of 5 elementary rules $\rho_i : \mathfrak{S} \to \mathfrak{S}$ that act on $\sigma$ in the following way:

(i) Elementary rule $\rho_1$: if $p_1 > 1$ and $m_1 > 0$, then $p_1$ is changed to $p_1 = 1$;
(ii) Elementary rule $\rho_2$: idem with $p_1, m_1$ exchanged;
(iii) Elementary rule $\rho_3$: if $p_1 > 0$, $m_1 > 0$, the pair $(p_1, m_1)$ is changed into $(p_1 - c, m_1 - c)$, where $c = p_1 \wedge m_1$;
(iv) Elementary rule $\rho_4$: if $p_1 > 0$, $m_1 > 0$, and if $p_2 > 0$, then $p_2$ is changed into $p_2 = 0$;
(v) Elementary rule $\rho_5$: idem with $m_2$ replacing $p_2$.

Hence, elementary rules delete terms only in nonassociative sequences, and leave the associative ones invariant.

A *(well-formed) computation rule* $R$ is a word built with the alphabet $\{\rho_1, \ldots, \rho_5\}$, i.e., $R \in \mathcal{L}(\rho_1, \ldots, \rho_5)$, such that $R(\sigma) \in \mathfrak{S} \setminus \mathfrak{S}_0$ for all $\sigma \in \mathfrak{S}$. The set of (well-formed) computation rules is denoted by $\mathfrak{R}$. Examples of rules are (words are read from left to right)

(i) $\langle\cdot\rangle_0 = \rho_3^*$, that corresponds to putting parentheses around each pair of maximal symmetric terms.
(ii) $\langle\cdot\rangle_= = (\rho_1\rho_2\rho_3)^*$, that corresponds to putting parentheses around terms with the same absolute value and sign, and then to putting parentheses around each each pair of maximal symmetric resulting terms.
(iii) $\langle\cdot\rangle_-^+ = (\rho_4\rho_5)^*\rho_1\rho_2\rho_3$, that corresponds to first putting parentheses around all positive terms and all negative terms, and then computing the symmetric maximum of the two results.

It is shown in [22] that each each computation rule $R \in \mathfrak{R}$ corresponds to an arrangement of parentheses together with a permutation on the terms of sequences. Thus each $R \in \mathfrak{R}$ turns the symmetric maximum into an associative operation $\mathbb{O}_R \colon \mathfrak{S}^* \to \mathfrak{S}$ defined by $\mathbb{O}_R = \mathbb{O} \circ R$, since $R(\sigma) \in \mathfrak{S} \setminus \mathfrak{S}_0$ for all $\sigma \in \mathfrak{S}$ [2]. Moreover, each computation rule has the form $R = T_1T_2\cdots$, where each $T_i$ has the form $\omega\rho_1^\alpha\rho_2^\beta\rho_3$, with $\omega \in \mathcal{L}(\rho_4, \rho_5)$ and $\alpha, \beta \in \{0, 1\}$ (factorization scheme)[3]

Now to compute $\mathbb{O}_R(\sigma)$, one needs to delete symbols in the sequence $\theta(\sigma)$ exactly as they are deleted in $\psi(\sigma)$. This entails an ordering of $\mathfrak{R}$ that is discussed below.

Let $R, R' \in \mathfrak{R}$ and, for each sequence $\sigma = (a_i)_{i\in I}$, let $J_\sigma \subseteq I$ and $J'_\sigma \subseteq I$, be the sets of indices of the terms in $\sigma$ deleted by $R$ and $R'$, respectively. Then, we write $R \leqslant R'$ if for all sequences $\sigma \in \mathfrak{S}$ we have $J_\sigma \supseteq J'_\sigma$. Clearly, it is reflexive and transitive, and thus it is a preorder. This induces an equivalence relation $\sim$ defined as follows: $R \sim R'$ if $R \leqslant R'$ and $R' \leqslant R$. The following proposition provides equivalent definitions of $\sim$.

**Proposition 3.** *Let $R, R' \in \mathfrak{R}$. Then the following assertions are equivalent.*

*(i)* $R \sim R'$.
*(ii)* $\mathbb{O}_R = \mathbb{O}_{R'}$.
*(iii)* $\mathrm{Ker}(\mathbb{O}_R) = \mathrm{Ker}(\mathbb{O}_{R'})$[4].

[2] For convenience, we assume that $\mathbb{O}_R(\varepsilon) = 0$ and $\mathbb{O}_R(a) = a$, for every $a \in \tilde{C}$
[3] Here, $\rho^0 = \varepsilon$ and $\rho^1 = \rho$.
[4] The *kernel* of $\mathbb{O}_R$ is defined by $\mathrm{Ker}(\mathbb{O}_R) = \{\sigma \in \mathfrak{S} \mid \mathbb{O}_R(\sigma) = 0\}$.

*Furthermore, they have exactly the same factorized irredundant form (see [22]).*

The structure of the poset $\mathfrak{R}/_\sim$ of equivalence classes endowed with the partial order induced by $\leqslant$ was investigated in [22] and shown to be highly complex. To give an idea, the subposet $\mathfrak{R}_{123}/_\sim$ of equivalence classes of rules $R \in \mathcal{L}(\rho_1, \rho_2, \rho_3)$ has infinitely many maximal elements, and $(\mathfrak{R}_{123}/_\sim, \leqslant)$ (and thus $(\mathfrak{R}/_\sim, \leqslant)$) embeds the powerset $(2^{\mathbb{N}}, \subseteq)$ of natural numbers, and hence it is of continuum cardinality. For further results on $\mathfrak{R}/_\sim$, see [22].

**Characterizations of symmetric maxima $\circledvee_R$** We now briefly describe the class of those integer functions $\phi : \tilde{L}^* \to \tilde{L}$ that coincide with symmetric maxima $\circledvee_R$, for $R \in \mathfrak{R}$, by making use of the tight connections to the theory of integer means.

A function $\phi : \tilde{C}^* \to \tilde{C}$ that verifies anonymity[5], internality[6], monotonicity[7], and decomposability[8] is called an *integer mean* in [1], where it was shown that all such functions are extremal[9]. However, the three last properties are too stringent for functions of the form $\phi = \circledvee_R$. This led us to considering some relaxations and variants of these properties, which culminated in the following characterizations of the class of symmetric maxima $\circledvee_R$, $R \in \mathfrak{R}$.

**Theorem 9.** *[23] Let $\phi : \tilde{L}^* \to \tilde{L}$ be an anonymous and weakly associative function that satisfies (C1), (C2), (C3) on $\tilde{L}^2$, and let $R \in \mathfrak{R}$. Then $\phi = \circledvee_R$ if and only if one (or, equivalently, all) of the following assertions hold:*

*(i)* $\mathrm{Ker}(\circledvee_R) \subseteq \mathrm{Ker}(\phi)$ *and $\phi$ is decomposable on every $K \subseteq J^R_\sigma$,*
*(ii) $\phi$ is extremal w.r.t. $R \in \mathfrak{R}$, i.e., $\phi(\sigma) = \Big(\min_{i \in I \setminus J^R_\sigma} \alpha_i\Big) \circledvee \Big(\max_{i \in I \setminus J^R_\sigma} \alpha_i\Big)$,*
*(iii) $\phi$ is retractive w.r.t. $R$, i.e., $\phi(\sigma) = \phi(\alpha_{I \setminus J^R_\sigma})$,*

*for every $\sigma = (\alpha_i)_{i \in I}$ in $\mathfrak{S}$.*

**Formulating the symmetric Sugeno integral** The question is how to define the Sugeno integral for functions which may take negative values, i.e., functions $f : [n] \to \tilde{L}$. We proceed by analogy with the Choquet integral. We recall its expression for a function $f : [n] \to [0, 1]$ w.r.t. a capacity $\mu$:

$$C_\mu(f) = \sum_{i=1}^{n} (f_{(i)} - f_{(i-1)}) \mu(A_{(i)}), \tag{24}$$

with $f_{(0)} := 0$. The usual way of defining the Choquet integral for functions taking negative values is the following one:

$$C_\mu(f) = C_\mu(f^+) - C_{\mu^c}(-f^-),$$

with $f^+, f^-$ the positive and negative parts of $f$, i.e., $f^+ = f \vee 0$ and $f^- = (-f)^+$, and $\mu^c$ is the conjugate or dual of $\mu$, defined when $\tilde{L} = \mathbb{R}$ by $\mu^c(S) = \mu([n]) - \mu(S^c)$. This is sometimes called the *asymmetric Choquet integral*, as it does *not* satisfy the property $C_\mu(-f) = -C_\mu(f)$. The motivation for its definition is that this is the only expression which is invariant by translation, i.e., $C_\mu(f + h) = C_\mu(f) + h\mu([n])$, $h$ being a constant function of value $h$. In our context of ordinal scales, translation has no meaning, and consequently mimicking the definition of the (asymmetric) Choquet integral for the Sugeno integral of $\tilde{L}$-valued functions is meaningless.

[5] $\phi$ is *anonymous* if for every $\sigma = (\alpha_i)_{i \in I} \in \mathfrak{S}$ and every permutation $\pi$ on $I$, $\circledvee_R(\sigma) = \circledvee_R(\sigma \circ \pi)$, where $\sigma \circ \pi = (\alpha_{\pi_i})_{i \in I}$
[6] $\phi$ is *internal* if for every $\sigma = (\alpha_i)_{i \in I} \in \mathfrak{S}$, $\min_{i \in I} \alpha_i \leqslant \phi(\sigma) \leqslant \max_{i \in I} \alpha_i$
[7] $\phi$ is *monotone* if $\phi(\sigma) \leqslant \phi(\sigma')$ whenever $\sigma = (\alpha_i)_{i \in I} \in \mathfrak{S}$ and $\sigma' = (a'_i)_{i \in I} \in \mathfrak{S}$ are such that $\alpha_i \leqslant \alpha'_i$ for every $i \in I$.
[8] $\phi$ is *decomposable* if for every $\sigma = (\alpha_i)_{i \in I}$ and $K \subseteq I$, $\phi(\alpha_K) = b$ implies $\phi(\alpha_K, \alpha_{I \setminus K}) = \phi(|K| \cdot b, \alpha_{I \setminus K})$, where $|K| \cdot b$ means $b, b, \ldots, b$ (repeated $|K|$ times).
[9] $\phi$ is *extremal* if for every $\sigma = (\alpha_i)_{i \in I} \in \mathfrak{S}$, $\phi(\sigma) = \phi\Big(\min_{i \in I} \alpha_i, \max_{i \in I} \alpha_i\Big)$.

On the other hand, the *symmetric Choquet integral* is defined by

$$\check{C}_\mu(f) = C_\mu(f^+) - C_\mu(f^-). \tag{25}$$

Its name comes from the fact that indeed this integral satisfies symmetry: $C_\mu(-f) = -C_\mu(f)$. Its explicit expression is found to be:

$$\begin{aligned}\check{C}_\mu(f) = \sum_{i=1}^{p-1}(f_{(i)} - f_{(i+1)})\mu(\{(1),\dots,(i)\}) + f_{(p)}\mu(\{(1),\dots,(p)\}) \\ + f_{(p+1)}\mu(\{(p+1),\dots,(n)\}) + \sum_{i=p+2}^{n}(f_{(i)} - f_{(i-1)})\mu(\{(i),\dots,(n)\}),\end{aligned} \tag{26}$$

where $f_{(1)} \leq \cdots \leq f_{(p)} < 0 \leq f_{(p+1)} \leq \cdots \leq f_{(n)}$.

Based on this, we define the *symmetric Sugeno integral* of $f : [n] \to \tilde{L}$ w.r.t. a capacity $\mu$ as follows:

$$\check{S}_\mu(f) = S_\mu(f^+) \mathbin{\textcircled{\vee}} (-S_\mu(f^-)), \tag{27}$$

where again $f^+, f^-$ denote the positive and negative parts of $f$, i.e., $f^+ = f \vee 0$ and $f^- = (-f)^+$. Observe that thanks to property (C1), the symmetric Sugeno integral extends the usual Sugeno integral on $\tilde{L}$. Moreover, thanks to (C2) and (C3), the symmetric Sugeno integral behaves like the symmetric Choquet integral on the real line. In particular, by (C3) we have that $-S_\mu(f) = S_\mu(-f)$, and thus the explicit expression of $\check{S}_\mu(f)$ is also close to the symmetric Choquet integral:

$$\check{S}_\mu(f) = \Big[\mathop{\textcircled{\vee}}_{i=p+1}^{n}(f_{(i)} \mathbin{\textcircled{\wedge}} \mu(\{(i),\dots,(n)\}))\Big] \mathbin{\textcircled{\vee}} \Big[\mathop{\textcircled{\vee}}_{i=1}^{p}(f_{(i)} \mathbin{\textcircled{\wedge}} \mu(\{(1),\dots,(i)\}))\Big], \tag{28}$$

with $f_{(1)} \leq \cdots \leq f_{(p)} < 0 \leq f_{(p+1)} \leq \cdots \leq f_{(n)}$ and $\textcircled{\wedge}$ is the symmetric minimum. Notice that there is no ambiguity due to lack of associativity in this formula.

Now, by using the computation rule $\langle\cdot\rangle_+^-$, (28) can be rewritten as:

$$\begin{aligned}\check{S}_\mu(f) = \textcircled{\vee}_{\langle\cdot\rangle_+^-}(\ & f_{(1)} \mathbin{\textcircled{\wedge}} \mu(\{(1)\}),\dots,f_{(p)} \mathbin{\textcircled{\wedge}} \mu(\{(1),\dots,(p)\}), \\ & f_{(p+1)} \mathbin{\textcircled{\wedge}} \mu(\{(p+1),\dots,(n)\}),\dots,f_{(n)} \mathbin{\textcircled{\wedge}} \mu(\{(n)\}))).\end{aligned} \tag{29}$$

This fact together with our general framework for computation rules motivates several other definitions of the symmetric Sugeno integral, each of which reflecting different tendencies (e.g., pessimistic vs optimistic views). Indeed, it is tempting to consider replacing $\langle\cdot\rangle_+^-$ by any computation rule $R$:

$$\begin{aligned}\check{S}_\mu(f) = \textcircled{\vee}_R(\ & f_{(1)} \mathbin{\textcircled{\wedge}} \mu(\{(1)\}),\dots,f_{(p)} \mathbin{\textcircled{\wedge}} \mu(\{(1),\dots,(p)\}), \\ & f_{(p+1)} \mathbin{\textcircled{\wedge}} \mu(\{(p+1),\dots,(n)\}),\dots,f_{(n)} \mathbin{\textcircled{\wedge}} \mu(\{(n)\}))).\end{aligned} \tag{30}$$

However, we must be careful due to the fact that not all computation rules are monotonic, i.e., increasing one of the elements of a sequence $\sigma$ cannot decrease the result. For example, the rule $\langle\cdot\rangle_=$ is not monotonic as the following example shows:

$$\textcircled{\vee}_{\langle\cdot\rangle_=}(5,-5,-5,4,3) = 4 \quad \text{whereas} \quad \textcircled{\vee}_{\langle\cdot\rangle_=}(5,-5,-4,4,3) = 3.$$

This naturally raises the question of determining those computation rules that are monotonic, which constitutes a topic of current research.

### 4.2 Deciding by evaluating pros and cons

A formal elementary framework for a bivariate bipolar multicriteria decision analysis requires a positive scale $L^+$ and a negative scale $L^-$ which are unipolar. We consider the symmetric case, namely there is an order-reversing bijection between $L^+$ and $L^-$. Basically they are copies of a finite totally ordered scale. In other words, the pair $(L^+, L^-)$ can be viewed as the two parts of a bipolar scale. In the simplest qualitative setting, criteria are valued on a bipolar scale $L_B = \{-, 0, +\}$, whose elements reflect negativity, neutrality and positivity respectively. Then $L^+ = \{0, +\}$, $L^- = \{-, 0\}$, which are copies of the Boolean scale $\{0, 1\}$.

The importance of criteria is evaluated on a scale $L$ that is commensurate with the two scales $L^+$ and $L^-$, in the sense that there is an order-preserving (for $L^+$) and an order-reversing (for $L^-$) injections preserving top and bottom between $L^+, L^-$ and $L$: for instance $L^+ \subseteq L$ and $L^- \subseteq \{-\lambda : \lambda \in L\}$.

In the following we assume that alternatives map $\mathcal{C}$ to $L \times L$; namely $f_i$ is of the form $(f_i^-, f_i^+)$ with $-f_i^- \in L^-$ and $f_i^+ \in L^+$, with the assumption that $\min(f_i^-, f_i^+) = 0$ (the rating of $f$ wrt a criterion is either positive or negative). Each value $f_i$ expresses that criterion $i$ brings an argument in favor of $f$ ($f_i^+ > 0$) or in disfavor of $f$ ($f_i^- > 0$) or yet is neutral to $f$ (when $f_i^+ = f_i^- = 0$). In other words we can split the alternative $f$ into positive $f^+$ and negative parts $f^-$, that can be independently evaluated on the scale $L$. In a nutshell, with respect to an alternative $f$, each criterion $i$ has

- *a polarity*: the criterion $i$ judges $f$ positively, or negatively in the wide sense.
- *a degree of importance* $\pi_i \in L$ that does not depend on the alternative. More generally, we can use a capacity for weighting dependent criteria.

The proposed framework is clearly of type II in the bipolarity typology.

**Example** Suppose that Luc has to choose a holiday destination and considers two options for which he has listed the pros and cons. Option $f$ is in a very attractive region (a strong pro), and hotel has a swimming pool ; but it is very expensive, and the plane company has a terrible reputation (two strong cons). Option $g$ is cheaper but it is in a non-democratic country, and Luc considers it a strong con. On the other hand, Option $g$ includes a tennis court and a swimming pool. These are three pros, but not very decisive: they do matter, but not as much as the other arguments.

Formally, let:

- $\mathcal{C} = \{$Attractiveness (1), price (2), democracy(3), sport facilities(4), airline (5)$\}$.
- $L^+ = \{0, +, ++\}, L^- = \{--, -, 0\}, L = \{0, \lambda, 1\}$.
- Available alternatives:
  $f$ gets $++$ on attractiveness, $--$ on price and airline, $+$ on sport facilities
  so $f^+ = (1, 0, 0, \lambda, 0)$ and $f^- = (0, 1, 0, 0, 1)$;
  $g$ gets $+$ on price $--$ on democracy and $++$ for sports
  so $g^+ = (0, \lambda, 0, 1, 0)$ and $g^- = (0, 0, 1, 0, 0)$.

A special case of this framework is studied at length in Bonnefon et al.[6, 7], where scales $L^+, L^-$ are Boolean, while $L$ is any bounded scale for assessing importance of criteria. Then an alternative $f$ is modelled by two disjoint subsets of $\mathcal{C}$: $F^+ = \{i : f_i^+ = 1\}$ and $F^- = \{i : f_i^- = 1\}$ collecting the pros and cons for and against $f$. The set $F = F^+ \cup F^-$ is the set of relevant criteria (i.e. those that matter) for $f$.

There are two approaches to deciding preference among alternatives.

- Either we consider for each alternatives positive and negative summaries of arguments via an aggregation operator, here Sugeno integral, and define a partial ordering between alternatives.
- Or we can build a preference relation via pairwise comparisons assuming that, when comparing $f$ and $g$, that criteria in disfavor of $g$ give reasons to prefer $f$ and conversely.

**Comparing pairs of positive and negative ratings by Pareto-dominance** Suppose the set of criteria is weighted by means of a capacity $\mu$ valued on $L$, and that $L^+ = L^- = L$. In the above bipolar setting, an alternative $f$ will then be evaluated by a pair $(S_\mu(f^-), S_\mu(f^+))$. Note that $S_\mu(f^-)$ is rated on a scale where 1 is interpreted as bad, 0 is neutral. If we wish to express it on a positive unipolar scale where 0 means bad, we have to replace $S_\mu(f^-)$ by $D_\nu(f^-) = S_{1-\mu^c}(1 - f^-)$ (using a monotone decreasing set-function $\nu = 1 - \mu^c$). $D_\nu(f^-)$ is named a desintegral in [55]. A first way of comparing alternatives $f$ and $g$ is to use a kind of Pareto-dominance between $(f^-, f^+)$ and $(g^-, g^+)$.

$$f \succeq^{PB} g \iff S_\mu(f^+) \geq S_\mu(g^+) \text{ and } S_\mu(f^-) \leq S_\mu(g^-) \tag{31}$$

This ordering is exactly Pareto dominance if we use the desintegral for $f^-$. It satisfies obvious monotonicity conditions, namely if $f$ and $g$ are such that $f_i^+ \geq g_i^+, \forall i \in [n]$ and $f_i^- \leq g_i^-, \forall i \in [n]$, then $f \succeq^{PB} g$.

In the special case of Boolean positive and negative ratings, Bonnefon et al.[6, 7] use a possibility assignment to weight criteria and a pair of possibility measures $(\Pi(F^-), \Pi(F^+))$ as a bipolar rating of this alternative. These evaluations [10] are justified by the focus effect according to which humans compare alternatives with respect to the most important criteria first, neglecting other ones. Note that the ordering $\succeq^{PB}$ collapses to Wald's pessimistic ordering if $F^+ = \emptyset$ (choosing based on the worst feature), and to its optimistic max-based counterpart if $F^- = \emptyset$.

This decision rule has some deficiencies:

- The bipolar outranking relation $\succeq^{PB}$ concludes to incomparability in some cases when a preference would sound more natural. When $f$ has both pros and cons, it is incomparable with the neutral alternative $\hat{f}$ such that $\hat{f}_i^+ = \hat{f}_i^- = 0, \forall i \in [n]$ if the importance of the cons in $f$ is negligible in front of the importance of its pros.
- Whenever two criteria $i$ and $j$ are such that $f_i^- = 1$ and $f_j^+ = 1$ (a very bad rating on criterion $i$ and a very good one on $j$) and these criteria have maximal importance $\mu_\#(\{i\}) = \mu_\#(\{j\}) = 1$, then $(S_\mu(f^-), S_\mu(f^+)) = (1, 1)$, so that such alternatives are considered equally preferred. There can be many of them, hence a lack of discrimination.

**A bipolar model in decision under uncertainty** The approach of Giang and Shenoy [68, 69] to decision under uncertainty is of the same vein. They have tried to obviate the need for making assumptions on the pessimistic or optimistic attitude of the decision-maker and thus, improve the discrimination power in the qualitative setting, by using, as a utility scale, a totally ordered set of possibility measures on a two element set $\{0, 1\}$ containing the values of the best and the worst consequences of acts. Each such possibility distribution represents a qualitative lottery in a set $L_\Pi = \{(\alpha, \beta), \max(\alpha, \beta) = 1, \alpha, \beta \in L\}$. Coefficient $\alpha$ represents the degree of possibility of obtaining the worst consequence, and coefficient $\beta$ the degree of possibility of obtaining the best. This set can be viewed as a bipolar value scale ordered by the following complete preordering relation expressing preference:

$$(\alpha, \beta) \geq (\gamma, \delta) \text{ if and only if } (\alpha \leq \gamma \text{ and } \beta \geq \delta)$$

The bottom of this utility scale is $(1, 0)$, its top is $(0, 1)$ and its neutral point $(1, 1)$ means "indifferent". The fact this relation is complete is due to the fact that pairs $(\alpha, \beta)$ and $(\gamma, \delta)$ such that $(\alpha, \beta) > (\gamma, \delta)$ and $(\gamma, \delta) > (\alpha, \beta)$ cannot both lie in $L_\Pi$ since then either $\max(\alpha, \beta) < 1$ or $\max(\gamma, \delta) < 1$. The canonical example of such a scale is the set of pairs $(\Pi(\overline{A}), \Pi(A))$ of degrees of possibility for event $A =$ "getting the best consequence", and its complement. The inequality $(\Pi(\overline{A}), \Pi(A)) > (\Pi(\overline{B}), \Pi(B))$ means that $A$ is more likely (certain or plausible) than $B$ (because it is equivalent to $\Pi(A) > \Pi(B)$ or $N(A) > N(B)$). In fact the induced likelihood ordering between events

$$A \succeq_{L\Pi} B \text{ if and only if } (\Pi(\overline{A}), \Pi(A)) \geq (\Pi(\overline{B}), \Pi(B))$$

[10] interpreted in terms of *order of magnitude* of importance, hence the notation $OM$ in [6].

is self-adjoint, that is, $A \succeq_{L\Pi} B$ is equivalent to $\overline{B} \succeq_{L\Pi} \overline{A}$.

Each consequence is supposed to have a utility value $(\alpha, \beta)$ in $L_\Pi$. The proposed preference functional maps acts, viewed as $n-$tuples $f = ((f_1^-, f_1^+), \ldots, (f_n^-, f_n^+))$ of values in $L_\Pi$, to $L_\Pi$ itself. The uncertainty is described by possibility weights $(\pi_1, \ldots, \pi_n)$ with $\max_{i=1,\ldots,n} \pi_i = 1$. The utility of an act $f$, called *binary possibilistic utility* is computed as the pair

$$W_{GS}(f) = (\max_{i=1,\ldots,n} \min(\pi_i, f_i^-), \max_{i=1,\ldots,n} \min(\pi_i, f_i^+)) \in L_\Pi.$$

Clearly, $W_{GS}(f)$ is of the form $(S_\mu(f^-), S_\mu(f^+))$ for $\mu = \Pi$. This form results from simple and very natural axioms on possibilistic lotteries, which are counterparts to the Von Neumann and Morgenstern axioms in decision under risk. Weng [104] proposed a Savage-style axiomatization of this binary possibilistic utility functional. It puts together the axiomatizations of the optimistic and the pessimistic possibilistic criteria by Dubois *et al.* [60], adding, to the axioms justifying Sugeno integral, two conditions: (i) the self-adjointness of the preference relation on binary acts, and (ii) an axiom enforcing axiom **OPT** on the subset of acts weakly preferred to a special act that plays the role of a neutral point separating favorable from unfavorable acts. Pessimistic and optimistic possibilistic criteria $SLMIN_\pi$ and $SLMAX_\pi$ are of course special cases of this bipolar criterion. They respectively correspond to either using the negative part of $L_\Pi$ only (not telling $(1,1)$ from $(0,1)$ in case of pessimism) or using the positive part of $L_\Pi$ only (not telling $(1,0)$ from $(1,1)$ in case of optimism).

**The bipolar possibility relation** The problem with the bipolar Pareto-dominance is that it does not account for the fact that the two evaluations *share a common importance scale* $L$. Another idea for comparing alternatives $f$ and $g$ is to focus on criteria in $F \cup G$, i.e., those that matter for both alternatives. The principle at work is simple: any argument against $f$ (resp. against $g$) is an argument pro $g$ (resp., pro $f$). The most supported decision is then preferred, by comparing global evaluations on $F^+ \cup G^-$ and $F^- \cup G^+$: Instead of comparing $f$ and $g$ we compare $f^+ \vee g^-$ with $g^+ \vee f^-$ with respective components $f_i^+ \vee g_i^-$ and $f_i^- \vee g_i^+$:

**Definition 5 (Bipolar Sugeno Dominance).** $f \succeq^{BS} g \iff S_\mu(f^+ \vee g^-) \geq S_\mu(f^- \vee g^+)$.

It would be worth studying this preference relation and compare it to the preference relation induced by the symmetric Sugeno integral (27). In particular Bipolar Sugeno Dominance does not require new operations on a bipolar scale since all computations are brought back to the positive part of the scale in Definition 5.

This kind of preference relation was first proposed in [6] when scales $L^+$ and $L^-$ are Boolean, $\mu$ is a possibility measure, and $F^- \cap G^+ = F^+ \cap G^- = \emptyset$. It yields the bipolar possibility relation

$$f \succeq^{B\Pi} g \iff \Pi(F^+ \cup G^-) \geq \Pi(G^+ \cup F^-)$$

This rule decides that $f$ is at least as good as $g$ as soon as there are important arguments either in favour of $f$ or attacking $g$ that are at least as strong as the best arguments in favour of $g$ or attacking $f$. Obviously, $\succeq^{B\Pi}$ collapses to Wald's pessimistic ordering if $F = F^-, G = G^-$ and to its optimistic counterpart when $G = G^+, F = F^+$. In some sense, this definition is the most straightforward generalisation of possibility relations [83, 34] to the bipolar case.

The bipolar possibility relation satisfies the following properties

(i) It is complete and its strict part is transitive.
(ii) The restriction of $\succeq^{B\Pi}$ to $f$ such that $F = \{i\}, i = 1, \ldots n$ is a weak order.
(iii) *Ground Monotony*: $\forall f, g, h, h'$ such that $H = \{i\}, H' = \{i'\}, F \cap \{i, i'\} = \emptyset$ and $h' \succeq^{B\Pi} h$:
$f \vee h \succ g \Rightarrow f \vee h' \succ g$; $f \vee h \sim g \Rightarrow f \vee h' \succeq g$;
$g \succ f \vee h' \Rightarrow g \succ f \vee h$; $g \sim f \vee h' \Rightarrow g \succeq f \vee h$.

(iv) *Positive Cancellation*: $\forall f, g, h$ such that $f_i^+ = 1$ and 0 otherwise, $g_j^+ = 1$ and 0 otherwise, $h_k^- = 1$ and 0 otherwise and denoting by 0 the alternative that receives 0 for each criterion: $f \vee h \sim 0$ and $g \vee h \sim 0 \Rightarrow f \sim h$.

(v) *Negative Cancellation*: $\forall f, g, h$ such that $f_i^- = 1$ and 0 otherwise, $g_j^- = 1$ and 0 otherwise, $h_k^+ = 1$ and 0 otherwise: $f \vee h \sim 0$ and $g \vee h \sim 0 \Rightarrow f \sim h$

(vi) *Strict negligibility*: $\forall f, gf', g' : f \succ g$ and $f' \succ g' \Rightarrow f \vee f' \succ g \vee g'$.

(vii) *Idempotent Negligibility* $\forall f, gf', g' : f \succeq g$ and $f' \succeq g' \Rightarrow f \vee f' \succeq g \vee g'$.

**Remarks:**

- Note that the weak relation $\succeq^{B\Pi}$ is generally not transitive.
- Properties (iv) and (v) express a form of anonymity. It is required when a positive argument blocks a negative argument of the same strength: this blocking effect should not depend on the arguments themselves, but on their position in the importance scale only.
- The two last properties are direct consequences of working with importance levels that are orders of magnitude. $f \succ g$ means that $f$ is much better than $g$, so much so as there is no way of overthrowing $f$ by sets of weaker arguments (property (vi)).
- The last property presupposes that several arguments of the same strength are worth just one.

The above properties turn out to be characteristic of the bipolar possibility rule [6]. They imply the existence of the importance scale, and the importance assignment to criteria as a possibility distribution.

**Comparison with Cumulative Prospect Theory (CPT)** There is a similarity between the bipolar possibility relation and the preference ordering of CPT. The latter assumes that the strength of reasons supporting an alternative $f$ and the strength of reasons against it can be measured by means of two numerical capacities $\sigma^+$ and $\sigma^-$ respectively mapping subsets $F^+$ and $F^-$ to the unipolar scale $[0, +\infty)$. The capacity $\sigma^+$ reflects the importance of the group of positive arguments for $f$, and $\sigma^-$ the importance of the group of negative arguments against it.

This approach moreover admits that it is possible to combine these evaluations by subtracting them and building a so-called "net predisposition" score expressed on a bipolar numerical scale (the real line):

$$\forall f, NP(f) = \sigma^+(F^+) - \sigma^-(F^-)$$

It is a special case of symmetric Choquet integral described in the previous section. Alternatives are then ranked according to this net predisposition: $f \succeq^{CPT} g \iff \sigma^+(F^+) - \sigma^-(F^-) \geq \sigma^+(G^+) - \sigma^-(G^-)$. The relation $\succeq^{B\Pi}$ can be viewed as the natural qualitative counterpart of $\succeq^{CPT}$ ; indeed, the bipolar possibility decision rule comes down to changing $+$ into $\max$ in the equivalent inequality $\sigma^+(F^+) + \sigma^-(G^-) \geq \sigma^+(G^+) + \sigma^-(F^-)$, that is, if $\sigma^+ = \sigma^-$ is additive, $\sigma^+(F^+ \cup G^-) \geq \sigma^+(G^+ \cup F^-)$.

So, there is a joint framework encompassing the CPT framework and the qualitative bipolar possibility relation, turning possibility measures into standard capacities $\kappa : 2^X \to L$ [40]:

$$f \succeq^\kappa g \iff \kappa(F^+ \cup G^-) \geq \kappa(G^+ \cup F^-)$$

adopting the view that an argument against alternative $f$ is an argument in favour of $g$ in the pairwise comparison of alternatives.

The following properties clearly hold for $\succeq^\kappa$: it is complete, and the restriction to single arguments is a weak order. However it is not clearly transitive, not even quasi-transitive in the general case. And while the non triviality, and both positive and negative monotony properties hold, the weak unanimity property, that would make $\succeq^\kappa$ a bipolar monotonic set relation, requires that $\kappa$ satisfy an additional property on top of inclusion-monotonicity of capacities [15]:

*Weak additivity*: Let $A, B, C, D \subseteq X$ such that $A \cap C = \emptyset, B \cap D = \emptyset$; if $\kappa(A) \geq \kappa(B)$ and $\kappa(C) \geq \kappa(D)$ then $\kappa(A \cup C) \geq \kappa(B \cup D)$.

This property is, for capacities, equivalent to the following property involving only three subsets $A, B, C(= D)$ [34]:

If $\kappa(A) \geq \kappa(B)$ then $\kappa(A \cup C) \geq \kappa(B \cup C)$, provided that $(A \cup B) \cap C = \emptyset$.

It implies that $\kappa$ is a decomposable measure [15], that is, there exists an operation $\star$ such that if $A \cap B = \emptyset$, $\kappa(A \cup B) = \kappa(A) \star \kappa(B)$. Due to compatibility with the underlying Boolean algebra of events, it is natural to consider that $\star$ is a co-norm. Choosing an Archimedean continuous co-norm on $L = [0, 1]$, it is clear that $\succeq^{\kappa}$ can verify additional properties:

- *Transitivity*: $\kappa(F^+ \cup G^-) \geq \kappa(G^+ \cup F^-)$ and $\kappa(G^+ \cup C^-) \geq \kappa(C^+ \cup G^-)$ imply $\kappa(F^+ \cup C^-) \geq \kappa(C^+ \cup F^-)$. Indeed the preconditions imply

$$\kappa(F^+) \star \kappa(G^-) \star \kappa(G^+) \star \kappa(C^-) \geq \kappa(G^+) \star \kappa(F^-) \star \kappa(G^-) \star \kappa(C^+)$$

  which yields the expected result by simplification (if $\star$ is a strict t-norm or $\kappa$ is properly normalized). This simplification cannot be made if $\star = \max$.
- *Ground monotony* holds under the same assumptions about $\star$.
- *Positive and negative cancellation* properties reduce to the trivial statement that $\kappa(\{x\}) = \kappa(\{y\})$ and $\kappa(\{z\}) = \kappa(\{y\})$ imply $\kappa(\{x\}) = \kappa(\{z\})$.

In fact, relation $f \succeq^{\kappa} g$ is a conjoint generalisation of $\succeq^{CPT}$ and $\succeq^{B\Pi}$ that either comes down to one of them ($\succeq^{CPT}$ is obtained, if $L = [0, 1]$, $\star$ is a nilpotent Archimedean t-norm and $\kappa$ is properly normalized, or a strict co-norm, taking the logarithm of $\kappa$) or a combination of them (if $\star$ is an ordinal sum of the basic conorms $\alpha + \beta - \alpha\beta$, $\min(1, \alpha + \beta)$, $\max$) up to a rescaling.

**Bipolar lexicographic outranking relations** The last property (Idempotent Negligibility) of the bipolar possibility rule is by far the most debatable feature of $\succeq^{B\Pi}$. It causes a drowning effect, usual in standard possibility theory. For instance, if the most important criteria satisfied by $f$ are of the same importance as most important criteria satisfied by $g$, but there are more of the latter, the two alternatives are judged equally.

A tempting way of refining $\succeq^{B\Pi}$, is to use a leximax relation instead. Then the number of arguments of equal strength on each side is then taken into account. Among the two basic axioms of qualitative modeling, it comes down to giving up Idempotent Negligibility, while retaining Strict Negligibility. Preference can then be based on counting arguments of the same strength, but we still do not allow an important argument to be superseded by several less important ones, however large their number be (focus effect). The criteria satisfied or violated in $f$ and $g$ are scanned top down, until a level is reached such that the numbers of positive and negative arguments pertaining to the two alternatives are different; then, the option with the least number of violated criteria and the greatest number of satisfied ones is preferred.

There are two such decision rules respectively called "Bivariate Levelwise Tallying" and (univariate) "Levelwise Tallying" [7], according to whether positive and negative arguments are treated separately or not.

For any importance level $\lambda \in L$, let $F_\lambda = \{i \in F, \pi_i = \lambda\}$ be the $\lambda$-section of $f$, the set of relevant criteria of strength $\lambda$ in $f$. Let $F_\lambda^+ = \{i \in F_\lambda : f_i^+ = 1\}$ (resp., $F_\lambda^- = \{i \in F_\lambda : f_i^- = 1\}$) be its positive (resp., negative) $\lambda$-section. Let $\delta(f, g)$ be the maximal level of importance where either the positive or the negative $\lambda$-sections of $f$ and $g$ differ, namely:

$$\delta(f, g) = \max\{\lambda : |F_\lambda^+| \neq |G_\lambda^+| \text{ or } |F_\lambda^-| \neq |G_\lambda^-|\}.$$

$\delta(f,g)$ is called the *decisive level* pertaining to $(f,g)$. The Bivariate Levelwise Tallying preference rule reads:

$$f \succeq^{BL} g \iff |F^+_{\delta(f,g)}| \geq |G^+_{\delta(f,g)}| \text{ and } |F^-_{\delta(f,g)}| \leq |G^-_{\delta(f,g)}|.$$

It is easy to show that $\succeq^{BL}$ is reflexive, transitive, refines the bipolar Pareto ordering but is not complete. Indeed, $\succeq^{BL}$ concludes to an incomparability if and only if there is a conflict between the positive view and the negative view at the decisive level. From a descriptive point of view, this range of incomparability is a good point in favour of $\succeq^{BL}$.

Now, if one can assume a compensation between positive and negative arguments *at each importance level*, one argument canceling another one on the other side, the following refinement of relation $\succeq^{BL}$, called Univariate Levelwise Tallying, can be obtained:

$$f \succeq^{ULT} g \iff \exists \lambda \in L \setminus 0_L \text{ s. t. } \begin{cases} \forall \theta > \lambda, |A^+_\theta| - |A^-_\theta| = |B^+_\theta| - |B^-_\theta| \\ \text{and} \quad |F^+_\lambda| - |F^-_\lambda| > |G^+_\lambda| - |G^-_\lambda| \end{cases}$$
$$\text{or } |A^+_\theta| - |A^-_\theta| = |B^+_\theta| - |B^-_\theta|, \forall \lambda \in L \setminus 0_L \text{ (the latter case is when } f \sim^{ULT} g).$$

Interestingly, relation $\succeq^{ULT}$ is closely related to the decision rule originally proposed two centuries ago by Benjamin Franklin [65].

The two decision rules proposed in this section obviously generate monotonic bipolar outranking relations. Each of them refines $\succeq^{BI\Pi}$. The most decisive one is $\succeq^{ULT}$, which is moreover complete and transitive. This relation is the refinement of $\succeq^{BI\Pi}$ that is a weak order and that satisfies the principle of preferential independence without introducing any bias on the importance order elementary criteria (that is, preserving the restriction of $\succeq^{BI\Pi}$ to single criteria). See [6] for such an axiomatisation. It turns out that Levelwise Tallying is the most likely decision rule to be used by people as an empirical study suggests [7].

## 5 Qualitative data analysis

As a general family of aggregation functions, it is of interest to identify the family of Sugeno integrals that are compatible with a dataset made of vectors of criteria values together with the corresponding global evaluations. When the family is non empty, it can be described by bracketing the data by means of a lower capacity and an upper capacity. Such a dataset can be also described by means of sets of selection and deletion rules, which then correspond to a combination of Sugeno utility functionals. In this section, we briefly discuss nonparametric methods based on Sugeno integral for learning rule-based models that are widely used in multicriteria decision aid and ordinal classification [79] tasks.

### 5.1 Approach by bracketing datasets with standard Sugeno integrals

The problem considered here is the elicitation of a family of Sugeno integrals that are compatible with a dataset. Here, a dataset is a collection of $L$-valued tuples $f = (f_1, \ldots, f_n)$ associated with a global rating $\delta \in L$.

**Definition 6.** *A pair $(f,\delta)$ is compatible with a Sugeno integral $S_\mu$ if and only if $S_\mu(f) = \delta$.*

In the following, we study the constraints induced by a pair $(f,\delta)$ on the Sugeno integrals compatible with it and we fully characterize this family following ideas first given in [95]. For convenience, we assume that the $f_i$'s are already increasingly ordered, i.e., $f_1 \leq \ldots \leq f_n$. Since

$$\bigwedge_{i=1}^{n} f_i \leq S_\mu(f_1, \ldots, f_n) \leq \bigvee_{i=1}^{n} f_i,$$

there exists a Sugeno integral that satisfies $S_\mu(f) = \delta$ if and only if $f_1 \leq \delta \leq f_n$. *We assume here that this consistency condition holds for the pairs* $(f, \delta)$ *considered*. For discussing the equation $S_\mu(f) = \delta$, it is useful to distinguish two cases.

**DIF Case :** $\forall i \in \mathcal{C}, f_i \neq \delta$.

Let $i$ be the index such that $f_1 \leq \ldots \leq f_{i-1} < \delta < f_i \leq \ldots \leq f_n$. We can then define two particular capacities $\check{\mu}_{f,\delta,DIF}$ and $\hat{\mu}_{f,\delta,DIF}$ :

**Definition 7.**

$$\forall X \in 2^{\mathcal{C}}, X \neq \emptyset, \mathcal{C} \quad \check{\mu}_{f,\delta,DIF}(X) = \begin{cases} \delta & \textit{if } \{i, \ldots, n\} \subseteq X \\ 0 & \textit{otherwise} \end{cases}$$

*and*

$$\forall X \in 2^{\mathcal{C}}, X \neq \emptyset, \mathcal{C} \quad \hat{\mu}_{f,\delta,DIF}(X) = \begin{cases} \delta & \textit{if } X \subseteq \{i, \ldots, n\} \\ 1 & \textit{otherwise} \end{cases}.$$

It can be shown that :

$$\forall \mu \text{ s.t. } S_\mu(f) = \delta \text{ we have } \check{\mu}_{f,\delta,DIF} \leq \mu \leq \hat{\mu}_{f,\delta,DIF}.$$

Thus $\check{\mu}_{f,\delta,DIF}$ and $\hat{\mu}_{f,\delta,DIF}$ are the lower and upper bounds of the lattice of capacities which define the family of Sugeno integrals compatible with the pair $(f, \delta)$ in the DIF case.

**EQU case :** $\exists i \in \mathcal{C}, f_i = \delta$**.**

Let $i$ and $j$ be the indices such that $f_1 \leq \ldots \leq f_{j-1} < f_j = \ldots = f_{i-1} = \delta < f_i \leq \ldots \leq f_n$. We can then define two particular capacities $\check{\mu}_{f,\delta,EQU}$ and $\hat{\mu}_{f,\delta,EQU}$ :

**Definition 8.**

$$\forall X \in 2^{\mathcal{C}},\ X \neq \emptyset, \mathcal{C} \quad \check{\mu}_{f,\delta,EQU}(X) = \begin{cases} \delta & \textit{if } \{j, \ldots, i-1, \ldots n\} \subseteq X \\ 0 & \textit{otherwise} \end{cases}$$

*and*

$$\forall X \in 2^{\mathcal{C}},\ X \neq \emptyset, \mathcal{C} \quad \hat{\mu}_{f,\delta,EQU}(X) = \begin{cases} \delta & \textit{if } X \subseteq \{i, \ldots, n\} \\ 1 & \textit{otherwise} \end{cases}.$$

It can be shown that :

$$\forall v \text{ s.t. } S_\mu(f) = \delta \text{ we have } \check{\mu}_{f,\delta,EQU} \leq \mu \leq \hat{\mu}_{f,\delta,EQU}.$$

Thus $\check{\mu}_{f,\delta,EQU}$ and $\hat{\mu}_{f,\delta,EQU}$ are the lower and the upper bounds of the lattice of capacities which define the family of Sugeno integrals compatible with the pair $(f, \delta)$ in the EQU case.

**Consistent family of a Sugeno integral with respect to a dataset** A dataset is consistent if there exists a non empty family of Sugeno integrals that are compatible with each pair $(f, \delta)$ in the dataset. Otherwise, it means that there is no representation of the dataset by a unique family of Sugeno integrals and that several families thereof are necessary, each covering a distinct subpart of the dataset. Let us consider a dataset $(f_i, \delta_i)_{i \in \{1,\ldots,p\}}$ that contains $p$ pairs. In order to simplify notations, we denote by $\check{\mu}_i$ the lower bound for $(f_i, \delta_i)$ and $\hat{\mu}_i$ the upper bound associated with $(f_i, \delta_i)$. Thus the lower and upper bounds of the family of compatible Sugeno integrals, if it exists, are respectively

$$\check{\mu} = \bigvee_{i=1}^{p} \check{\mu}_i \quad \text{and} \quad \hat{\mu} = \bigwedge_{i=1}^{p} \hat{\mu}_i.$$

Thus, when a new piece of information $(f, \delta)$ is considered, $\check{\mu}$ and $\hat{\mu}$ are then revised by

$$\check{\mu}_{revised} = \check{\mu} \vee \check{\mu}_{f,\delta} \quad \text{and} \quad \hat{\mu}_{revised} = \hat{\mu} \wedge \hat{\mu}_{f,\delta}.$$

These results have been applied to a case study [91] on mental workload data where the global evaluation rely on six criteria, where several families of Sugeno integrals were necessary for recovering the whole data set, thus revealing different aggregation attitudes with respect to mental workload. The families of Sugeno integral were identified thanks to a simulated annealing method [91]. Besides, it has been shown [89] that the bracketing procedure can be viewed as a graded extension of the version space approach in machine learning [90].

### 5.2 Approach by best approximation using Sugeno utility functionals

We consider datasets that can be accurately modeled by a nondecreasing function. Since a SUF uses utility functions as arguments of the Sugeno integral, it can model rules with different thresholds. We consider sets of (selection) rules of the form

$$\textbf{if } f_1 \geq \alpha_1 \textbf{ and } \ldots \textbf{ and } f_n \geq \alpha_n \textbf{ then } y \geq \delta \tag{32}$$

where $(\alpha_1, \ldots, \alpha_n) \in L^n$. Using results in Subsection 2.4, a SUF induces single-thresholded rules of the form

$$\textbf{if } \varphi_1(f_1) \geq \delta \textbf{ and } \ldots \textbf{ and } \varphi_n(f_n) \geq \delta \textbf{ then } y \geq \delta$$

which can be turned into the form (32), if we choose monotone utility functions such that $\varphi_i(\alpha_i) \geq \delta$.

Let $R$ be a set of rules of the form (32). There may be several functions that are compatible with $R$. We denote by $\Phi_R$ the smallest function compatible with $R$, defined by $\Phi_R = \max_{r \in R} \Phi_r$ such that for each rule $r$:

$$\Phi_r(f) = \delta^r, \textbf{ if } \forall i \in \mathcal{C},\ f_i \geq \alpha_i^r, \textbf{ and } 0 \textbf{ otherwise.}$$

We will say that a function $\Phi$ is *equivalent* to $R$ if $\Phi = \Phi_R$. It was shown in [19, 11] that:

(i) Any SUF $S_{\mu,\varphi}$ is equivalent to a rule set.
(ii) Any single rule is equivalent to a SUF.
(iii) Some rule sets are not equivalent to a single SUF.

Any SUF $S_{\varphi,\mu}$ is equivalent to the rule set

$$\bigcup_{I \subseteq [n]} \bigcup_{\delta \leq \mu(I)} \{\forall i \in I,\ f_i \geq \alpha_i \Rightarrow y \geq \delta\}, \tag{33}$$

where $\alpha_i = \min\{\lambda \in L_i \mid \varphi_i(\lambda) \geq \delta\}$. Note that this set is likely to contain redundant rules.

So each multiple-thresholded rule induces constraints on the utility functions. But the constraints induced by two or more rules can be inconsistent [19]. In other words, some combinations of rules cannot be expressed by a single SUF. Nonetheless, the second assertion shows that any rule set is equivalent to some function $M_{\boldsymbol{S}} : L^n \to L$ defined by

$$M_{\boldsymbol{S}}(f) = \max\{S_{\mu,\varphi}(f) \mid S_{\mu,\varphi} \in \boldsymbol{S}\},$$

where $\boldsymbol{S}$ is a set of SUFs. We call such a function a *max-SUF*.

In [11] a method of translation of a rule set $R$ into a SUF is provided:

(i) Initialize $\mu$ and $\varphi = (\varphi_1, \ldots, \varphi_n)$ with minimal values.
(ii) For each rule $f_1 \geq \alpha_1, \ldots, f_n \geq \alpha_n \Rightarrow y \geq \delta$ in $R$:
  (a) let $A = \{i \in [n] \mid \alpha_i > 0\}$,

(b) increase $\mu(A)$ up to $\delta$,
(c) for each $i \in A$, increase $\varphi_i(\alpha_i)$ up to $\delta$

After these steps we always have $S_{\mu,\varphi} \geq \Phi_R$. When $S_{\mu,\varphi} > \Phi_R$, no SUF is equivalent to $R$.

In some cases, it is not problematic that $S_{\mu,\varphi} > \Phi_R$. For example, if $\Phi_R$ is a model of a dataset $\mathcal{D}$, we may want to find an SUF that best fits $\mathcal{D}$. Obtaining $S_{\mu,\varphi} = \Phi_R$ is not always possible since SUFs are not expressive enough. However, equality can be always achieved using a max-SUF [11]. The method presented in the next section relies on this fact.

**Learning rules from empirical data** Now there is no reason to think that a max-SUFis more interpretable than its equivalent rule set. Thus, an interesting question is whether SUFs can serve as an intermediary model that helps guiding the learning process of a rule based model. Indeed, in [11], such a learning algorithm is proposed. Let $\mathcal{D}$ be a dataset. The following three steps provide a method for modeling $\mathcal{D}$ by a max-SUF.

**1. Selection of an order-preserving subset of data.** Two data items $(f, \delta)$ and $(g, \gamma)$ can be *anti-monotonic* together, i.e, $f \leq g$ and $\gamma \leq \delta$. We iteratively remove instances from $\mathcal{D}$, starting from those that are anti-monotonic with the highest number of other instances, until no anti-monotonic pair remains. We denote by $\mathcal{D}^-$ the dataset obtained in this way.

**2. Modeling $\mathcal{D}^-$ by a rule set $R$.** Initialize $R$ to $\emptyset$. For each instance $((\alpha_1, \ldots, \alpha_n), \delta)$ in $\mathcal{D}^-$, search for $A \subseteq [n]$ with minimal cardinality, such that the rule

$$\forall i \in A,\ f_i \geq \alpha_i \Rightarrow y \geq \delta, \tag{34}$$

is not contradicted by any instance in $\mathcal{D}^-$. Add the rule (34) to $R$. At the end of this step, the class of each instance in $\mathcal{D}^-$ is exactly predicted by $\Phi_R$.

**3. Translation of $R$ into a max-SUF.** See Algorithm 1. The obtained max-SUF is not necessarily equivalent to $R$, but it fits $\mathcal{D}^-$ precisely.

**Algorithm 1:** Makes a partition **P** of $R$ such that the max-SUF $M_{\boldsymbol{S}}$ verifies $M_{\boldsymbol{S}}(f) = y$ for each instance $(f, y)$.

```
P ← {}
for each r ∈ R do
    affected ← false
    for each P ∈ P do
        translate P into a SUF S_{μ,φ}
        if S_{μ,φ}(f) ≤ y for all instance (f, y) in D⁻ then
            add r to P
            affected ← true
            break loop
    if affected = false then
        add {r} to P
```

Note that the max-SUF given by this method can be translated back into a rule set, which constitutes an equivalent model and is easier to interpret.

**Empirical study** The VC-DomLEM algorithm [4] is another method that can learn such a set of rules, which yields a good accuracy compared to other interpretable models. This method requires the tuning

of hyperparameters, contrary to ours. Our method is competitive with VC-DomLEM in terms of accuracy. Moreover, this method raised new questions about the relevance of capacities (i.e., monotonically increasing set functions) in data-modeling.

The method in [4] was compared to VC-DomLEM on the 12 datasets. In order to get an idea of the importance of Step 3 in our method, we separately evaluated the rule set given by steps 1 and 2 alone, and the max-SUF given by steps 1,2, and 3. We see that Step 3 does not increase the accuracy on average.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | avg. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Steps 1,2. | 74.4 | 95.8 | 97.7 | 93.6 | 91.7 | 65.6 | 83.2 | 27 | 67 | 63.6 | 58.2 | 51.4 | 72.4 |
| Steps 1,2,3. | 76 | 95.3 | 97.2 | 89.3 | 92.4 | 65.2 | 84.5 | 26.4 | 69.4 | 63 | 56.7 | 53.2 | 72.4 |
| VC-DomLEM | 76.7 | 96.3 | 97.1 | 91.7 | 95.4 | 67.5 | 87.7 | 26.9 | 66.7 | 55.6 | 56.4 | 54.6 | 72.7 |

Table 1: Accuracy obtained with each method on each dataset. Datasets are numbered as in [4]

Therefore, the good results of this method are not due to the use of SUFs, but to the 2 first steps.

Now, the *length* of a rule is the number of attributes $i$ where $\alpha_i > 0$ (since the condition $\alpha_i \geq 0$ is trivial). Shorter rules are easier to interpret and constitute more concise models. In [11] it was presented an empirical study of rule length distributions obtained after steps 1, 2, and 3.

Moreover, it was compared to an analogous method using the dual max-SUFs. The dual of max-SUFs are the min-SUFs that correspond to sets of (rejection) rules of the form

$$\textbf{if } f_1 \leq \alpha_1 \textbf{ and } \ldots \textbf{ and } f_n \leq \alpha_n \textbf{ then } y \leq \delta.$$

When learning min-SUFs by a dual method, the rule-length distribution differs from that obtained by learning max-SUFs. Long rules of one type sometimes go along with short rules of the other type. This empirical result hints at a combined method for rule length improvement.

## 6 Conclusion

This chapter has tried to advocate the merits of Sugeno integral as a tool for the qualitative evaluation of alternatives when utility values are not supposed to be numerical. After recalling the algebraic nature of Sugeno integral and its close links to the notion of median, recent developments of the approach, overcoming some of its limitations, have been surveyed: namely the enhancement of its expressiveness for the pairwise comparison of alternatives via lexicographic refinements, the use of utility functions when several attribute scales need to be reconciled, and the extension of the role of attribute weights in the aggregation process. The question of borrowing concepts from cumulative prospect theory for the distinct treatment of pros and cons in Sugeno integral has been also discussed.

Finally, methods for representing a qualitative dataset by means of one or several Sugeno integrals have been outlined, questioning the possibility of learning them from data. These results provide an extended range of tools for processing qualitative preference data in a non-trivial way, thus obviating the need to use numerical scales for attributes that are not easily and meaningfully measured. Besides, the close connection between Sugeno integral and fuzzy decision rules suggest a way to extract meaning from data. However, the bracketting method may yield imprecise results, the SUFs are not expressive enough, and the max-SUFs may be complex. So, the quest for concise and faithful models devoted to qualitative data looks challenging.

## References

1. Bennett C.D., Holland W.C., Székely G.J.: Integer valued means. *Aequationes Mathematicae*, **88**: 137–149 (2014)

2. Benferhat S., Dubois D., Prade H.: Possibilistic and standard probabilistic semantics of conditional knowledge bases. *Journal of Logic and Computation*, **9**: 873–895 (1999)
3. Benferhat S., Dubois D., Kaci S., Prade H.: Bipolar possibility theory in preference modeling: Representation, fusion and optimal solutions. Information Fusion **7**: 135–150 (2006)
4. Blaszczyński J., Slowiński R., Szelag M.: Sequential covering rule induction algorithm for variable consistency rough set approaches. *Information Sciences*, **181**(5): 987–1002 (2011)
5. Boczek M., Kaluszka M.: On conditions under which some generalized Sugeno integrals coincide: A solution to Dubois' problem. *Fuzzy Sets and Systems* **326**: 81–88 (2017)
6. Bonnefon J.-F, Dubois D., Fargier H.: On the qualitative comparison of decisions having positive and negative features. *J. Artificial Intelligence Research*, **32**: 385–417 (2008)
7. Bonnefon J.-F, Dubois D., Fargier H., Leblois S.: Qualitative heuristics for balancing the pros and cons. *Theory and Decision*, **65**: 71–85 (2008)
8. Borzová-Molnárová J.,Halčinová L., Hutník O.: The smallest semicopula-based universal integrals, part I. *Fuzzy Sets and Systems*, **271**: 1–17 (2015)
9. Boutilier C., Brafman R. I., Domshlak C., Hoos H. H., Poole D.: CP-nets: A tool for representing and reasoning with conditional ceteris paribus preference statements. *J. Artificial Intelligence Research*, **21**: 135–191 (2004)
10. Bouyssou D., Marchant T. and Pirlot M.: A conjoint measurement approach to the discrete Sugeno integral. In: *The Mathematics of Preference, Choice and Order*; Essays in Honor of Peter C. Fishburn (S. Brams, W. V. Gehrlein, and F. S. Roberts, Eds), Berlin; London: Springer, 85–109 (2009)
11. Brabant Q., Couceiro M., D. Dubois, Prade H., A. Rico.: Extracting Decision Rules from Qualitative Data via Sugeno Utility Functionals. In: Proc. Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU 2018), (Jesus Medina et al.,Eds.), Springer, CCIS 853, pp. 253–265 (2018)
12. Cacioppo J. T., Berntson G. G.: Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin*, **115**: 401–423 (1994)
13. de Campos L. M, Bolaños M. J.: Characterization and comparison of Sugeno and Choquet integrals, *Fuzzy Sets and Systems*, **52**: 61–67 (1992)
14. de Campos L. M, Lamata M. T., Moral S.: A unified approach to define fuzzy integrals *Fuzzy Sets and Systems*, **39**(1): 75–90 (1991)
15. Chateauneuf A.: Decomposable capacities, distorted probabilities and concave capacities. *Mathematical Social Sciences*, **31**(1):19–37 (1996)
16. Chateauneuf A., Grabisch M., Rico A.: Modeling attitudes toward uncertainty through the use of the Sugeno integral. *J. Math. Econ.*, **44**(11): 1084–1099 (2008)
17. Chateauneuf A.: Modeling attitudes towards uncertainty and risk through the use of Choquet integral. *Annals of Operations Research*, **52**: 3–20 (1994)
18. Cohen M., Jaffray J.Y.: Rational behavior under complete ignorance. *Econometrica*, **48**(5): 1281–1299 (1980)
19. Couceiro M., Dubois D., Prade H., Rico A.: Enhancing the Expressive Power of Sugeno Integrals for Qualitative Data Analysis. In: Kacprzyk J.et al. (eds) Advances in Fuzzy Logic and Technology 2017 (Proc. EUSFLAT 2017). Advances in Intelligent Systems and Computing, vol 641. Springer, 534–547 (2017)
20. Couceiro M., Dubois D., Prade H., Waldhauser T.: Decision-Making with Sugeno Integrals - Bridging the Gap Between Multicriteria Evaluation and Decision Under Uncertainty. *Order* **33**(3): 517–535 (2016)
21. Couceiro M., Foldes S., Lehtonen E.: Composition of Post classes and normal forms of Boolean functions. *Discrete Mathematics* **306**: 3223–3243 (2006)
22. Couceiro M., Grabisch M.: On the poset of computation rules for nonassociative calculus. *Order* **30**:1 269–288 (2013)
23. Couceiro M., Grabisch M.: On integer-valued means and the symmetric maximum. *Aequationes Mathematicae* **91**:2 353–371 (2017)
24. Couceiro M., Lehtonen E., Marichal J.-L., Waldhauser T.: An algorithm for producing median normal form representations for Boolean functions. In the proc. of the Reed-Muller Workshop 2011, 49–54 (2011)
25. Couceiro M., Marichal J.-L.: Characterizations of discrete Sugeno integrals as polynomial functions over distributive lattices. *Fuzzy Sets and Systems* **161**(5): 694–707 (2010)
26. Couceiro M., Marichal J.-L.: Representations and characterizations of polynomial functions on chains. *Journal of Multiple-Valued Logic and Soft Computing*, **16**:1-2, 65–86 (2010)
27. Couceiro M., Marichal J.-L.: Polynomial functions over bounded distributive lattices. *Journal of Multiple-Valued Logic and Soft Computing* **18**: 247–256 (2012)
28. Couceiro M., Mercuriali P., Péchoux R., Saffidine A.: Median based calculus for lattice polynomials and monotone Boolean functions. *47th IEEE International Symposium on Multiple-Valued Logic* (ISMVL 2017), IEEE Computer Society, 37–42 (2017)
29. Couceiro M., Mercuriali P., Péchoux R., Saffidine A.: On the complexity of minimizing median normal forms of monotone Boolean functions and lattice polynomials, 23 pages. https://hal.inria.fr/hal-01905491
30. Couceiro, M., Waldhauser, T.: Axiomatizations and factorizations of Sugeno utility functionals. *Int. J. Uncertain. Fuzziness Knowledge Based Systems*, **19**(4): 635–658 (2011)

31. Couceiro M., Waldhauser T.: Pseudo-polynomial functions over finite distributive lattices. *Fuzzy Sets and Systems*, **239**: 21–34 (2014)
32. Deschamps R., Gevers L.: Leximin and utilitarian rules: a joint characterization. *J. of Economic Theory*, **17**: 143–163 (1978)
33. Doyle J., Thomason R.: Background to qualitative decision theory. *The AI Magazine*, **20**(2): 55–68 (1999)
34. Dubois D.: Belief structures, possibility theory and decomposable confidence measures on finite sets. *Computers and Artificial Intelligence*, **5**(5): 403–416 (1986)
35. Dubois D., Durrieu C., Prade H., Rico A., Ferro Y.: Extracting Decision Rules from Qualitative Data Using Sugeno Integral: A Case-Study. ECSQARU 2015: 14–24 (2015)
36. Dubois D., Fargier H.: Making Discrete Sugeno Integrals More Discriminant. *International Journal of Approximate Reasoning*, **50**: 880–898 (2009)
37. Dubois D., Fargier H.: Capacity refinements and their application to qualitative decision evaluation. In: Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2009), C. Sossai, G. Chemello (Eds.), Springer, LNAI 5590, 311–322 (2009)
38. Dubois D., Fargier H.: Qualitative Bipolar Decision Rules: Toward More Expressive Settings. Dans: Preferences and Decisions - Models and Applications. Salvatore Greco, Ricardo Marquez Peirera, Massimo Squillante, Ronald R. Yager, Janusz Kacprzyk (Eds.), Springer, 139–158, Vol. 257, Studies in Fuzziness and Soft Computing, 2010.
39. Dubois D., Fargier H., Prade H.: Refinements of the maximin approach to decision-making in a fuzzy environment. *Fuzzy Sets Systems*, **81**(1): 103–122 (1996)
40. Dubois D., Fargier H., Prade H., Sabbadin R.: A survey of qualitative decision rules under uncertainty. In D. Bouyssou, D. Dubois, M. Pirlot, and Prade H., editors, *Decision-Making Process - Concepts and Methods*, chapter 11, 435–473. Wiley, 2009
41. Dubois D., Fortemps P.: Computing improved optimal solutions to max-min flexible constraint satisfaction problems. *European Journal of Operational Research*, **118**: 95–126 (1999)
42. Dubois D., Le Berre D., Prade H., Sabbadin R.: Using possibilistic logic for modeling qualitative decision: ATMS-based algorithms. *Fundamenta Informaticae*, **37**: 1–30 (1999)
43. Dubois D., Marichal J-L., Prade H., Roubens M., Sabbadin R.: The use of the discrete Sugeno integral in decision making: a survey. *Int. J. of Uncertainty, Fuzziness and Knowledge-based Systems*, **9** 539–561 (2001)
44. Dubois D., Prade H.: *Fuzzy Sets and Systems, Theory and Applications,* Academic Press (1980)
45. Dubois D., Prade H.: A theorem on implication functions defined from triangular norms. *Stochastica*, **8**: 267–279 (1984)
46. D. Dubois and Prade H..: Evidence measures based on fuzzy information. Automatica, **21**: 547–562 (1985)
47. Dubois D., Prade H.: Weighted minimum and maximum operations. *Information Sciences*, **39**: 205–210 (1986)
48. Dubois D., Prade H.: The mean value of a fuzzy number. *Fuzzy Sets and Systems*, **24**: 279–300 (1987)
49. Dubois D., Prade H.: *Possibility Theory: An Approach to Computerized Processing of Uncertainty.* Plenum Press, New York (1988)
50. Dubois D., Prade H.: Possibilistic logic: a retrospective and prospective view. *Fuzzy Sets and Systems*, **144**(1): 3–23 (2004)
51. Dubois D., Prade H.: An introduction to bipolar representations of information and preference. *Int. J. Intelligent Systems.*, **23**(8): 866–877 (2008)
52. Dubois D., Prade H.: Possibility theory and its applications: Where do we stand ? *Springer Handbook of Computational Intelligence* (J. Kacprzyk, W. Pedrycz, Eds), Springer, 31–60 (2015)
53. Dubois D., Prade H., Rico A.: The logical encoding of Sugeno integrals. *Fuzzy Sets and Systems*, **241**: 61–75 (2014)
54. Dubois D., Prade H., Rico A.: Representing qualitative capacities as families of possibility measures. *Int. J. Approx. Reasoning*, **58**: 3–24 (2015)
55. Dubois D., Prade H., Rico A.: Residuated variants of Sugeno integrals. *Information Sciences*, **329**: 765–781 (2016)
56. Dubois D., Prade H., Rico A., Teheux B.: Generalized Sugeno integrals. Proc. *16th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems* (IPMU 2016), vol. 1, Communications in Computer and Information Science 610, Springer, 363–374 (2016)
57. Dubois D., Prade H., Rico A., Teheux B.: Generalized qualitative Sugeno integrals. To appear in *Information Sciences* (2017)
58. Dubois D., Prade H., Sabbadin, R.: Qualitative decision theory with Sugeno integrals. *Proc. of 14th Conference on Uncertainty in AI*, 121–128 (1998)
59. Dubois D., Prade H. Sabbadin, R.: Qualitative decision theory with Sugeno integrals. In: (M. Grabisch, T. Murofushi, M. Sugeno, Eds. *Fuzzy Measures and Integrals. Theory and Applications*,Studies in Fuzziness and Soft Computing, Physica-Verlag, 314–322 (2000)
60. Dubois D., Prade H. Sabbadin, R.: Decision theoretic foundations of qualitative possibility theory. *European Journal of Operational Research*, **128**: 459–478 (2001)
61. Dubois D., Rico A.: New axiomatisations of discrete quantitative and qualitative possibilistic integrals. *Fuzzy Sets and Systems*, **343**: 3–19 (2018)
62. Dvořák A., Holčapek M.: Fuzzy measures and integrals defined on algebras of fuzzy subsets over complete residuated lattices. *Information Sciences*, **185**: 205–229 (2012)

63. Fargier H., Lang J., Schiex T.: Selecting preferred solutions in fuzzy constraint satisfaction problems. In Proc. *1st European Congress on Fuzzy and Intelligent Technologies* (EUFIT'93), 1128–1134, Aachen, Germany (1993)
64. Fargier H., Sabbadin R.: Qualitative decision under uncertainty: Back to expected utility. *Artificial Intelligence*, **164**: 245–280 (2005)
65. Franklin B.: Letter to J. B. Priestley, 1772, in The Complete Works, J. Bigelow, ed. New York: Putnam (1887)
66. Gérard R., Kaci S., Prade H.: Ranking alternatives on the basis of generic constraints and examples - A possibilistic approach. Proc. *20th Inter. Joint Conf. on Artificial Intelligence* (IJCAI 2007), Hyderabad, Jan. 6-12, (M. M. Veloso, ed.), 393–398 (2007)
67. R. L. Goodstein. The solution of equations in a lattice. *Proc. Roy. Soc. Edinburgh Sect. A*, **67**:231–242 (1965/1967)
68. Giang P. H., Shenoy P. P.: A qualitative utility theory for Spohn's theory of epistemic beliefs. In Proc. of the 16th Conf. on Uncertainty in Artificial Intelligence, 220–227 (2000)
69. Giang P. H., Shenoy P. P.: Two axiomatic approaches to decision-making using possibility theory. *European Journal of Operational Research*, **162**: 450–467 (2005)
70. Gigerenzer G., Todd P. M., the ABC group.: Simple heuristics that make us smart. Oxford Univ. Press (1999)
71. Grabisch M.: The application of fuzzy integrals in multicriteria decision making. *Europ. J. of Operat. Res.*, **89**(3): 445–456 (1996)
72. M. Grabisch. The symmetric Sugeno integral. *Fuzzy Sets and Systems*, **139**: 473–490 (2003)
73. Grabisch M.: The Moebius transform on symmetric ordered structures and its application to capacities on finite sets. *Discrete Mathematics*, **287**: 17–34 (2004)
74. Grabisch, M.: *Set functions, Games and Capacities in Decision-Making*. Springer-Verlag (2016)
75. Grabisch M., Labreuche C.: A decade of application of the Choquet and Sugeno intégrals in multi-criteria décision aid. *Annals of Operations Research*, **175**: 247-286 (2010)
76. Grabisch, M. Marichal J.-L., Mesiar R., Pap E.: *Aggregation Functions*, Cambridge University Press, (2009)
77. Grabisch, M. Murofushi, T. Sugeno M.: Fuzzy measure of fuzzy events defined by fuzzy integrals, *Fuzzy Sets and Systems*, **50**(3): 293–313 (1992)
78. Greco S., Matarazzo B., Slowinski R.: Axiomatic characterization of a general utility function and its particular cases in terms of conjoint measurement and rough-set decision rules. *Eur. J. Oper. Res.*, **158**: 271–292 (2004)
79. Gutiérrez P. A., Pérez-Ortiz M., Sánchez-Monedero J., Fernández-Navarro F., Hervás-Martínez C.: Ordinal regression methods: Survey and experimental study. *IEEE Trans. on Knowledge and Data Engineering*, **28**(1): 127–146 (2016)
80. Heilpern S.: The expected value of a fuzzy number. *Fuzzy Sets and Systems*, **47**: 81–87 (1992)
81. Kandel A., Byatt W. J.: Fuzzy sets, fuzzy algebra and fuzzy statistics. *Proceedings of the IEEE*, **68**: 1619–1639 (1978)
82. Kaufmann A.: Le calcul des admissibilités. Une idée nouvelle à partir de la théorie des sous-ensembles flous. *Proc. Colloque International sur la Théorie et les Applications des Sous-Ensembles Flous*, Vol. I, (14 p.), Marseilles (1978)
83. Lewis D.: Counterfactuals and comparative possibility. *Journal of Philosophical Logic* **2**(4): 418–446 (1973)
84. Marichal J.-L.: On Sugeno integral as an aggregation function. *Fuzzy Sets and Systems* **114**: 347–365 (2000)
85. Marichal J.-L.: Weighted Lattice Polynomials. *Discrete Mathematics*, 309(4): 814–820 (2009)
86. Mesiar R.: $k$-order pan-discrete fuzzy measures. Proc. *7th IFSA world Congress*, vol. 1, 488–490, Prague (1997)
87. Osgood C. E., Suci G. J., Tannenbaum P. H.: *The Measurement of Meaning*. Univ. of Illinois Press, Chicago (1957)
88. Öztürk M., Tsoukiás A.: Bipolar preference modeling and aggregation in decision support. *Int. J. Intelligent Systems*, **23**(9): 970–984 (2008)
89. Prade H., Rico A., Serrurier M.: Elicitation of Sugeno integrals: A version space learning perspective. Proc. *18th Int. Symp. on Foundations of Intelligent Systems* (ISMIS'09), (J. Rauch, Z. W. Ras, P. Berka, T. Elomaa, eds.), Prague, Czech Rep., Sept. 14-17, Springer LNCS 5722, 392–401 (2009)
90. Mitchell, T.: Generalization as search. *Artificial intelligence* **18**: 203–226 (1982)
91. Prade, H., Rico, A., Serrurier, M., Raufaste, E.: Elicitating Sugeno integrals: Methodology and a case study. Proc. *10th Europ. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty* (ECSQARU'09), (C. Sossai, G. Chemello, eds.), Verona, Italy, July 1-3, LNCS 5590 (2009)
92. Ralescu D., Sugeno M.: Fuzzy integral representation, *Fuzzy Sets and Systems* **84**: 127–133 (1996)
93. Rébillé Y.: Decision-making over necessity measures through the Choquet integral criterion. *Fuzzy Sets and Systems*, **157**: 3025–3039 (2006)
94. Rico, A.: *Modélisation des Préférences pour l'Aide à la Décision par l'Intégrale de Sugeno*, Ph. D. Thesis, Université Paris I Sorbonne (2002)
95. Rico, A., Grabisch, M., Labreuche, C., Chateauneuf, A.: Preference modeling on totally ordered sets by the Sugeno integral. *Discrete Applied Mathematics*, **147**: 113–124 (2005)
96. Roy B.: *Multicriteria Methodology for Decision Aiding. Nonconvex Optimization and its Applications*, vol. 12. Kluwer Academic Publishers, Dordrecht (1996)
97. Schmeidler D.: Cores of exact games. *J. of Math. Anal. and Appl.*, **40**(1): 214–225 (1972)
98. Shilkret N.: Maxitive measure and integration. *Indag. Math.*, **33**: 109–116 (1971)
99. Slovic P., Finucane M., Peters E., MacGregor D.G.: Rational actors or rational fools? implications of the affect heuristic for behavioral economics. *The J. of Socio-Economics*, **31**: 329–342 (2002)

100. Snow P.: Diverse confidence levels in a probabilistic semantics for conditional logics. *Artificial Intelligence*, **113**(1-2): 269–279 (1999)
101. Sugeno M.: *Theory of Fuzzy Integrals and its Applications.* PhD Thesis, Tokyo Institute of Technology (1974)
102. Sugeno M.: Fuzzy measures and fuzzy integrals: A survey. In: M. M. Gupta, *et al.*, eds.*Fuzzy Automata and Decision Processes,* North-Holland, 89–102 (1977)
103. Tversky A., Kahneman D.: Advances in prospect theory: Cumulative representation of uncertainty. *J. of Risk and Uncertainty*, **5**: 297–323 (1992)
104. Weng P.: An axiomatic approach in qualitative decision theory with binary possibilistic utility. In Proc. of the *17th European Conference on Artificial Intelligence* (ECAI 2006), Riva del Garda, Italy, 467–471. IOS Press (2006)
105. Zadeh L. A.: Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, **1**: 3–28 (1978)