Making Discrete Sugeno Integrals More Discriminant

Didier Dubois [a] and Hélène Fargier [a]

[a] *IRIT - Université Paul Sabatier - 31062 Toulouse Cedex, France*

Abstract

This paper deals with qualitative evaluation processes when the worth of items is computed by means of Sugeno integral. One limitation of this approach is the coarse ranking of items it produces. In order to refine this ranking, generalizations of leximin and leximax to Sugeno integrals are studied. Numerical encodings of such generalized lexicographic methods are described by means of mappings from the qualitative value scale to the reals. In some of these transformations Sugeno integral is changed into a Choquet integral. The issue of refining the capacity at work in Sugeno integral also receives a preliminary examination. This work relies on a previous similar attempt at refining prioritized minimum and maximum aggregations (in the setting of decision under uncertainty) into a so-called big-stepped weighted average, encoding a very refined qualitative lexicographic ordering of items.

Key words: Sugeno integral, Choquet integral, qualitative decision theory, lexicographic ordering,

1 INTRODUCTION

Qualitative decision theory is a framework that suits situations where the evaluation of complex objects cannot rely on the availability of full-fledged numerical ratings. This is typical of electronic commerce, or recommender systems (that provide advice or suggestions) for instance. In many cases, it sounds more satisfactory to implement a choice method that is fast, and based on rough information about the user preferences and knowledge. Two research lines can then be followed in

Email addresses: dubois@irit.fr (Didier Dubois), fargier@irit.fr (Hélène Fargier).

the qualitative framework: the non-commensurable approach and the absolute approach.

Following the first approach, the various aspects involved in the evaluation process (e.g., uncertainty, utility, importance of criteria) are rated on different value scales that are unrelated to one another. This view is close to the framework of voting theories after Arrow [2], Sen [35] etc. It may lead to the same technical difficulties. In the case of decision under uncertainty, various authors [4,39,6–8,40] propose to compare the merits of acts on the basis of their tuples of utilities restricted to the set of most plausible states: degrees of utility are never compared to degrees of plausibility, but only to degrees of utility.

The absolute approach presupposes the existence of a common totally ordered value scale (typically a finite one), for all kinds of local ratings. For instance grading both likelihood and utility on the same scale. This is based on the idea that any decision involving uncertainty can be compared in terms of preference to a sure gain or a sure loss (involving utility only). In multifactorial evaluation, it corresponds to the assumption of adopting a common value scale for various criteria involved and their importance weights. Decision rules generalizing maximin and maximax criteria can be defined on this ordinal scale[45,43,16]. They are special cases of Sugeno integral [37,38], a general qualitative aggregation that can be used as a decision criterion under uncertainty [18], and a tool for multicriteria aggregation [32]. The rationality of these qualitative aggregation tools was established using an axiomatic approach in the style of Savage [19], or in the style of conjoint measurement [25,5].

Such qualitative criteria can be instrumental to solve discrete decision problems involving finite state spaces, or problems where it is not natural, or very difficult, to elicitate numerical utility functions or probabilities. Namely,

- when the problem is located in a dynamic environment involving a large state space, a non-quantifiable goal to be reached, and partial information on the current state. This case can be found in robotic planification problems;
- when only a very high level description of a decision problem is available, where states and consequences of decisions are coarsely defined (for instance in some kinds of strategic decision-making);
- or yet when there is no time to quantify utilities and probabilities because a fast advice is requested (like in recommender systems).

A number of natural properties any realistic decision theory should satisfy in such applications can be laid bare:

(1) *Faithfulness* to available information supplied by decision-makers, as poor as it be : an ordinal declarative approach sounds closer to human capabilities.
(2) *Cognitive Relevance* : The number of levels in the value scale must be small enough (according to well-known psychological studies, not more than seven).

(3) *Good Discrimination* : especially respecting the strict Pareto-dominance.
(4) *Decisive Power* : avoiding incomparability and favor linear rankings.
(5) *Exhaustivity*: Taking into account all available information, especially the importance of criteria or the plausibility of states of affairs.

These requirements are often conflicting. Weighted averages are information-demanding, and hardly compatible with the limited perception capabilities of human decision-makers. The maximin criterion is too extreme and neglects available information. Approaches based on ordinal preference relations either leave room to incomparability to a large extent, or focus too much on the most important aspects. Approaches based on an absolute value scale improve the expressivity of maximin and maximax criteria by accounting for the respective plausibility of states or the importance of criteria. They provide rankings of decisions but lack discrimination power because the set of objects to be ranked contains just as many classes of equally preferred items as the number of steps in the value scale. There is some inconsistency between the requirement of a fine-grained discrimination (respecting Pareto-dominance) and the requirement of a total (especially transitive) ranking of alternatives in the qualitative framework.

In order to cope with this limitation, refinements of the final ranking of decisions have been devised, in the restricted case of prioritized minimum and maximum [20]. Following this approach, the final ranking of decisions is not only qualitative (it relies on the use of leximin and leximax procedures [10]) but it also satisfies all the properties of a weighted average (like in expected utility theory). And it can indeed be represented as a weighted average, where the utility functions and the weight functions are big-stepped, i.e. form superincreasing (or decreasing) sequences.

In the present paper, we try to extend this approach to the case where interaction between criteria exist. The natural criteria aggregation tool is then Sugeno integral. The idea is to refine Sugeno integral-based rankings using similar leximin and leximax ingredients. Beforehand it should be noticed that the refinement of the prioritized minimum and maximum by a weighted average is made possible by the fact that these criteria do not strongly violate the preferential independence axiom obeyed by the latter: only a blurring effect is observed, which causes the lack of discrimination. But due to the strong violation of the independence by Sugeno integral, the latter cannot be refined by means of a weighted average. In fact, due to the role of comonotonicity in the representation of Sugeno integral, the natural numerical criterion refining the latter is *Choquet integral*, now used in decision analysis for some time [23].

Section 2 presents properties of Sugeno integral. Section 3 explains why Sugeno integral lacks discrimination power. Section 4 recalls basic results on the refinement of qualitative prioritized maximum and minimum by means of a weighted average. They are instrumental for the rest of the paper. Sections 5, 6, and 7 contain the main results of the paper. First, refinements of Sugeno integral are proposed that preserve the set-function representing the importance of features. Another approach based on qualitative Moebius transforms is proposed, where the capacity is changed into a belief function. Section 7 provides some insight into the problem of refining a non-additive set-function, since part of the lack of discrimination is due to the non-additivity of the set-function weighting criteria or states of nature. A preliminary version of the first six sections was presented at the ECSQARU 2007 conference and published in its proceedings[13].

2 Sugeno integral as a qualitative decision rule

A decision evaluation problem will be cast in the usual framework: we consider a set $\mathcal{F}$ of n features or criteria (denoted by integers i), and a set Ω of objects or items to be rated according to these points of view. For rating the merit of objects, there is a totally ordered value scale $(L, \leq)$, supposed to be common to all features, with top $\top$ and bottom $\perp$. In the numerical case, $L = [0, 1]$ for instance. In the qualitative case, it is a finite chain. We will then denote by λ_j the elements of L, with $\lambda_0 = \perp < \lambda_1 < \ldots < \lambda_m = \top$. Moreover, L is equipped with its involutive order-reversing map ν; in particular $\nu(\top) = \perp, \nu(\perp) = \top$. The rating of objects $\omega \in \Omega$ according to feature i are denoted by Greek letters $\alpha_i, \beta_i, \dots \in L$. The weight of a feature will be denoted by p_i, when numerical, and π_i when qualitative.

The set Ω of objects will be identified with the set L^n of n-tuples $\vec{\alpha}$ of values of L. The idea is that objects having the same description cannot be distinguished. We denote by $\underline{\lambda}$ constant tuples containing the same rating λ for each feature. The top and bottom tuple are such that $\lambda = \top$ and $\lambda = \perp$, respectively, and denoted by $\underline{\top}$ and $\underline{\perp}$. They are respectively the best rated and the worst rated objects. If $\vec{\alpha}, \vec{\beta} \in L^n$, and A is a subset of features, $\vec{\alpha}A\vec{\beta}$ denotes the tuple such that $(\vec{\alpha}A\vec{\beta})_i = \alpha_i$ if $i \in A$, and β_i otherwise. In particular, a binary tuple is denoted by $\underline{\alpha}A\underline{\beta}$ and is such that $(\underline{\alpha}A\underline{\beta})_i = \alpha \in L$ if $i \in A$, and $\beta \in L$ otherwise. A Boolean tuple is of the form $\underline{\top}A\underline{\perp}$.

This framework covers not only multifactorial evaluation but decision under uncertainty as well. Then $\mathcal{F}$ is a set of states of nature, ω is an act, understood as a mapping from $\mathcal{F}$ to a set X of consequences, and α_i is the degree of utility of the consequence of this act when the state is i. Then weight p_i is the degree of probability of a state and weight π_i its degree of possibility. Whatever the chosen framework, the problem is to evaluate and compare tuples of ratings of the form $\vec{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in L^n$.

The most usual numerical aggregation rule in multifactorial evaluation (assuming

$L \subset [0,1]$) as well as decision under uncertainty is based on the weighted average :

$$\text{WA}^p(\vec{\alpha}) = \sum_{i=1}^{n} p_i \cdot \alpha_i. \tag{1}$$

A tuple $\vec{\alpha}$ is then strictly preferred to another tuple $\vec{\beta}$ if and only if $\text{WA}^p(\vec{\alpha}) > \text{WA}^p(\vec{\beta})$. Such a weigthed average implicitly assumes that the features are preferentially independent with respect to each other. When dependencies between features have to be taken into account, the decision making procedure had better rely on a Choquet integral aggregation:

$$Ch_v(\vec{\alpha}) = \sum_{j=1}^{m} v(A_{\lambda_j}).(\lambda_j - \lambda_{j-1})). \tag{2}$$

where $A_{\lambda_j} = \{i : 1 \leq i \leq n, \alpha_i \geq \lambda_j\}$. In this approach, the importance of groups of features is assumed to be directly captured by means of a monotonic set-function $v : 2^{\mathcal{F}} \to [0,1]$ (also called a *capacity*), such that:

$$v(\emptyset) = 0, v(\mathcal{F}) = 1, A \subseteq B \Rightarrow v(A) \leq v(B).$$

The use of such a set-function is very general and natural in this context. It includes additive measures (hence, the weighted average is a particular Choquet integral) and most other well-known set-functions (including belief and plausibility functions, necessity and possibility measures...).

2.1 Sugeno integral and its special cases

In the following, we assume the value scale is qualitative, and a qualitative capacity is denoted by κ. In this case the most general type of aggregation operation is Sugeno integral (see [24]). The global evaluation of the merits of an object is based on the comparison of ratings of the object with respect to the evaluation scale, and the importance of groups of features is evaluated on the same scale (it is modelled by their κ values). Sugeno integral is often defined as follows:

$$S_\kappa(\vec{\alpha}) = \max_{\lambda \in L} \min(\lambda, \kappa(A_\lambda)), \tag{3}$$

where $A_\lambda = \{i : 1 \leq i \leq n, \alpha_i \geq \lambda\}$ is the set of features having best ratings for object ω, down to utility threshold λ, and $\kappa(A)$ is the degree of importance of feature set A.

If the set of features is rearranged in decreasing order in such a way that $\alpha_1 \geq \cdots \geq \alpha_n$, then denoting $A_i = \{1, 2, \ldots, i\}$, $S_\kappa(\vec{\alpha})$ can be expressed in terms of features as follows:

$$S_\kappa(\vec{\alpha}) = \max_{i=1,\cdots,n} \min(\alpha_i, \kappa(A_i)). \tag{4}$$

It turns out that $S_\kappa(\vec{\alpha})$ is the median of the set $\{\alpha_1, \ldots, \alpha_n\} \cup \{\kappa(A_1), \ldots, \kappa(A_{n-1})\}$ whose cardinality is ever odd. For a binary tuple $\alpha A \beta$ where $\alpha \geq \beta$, $S_\kappa(\alpha A \beta)$ is the median value in the set $\{\alpha, \beta, \kappa(A)\}$. The original definition of Sugeno integral [37] actually had the following form:

$$S_\kappa(\vec{\alpha}) = \max_{A \subseteq S} \min(\kappa(A), \min_{i \in A} \alpha_i). \tag{5}$$

This expression shows a trade-off between the degrees of importance of feature sets and their worst ratings in such sets.

The prioritized maximum and minimum aggregations are particular cases of Sugeno integrals (e.g. [27]). These aggregations are based on an L-valued possibility distribution π [46] on $\mathcal{F}$ measuring the importance of *individual* features: the ordinal value π_i represents the importance of feature $\mathcal{F}$. The prioritized maximum W_π^+ is retrieved when κ is a possibility measure based on distribution π ($\kappa(A) = \max_{i \in A} \pi_i$):

$$W_\pi^+(\vec{\alpha}) = \max_{i=1,\ldots,n} \min(\pi_i, \alpha_i). \tag{6}$$

This optimistic aggregation proposed in [46,45] is a qualitative counterpart to the weighted convex sum, where the sum is replaced by a *sup* (a *max* in the finite case) and the product by an *inf* (a *min* in the finite case). It is an extension of the maximum aggregation: $W^+(\vec{\alpha}) = \max_{i=1,\ldots,n} \alpha_i$.

The prioritized minimum W_π^- is obtained when κ is a necessity measure ($\kappa(A) = \min_{i \notin A} \nu(\pi_i)$, where ν is the order-reversing map on L). It is a pessimistic criterion proposed in [43,17], of the form :

$$W_\pi^-(\vec{\alpha}) = \min_{i=1,\ldots,n} \max(\nu(\pi_i), \alpha_i). \tag{7}$$

So, $\nu(\pi_i)$ represents the degree of negligibility of feature i. In particular, $\nu(\pi_i) = \top$ for fully neglected features. The value of $W_\pi^-(\vec{\alpha})$ is small as soon as there exists a highly important feature ($\nu(\pi_i) = \bot$) with low utility rating for the object. This aggregation is actually a prioritized extension of the Wald maximin criterion

$$W^-(\vec{\alpha}) = \min_{i=1,\ldots,n} \alpha_i. \tag{8}$$

This rule rates objects on the basis of their least preferred marginal ratings. It was advocated and axiomatized by Arrow and Hurwicz [1]. It is recovered in case of equally important features, i.e. when $\pi_i = \top$ for all $i = 1, \ldots, n$ in (7). In the prioritized version, decisions are made according to the merits of objects using the worst rated among the most important features. The set of important features $A^\star = \{i : \pi_i \geq \nu(W_\pi^-(\vec{\alpha}))\}$) achieves a trade-off between importance and local ratings as expressed in the min-max expression.

2.2 Properties of Sugeno integrals and the induced ordering

The basic properties of Sugeno integrals exploit *disjunctive and conjunctive combinations* of ratings. Define a tuple $\vec{\alpha}\wedge\vec{\beta}$ as the one that always gets the worst ratings of $\vec{\alpha}$ and $\vec{\beta}$ for each feature, while $\vec{\alpha} \vee \vec{\beta}$ always gets the best of them:

$$(\vec{\alpha} \wedge \vec{\beta})_i = \alpha_i \text{ if } \beta_i \geq \alpha_i \text{ and } \beta_i \text{ otherwise;} \tag{9}$$

$$(\vec{\alpha} \vee \vec{\beta})_i = \alpha_i \text{ if } \alpha_i \geq \beta_i \text{ and } \beta_i \text{ otherwise.} \tag{10}$$

They are respectively intersection and union of fuzzy sets viewed as n-tuples of values. Obviously $S_\kappa(\vec{\alpha}\wedge\vec{\beta}) \leq \min(S_\kappa(\vec{\alpha}), S_\kappa(\vec{\beta}))$ and $S_\kappa(\vec{\alpha}\vee\vec{\beta}) \geq \max(S_\kappa(\vec{\alpha}), S_\kappa(\vec{\beta}))$. The first one holds with equality for the possibilistic pessimistic criterion W_π^- and the second one likewise for its optimistic counterpart W_π^+. These properties hold with equality whenever $\vec{\alpha}$ or $\vec{\beta}$ is a constant tuple, i.e., noticing that $S_\kappa(\underline{\lambda}) = \lambda$,

$$S_\kappa(\vec{\alpha} \wedge \underline{\lambda}) = \min(S_\kappa(\vec{\alpha}), \lambda) \text{ and } S_\kappa(\vec{\alpha} \vee \underline{\lambda}) = \max(S_\kappa(\vec{\alpha}), \lambda).$$

These properties are in fact characteristic of Sugeno integrals for monotonic aggregation operators (e.g. [27]).

Let us now denote by $\succeq$ a preference relation among objects. Its strict part is denoted by $\succ$ and defined by $\vec{\alpha} \succ \vec{\beta} \iff \vec{\alpha} \succeq \vec{\beta}$ and $\neg(\vec{\beta} \succeq \vec{\alpha})$. Finally, $\simeq$ denotes its symmetric part ($\vec{\alpha} \simeq \vec{\beta} \iff \vec{\alpha} \succeq \vec{\beta}$ and $\vec{\beta} \succeq \vec{\alpha}$). Sugeno integral defines such a preference relation that is a weak order on L^n (i.e. a complete and transitive relation):

$$\vec{\alpha} \succeq_\kappa^{sug} \vec{\beta} \Leftrightarrow S_\kappa(\vec{\alpha}) \geq S_\kappa(\vec{\beta}). \tag{11}$$

When there is no ambiguity, we simply use the notation $\succeq^{sug}$. We also write $\vec{\alpha} \succeq_P \vec{\beta}$ when $\vec{\alpha}$ weakly dominates $\vec{\beta}$ in the sense of Pareto, namely:

$$\vec{\alpha} \succeq_P \vec{\beta} \iff \alpha_i \geq \beta_i, \forall i = 1, \ldots, n. \tag{12}$$

Sugeno integral is *weakly* Pareto-monotonic, i.e., it obeys:

Axiom WPAR: $\forall \vec{\alpha}, \vec{\beta}, \vec{\alpha} \succeq_P \vec{\beta} \implies \vec{\alpha} \succeq_\kappa^{sug} \vec{\beta}$.

A tuple $\vec{\alpha}$ is said to κ-dominate $\vec{\beta}$ whenever $\forall \lambda \in L, \kappa(A_\lambda) \geq \kappa(B_\lambda)$. This is a general form of *Stochastic dominance*. From its expression as in (3), Sugeno integral is obviously in agreement with this kind of comparison:

Axiom WGSD : If $\vec{\alpha}$ κ-dominates $\vec{\beta}$ then then $S_\kappa(\vec{\alpha}) \geq S_\kappa(\vec{\beta})$.

2.3 Null sets

Some features may be considered totally useless in the evaluation process.

Null Sets A set of features A is said to be null with respect to a preference relation $\succeq$ on tuples if and only if $\forall \vec{\alpha}, \vec{\beta}, \vec{\gamma} \in L^n, \vec{\alpha} A \vec{\gamma} \succeq \vec{\beta} A \vec{\gamma}$.

In other words, the preference pattern between two objects does not depend on ratings according to features inside set A. These features have thus no importance. If $\succeq$ is defined by a weighted average, null sets A are characterized by $p_i = 0, \forall i \in A$. This is different when Sugeno integral is used.

Proposition 1 *When the preference relation is defined by Sugeno integral S_κ, A is null if and only if $\kappa(A \cup B) = \kappa(B), \forall B$.*

PROOF. If A is null, then let $\vec{\alpha} = \underline{\perp}, \vec{\beta} = \underline{\top}$, and for any set B, $\vec{\gamma} = \underline{\top} B \underline{\perp}$. Then $S_\kappa(\vec{\alpha} A \vec{\gamma}) \geq S_\kappa(\vec{\beta} A \vec{\gamma})$ reads $\kappa(B \setminus A) \geq \kappa(A \cup B)$, hence $\kappa(B) = \kappa(A \cup B)$. Conversely, assume $\kappa(A \cup B) = \kappa(B), \forall B$. Then $S_\kappa(\vec{\alpha} A \vec{\gamma}) = \max(\theta_1, \theta_2)$ where
$\theta_1 = \max_{E \subseteq A^c} \min(\kappa(E), \min_{i \in E} \gamma_i)$, and
$\theta_2 = \max_{E \not\subseteq A^c} \min(\kappa(E), \min_{i \in E \cap A^c} \gamma_i, \min_{i \in E \cap A} \alpha_i)$
where A^c is the complement of A. When $E \not\subseteq A^c$ let $C = E \cap A^c, D = E \cap A \neq \emptyset$, and notice that by assumption, $\kappa(E) = \kappa(C)$; it leads to:
$\theta_2 = \max_{C \subseteq A^c, \emptyset \neq D \subseteq A} \min(\kappa(C), \min_{i \in C} \gamma_i, \min_{i \in D} \alpha_i)$.
This also writes :
$\theta_2 = \min(\max_{\emptyset \neq D \subseteq A} \min_{i \in D} \alpha_i, \max_{C \subseteq A^c} \min(\kappa(C), \min_{i \in C} \gamma_i))$
$= \min(\max_{\emptyset \neq D \subseteq A} \min_{i \in D} \alpha_i, \theta_1) \leq \theta_1$.
In consequence, $S_\kappa(\vec{\alpha} A \vec{\gamma}) = \theta_1 = \max_{E \subseteq A^c} \min(\kappa(E), \min_{i \in E} \gamma_i)$, which does not depend on the features of $\vec{\alpha}$ in A. Hence A is null. QED

This characteristic property of null sets was proposed by Murofushi and Sugeno [30] who proved its equivalence with our definition for Sugeno integral. Remark that if A and B are null, so is $A \cup B$ and conversely (as also proved by Murofushi and Sugeno). If the preference relation $\succeq$ is defined by a Sugeno integral, A null obviously implies $\kappa(A) = \perp$, but $\kappa(A) = \perp$ does not imply that A is null. For instance assume there are three features and let κ be the necessity measure built on the possibility distribution $\pi_i = \top, \forall i$. Consider the tuples $\underline{\top}$ and $\underline{\perp}\{1\}\underline{\top}$. Obviously, $\kappa(\{1\}) = \perp$. But feature 1 is not null; indeed, $S_\kappa(\underline{\top}) = \top > S_\kappa(\underline{\perp}\{1\}\underline{\top}) = \perp$ (it is the median of $\{\perp, \perp, \top\}$). However, if κ is a possibility measure, then $\kappa(A) = \perp$ implies that A is null.

3 The weak discrimination power of qualitative preference functionals

Sugeno integrals, like other more specialized qualitative criteria, suffer from a lack of decisiveness and fail to satisfy strict monotonicity, i.e. the Pareto principle of efficiency:

Axiom SPAR: $\vec{\beta}, \vec{\alpha} \succ_P \vec{\beta}$ implies $\vec{\alpha} \succ \vec{\beta}, \forall \vec{\alpha}$.

where $\succ_P$ denotes the strict part of the Pareto dominance: $\vec{\alpha} \succ_P \vec{\beta}$ if and only if $\vec{\alpha} \succeq_P \vec{\beta}$, and $\exists i$ not null such that $\alpha_i > \beta_i$. This general principle says that, if $\vec{\alpha}$ is as least as good as $\vec{\beta}$ on each feature, and better than $\vec{\beta}$ on some non null feature [1], then $\vec{\alpha}$ should be strictly preferred to $\vec{\beta}$. However, even if $\vec{\alpha} \succ_P \vec{\beta}$, it may be that $S_\kappa(\vec{\alpha}) = S_\kappa(\vec{\beta})$, so that SPAR is not satisfied. This lack of discrimination is due to the so-called drowning effect.

3.1 Several drowning effects

The "drowning effect", is related to the use of idempotent operations — max and min. In particular, when two objects have identical good consequences for some important features, they may globally rate the same, although they may have significantly different ratings for the other features. As a consequence the principle of strict Pareto dominance is not satisfied, as already noticed.

For instance, let $n = 2$ features, $m = 10, \lambda_j = j$. Let $\vec{\alpha}$ and $\vec{\beta}$ be two objects whose ratings according to features 1 and 2 are listed below.

Feature	1	2
α_i	7	9
β_i	7	8

Consider the capacities

	$\emptyset$	$\{1\}$	$\{2\}$	$\{1,2\}$
κ_1	$\bot$	8	2	$\top$
κ_2	$\bot$	$\top$	2	$\top$
κ_3	$\bot$	8	$\bot$	$\top$

κ_1 is a classical probability measure, κ_2 and κ_3 are respectively the possibility and the necessity measure built on the possibility distribution $\pi_1 = \top$, $\pi_2 = 2$. All these measures contain the same ordinal information with respect to the relative importance of the features (1 is more important than 2 that is not null).

[1] Axiom SPAR does not apply to features forming null sets, which by definition do not play any role in the preference between acts.

One can check that, for each κ_i:

$$S_{\kappa_i}(\vec{\alpha}) = \max(\min(9, \kappa_i(\{2\})), \min(7, \kappa_i(\{1,2\})) = 7;$$
$$S_{\kappa_i}(\vec{\beta}) = \max(\min(8, \kappa_i(\{2\})), \min(7, \kappa_i(\{1,2\})) = 7.$$

So, $S_{\kappa_i}(\vec{\alpha}) = S_{\kappa_i}(\vec{\beta}), \forall i$, although $\vec{\alpha}$ strictly dominates $\vec{\beta}$ (as $\alpha_1 = \beta_1$ and $\alpha_2 > \beta_2$). The drowning effect is here due to the maximum operator: in both $S_{\kappa_i}(\vec{\alpha})$ and $S_{\kappa_i}(\vec{\beta})$ the external maximum is driven by the term $\min(7, \kappa_i(\{1,2\})$, which is equal to 7. The second term ($\min(9, \kappa_i(\{2\}))$ for $\vec{\alpha}$, $\min(8, \kappa_i(\{2\}))$ for $\vec{\beta}$) is not taken into account.

A second drowning effect may exist, driven by the minimum operator: suppose that the rating of both acts on feature 1 is the least possible – $\perp$. Then the term $\min(7, \kappa_i(\{1,2\})$ should be replaced by $\min(\perp, \kappa_i(\{1,2\})$, which is equal to $\perp$. For each of the capacities, the maximal term is the first one, equal to $\min(9, \kappa_i(\{2\})) = \kappa_i(\{2\})$ for $\vec{\alpha}$ and $\min(8, \kappa_i(\{2\})) = \kappa_i(\{2\})$ since $\kappa_i(\{2\}) < 8$. Hence, the Pareto dominance of $\vec{\beta}$ by $\vec{\alpha}$ on feature 2 is drowned by the fact that the weight $\kappa_i(\{2\})$ is very low.

Finally, one should notice that a third drowning effect is present, inherent to the capacity itself. Indeed, the capacities are not required to satisfy the strict Pareto principle. Applied to sets of features, this condition writes $\forall A, B, \kappa(B) > \perp \implies \kappa(A \cup B) > \kappa(A)$. Probabilities obviously satisfy it. But possibility measures, and plausibility measures in the sense of Shafer[36] fail to satisfy this property. Overcoming these drowning effects is the major motivation of the results presented in the paper.

3.2 Pareto-dominance and preferential independence

The drowning effect is also often understood as an incapacity to obey preferential independence (a form of Savage's Sure-Thing Principle):

Axiom PI $\vec{\alpha}A\vec{\gamma} \succeq \vec{\beta}A\vec{\gamma}$ if and only $\vec{\alpha}A\vec{\delta} \succeq \vec{\beta}A\vec{\delta}$.

It can be severely violated by Sugeno integral. It is easy to show that there may exist four tuples such that $\vec{\alpha}A\vec{\gamma} \succ \vec{\beta}A\vec{\gamma}$ while $\vec{\beta}A\vec{\delta} \succ \vec{\alpha}A\vec{\delta}$. It is enough to consider Boolean tuples (subsets) and notice that, generally if A is disjoint from $B \cup C$, nothing forbids a fuzzy measure κ to satisfy $\kappa(B) > \kappa(C)$ along with $\kappa(A \cup C) > \kappa(A \cup B)$ (for instance, belief functions are such). The prioritized maximum and minimum W_π^+ and W_π^- violate independence to a lesser extent since they obey the following weak form of PI:

Axiom WPI: $\forall A, \forall \vec{\alpha}, \vec{\beta}, \vec{\gamma}, \vec{\delta}, \vec{\alpha}A\vec{\gamma} \succ \vec{\beta}A\vec{\gamma} \Rightarrow \vec{\alpha}A\vec{\delta} \succeq \vec{\beta}A\vec{\delta}$.

It has been shown by Marichal [26] that axiom PI is generally not compatible with Sugeno integrals. We can moreover prove that Sugeno integrals are almost incompatible not only with PI, but also with the less demanding principle of Pareto efficiency.

Theorem 2 *Relation $\succeq^{sug}_{\kappa,u}$ is Pareto-efficient if and only if there exists a unique feature i^* such that $\forall A, \kappa(A) = \top$ if $i^* \in A, \kappa(A) = \bot$ if $i^* \notin A$.*

PROOF. Suppose i^* such that $\forall A, \kappa(A) = \top$ if $i^* \in A, \kappa(A) = \bot$ if $i^* \notin A$. Then, $\forall A, i^* \notin A$, A is a null set since $\kappa(A \cup B) = \kappa(B) = \bot$ if $i^* \notin B$, and $\kappa(A \cup B) = \kappa(B) = \top$ if $i^* \in B$. So, all features are null, but i^*. So, $S_\kappa(\omega) = (\omega)_{i^*}$ satisfies SPAR.

Conversely, suppose $\exists A, \top > \kappa(A) = \lambda > \bot$. Then consider the constant tuple $\underline{\lambda}$ and tuple $\underline{\top} A \underline{\lambda} \succ_P \underline{\lambda}$. Obviously $S_\kappa(\underline{\lambda}) = \lambda = \kappa(A)$ and $S_\kappa(\underline{\top} A \underline{\lambda}) = \max(\min(\lambda, \top), \min(\top, \kappa(A))) = \kappa(A)$. Since A is not null, this result constitutes a violation of the principle of efficiency. Hence we must restrict to Boolean set-functions such that $\kappa(A) \in \{\bot, \top\}$. Now let A, B be disjoint subsets such that $\kappa(A) = \kappa(B) = \top$. Consider objects $\underline{\alpha} A \underline{\beta}$ and $\underline{\alpha}(A \cup B)\underline{\beta}$, with $\alpha > \beta$. Then $S_\kappa(\underline{\alpha} A \underline{\beta}) = S_\kappa(\underline{\alpha}(A \cup B)\underline{\beta}) = \alpha$ while clearly $\underline{\alpha}(A \cup B)\underline{\beta} \succ_P \underline{\alpha} A \underline{\beta}$. So the set-functions obeying strict Pareto efficiency must be such that $\kappa(A) = \kappa(B) = \top$ implies $A \cap B \neq \emptyset$. So the minimal sets A such that $\kappa(A) = \top$ cannot be disjoint. If one such minimal set A contains at least two features, then consider the n-tuple $\vec{\delta}$ such that $\delta_i = \alpha$ for some $i \in A$, $\delta_k = \beta$, for $k \in A$, $k \neq i$, and $\bot$ otherwise. Clearly, $\vec{\delta} \succ_P \underline{\beta} A \bot$. It is clear that $S_\kappa(\vec{\delta}) = S_\kappa(\underline{\beta} A \bot) = \min(\kappa(A), \beta) = \beta$. Hence A must be a singleton $\{i^*\}$. Hence it is unique. QED

Note that theorem could be reformulated as follows: $\succeq^{sug}$ is Pareto-efficient if and only if there exists a unique essential feature according to which objects are compared. This result means that Sugeno integral involving more than one feature cannot be efficient. Such impossibility results are not necessarily damning. It indeed remains possible to look for a refinement of the weak ordering induced by Sugeno integral, i.e. get a decision rule coherent with Sugeno integral (i.e. following the strict preference induced by latter, if any) but possibly overcoming the drowning situations, and thus being more discriminant than Sugeno integral.

The problem of refining the weak ordering induced by Sugeno integrals was actually studied by Murofushi [29]. This author showed a result similar to the one above considering a weaker condition that SPAR, namely, if $\alpha_i > \beta_i, \forall i = 1, \dots, n$ then $\vec{\alpha} \succ \vec{\beta}$. He noticed the lack of discrimination of Sugeno by proving that $\succeq^{sug}$ satisfies the latter property only if the capacity κ takes values in $\{\bot, \top\}$. He then proposed to refine it by means of several capacities $\kappa_1, \dots, \kappa_q$ inducing a tuple of global evaluations $(S_{\kappa_1}(\vec{\alpha}), \dots, S_{\kappa_q}(\vec{\alpha}))$ for each $\vec{\alpha}$. Murofushi then proposed to refine the $\succeq^{sug}$ ordering by a lexicographic use of the tuples of global evaluations, showing conditions to recover the SPAR axiom.

However, it has been proved more recently [21] that two special Sugeno integrals, namely those defined from a possibility distribution, can be *refined* by a weighted average, which corresponds to a generalization of the leximax and leximin procedures. So there is hope to refine the ordering $\succeq^{sug}$ by exploiting these recent results.

4 Refined qualitative prioritized maximum and minimum

The basic natural way to overcome the lack of discrimination power of Sugeno integrals consists of refining the ordering $\succeq^{sug}$. Recall that for any preference relation $\succeq$, a refinement of $\succeq$ is a different relation $\succeq'$ on the same universe such that:

$$\vec{\alpha} \succ \vec{\beta} \implies \vec{\alpha} \succ' \vec{\beta}. \tag{13}$$

Lexicographic refinements are a natural way to go in a qualitative setting. They can refine both the Pareto-ordering and the pessimistic ordering based on Wald criterion. They were recently successfully extended to overcome the lack of discrimination of the possibilistic qualitative decision rules (i.e. for criteria W_π^- and W_π^+) [21]. Since these results are the basis of the findings in the present paper, they are recalled in the remainder of the Section.

4.1 Additive refinements of minimum and maximum

When comparing tuples, the drowning effect of the minimum aggregation can be fixed by the so-called leximin ordering. Symmetrically, the leximax ordering overcomes the lack of discrimination of the maximax prioritized maximum criterion. Practically, the leximin procedure (resp. the leximax procedure) consists in ordering both tuples in increasing (resp. decreasing) order and then lexicographically comparing them [10].

Definition 3 (leximax, leximin) *Let $\vec{\alpha}, \vec{\beta} \in L^n$. Then*

- $\vec{\alpha} \succ_{lmax} \vec{\beta} \Leftrightarrow \exists i, \forall j < i, \alpha_{(j)} = \beta_{(j)}$ *and* $\alpha_{(i)} > \beta_{(i)}$;
- $\vec{\alpha} \succ_{lmin} \vec{\beta} \Leftrightarrow \exists i, \forall j > i, \alpha_{(j)} = \beta_{(j)}$ *and* $\alpha_{(i)} > \beta_{(i)}$;
- $\vec{\alpha} \sim_{lmax} \vec{\beta} \Leftrightarrow \vec{\alpha} \sim_{lmin} \vec{\beta} \Leftrightarrow \forall j, \alpha_{(j)} = \beta_{(j)}$,

where, for any $\vec{w} \in L^n$, $w_{(k)}$ is the k-th greatest *element of $\vec{w}$ (i.e. $w_{(1)} \geq \ldots \geq w_{(n)}$).*

Both rules conclude to indifference if and only if the corresponding reordered tuples are the same. The leximin-ordering is a refinement of both the Pareto-ordering and the maximin-ordering [14]: $\min_{i=1,\ldots,n} \alpha_i > \min_{i=1,\ldots,n} \beta_i$ implies $\vec{\alpha} \succ_{lmin} \vec{\beta}$ and $\vec{\alpha} \succ_P \vec{\beta}$ implies $\vec{\alpha} \succ_{lmin} \vec{\beta}$. leximin optimal decisions are always indeed min-optimal and Pareto-maximal: $\succeq_{lmin}$ is the most selective among these preference relations. The leximin ordering can discriminate more than any symmetric aggregation function: for instance, the reordered tuples can be different (and thus the leximin criterion is capable of discriminating) even when (assuming numbers) the sum of α_i's equals the sum of β_i's. Similar remarks apply to the leximax ordering with respect to the maximax criteria.

Interestingly, the qualitative leximin and leximax rules can be simulated by means of a sum of numerical values 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 ϕ from L to the reals. It is possible to define this mapping in such a way as to refine the maximax ordering:

$$\max_{i=1,\ldots,n} \alpha_i > \max_{i=1,\ldots,n} \beta_i \text{ implies } \sum_{i=1}^{n} \phi(\alpha_i) > \sum_{i=1}^{n} \phi(\beta_i). \tag{14}$$

For instance, the transformation $\phi(\lambda_j) = N^j$ with $N > n$ achieves this goal. It is a super-increasing mapping in the sense that $\phi(\lambda_j) > \sum_{k<j} \phi(\lambda_k), \forall j = 1, \ldots, m$. In order to map L to [0, 1] so that $\phi(\lambda_0) = 0$ and $\phi(\lambda_m) = 1$ just take $\phi(\lambda_j) = \frac{N^j - 1}{N^m - 1}$.

It can actually be checked that the leximax ordering is retrieved by means of this refinement, based on the sum:

$$\vec{\alpha} >_{lmax} \vec{\beta} \text{ if and only if } \sum_{i=1}^{n} \phi(\alpha_i) > \sum_{i=1}^{n} \phi(\beta_i). \tag{15}$$

Function $\phi(.)$ is convex, which is in line with the optimistic behavior of W^+.

A similar encoding of the leximin procedure can be achieved by means of a sum, using another super-increasing mapping of the form $\psi(\lambda_j) = k - \phi(\nu(\lambda_j))$, (for instance, with $k = 1$, the transformation $\psi(\lambda_j) = \frac{1-N^{-j}}{1-N^{-m}}$):

$$\min_{i=1,\ldots,n} \alpha_i > \min_{i=1,\ldots,n} \beta_i \text{ implies } \sum_{i=1}^{n} \psi(\alpha_i) > \sum_{i=1}^{n} \psi(\beta_i). \tag{16}$$

It can actually be checked that the leximin ordering is retrieved by means of this refinement, function $\psi(.)$ being concave, which is in line with the pessimistic behavior of W^-.

The qualitative pessimistic and optimistic Wald criteria are thus refined by means of a numerical criterion with respect to a risk-averse and risk-prone utility function respectively, as can be seen by plotting L against numerical values in $\phi(L)$ and

$\psi(L)$. Notice that these transformations are not possible when L is not finite [28] although the leximin and leximax procedures make mathematical sense even in this case.

4.2 Additive refinements of possibilistic preference functionals

Prioritized maximum and minimum W_π^+ and W_π^- can be refined by by means of weighted averages, thus recovering Savage's five first axioms[20]. Consider first the prioritized maximum W_π^+ under a given possibility distribution π. We can again define an increasing mapping χ from L to the reals such that $\chi(\lambda_0) = 0$ and especially:

$$\begin{array}{l} \max_{i=1,\ldots,n} \min(\pi_i, \alpha_i) > \max_{i=1,\ldots,n} \min(\pi_i, \beta_i) \\ \text{implies} \\ \sum_{i=1}^n \chi(\pi_i) \cdot \chi(\alpha_i) > \sum_{i=1}^n \chi(\pi_i) \cdot \chi(\beta_i). \end{array} \tag{17}$$

A sufficient condition is that:

$$\forall j \in \{1, \ldots, m\}, \chi(\lambda_j)^2 \geq N\chi(\lambda_{j-1}) \cdot \chi(\top), \tag{18}$$

for $N > n$. The increasing mapping is such that :

$$\chi(\lambda_m) = 1, \chi(\lambda_0) = 0, \chi(\lambda_j) = \frac{N}{N^{2^{m-j}}}, j = 1, m-1, \tag{19}$$

with $N = n + 1$ can be chosen, with $m = |L|$. Moreover, let $\{E_0, \ldots, E_k\}$ be the partition of $\{1, 2, \ldots, n\}$ induced by π, such that $\forall i, i' \in E_j, \pi(i) = \pi(i')$ and whenever $j > j', i \in E_j, i' \in E_{j'}, \pi(i) > \pi(i')$. E_k contains the most important features, and E_0 the null features. Let $K = \frac{1}{\sum_{l=1,k} |E_l| \cdot \chi(\pi_l)}$. Define $\chi^*(\lambda_j) = K\chi(\lambda_j)$, it holds that:

- $p = \chi^*(\pi(\cdot))$ is a probability assignment respectful of the possibilistic ordering of states. In particular, p is uniform on equi-possible states (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 states, that is, $p_i > P(E_{i-1} \cup \cdots \cup E_0)$. Such probabilities generalize the linear big-stepped probabilities that form a super-increasing sequence [3] (recovered when the E_i's are singletons) and are simply called big-stepped probabilities here.
- The $\chi(\lambda_j)$'s coefficients form a super-increasing sequence of reals $r_m > \cdots > r_1$ such that $\forall m \geq j > 1, r_j > n \cdot r_{j-1}$ that can be encoded by a convex real mapping.

- The preference functional

$$\mathrm{WA}^{+}_{\chi(\pi)}(\vec{\alpha}) = \sum_{i=1}^{n} \chi^{*}(\pi_i) \cdot \chi(\alpha_i) \tag{20}$$

is a big-stepped weighted average function (e.g. an expected utility criterion) for a risk-seeking decision-maker, and $W^{+}_{\pi}(\vec{\alpha}) > W^{+}_{\pi}(\vec{\beta})$ implies $\mathrm{WA}^{+}_{\chi(\pi)}(\vec{\alpha}) > \mathrm{WA}^{+}_{\chi(\pi)}(\vec{\beta})$. Namely this is precisely equation (17) up to the multiplicative constant K. i.e., the weighted average criterion so-obtained refines the prioritized maximum criterion. As a refinement, it is perfectly compatible with but more decisive than the latter. Since it is a weighted average, it obviously satisfies preferential independence PI as well as strict Pareto dominance SPAR. Moreover, it does not use any other information but the original ordinal one. It can be shown that it is not the only criterion in this family of sound "unbiased" refinements, but it is the most efficient among them (up to an equivalence relation), since it refines any unbiased refinement of the prioritized maximum criterion (see [21] for more details).

The prioritized minimum criterion can be similarly refined. Notice that $W^{-}_{\pi}(\vec{\alpha}) = \nu(W^{+}_{\pi}(\nu(\vec{\alpha})))$, with $\nu(\vec{\alpha})_i = \nu(\alpha_i)$ using the order-reversing map ν of L. Then, choosing the same mapping χ^{*} as above, one may have that

$$\begin{array}{l} \min_{i=1,\dots,n} \max(\nu(\pi_i), \alpha_i) > \min_{i=1,\dots,n} \max(\nu(\pi_i), \beta_i) \\ \text{implies} \\ \sum_{i=1}^{n} \chi^{*}(\pi_i) \cdot \phi(\alpha_i) > \sum_{i=1}^{n} \chi^{*}(\pi_i) \cdot \phi(\beta_i), \end{array} \tag{21}$$

where $\phi(\lambda_j) = 1 - \chi(\nu(\lambda_j))$ (it is equal to $1 - \frac{N}{N^{2j}}$, for $j < m$, and 1 if $j = m$, with the same value of N as for the prioritized maximum). Coefficients $\phi(\lambda_j)$ form a super-increasing sequence that can be encoded by means of a concave real mapping, and the weighted average criterion

$$\mathrm{WA}^{-}_{\chi(\pi)}(\vec{\alpha}) = \sum_{i=1}^{n} \chi^{*}(\pi(i)) \cdot \phi(\alpha_i) \tag{22}$$

is a risk-averse one, that refines W^{-}_{π} in the sense that $W^{-}_{\pi}(\vec{\alpha}) > W^{-}_{\pi}(\vec{\beta})$ implies $\mathrm{WA}^{-}_{\chi(\pi)}(\vec{\alpha}) > \mathrm{WA}^{-}_{\chi(\pi)}(\vec{\beta})$.

4.3 Weighted leximax / leximin criteria

The orderings induced by $\mathrm{WA}^{+}_{\chi(\pi)}(\vec{\alpha})$ and $\mathrm{WA}^{-}_{\chi(\pi)}(\vec{\alpha})$ actually correspond to generalizations of leximin and leximax to prioritized minimum and maximum aggregations, thus bridging the gap between prioritized maximum and minimum and classical decision theory. To make this generalization clear, let us simply consider that leximin and leximax orderings are defined on sets of tuples whose components belong to a totally ordered set $(\Lambda, \trianglerighteq)$, say $leximin(\trianglerighteq)$ and $leximax(\trianglerighteq)$. Now, suppose $(\Lambda, \trianglerighteq) = (L^l, \geq_{lmin})$ or $(\Lambda, \trianglerighteq) = (L^l, \geq_{lmax})$, with any $l \in \mathbb{N}$. Then, nested lexicographic ordering relations can be recursively defined by nesting procedures such as $leximin(\geq_{lmin})$, $leximax(\geq_{lmin})$, $leximin(\geq_{lmax})$, and finally $leximax(\geq_{lmax})$, that can compare L-valued matrices.

Consider the procedure $leximax(\geq_{lmin})$ defining the relation $\succeq_{lmax(\succeq lmin)}$, for instance. It applies to matrices $[\alpha]$ of dimension $q_1 \times q_2$ with coefficients α_{ij} in $(L, \geq)$. These matrices can be totally ordered in a very refined way by this relation. Denote by $\alpha_{i\cdot}$ row i of $[\alpha]$. Let $[\alpha^\star]$ and $[\beta^\star]$ be rearranged matrices $[\alpha]$ and $[\beta]$ such that terms in each row are reordered increasingly and rows are arranged lexicographically top-down in decreasing order. $[\alpha] \succ_{lmax(\succeq lmin)} [\beta]$ is defined as follows :

$$\exists k \leq q_1 \text{ s.t. } \forall i < k, \alpha^\star_{i\cdot} =_{lmin} \beta^\star_{i\cdot} \text{ and } \alpha^\star_{k\cdot} >_{lmin} \beta^\star_{k\cdot}$$

Relation $\succeq_{lmax(\succeq lmin)}$ is a complete preorder. $[\alpha] \simeq_{lmax(\succeq lmin)} [\beta]$ 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 prioritized maximum criterion:

$$\max_{i=1,\ldots,q_1} \min_{j=1,\ldots,q_2} \alpha_{ij} > \max_{i=1,\ldots,q_1} \min_{j=1,\ldots,q_2} \beta_{ij} \text{ implies } [\alpha] \succ_{lmax(\succeq lmin)} [\beta].$$

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

Comparing tuples $\vec{\alpha}$ and $\vec{\beta}$ in the context of a possibility distribution π can be done using relations $\succeq_{lmax(\succeq lmin)}$ applied to $n \times 2$ matrices with coefficients in $(L, \leq)$, n being the number of features, namely on the matrices $[\alpha^\pi]$ and $[\beta^\pi]$ with coefficients $\alpha^\pi_{i1} = \pi_i$ and $\alpha^\pi_{i2} = \alpha_i$, $\beta^\pi_{i1} = \pi_i$ and $\beta^\pi_{i2} = \beta_i$.

The weighted average $\mathrm{WA}^+_{\chi(\pi)}(\vec{\alpha})$ defined in the previous section precisely encodes the relation $\succeq_{lmax(\succeq lmin)}$:

Theorem 4 *[20]: $WA^+_{\chi(\pi)}(\vec{\alpha}) \geq WA^+_{\chi(\pi)}(\vec{\beta})$ if and only if $[\alpha^\pi] \succeq_{lmax(\succeq lmin)} [\beta^\pi]$.*

In other terms, $\mathrm{WA}^+_{\chi(\pi)}$ applies a leximax procedure to utility degrees weighted by possibility degrees. Similarly, $\mathrm{WA}^-_{\chi(\pi)}$ applies a leximin procedure to utility degrees weighted by “impossibility degrees”:

Theorem 5 *[20]: $WA^-_{\chi(\pi)}(\vec{\alpha}) \geq WA^-_{\chi(\pi)}(\vec{\beta})$ if and only if $[\alpha^{\nu(\pi)}] \succeq_{lmin(\succeq lmax)} [\beta^{\nu(\pi)}]$.*

i.e., the weighted average $\mathrm{WA}^-_{\chi(\pi)}(\vec{\alpha})$ just encodes the application of a procedure *leximin(leximax)* not directly on $[\alpha^\pi]$ and $[\beta^\pi]$ but on the corresponding matrices $[\alpha^{\nu(\pi)}]$ and $[\beta^{\nu(\pi)}]$ with coefficients $[\alpha^{\nu(\pi)}]_{i1} = \nu(\pi_i)$ and $[\alpha^{\nu(\pi)}]_{i2} = \alpha_i$, $[\beta^{\nu(\pi)}]_{i1} =$

$\nu(\pi_i)$ and $[\beta^{\nu(\pi)}]_{i2} = \beta_i$.

As a consequence, the additive preference functionals $\mathrm{WA}^+_{\chi(\pi)}(\vec{\alpha})$ and $\mathrm{WA}^-_{\chi(\pi)}(\vec{\alpha})$ refining the prioritized maximum and minimum are qualitative despite their numerical encoding. Moreover, the two orderings $\succeq_{lmax(\succeq lmin)}$ and $\succeq_{lmin(\succeq lmax)}$ of acts are defined even on coarse ordinal scales L while *obeying preferential independence*.

The two relations coincide if the utility functions are Boolean. This is not surprising since $\succeq_{lmin}$ and $\succeq_{lmax}$ are conjugate ($\vec{\alpha} \succeq_{lmin} \vec{\beta}$ if and only if $(\nu(\beta_1), \ldots, \nu(\beta_k))$ $\succeq_{lmax}$ $(\nu(\alpha_1), \ldots, \nu(\alpha_k))$. Another formulation of this result consists in noticing that $\mathrm{WA}^+_{\chi(\pi)}(\vec{\alpha})$ and $\mathrm{WA}^-_{\chi(\pi)}(\vec{\alpha})$ share the same big-stepped probability function. This representation is probabilistic, although qualitative, and is precisely the lexi-refinement of both possibility and necessity orderings identified by [15]:

$$A \succeq_{\Pi_{Lex}} B \text{ if and only if } \vec{\pi_A} \succeq_{lmax} \vec{\pi_B}, \tag{23}$$

where $\vec{\pi_A}$ is the tuple $(a_1, \ldots, a_n)$ such that $a_i = \pi_i$ if $i \in A$ and $a_i = \perp$ otherwise. This importance relation among sets of features is called "leximax" likelihood [15,12]. It is a complete preordering which refines the possibilistic ordering of sets ($A \succeq_\Pi B \iff \Pi(A) \geq \Pi(B)$, where $\forall A \subseteq S, \Pi(A) = \max_{i \in A} \pi_i$) together its conjugate necessity ordering ($A \succeq_N B \iff N(A) \geq N(B)$, where $\forall A \subseteq S, N(A) = \nu(\Pi(A^c))$). The leximax refinement $\succeq_{\Pi_{Lex}}$ of a possibility ordering induced by a uniform possibility distribution on features coincides with the comparative probability relation induced by the uniform probability distribution. This is not surprising in view of the fact that the leximax likelihood relation is really a comparative probability relation in the usual sense, representable by a big-stepped probability function.

4.4 Lessons learnt for refining Sugeno integral

The results presented above obviously suggest that Sugeno integral could be refined in a similar way. Some preliminary remarks provide some insight on the possible extension of this criterion.

First, the reason why the prioritized maximum and minimum could be refined by means of a weighted average is because these qualitative aggregation rules satisfy a weak form of preferential independence, namely WPI. However, there is no hope of refining Sugeno integral by means of a weighted average since the former strongly violates axiom PI. However the form of Sugeno integral:

$$S_\kappa(\vec{\alpha}) = \max_{j=1,\ldots,m} \min(\lambda_j, \kappa(E_{\lambda_j}))$$

strongly suggests to refine a Sugeno integral by means of a Choquet integral with respect to a numerical capacity v encoded by the qualitative one κ:

$$Ch_v(\vec{\alpha}) = \sum_{j=1}^{m} v(E_{\lambda_j}).(u_j - u_{j-1})$$

where u_j is a numerical encoding of λ_j. Choquet integral is additive on co-monotonic tuples $\vec{u}$ and $\vec{u'}$, and, for such tuples, the sure-thing principle is valid [34,26]. Sugeno integral is minitive and maxitive for co-monotonic tuples and obeys, for such acts, the weak form of independence (WPI) satisfied by the prioritized minimum and maximum:

Proposition 6 *If $\vec{\alpha}A\vec{\gamma}, \vec{\beta}A\vec{\gamma}, \vec{\alpha}A\vec{\delta}$ and $\vec{\beta}A\vec{\delta}$ induce the same ordering of features, then the following property holds: $S_\kappa(\vec{\alpha}A\vec{\gamma}) > S_\kappa(\vec{\beta}A\vec{\gamma})$ implies $S_\kappa(\vec{\alpha}A\vec{\delta}) \geq S_\kappa(\vec{\beta}A\vec{\delta})$.*

PROOF. Let $\{1, \ldots, n\}$ be the joint ordering of features induced by the four tuples. Then let $A_i = \{1, \ldots, i\}$ and $\pi_i = \kappa(A_i), \forall i = 1 \ldots, n$. The one can write $S_\kappa(\vec{\alpha}A\vec{\gamma}) = W_\pi^+(\vec{\alpha}A\vec{\gamma})$, $S_\kappa(\vec{\beta}A\vec{\gamma}) = W_\pi^+(\vec{\beta}A\vec{\gamma})$, $S_\kappa(\vec{\beta}A\vec{\delta}) = W_\pi^+(\vec{\beta}A\vec{\delta})$, and $S_\kappa(\vec{\alpha}A\vec{\delta}) = W_\pi^+(\vec{\alpha}A\vec{\delta})$. Hence WPI holds for such four acts.

Sugeno integral also respects stochastic dominance in the wide sense, which is one of the key axioms proposed in [33] to axiomatize Choquet integral in Savage style.

Actually, restricting to tuples of utilities that rank features in a prescribed order, Choquet integral behaves like a weighted average and Sugeno integral behaves like a prioritized minimum or maximum. So refining a Sugeno integral by means of a Choquet integral looks like the right way to go, relying on the method for refining the prioritized minimum and maximum by means of a weighted average.

However, as Sugeno integral takes various equivalent forms, the result of the refinement will depend on the chosen form to which a big-stepped transformation is applied. Hence, there are two approaches one might think of for achieving this program.

- Applying a super-increasing transformation directly on the original definition of Sugeno integral, thus preserving the nature of the original capacity. This approach preserves the potential lack of discrimination due to the set-function. The latter can be refined in turn if needed. This approach can be used on the forms (3) or (4) of Sugeno integral.
- Applying a super-increasing transformation to the expression (5) of Sugeno integral, involving all subsets of features. A representation of the capacity by means of an ordinal counterpart to the Moebius transform is used to reduce the redundancy of expression (5). The questionable point in this method is that the nature

of the capacity changes in the transformation since it becomes a belief function. But the method retrieves the weighted average refinement of the prioritized maximum criterion as a special case.

5 Capacity-preserving refinements

In its standard expression $S_\kappa(\vec{\alpha}) = \max_{\lambda_j \in L} \min(\lambda_j, \kappa(A_{\lambda_j}))$, the two operators $\max$ and $\min$ are monotonic but not strictly, hence two nested drowning effects. The simplest idea to refine Sugeno integral is to consider a *leximax($\succeq_{lmin}$)* refinement of this maxmin expression. However we can also use expression (4) of Sugeno integral where we maximize over the feature set, yielding another refinement. The reconciliation of the two approaches is discussed.

5.1 Refinements respecting stochastic dominance

Consider the following decision rule, based on a straightforward lexicographic refinement of the standard expression (3):

$$\vec{\alpha} \succeq_L^{lsug} \vec{\beta} \iff [\vec{\alpha}^\kappa]_L \succeq_{lmax(\succeq lmin)} [\vec{\beta}^\kappa]_L, \tag{24}$$

where $[\vec{\alpha}^\kappa]_L$ is a $m \times 2$ matrix on $(L, \leq)$ with coefficients $\vec{\alpha}^\kappa_{j1} = \lambda_j$ and $\vec{\alpha}^\kappa_{j2} = \kappa(A_{\lambda_j}), i = 1, \ldots, m$. Note that $\kappa(A_{\lambda_0}) = \top$ always, and we do not need row$(\perp, \top)$ in the matrix. The properties of $\succeq_{lmax(\succeq lmin)}$ are thus inherited:

Corollary 7
$\succeq_L^{lsug}$ is a complete and transitive relation.
It refines the ranking of acts $\succeq^{sug}$ provided by Sugeno integral S_κ.

Moreover, since the maximum operator in the standard expression is taken over elements of the scale L, we are fully in agreement with stochastic dominance:

Proposition 8
$\vec{\alpha} \sim_L^{lsug} \vec{\beta} \iff \forall \lambda, \kappa(A_\lambda) = \kappa(B_\lambda)$;
If $\vec{\alpha}$ κ-dominates $\vec{\beta}$ ($\forall \lambda, \kappa(A_\lambda) \geq \kappa(B_\lambda)$ and $\kappa(A_\lambda) > \kappa(B_\lambda)$ for some λ) then $\vec{\alpha} \succ_L^{lsug} \vec{\beta}$.

Example 9 *Consider tuples $\vec{\alpha}$ such that $\alpha_i = 5$ if $i \in A$ and 2 otherwise, and $\vec{\beta}$ such that $\beta_i = 7$ if $i \in B$, 4 if $i \in C$ and 2 otherwise, where B and C are disjoint sets of features. Assume $\kappa(A) = 4$, $\kappa(B) = 2$, $\kappa(B \cup C) = 5$. Then the following matrices $[\vec{\alpha}^\kappa]_L$ and $[\vec{\beta}^\kappa]_L$ with rows $(\lambda_j, \kappa(A_{\lambda_j}))$ and $(\lambda_j, \kappa(B_{\lambda_j}))$ can be devised:*

$$[\vec{\alpha}^\kappa]_L = \begin{pmatrix} \top & \bot \\ 7 & \bot \\ 5 & 4 \\ 4 & 4 \\ 2 & \top \end{pmatrix}; [\vec{\beta}^\kappa]_L = \begin{pmatrix} \top & \bot \\ 7 & 2 \\ 5 & 2 \\ 4 & 5 \\ 2 & \top \end{pmatrix}.$$

It is clear that $S_\kappa(\vec{\alpha}) = S_\kappa(\vec{\beta}) = 4$, but $\vec{\alpha} \succ_L^{lsug} \vec{\beta}$ since the maximal $leximin$-pair on each side is $(4,5) \sim_{lmin} (5,4)$ and then $(4,4)$ is the next dominating pair.

Now, being a *leximax($\geq_{lmin}$)* procedure, $\succeq_L^{lsug}$ can be encoded by a sum of products. We can for instance use a super-increasing function χ similar to the previous one, built with respect to the number of levels in the scale L rather than with respect to the number of features. Here, the $\max$ operator applies to the m positive levels in L rather than to the n features of $\mathcal{F}$, hence we choose constant $N = m + 1$ in the definition (19) of function χ. We can now immediately derive:

Theorem 10 $\vec{\alpha} \succeq_L^{lsug} \vec{\beta} \iff \sum_{\lambda \in L} \chi(\lambda) \cdot \chi(\kappa(A_\lambda)) \geq \sum_{\lambda \in L} \chi(\lambda) \cdot \chi(\kappa(B_\lambda))$.

So, we define a new evaluation function E_L^{lsug}, that refines the ranking provided by S_κ, in agreement with $\succeq_L^{lsug}$:

$$E_L^{lsug}(\vec{\alpha}) = \sum_{\lambda \in L} \chi(\lambda) \cdot \chi(\kappa(A_\lambda)). \tag{25}$$

It should be noticed that $E_L^{lsug}(\top A \bot)$ is proportional to $\chi(\kappa(A))$ i.e. when utility degrees are Boolean, the comparison of tuples in terms of E_L^{lsug} is perfectly equivalent to the comparison in terms of κ — that is why we say that this refinement preserves the capacity. However, the aggregated evaluation E_L^{lsug} is not idempotent since $E_L^{lsug}(\underline{\lambda_j}) = \sum_{k \leq j} \chi(\lambda_k) \neq \lambda_j$. The numerical representation we look forward to is a Choquet integral (2), which preserves idempotence.

Notice that Sugeno integral is of the form $\max_{j=1}^m \min(\lambda_j, \gamma_j)$ with $\top \geq \gamma_1, \geq \ldots, \geq \gamma_m \in L$, letting $\gamma_j = \kappa(A_{\lambda_j})$. Then the following result is instrumental:

Lemma 11 *Consider three groups of coefficients $\top \geq \gamma_1 \geq \cdots \geq \gamma_m \in L$, $\top \geq \delta_1 \geq \cdots \geq \delta_m \in L$, and $\lambda_1 < \cdots < \lambda_m = \top \in L$, there exists an increasing mapping $\Phi : L \to [0, 1]$ such that $\Phi(\bot) = 0, \Phi(\top) = 1$ and:*

$$\max_{j=1,\ldots,m} \min(\lambda_j, \gamma_j) > \max_{j=1,\ldots,m} \min(\lambda_j, \delta_j)$$

implies

$$\sum_{j=1}^m \Phi(\gamma_j) \cdot (\Phi(\lambda_j) - \Phi(\lambda_{j-1})) > \sum_{j=1}^m \Phi(\delta_j) \cdot \Phi(\lambda_j).$$

PROOF. Increasing mapping $\Phi : L \to [0,1]$ such that $\Phi(\bot) = 0, \Phi(\top) = 1$ clearly exist. The most demanding situation for ensuring that the above strict inequality between maxmin qualitative expressions enforces the other quantitative inequality side is when $\max_{j=1,\dots,m} \min(\lambda_j, \gamma_j) = \lambda_k$ and $\max_{j=1,\dots,m} \min(\lambda_j, \delta_j) = \lambda_{k-1}$, with moreover, $\gamma_j = \bot, \forall j > k$ and $\gamma_j = \lambda_k, \forall j \leq k$, while $\delta_j = \top, \forall j < k-1$, and $\delta_j = \lambda_{k-1}, \forall j \geq k-1$. Then the quantitative inequality reads:

$$\sum_{j=1}^{k} \Phi(\lambda_k) \cdot (\Phi(\lambda_j) - \Phi(\lambda_{j-1})) > \sum_{j=1}^{k-2} \Phi(\top) \cdot \Phi(\lambda_j) + \sum_{j=k-1}^{m} \Phi(\lambda_{k-1}) \cdot \Phi(\lambda_j)$$

In order to ensure the above inequality, noticing that

- $\sum_{j=1}^{k} \Phi(\lambda_k) \cdot (\Phi(\lambda_j) - \Phi(\lambda_{j-1})) = \Phi(\lambda_k)^2$
- $\sum_{j=1}^{k-2} \Phi(\top) \cdot \Phi(\lambda_j) + \sum_{j=k-1}^{m} \Phi(\lambda_{k-1}) \cdot \Phi(\lambda_j) < \Phi(\lambda_{k-1}) \cdot ((k-2)\Phi(\top) + \sum_{j=k-1}^{m} \Phi(\lambda_j)) < N\Phi(\top)\Phi(\lambda_{k-1})$,

we can require a stronger sufficient condition : $\Phi(\lambda_k)^2 \geq N\Phi(\top)\Phi(\lambda_{k-1})$ with $N > m$. It is thus sufficient to define Φ such that $\Phi(\lambda_{k-1}) \leq \frac{\Phi(\lambda_k)^2}{N+1}, \forall k = 1, \dots, m$, since $\Phi(\top) = 1$. QED

However the above results show the existence of Choquet-integral-based refinements of Sugeno integral orderings, but not their unicity. This lemma implies that Sugeno integral can be refined by a Choquet integral using the same mapping as the one used in (25) for representing $\succeq_L^{lsug}$ by a sum of products, choosing the constant N large enough (as shown in the above proof). Hence the following result:

Theorem 12 *$\vec{\alpha} \succ^{sug} \vec{\beta}$ implies $Ch_{\Phi\circ\kappa}(\Phi(\vec{\alpha})) > Ch_{\Phi\circ\kappa}(\Phi(\vec{\beta}))$, where $\Phi(\vec{\alpha})$ is the tuple with components $\Phi(\alpha_i)$.*

PROOF. Suppose that
$S_\kappa(\vec{\alpha}) = \max_{j=1,\dots,m} \min(\lambda_j, \kappa(A_{\lambda_j})) > S_\kappa(\vec{\beta}) = \max_{j=1,\dots,m} \min(\lambda_j, \kappa(B_{\lambda_j}))$.
Using the above lemma, it follows that $\sum_{j=1}^{m} \Phi(\kappa(A_{\lambda_j})) \cdot (\Phi(\lambda_j) - \Phi(\lambda_{j-1})) > \sum_{j=1}^{m} \Phi(\kappa(B_{\lambda_j})) \cdot \Phi(\lambda_j)$ The latter term is clearly larger than $\sum_{j=1}^{m} \Phi(\kappa(B_{\lambda_j})) \cdot (\Phi(\lambda_j) - \Phi(\lambda_{j-1}))$. QED

Denote this refinement of Sugeno integral ordering as $\succeq^{ch}$. It is clear that :

Corollary 13 *$\succeq^{ch}$ satisfies weak preferential independence (WPI) restricted to co-monotonic tuples.*

The pending question is then whether the latter refinement $\succeq^{ch}$ defined by a Choquet integral coincides with $\succeq_L^{lsug}$. The answer is no in the general case. It may happen that $\vec{\alpha} \succ_L^{lsug} \vec{\beta}$ while $\vec{\beta} \succ^{ch} \vec{\alpha}$. For instance in the above example note that

- $Ch_{\Phi\circ\kappa}(\vec{\alpha}) = \Phi(\top)\cdot\Phi(2) + \Phi(4)\cdot(\Phi(4) - \Phi(2)) + \Phi(4)\cdot(\Phi(5) - \Phi(4)) = \Phi(2) + \Phi(4)\cdot(\Phi(5) - \Phi(2))$
- $Ch_{\Phi\circ\kappa}(\vec{\beta}) = \Phi(2)+\Phi(5)\cdot(\Phi(4)-\Phi(2))+\Phi(2)\cdot(\Phi(5)-\Phi(4))+\Phi(2)\cdot(\Phi(7)-\Phi(5)) = \Phi(2) + \Phi(4)\cdot(\Phi(5) - \Phi(2)) + \Phi(2)(\Phi(7) - \Phi(5)) > Ch_{\Phi\circ\kappa}(\vec{\alpha})$.

while $\vec{\alpha} \succ_L^{lsug} \vec{\beta}$. The point is that the original expression (3) of the Sugeno integral involves redundant pairs of the form $(\lambda_i, \kappa(A_{\lambda_j}))$ and $(\lambda_{j+1}, \kappa(A_{\lambda_{j+1}}))$, with $\kappa(A_{\lambda_j}) = \kappa(A_{\lambda_{j+1}})$ (like pairs $(4,4)$ and $(5,4)$ in the example). The quantity E_L^{lsug} (likewise $\succeq_L^{lsug}$) can be seen as problematic for the following reasons:

- It depends on the number of elements in the scale L we consider. Namely, if we introduce an additional level λ, between λ_j and λ_{j+1}, all other things being the same, $E_L^{lsug}(\vec{\alpha})$ will change (the term $\Phi(\kappa(A_{\lambda_j}))\cdot\Phi(\lambda_j)+\Phi(\kappa(A_{\lambda_{j+1}}))\cdot\Phi(\lambda_{j+1})$ becomes $\Phi(\kappa(A_{\lambda_j}))\cdot\Phi(\lambda_j) + \Phi(\kappa(A_{\lambda_j}))\cdot\Phi(\lambda) + \Phi(\kappa(A_{\lambda_{j+1}}))\cdot\Phi(\lambda_{j+1})$ (as $\kappa(A_{\lambda_j}) = \kappa(A_\lambda)$).
- It counts the contribution of the same set twice (computing $\chi(\kappa(A_{\lambda_{j+1}}))\cdot(\chi(\lambda_j)+\chi(\lambda_{j+1}))$, when $\kappa(A_{\lambda_j}) = \kappa(A_{\lambda_{j+1}})$), while the Choquet integral avoids such a double counting (using a single term $\chi(\kappa(A_{\lambda_{j+1}}))\cdot(\chi(\lambda_{j+1} - \chi(\lambda_{j-1}))$.

So, it seems reasonable to strip matrices $[\vec{\alpha}^\kappa]_L$ from all pairs $(\lambda_i, \kappa(A_{\lambda_j}))$ which never affect the value of Sugeno integral (3). These are pairs where $\kappa(A_{\lambda_j}) = \kappa(A_{\lambda_{j+1}})$ and likewise rows $(\lambda_i, \kappa(A_{\lambda_j}))$ for which $A_{\lambda_j} = \emptyset$. Let $J(\vec{\alpha}) = \{j : A_{\lambda_j} \neq \emptyset, \kappa(A_{\lambda_j}) \neq \kappa(A_{\lambda_{j+1}}), j = 1, m\}$ be the set of non-redundant indices for $\vec{\alpha}$. Sugeno integral can be equivalently expressed as $S_\kappa(\vec{\alpha}) = \max_{j\in J(\vec{\alpha})} \min(\lambda_j, \kappa(A_{\lambda_j}))$. Let $[\vec{\alpha}^\kappa]$ and $[\vec{\beta}^\kappa]$ be the non-redundant matrices so-constructed. They have respectively $|J(\vec{\alpha})|$ and $|J(\vec{\beta})|$ rows, and missing rows of the form $(\bot, \bot)$ can be artificially added to the smallest matrix so as to let them have the same size. A new relation $\succ^{lsug}$ is defined by comparing such matrices $[\vec{\alpha}^\kappa]$ and $[\vec{\beta}^\kappa]$ using $leximax(\succeq_{lmin})$.

In the above example, it comes down to removing rows $(4,4), (7,\bot), (\top,\bot)$ from $[\vec{\alpha}^\kappa]_L$ and $(5,2), (\top,\bot)$ from $[\vec{\beta}^\kappa]_L$. Namely:

$$[\vec{\alpha}^\kappa] = \begin{pmatrix} 5 & 4 \\ 2 & \top \end{pmatrix}; [\vec{\beta}^\kappa] = \begin{pmatrix} 7 & 2 \\ 4 & 5 \\ 2 & \top \end{pmatrix}.$$

Then, with such reduced matrices, $\vec{\beta} \succ^{lsug} \vec{\alpha}$ because $(7,2) \succ_{lmin} (2,\bot)$, and $\vec{\beta} \succ^{ch} \vec{\alpha}$ as well.

Note that this deletion process does not affect the result of the Choquet integral transform, as can be checked on the example, by recomputing $Ch_{\Phi\circ\kappa}(\vec{\alpha})$ and $Ch_{\Phi\circ\kappa}(\vec{\beta})$ on the above matrices. However, even after deletion of redundant pairs as proposed above, both orderings $\succeq^{lsug}$ and $\succeq^{ch}$ do not coincide. In particular, one may have $[\vec{\alpha}^{\kappa}] \sim^{lmax(lmin)} [\vec{\beta}^{\kappa}]$ while $\vec{\beta} \succ^{ch} \vec{\alpha}$. To see it consider tuples of the form $\vec{\alpha} = \lambda A\gamma$ and $\vec{\beta} = \mu B\gamma$ with $\top \geq \lambda > \gamma$ and $\top \geq \mu > \gamma$, $\kappa(A) = \mu$ and $\kappa(B) = \lambda$. The corresponding matrices are :

$$[\vec{\alpha}^{\kappa}] = \begin{pmatrix} \lambda & \mu \\ \gamma & \top \end{pmatrix}; [\vec{\beta}^{\kappa}] = \begin{pmatrix} \mu & \lambda \\ \gamma & \top \end{pmatrix}$$

(e.g. delete row $(7, 2)$ in the previous example matrix $[\vec{\beta}^{\kappa}]$). It is clear that $\vec{\alpha} \sim^{lsug} \vec{\beta}$. However, $Ch_{\Phi\circ\kappa}(\vec{\alpha}) = \Phi(\mu) \cdot (\Phi(\lambda) - \Phi(\gamma)) + \Phi(\gamma)$, while $Ch_{\Phi\circ\kappa}(\vec{\beta}) = \Phi(\lambda) \cdot (\Phi(\mu) - \Phi(\gamma)) + \Phi(\gamma) > Ch_{\Phi\circ\kappa}(\vec{\alpha})$ if and only if $\mu > \lambda$. The issue of whether $\succ^{ch}$ refines $\succ^{lsug}$ or disagrees with it on the big-stepped quantitative scale remains open, even if it is possible to find matrices of real numbers where $\vec{\alpha} \sim^{ch} \vec{\beta}$ while $\vec{\alpha} \sim^{lsug} \vec{\beta}$ does not hold. For instance consider the last example above where matrix $[\vec{\beta}^{\kappa}]$ has first line (μ, δ) such that $\Phi(\delta) = \Phi(\mu)\frac{\Phi(\lambda)-\Phi(\gamma)}{\Phi(\mu)-\Phi(\gamma)}$. It ensures $\vec{\alpha} \sim^{ch} \vec{\beta}$ but $\vec{\alpha} \sim^{lsug} \vec{\beta}$ does not hold since $\delta \neq \lambda$, generally. However the existence of such a super-increasing mapping and a value δ in a finite scale is not guaranteed since if $\Phi(\gamma)$ is very small in front of $\Phi(\lambda)$, then $\Phi(\delta)$ and $\Phi(\lambda)$ should be of a similar order of magnitude even if not equal. This possibility makes sense on a continuous value scale. This is not what is assumed with discrete qualitative scales were successive steps are far away for one another. At this stage one should either prove that $\succ^{ch}$ refines $\succ^{lsug}$ or find a pair of tuples where the two orderings are conflict. This is left for further investigation.

It should be noticed that, when the capacity is a possibility measure Π (resp. a necessity measure N), none of the above refinements recovers the ranking of tuples provided by weighted average $\mathrm{WA}^{+}_{\chi(\pi)}$ (resp. $\mathrm{WA}^{-}_{\chi(\pi)}$). Hence none of them is the generalization of the $\mathrm{WA}^{+}_{\chi(\pi)}$ ranking nor of the $\mathrm{WA}^{-}_{\chi(\pi)}$ ranking. Actually, $\succeq^{lsug}_{L}$ can be viewed as using the *leximax($\geq_{lmin}$)* refinement on the standard expression of Sugeno integral (3) while $\succeq^{(WA)+}$ applies it to an expression involving a possibility distribution (since $S_{\Pi}(\vec{\alpha}) = \max_{i=1,\ldots,n} \min(\alpha_i, \pi_i)$), which is turned into a probability distribution. So, $\succeq^{lsug}_{L}$ and $\succeq^{ch}$ preserve the capacity while $\succeq^{(WA)+}$ refines it. It should be noticed that for Boolean tuples of the form $\top A \bot$ where A is a subset of features, $Ch_{\Phi(\kappa)}(\top A\bot) = \Phi(\kappa(A))$ and $E^{lsug}_{L}(\top A\bot) = \Phi(\kappa(A)) \cdot (\sum_{\lambda>\bot} \Phi(\lambda))$, which shows that, when $\kappa = \Pi$, $\top A\bot \succeq^{ch} \top B\bot \iff \top A\bot \succeq^{lsug}_{L} \top B\bot \iff \Pi(A) \geq \Pi(B)$, while it was shown that $\mathrm{WA}^{+}_{\chi(\pi)}(\top A\bot) \geq \mathrm{WA}^{+}_{\chi(\pi)}(\top B\bot) \iff A \succeq_{\Pi_{lex}} B$. In other terms, $\mathrm{WA}^{+}_{\chi(\pi)}$ purposedly overcomes the drowning effect inherent to the capacity, while neither $\succeq^{lsug}_{L}$ nor $\succeq^{ch}$ do, considering that the capacity supposedly contains all the information about the importance of features.

The formulation (4) of Sugeno integral, i.e., $S_\kappa(\vec{\alpha}) = \max_{i=1,\ldots,n} \min(\alpha_i, \kappa(A_i))$ where $A_i = \{1, 2, \ldots, i\}$, presupposes that the feature ratings α_i are ranked in decreasing order. It leads to a different refinement. In this case the maximum is performed over features, not levels in the scale L. We can still use the transformation refining the prioritized maximum (with n features, instead of the m levels of the scale L). If the α_i's are totally ordered ($\alpha_1 > ... > \alpha_n$), the following expression is obtained:

$$E\mathcal{F}^{lsug}(\vec{\alpha}) = \sum_{i=1}^{n} \chi(\alpha_i) \cdot \chi(\kappa(A_i)). \tag{26}$$

When some α_i's are equal, it is no longer well-defined and the following equivalent formulation is a natural way of extending it, rewriting (4) under the form $S_\kappa(\vec{\alpha}) = \max_{i=1,\ldots,n} \min(\alpha_i, \kappa(A_{\alpha_i}))$:

$$E\mathcal{F}^{lsug}(\vec{\alpha}) = \sum_{i=1}^{n} |\{j : \alpha_i = \lambda\}| \cdot \chi(\lambda) \cdot \chi(\kappa(A_\lambda)). \tag{27}$$

Let $\succeq^{lf}$ be the preference ordering induced by $E\mathcal{F}^{lsug}$. It is a refinement of $\succeq^{sug}$. $E\mathcal{F}^{lsug}$ also refines the ordering encoded by κ. Indeed, $E\mathcal{F}^{lsug}(\top A \bot)$ is proportional to $|A|.\chi(\kappa(A))$. So, $A \sim^{lf} B \iff (\kappa(A) = \kappa(B)$ and $|A| = |B|)$, and $A \succ^{lf} B \iff \kappa(A) > \kappa(B)$ or $(\kappa(A) = \kappa(B)$ and $|A| > |B|)$. We get a refinement of the κ-ordering of sets by their cardinality. It is clear that comparing tuples by means of $\succeq^{lf}$ comes down to comparing matrices $[\vec{\alpha}^\kappa]_f$ with n rows $(\alpha_i, \kappa(A_{\alpha_i}))$ by means of the *leximax($\succeq_{lmin}$)* ordering.

Moreover, it turns out that $\succeq^{lsug}$ and $\succeq^{lf}$ are not comparable: $\succeq^{lf}$ is not a refinement of $\succeq^{lsug}$, nor is $\succeq^{lsug}$ a refinement of $\succeq^{lf}$, as shown by the following counterexample.

Example 14
Let $\kappa = \Pi$ be a possibility measure on three features denoted by $1, 2, 3$. Consider two objects $\vec{\alpha}$ and $\vec{\beta}$, and the following possibility distribution π:

$Features$	$\vec{\alpha}$	$\vec{\beta}$	π
1	8	8	8
2	6	8	2
3	7	6	$\top$

Then $S_\Pi(\vec{\alpha}) = S_\Pi(\vec{\beta}) = 8$. The tuples of pairs $(\lambda_j, \Pi(A_{\lambda_j})), j > 0$ are:
$(6, \top), (7, \top), (8, 8), (\top, \bot)$ *for $\vec{\alpha}$ and* $(6, \top), (7, 8), (8, 8), (\top, \bot)$ *for $\vec{\beta}$. The non-redundant pairs are* $(7, \top), (8, 8)$ *for $\vec{\alpha}$ and* $(6, \top), (8, 8)$ *for $\vec{\beta}$. Hence $\vec{\alpha} \succ^{lsug}$*

$\vec{\beta}$. Now use pairs $(\alpha_i, \Pi(A_{\alpha_i})), i = 1,2,3$. We get $(8,8),(7,\top),(6,\top)$ for $\vec{\alpha}$ and $(8,8),(8,8),(6,\top)$ for $\vec{\beta}$. Hence $\vec{\beta} \succ^{lf} \vec{\alpha}$. Opposite rankings are found.

On this example, the choice of $\succeq^{lsug}$ is closer to the intuition than the one of $\succeq^{lf}$, because $\vec{\beta}$ is better than $\vec{\alpha}$ only on one feature of low importance (it could even be a null feature), while $\vec{\alpha}$ is better than $\vec{\beta}$ on each important feature. The questionable point about $\succeq^{lf}$ is that it again involves redundant information. Namely, it is clear that $\kappa(A_{\alpha_i}) = \kappa(A_{\alpha_k})$ may occur when $\alpha_i \neq \alpha_k$ and moreover, identical rows appear when $\alpha_i = \alpha_k$. The following algorithm constructs reduced non-redundant matrices $[\vec{\alpha}^{\kappa}]^*$:

Algorithm: Constructing $[\vec{\alpha}^{\kappa}]^*$ from $[\vec{\alpha}^{\kappa}]_f$

(1) Rank features such that $i < k$ implies $\alpha_i \geq \alpha_k$;
(2) For $i = 2, \ldots, n$ do :
- If $\alpha_i = \alpha_{i-1}$, then delete row $i-1$ (it is the same as row i),
- else if $\kappa(A_{\alpha_i}) = \kappa(A_{\alpha_{i-1}})$ then delete row i.

Matrices $[\vec{\alpha}^{\kappa}]_f$ and $[\vec{\beta}^{\kappa}]_f$ in the example again contain redundant rows. For instance, $\Pi(\{1,3\}) = \top$, so that line $(6,\top)$ is redundant in $[\vec{\alpha}^{\kappa}]_f$, while $(8,8)$ is once too many in $[\vec{\beta}^{\kappa}]_f$. Now the remaining matrices are $[\vec{\alpha}^{\kappa}]^* \succ_{lmax(lmin)} [\vec{\beta}^{\kappa}]^*$. These matrices $[\vec{\alpha}^{\kappa}]^*$ and $[\vec{\beta}^{\kappa}]^*$ are the same as $[\vec{\alpha}^{\kappa}]$ and $[\vec{\beta}^{\kappa}]$ after deletion of redundant rows. This is no coincidence.

Proposition 15 $[\vec{\alpha}^{\kappa}]^* = [\vec{\alpha}^{\kappa}]$.

PROOF. To form matrix $[\vec{\alpha}^{\kappa}]$, all rows $(\lambda, \kappa(A_\lambda))$ such that $\kappa(A_{\alpha_i}) = \kappa(A_\lambda), \lambda < \alpha_i$ and $A_\lambda = \emptyset$ are deleted from $[\vec{\alpha}^{\kappa}]_L$. Remaining rows are thus of the form $(\alpha_i, \kappa(A_{\alpha_i}))$ for some feature i and, by construction, all remaining α_i's are distinct and such that if $\alpha_k < \alpha_i$ then $\kappa(A_{\alpha_k}) > \kappa(A_{\alpha_i})$. They belong to $[\vec{\alpha}^{\kappa}]_f$. The above algorithm applied to $[\vec{\alpha}^{\kappa}]_f$ first keeps only one row among the identical ones. The matrix obtained at this point still contains all rows in $[\vec{\alpha}^{\kappa}]$. Next, all rows $(\alpha_k, \kappa(A_{\alpha_k}))$ such that there is a feature i, for which $\alpha_k > \alpha_i$ and $\kappa(A_{\alpha_k}) = \kappa(A_{\alpha_i})$ are deleted. So $[\vec{\alpha}^{\kappa}]^*$ contains all rows of $[\vec{\alpha}^{\kappa}]$. But converse is true as well since if not then the corresponding row of $[\vec{\alpha}^{\kappa}]^*$ not in $[\vec{\alpha}^{\kappa}]$, say $(\alpha_k, \kappa(A_{\alpha_k}))$ would be such that if $\alpha_k < \alpha_i$ then $\kappa(A_{\alpha_k}) > \kappa(A_{\alpha_i})$. By construction, this row is present in $[\vec{\alpha}^{\kappa}]_L$ and would never be deleted when constructing $[\vec{\alpha}^{\kappa}]$. So this case is impossible.

So the relation $\succeq^{lsug}$ is the same whether we use features or steps in the value scale, after deleting redundant rows.

6 Refinement of Sugeno integral based on qualitative Moebius transforms

The information contained in a capacity κ can be expressed in a non redundant way by means of its qualitative Moebius transform; it is another set-function $\kappa_\#$ defined in [22] by $\kappa_\#(A) = \kappa(A)$ if $\kappa(A) > \max\{\kappa(B) : B \subset A\}$ and $\kappa_\#(A) = \bot$ otherwise. It is clear that $\kappa_\#$ contains the minimal information to reconstruct κ as:

$$\kappa(A) = \max_{B \subseteq A} \kappa_\#(B). \tag{28}$$

Function $\kappa_\#$ plays the role of a "qualitative" basic probability assignment instrumental in Shafer's theory of evidence and obtained via Moebius transform. The subsets B that receive a positive support in terms of $\kappa_\#$ play the same role for κ as the focal elements in Shafer's theory of evidence[36]: they are the primitive items of knowledge. Equation (28) appears as the qualitative counterpart of the definition of a belief function (even if κ may fail to satisfy axiom BEL) or an inner measure. The set-function $\kappa_\#$ can also be viewed as a possibilistic mass assignment, a possibility distribution over the power set $2^{\mathcal{F}}$. Indeed, (28) is also a generalization of the definition of the degree of possibility of a set in terms of a possibility distribution on $\mathcal{F}$. Indeed, the function $\Pi_\#(E) = \bot$ as soon as E is not a singleton, and $\Pi_\#(\{i\}) = \pi_i, \forall i \in \mathcal{F}$.

In the third expression (5) of Sugeno integral, the set-function κ can be replaced without loss of information by $\kappa_\#$. We now get another expression of Sugeno integral, maximizing over the family $\mathcal{P}_\#(\mathcal{F})$ of subsets of features A with $\kappa_\#(A) \neq \bot$:

$$S_\kappa(\vec{\alpha}) = \max_{A \in \mathcal{P}_\#(\mathcal{F})} \min(\kappa_\#(A), \alpha_A), \tag{29}$$

where $\alpha_A = \min_{i \in A} \alpha_i$. The above expression of Sugeno integral has the standard maxmin form viewing $\kappa_\#$ as a possibility distribution over the power set of $\mathcal{F}$, since $\max_{A \subseteq \mathcal{F}} \kappa_\#(A) = \top$. Moreover the use of $\kappa_\#$ instead of κ avoids a lot of potential redundant terms that appear in the other formulations and created difficulties when refining Sugeno integral. The above expression is optimally non-redundant in this sense. Moreover, the form (29) is very similar to the optimistic possibilistic criterion W_π^+ because $\kappa_\#$ is an extension of the possibility distribution explicitly appearing in (6). Hence it is tempting to apply the super-increasing transform χ to (29). Doing so changes a maxmin form into a sum of products :

$$E_\#^{lsug}(\vec{\alpha}) = \sum_{A \in 2^{\mathcal{F}}} \chi(\alpha_A) \cdot \chi^*(\kappa_\#(A)).$$

Ranking tuples by $E_\#^{lsug}(\vec{\alpha})$ comes down to a *leximax($\geq_{lmin}$)* comparison of $(2^n \times 2)$ matrices with rows of the form $(\kappa_\#(A), \alpha_A)$. Notice that here the referential is not $\mathcal{F}$ nor L, but $2^{\mathcal{F}}$ and $\kappa_\#(\emptyset) = \bot$; so, in the definition of χ, we set $N = 2^{|\mathcal{F}|}$. Function χ^* is the normalization of χ in such a way that $\sum_{A \in 2^{\mathcal{F}}} \chi^*(\kappa_\#(A)) = 1$.

So, the function $m_\# : 2^{\mathcal{F}} \mapsto [0,1]$:

$$m_\#(A) = \chi^*(\kappa_\#(A))$$

is a mass assignment in the sense of Shafer[36]; in particular, $m_\#(\emptyset) = 0$. Note that $m_\#$ is a big-stepped mass function in the sense that:

$$m_\#(A) > 0 \implies m_\#(A) > \sum_{B \subseteq \mathcal{F},\ \text{s. t.} m_\#(B) < m_\#(A)} m_\#(B).$$

A consequence of this property is that if $\kappa_\#(A) > \bot$ then $m_\#(A) > \max_{B \subset A} m_\#(B)$ since when $A \subset B$ and $\kappa_\#(A) > \bot, \kappa_\#(B) > \bot$, then $\kappa_\#(A) < \kappa_\#(B)$. Now, it is easy to show that $\chi(\alpha_A) = \chi(\min_{i \in A} \alpha_i) = \min_{i \in A} \chi(\alpha_i)$. Then:

$$E_\#^{lsug}(\vec{\alpha}) = \sum_{A \subseteq \mathcal{F}} m_\#(A) \cdot \min_{s \in A} \chi(\alpha_i)$$

is a Choquet integral with respect to a belief function which refines the original Sugeno integral, noticing that the expression of a Choquet integral of a tuple of ratings in terms of the Moebius transform m_v of a numerical capacity v is of the form

$$Ch_v(\vec{\alpha}) = \sum_{A \subseteq \mathcal{F}} m_v(A) \cdot \min_{s \in A} \chi(\alpha_i).$$

Letting $Bel_\#(A) = \sum_{B \subseteq A} m_\#(B)$ be the belief function induced by $m_\#$, the Choquet integral $E_\#^{lsug}$ also reads :

$$E_\#^{lsug}(\vec{\alpha}) = Ch_{Bel_\#}(\vec{\alpha}) = \sum_{j=1}^{m} Bel_\#(A_{\lambda_j}) \cdot (\chi(\lambda_j) - \chi(\lambda_{j-1})).$$

This shows that any Sugeno integral can be refined by a Choquet integral *w.r.t a belief function*. In summary:

Theorem 16 *For any Sugeno integral S_κ, there exist a Choquet integral Ch_{Bel} with respect to a belief function Bel and a utility function u such that:*

$$S_\kappa(\vec{\alpha}) > S_\kappa(\vec{\beta}) \implies Ch_{Bel}(u(\vec{\alpha})) > Ch_{Bel}(u(\vec{\beta})).$$

Contrary to the solution obtained in the previous section, the capacity κ is generally not preserved under the present transformation. The resulting Choquet integral is always pessimistic, and sometimes not more discriminant than the original criterion.

Two particular cases are interesting to consider:

- If κ is a possibility measure Π, then $\kappa_\#(A)$ is positive on singletons of positive possibility only. In other words, $\kappa_\#$ coincides with the possibility distribution

of Π and the mass function obtained by the super-increasing transformation is a probability assignment on $\mathcal{F}$. Then the Moebius expression of Sugeno integral coincides with the expression of the prioritized maximum. So $m_{\#}$ is a regular big-stepped probability function and Choquet integral reduces to a regular weighted average. We retrieve the maximal refinement $\mathrm{WA}^{+}_{\chi(\pi)}$ of the prioritized maximum presented in Section 4.2.
- On the contrary if κ is a necessity measure N, $Ch_{Bel_{\#}}$ does not collapse at all with the pessimistic expected utility $\mathrm{WA}^{-}_{\chi(\pi)}$. Indeed, if κ is a necessity measure N, $\kappa_{\#}(A)$ is positive on alpha-cuts of the possibility distribution only. So the mass assignment $m_{\#}$ is positive on the nested family of sets A_i, and the belief function $Bel_{\#}$ is a necessity measure ordinally equivalent to the original one. In this case, the resulting Choquet integral is one with respect to a necessity measure. Only the "max-min" framing of Sugeno integral has been turned into a "sum-product" framing: the transformation has preserved the nature of the original capacity and the capacity-preserving refinement $\succeq^{ch}$ identified in first part of Section 5 is retrieved.

7 Refining capacities

The above results motivate an investigation into the conditions under which a capacity can be refined. When utility tuples are of the zero-one type, capacity-preserving refinements are totally useless since Sugeno integral then coincides with $\kappa(A)$ for some set A. In some situations, the full-fledged refinement of a Sugeno integral should refine the capacity itself, as shown in the case of prioritized minimum and maximum. In this section, some preliminary definitions and results are presented to this aim. The ultimate goal is to get as close as possible to enforcing strict Pareto dominance (Axiom SPAR). Examples of likelihood ordering achieving this goal are discrimin and leximin lexicographic refinements of possibility measures [15,12]. We show how to extend these types of refinements to capacities.

7.1 Coping with the strict Pareto principle

Axiom SPAR in the Boolean setting reads: $\forall A, B$ disjoint: B not null implies $\kappa(A \cup B) > \kappa(A)$. Indeed, as the set A, viewed as a tuple, is Pareto-dominated by $A \cup B$, the latter should be more important than B. When there are no non-empty null sets, it comes down to requiring $\kappa(A) > \kappa(B)$ whenever the strict inclusion $B \subset A$ holds. A weaker requirement is as follows

S: $\forall A, B$ disjoint: $\kappa(B) > \bot \implies \kappa(A \cup B) > \kappa(A)$ (Strictness)

since $\kappa(B) > \bot$ implies that B is not null. A capacity exhibits limited discrimination power if there exist two **disjoint** sets A, B such that $\kappa(B) > \bot$ and $\kappa(A \cup B) = \kappa(A)$. So, we aim at defining for any original capacity κ, another capacity κ' refining the ordering induced by the former, and satisfying axiom S, at least. This axiom seems to have been first proposed by Z. Wang[42] under the name *converse null-additivity*. The converse of axiom **S** is :

NA : $\forall A, B$ disjoint: $\kappa(A \cup B) > \kappa(A) \implies \kappa(B) > \bot$

This axiom also writes $\kappa'(B) = \bot \implies \kappa'(A \cup B) = \kappa'(A)$ and is called *null-additivity* by Z. Wang [41] (see also Pap[31]). When NA or S holds with equivalence, the corresponding property is denoted by NAS.

Proposition 17 *For capacities, NAS implies SPAR*

PROOF. Assume that $\forall A, B$ disjoint: $\kappa(A \cup B) > \kappa(A) \iff \kappa(B) > \bot$. If $\kappa(B) > \bot$ then B is not null and $\kappa(A \cup B) > \kappa(A)$. If $\kappa(B) = \bot$, then $\forall A$ disjoint from B, $\kappa(A \cup B) = \kappa(A)$, hence B is null, and SPAR does not apply.

Notice that axiom **S** is a weak form of a property of the ordering induced by Shafer's belief functions [44], namely:

BEL: $\forall A, B, C$ disjoint sets: $\kappa(A \cup B) > \kappa(A) \implies \kappa(A \cup B \cup C) > \kappa(A \cup C)$

In fact likelihood relations that are monotonic with inclusion and obey this property can always be represented by belief functions Bel [44]. In particular, necessity measures satisfy it. Axiom **S** is retrieved when assuming $A = \emptyset$. The converse implication is a property of the ordering induced by Shafer's plausibility functions $Pl(A) = 1 - Bel(A^c)$, hence also satisfied by possibility measures (even if A and B are not disjoint):

PL: $\forall A, B, C$ disjoint sets: $\kappa(A \cup B \cup C) > \kappa(A \cup C) \implies \kappa(A \cup B) > \kappa(A)$.

Note that **BEL** and **PL** are just slight reinforcements of the property $\forall A, B, C$ disjoint sets: $\kappa(A \cup B) > \kappa(A) \implies \kappa(A \cup B \cup C) \geq \kappa(A \cup C)$, which trivially holds whenever κ is monotonic with inclusion. Joining **BEL** and PL, the following property stronger than NAS can be considered:

BELPL: $\forall A, B, C$ disjoint : $\kappa(A \cup B) > \kappa(A) \iff \kappa(A \cup B \cup C) > \kappa(A \cup C)$.

Axiom **BELPL** is a weak form of the classical *preadditivity axiom*, denoted by **PRAD**, restricting preferential independence to sets, that underlies comparative probabilities.

PRAD : $\forall A, B$, and C disjoint from $A \cup B$: $\kappa(B) \geq \kappa(A) \iff \kappa(B \cup C) \geq \kappa(A \cup C)$.

PRAD implies **BELPL** that implies **SPAR**. The converse is not true. Remembering that due to the strong violation of independence PI, not all capacities can be refined by a comparative probability; capacities whose induced orders could be refined in this way are the ones such that

$$\forall A, B, \text{and } C \text{ disjoint from } A \cup B\text{: } \kappa(B) \geq \kappa(A) \implies \kappa(B \cup C) \geq \kappa(A \cup C).$$

These functions are studied by Dubois[11] and Chateauneuf [9]. They are decomposable measures in the sense that there exists an operation $\star$ on L such that if $A \cap B = \emptyset$ then $\kappa(A \cup B) = \kappa(A) \star \kappa(B)$.

A straightforward way to construct an ordering on events satisfying **BELPL**, hence Pareto-dominance, in the absence of null elements, is to refine the ranking induced by κ by means of the inclusion relation:

$$B \succ_{\kappa}^{\subset} A \iff \kappa(B) > \kappa(A) \text{ or } A \subset B.$$

$\succ_{\kappa}^{\subset}$ is obviously a transitive but partial ordering. Basically, each equivalence class $\mathcal{C}_\kappa$ of equally important sets in the sense of κ is internally partially ordered by the relation $\subset$. The partial ordering $\succ_{\kappa}^{\subset}$ restricted to each $\mathcal{C}_\kappa$ can be embedded into a weak order, for instance considering cardinality (as obtained earlier in the refinement of Sugeno integral based on feature ratings of Section 5.2):

$$B \succ_{\kappa}^{card} A \iff \kappa(B) > \kappa(A) \text{ or } (\kappa(B) = \kappa(A) \text{ and } |A| < |B|.)$$

This relation can be represented by a capacity κ^{card} refining κ (e.g. $\kappa^{card}(A) = |A| \cdot \chi(\kappa(A))$, as already done in Section 5.2). For any capacity κ, the relations $\succ_{\kappa}^{\subset}$ and $\succ_{\kappa}^{card}$ satisfy axiom **BELPL**. Indeed, by construction, $B \succ_{\kappa}^{\subset} \emptyset$ always holds if $B \neq \emptyset$, and so does $A \cup B \succ_{\kappa}^{\subset} A$. The capacity κ^{card} is one among other possible refinements of κ. Since it satisfies **BELPL**, we get a ranking that can be represented by both a plausibility and a belief function. The study of such measures of uncertainty characterized by Axiom **BELPL** is out of the scope of the present paper.

7.2 Discri- and Lexi- refinements of capacities

A natural refinement of a possibility measure Π is called discrimax refinement [15], and it is defined by $A \succ_{dmax}^{\Pi} B \iff \Pi(A \setminus B) > \Pi(B \setminus A)$. The discrimax refinement of a capacity κ is then defined likewise using the qualitative Moebius transform $\kappa_\#$ of κ introduced in Section 6 :

$$A \succ_{dcap}^{\kappa} B \iff \max_{E, E \subseteq A, E \not\subseteq B} \kappa_\#(E) > \max_{E, E \subseteq B, E \not\subseteq A} \kappa_\#(E). \tag{30}$$

In this definition, all subsets common to A and B play the same role in the expressions of $\kappa(A)$ and $\kappa(B)$ and are cancelled since they cannot discriminate between them. It is easy to check that this partial order refines the ranking induced by κ, since if $\kappa(A) > \kappa(B)$, there is a set $E \subseteq A$ such that $\kappa_\#(E) > \max_{F \subseteq B} \kappa_\#(F)$. Moreover the *dcap* relation is also of the **BEL** type:

Proposition 18 *Relation* $\succ^\kappa_{dcap}$ *satisfies* **BEL***.*

PROOF. Suppose $A \cap B = \emptyset$. Then suppose $A \cup B \succ_{dcap} A$. Since $\{E : E \subseteq A, E \not\subseteq A \cup B\}) = \emptyset$, it comes down to $\max_{E \subseteq A \cup B; E \not\subseteq A} \kappa_\#(E) > \bot$. So there is $E^* \subseteq A \cup B; E^* \not\subseteq A$ with $\kappa_\#(E^*) > \bot$. Now $\{E : E \subseteq A \cup B, E \not\subseteq A \cup B \cup C) = \emptyset$ again. But clearly $E^* \in \{E : E \subseteq A \cup B \cup C, E \not\subseteq A \cup C\}$, so $\max_{E \subseteq A \cup B \cup C; E \not\subseteq A \cup C} \kappa_\#(E) > \bot$. So $A \cup B \cup C \succ_{dcap} A \cup C$.

Note that when κ is a possibility measure, then $\succ^\Pi_{dcap}$ satisfies even axiom PRAD (hence BELPL), and is self-conjugate ($A \succ^\Pi_{dcap} B \iff B^c \succ^\Pi_{dcap} A^c$). It refines the conjugate necessity measure as well. However, in general $\succ^\kappa_{dcap}$ is not self-conjugate, and is generally not of the **PL** type as the existence of $E^* \subseteq A \cup B \cup C$ while $E^* \not\subseteq A \cup C$ does not ensure that $E^* \subseteq A \cup B$ and $E^* \not\subseteq A$ (for instance if $E^* = A \cup B \cup D$ with $D \subseteq C$ not empty).

The lexicographic refinement $\succeq^\kappa_{lcap}$ of $\succeq^\kappa_{dcap}$ is a ranking defined likewise:

$$A \succeq^\kappa_{lcap} B \iff \vec{A} \succeq_{lmax} \vec{B}, \tag{31}$$

where $\vec{A}$ (resp. $\vec{B}$) is the tuple with size $2^{\mathcal{F}}$ containing all values $\kappa_\#(E)$, $\forall E \subseteq A$ (resp. $\forall E \subseteq B$), and $\bot$ if $E \not\subseteq A$. It is clear that if κ is a possibility measure, then $\succeq^\kappa_{lcap}$ boils down to the leximax possibility ordering $\succeq_{\Pi_{Lex}}$, encountered in previous sections. It is possible to construct a capacity κ_{lmax} on a refined ordinal scale Λ encoding this refinement. Using a super-increasing transformation, it is possible to turn it into a big-stepped belief function with mass function $\chi(\kappa_\#(\cdot))$, as shown in the previous section. So, $\succeq^\kappa_{lcap}$ is also a complete preordering of the **BEL** type.

However the above refinements are ineffective on necessity measures. Indeed, consider a possibility distribution π such that $\pi_1 > \cdots > \pi_n \geq \pi_{n+1} = \bot$. Then let $E_i = \{1, \ldots, i\}$. The qualitative Moebius transform of a necessity measure N based on π is of the form $N_\#(A) = \nu(\pi_{i+1})$ if $A = E_i$ and $\bot$ otherwise. Moreover, $N(A) = \max_{E_i \subseteq A} N_\#(E_i)$. Suppose $N(A) = N(B)$. It means that $N(A) = N(B) = N_\#(E_i)$ for some $E_i \subseteq A \cap B$. But clearly, $\{E : N_\#(E) > \bot, E \subseteq A, E \not\subseteq B\} = \emptyset$ since if not, then it is E_j for some $j > i$, but then it would mean $N(A) = \nu(\pi_{j+1}) > \nu(\pi_{i+1})$. So the sets $E \subseteq A$ while $E \not\subseteq B$ are such that $N_\#(E) = \bot$, and likewise exchanging A and B. So, none of the relations $\succeq^N_{dcap}$ nor $\succeq^N_{lcap}$ can refine a necessity measure. As a consequence, relation $\succ^N_{dcap}$ does not refine $\succ_N$.

In order to directly refine a necessity measure, another qualitative representation of a capacity κ, a set function denoted by $\kappa^{\#}$, can be used, the knowledge of which is enough to reconstruct the capacity:

$\kappa^{\#}(A) = \kappa(A)$ if $\kappa(A) < \min\{\kappa(F) : A \subset F\}$

and $\kappa^{\#}(A) = \top$ otherwise. The original capacity is then retrieved as $\kappa(A) = \min_{A \subseteq F} \kappa^{\#}(F)$, which reminds of outer measures. Function $\kappa^{\#}$ can be called outer qualitative mass function of κ, as $\kappa(A)$ is recovered from $\kappa^{\#}$ via weights assigned to supersets of set A, while $\kappa_{\#}$ stands for an inner qualitative mass function. So we could consider refining the κ ordering as follows:

$$A \succ_{\kappa}^{dcap} B \iff \min_{E: A \subseteq E, B \not\subseteq E} \kappa^{\#}(E) > \min_{E: B \subseteq E, A \not\subseteq E} \kappa^{\#}(E). \qquad (32)$$

Proposition 19 *Relation* $\succ_{\kappa}^{dcap}$ *satisfies* **PL**.

PROOF. Suppose A, B, C disjoint. Then suppose $A \cup B \cup C \succ_{\kappa}^{dcap} A \cup C$. Since $\{E : A \cup B \cup C \subseteq E, A \cup C \not\subseteq E\} = \emptyset$, it comes down to the inequality $\min_{A \cup C \subseteq E; A \cup B \cup C \not\subseteq E} \kappa_{\#}(E) < \top$. So there is $E^*, A \cup C \subseteq E^*; A \cup B \cup C \not\subseteq E^*$ with $\kappa^{\#}(E^*) < \top$. Now $\{E : A \cup B \subseteq E, A \not\subseteq E\} = \emptyset$ again. But clearly E^* satisfies $A \subseteq E^*$ and $A \cup B \not\subseteq E^*$, so $\min_{E A \subseteq E; A \cup B \not\subseteq E} \kappa_{\#}(E) < \top$. So $A \cup B \succ_{\kappa}^{dcap} A$.

However, this relation is generally not of the **BEL** type. Interestingly, the inner qualitative mass function $\kappa^c_{\#}$ of κ^c is related to the outer qualitative mass function $\kappa^{\#}$:

$$\kappa^{\#}(A) = \nu(\kappa^c_{\#}(A^c)).$$

Indeed, $\kappa(A) < \min\{\kappa(F) : A \subset F\}$ also writes $\kappa^c(A^c) > \max\{\kappa^c(F^c) : F^c \subset A^c\}$. For instance, $N^{\#}(E) \neq \top$ only if $E = \mathcal{F} \setminus \{i\}$ for some $i \in \mathcal{F}$, and then $N^{\#}(\mathcal{F} \setminus \{i\}) = \nu(\pi_i)$. As a consequence,

$$A \succ_{\kappa}^{dcap} B \iff \min_{E: A \subseteq E, B \not\subseteq E} \nu(\kappa^c_{\#}(E^c)) > \min_{E: B \subseteq E, A \not\subseteq E} \nu(\kappa^c_{\#}(E^c)).$$

But $E \subseteq B, E \not\subseteq A$ also write $B^c \subseteq E^c, A^c \not\subseteq E^c$, so,

$$A \succ_{\kappa}^{dcap} B \iff \max_{E: E \subseteq A^c, E \not\subseteq B^c} \kappa^c_{\#}(E) < \max_{E: E \subseteq B^c E \not\subseteq A^c} \kappa^c_{\#}(E) \iff B^c \succ_{dcap}^{\kappa^c} A^c.$$

If κ is a necessity measure, then we get $A \succ_{N}^{dcap} B \iff B^c \succ_{dcap}^{\Pi} A^c$ which is equivalent to $A \succ_{dcap}^{\Pi} B$. However this is not true in general so that the four relations made of $\succ_{\kappa}^{dcap}$, $\succ_{\kappa}^{dcap}$ and their conjugates will differ from one another.

Relation	is the same as	Type	Case $\kappa = \Pi$
$A \succ_{dcap}^{\kappa} B$	$B^c \succ_{\kappa^c}^{dcap} A^c$	BEL	$\succ_{dmax}^{\Pi}$
$A \succ_{dcap}^{\kappa^c} B$	$B^c \succ_{\kappa}^{dcap} A^c$	BEL	does not refine N
$B^c \succ_{dcap}^{\kappa} A^c$	$A \succ_{\kappa^c}^{dcap} B$	PL	$\succ_{dmax}^{\Pi}$
$B^c \succ_{dcap}^{\kappa^c} A^c$	$A \succ_{\kappa}^{dcap} B$	PL	does not refine Π

Table 1
Comparison of refined capacities

Altogether this study lays bare two possible lines of refinements of a capacity κ and its conjugate, using the outer and inner Moebius transforms. There is no unique capacity which refines both a prescribed capacity and its conjugate, except for special cases like probability measures. In the case of possibility measures, $\Pi(A) > \Pi(B)$ does imply $\Pi(B^c) \geq \Pi(A^c)$ which allows for such a conjoint refinement : $\succ_{dcap}^{\kappa}$ and $\succ_{\kappa^c}^{dcap}$ coincide when κ (resp. κ^c) is a possibility (resp. a necessity) measure. However in the general case, we may have $\kappa(A) > \kappa(B)$ and $\kappa^c(B) > \kappa^c(A)$. Moreover, $\succ_{dcap}^{\kappa}$ and $\succ_{\kappa}^{dcap}$ may produce conflicting rankings (if $\kappa(A) = \kappa(B)$ one may get $A \succ_{dcap}^{\kappa} B$ and $B \succ_{\kappa}^{dcap} A$ as each ordering is obtained from distinct sets of values). So one may get up to four refinements, two obeying axiom **BEL**, the others obeying the axiom **PL** (see Table 1). A complete comparison of these variants is a matter of further research.

8 Conclusion

This paper tries to bridge the gap between qualitative and quantitative criteria for decision-making with a view to increasing the discrimination power of the latter, especially to respect Pareto-dominance in the strict sense. We provide preliminary results when the weight function, expressing the importance of features or the likelihood of states, is encoded by a qualitative capacity or fuzzy measure and the aggregation is performed by means of a Sugeno integral. The paper shows how to refine weak orders induced by Sugeno integral by mean of lexicographic schemes extending leximin and leximax. It also shows the existence of Choquet integrals that characterize refined rankings. Two approaches have been proposed: one that preserves the capacity at work in Sugeno integral, the other focusing on the basic information sufficient to generate the capacity. Moreover, we show that the issue of addressing the lack of discrimination due to the max-min form of Sugeno integral is distinct from the problem of enhancing the discrimination power of the capacity itself. The possibility of refining the rankings of decisions induced by Sugeno integral enhances its applicability in identification problems where the underlying capacity must be learned from preference data containing more classes than the qualitative value scale can allow.

Note that these results rely on the finiteness of the setting. Extending these results to infinite spaces looks hopeless because lexicographic schemes cannot be simulated by continuous operations, generally. Several questions remain open.

(1) More work is needed to compare the the $leximax(\succeq_{lmin})$ ranking refining Sugeno integral in its standard form with the Choquet integrals refinements.
(2) A detailed study of lexicographic refinements of a capacity is needed. The refined capacity can be used so as to improve capacity-preserving refinements of Sugeno integral, in case of a tie with respect to stochastic dominance. The qualitative Moebius transform approach could also benefit from the obtained results on capacity refinement, especially the use of outer qualitative mass functions looks promising to fully retrieve the canonical refinement of the prioritized minimum in the qualitative Moebius transform approach to the refinement of Sugeno integral.
(3) Lastly, one may consider finding a system of axioms characterizing the refined decision rules proposed here, by putting together Savage axioms and Sugeno integral axioms in some way.

Acknowledgements : The authors are extremely grateful to referees for their careful check of the manuscript, especially one of them who pointed out some technical mistakes, hopefully removed in this version.

References

[1] K. Arrow and L. Hurwicz. An optimality criterion for decision-making under ignorance. In C.F. Carter and J.L. Ford, editors, *Uncertainty and Expectations in Economics*. Basil Blacwell and Mott Ltd, 1972.

[2] K. J. Arrow. *Social Choice and Individual Values*. Cowles Foundations and Wiley - New-York, 1951.

[3] S. Benferhat, D. Dubois, and H. Prade. Possibilistic and standard probabilistic semantics of conditional knowledge bases. *Journal of Logic and Computation*, 9:873–895, 1999.

[4] C. Boutilier. Toward a logic for qualitative decision theory. In *Proceedings of KR'94*, pages 75–86, 1994.

[5] D. Bouyssou, T. Marchant, and M. Pirlot. A conjoint measurement approach to the discrete Sugeno integral. In S. Brams, W. Gehrlein, and F. Roberts, editors, *The Mathematics of Preference, Choice and Order (Essays in Honor of Peter C. Fishburn)*, pages 1–25. 2008.

[6] R. Brafman and M. Tennenholtz. On the foundations of qualitative decision theory. In *Proceedings of AAAI'96*, pages 1291–1296, 1996.

[7] R. Brafman and M. Tennenholtz. On the axiomatization of qualitative decision theory. In *Proceedings of AAAI'97*, pages 76–81, 1997.

[8] R. Brafman and M. Tennenholtz. An axiomatic treatment of three qualitative decision criteria. *Journal of the ACM*, 47(3):452–482, 2000.

[9] A. Chateauneuf. Decomposable capacities, distorted probabilities and concave capacities. *Math. Soc. Sciences*, 31:19–37, 1996.

[10] R. Deschamps and L. Gevers. Leximin and utilitarian rules: A joint characterization. *Journal of Economic Theory*, 17:143–163, 1978.

[11] D. Dubois. Belief structures, possibility theory and decomposable confidence measures on finite sets. *Computers and Artificial Intelligence*, 5(5):403–416, 1986.

[12] D. Dubois and H. Fargier. An axiomatic framework for order of magnitude confidence relations. In *Proceedings of UAI'04*, Banff, CA, 2004.

[13] D. Dubois and H. Fargier. Lexicographic refiniments of Sugeno integrals. In K. Mellouli, editor, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty (European Conference ECSQARU 2007)*, volume 4724 of *Lecture Notes in Artificial Intelligence*, pages 611–622. Springer, 2007.

[14] D. Dubois, H. Fargier, and H. Prade. Refinements of the maximin approach to decision-making in a fuzzy environment. *Fuzzy Sets Systems*, 81(1):103–122, 1996.

[15] D. Dubois, H. Fargier, and H. Prade. Possibilistic likelihood relations. In *Proceedings of IPMU'98*, pages 1196–1203. EDK, Paris, 1998.

[16] D. Dubois and H. Prade. Possibility theory as a basis for qualitative decision theory. In *Proceedings of IJCAI'95*, pages 1925–1930, 1995.

[17] D. Dubois and H. Prade. Possibility theory as a basis for qualitative decision theory. In *Proceedings of the Int. Joint Conf on Artificial Intelligence (IJCAI'95)*, pages 1925–1930, 20-25 August 1995.

[18] D. Dubois, H. Prade, and R. Sabbadin. Qualitative decision theory with Sugeno integrals. In *Proceedings of UAI'98*, pages 121–128, 1998.

[19] D. Dubois, H. Prade, and R. Sabbadin. Decision-theoretic foundation of qualitative possibility theory. *European Journal of Operations Research*, 128:478–495, 2001.

[20] H. Fargier and R. Sabbadin. Qualitative decision under uncertainty: Back to expected utility. In G. Gottlob and T. Walsh, editors, *Proceedings of IJCAI'03*, pages 303–308, Acapulco, Mexico, 2003. Morgan Kaufmann.

[21] H. Fargier and R. Sabbadin. Qualitative decision under uncertainty: Back to expected utility. *Artificial Intelligence*, 164:245–280, 2005.

[22] M. Grabisch. The Moebius transform on symmetric ordered structures and its application to capacities on finite sets. *Discrete Mathematics*, 287:17–34, 2004.

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

[24] M. Grabisch, T. Murofushi, and M. Sugeno (eds.). *Fuzzy Measures and Integrals - Theory and Applications*. Physica Verlag, Heidelberg, 2000.

[25] S. Greco, B. Matarazzo, and R. Slowinski. Axiomatic characterization of a general utility function and its particular cases in terms of conjoint measurement and rough-set decision rules. *European Journal of Operational Research*, 158(2):271–292, 2004.

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

[27] J.-L. Marichal. On Sugeno integrals as an aggregation function. *Fuzzy Sets and Systems*, 114(3):347–365, 2000.

[28] H. Moulin. *Axioms of Cooperative Decision Making*. Wiley, New-York, 1988.

[29] T. Murofushi. Lexicographic use of Sugeno integrals and monotonicity conditions. *IEEE trans. Fuzzy Systems*, 9:785–794, 2001.

[30] T. Murofushi and M. Sugeno. Null sets with respect to fuzzy measures. In *Proc. 3rd Intern. Fuzzy System Association World Congress (IFSA 1989), Seattle, Wash.*, pages 172–175, 1989.

[31] E. Pap. *Null-Additive Set-Functions*. Kluwer Academic, Dordrecht, The Netherlands, 1995.

[32] A. Rico, M. Grabisch, C. Labreuche, and A. Chateauneuf. Preference modeling on totally ordered sets by the Sugeno integral. *Discrete Applied Mathematics*, 147(1):113–124, 2005.

[33] R. Sarin and P. P. Wakker. A simple axiomatization of nonadditive expected utility. *Econometrica*, 60(6):1255–1272, 1992.

[34] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57:571–587, 1989.

[35] A. K. Sen. *Social Choice Theory.* In: K.J. Arrow, M.D. Intrilligator (Eds), *Handbook of Mathematical Economics*, chapter 22, pages 1073–1181. Elsevier Sciences Publishers, North-Holland, 1986.

[36] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, 1976.

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

[38] M. Sugeno. *Fuzzy Automata and Decision Processes*, chapter Fuzzy measures and fuzzy integrals-a survey, pages 89–102. M.M. Gupta, G.N. Saridis and B.R. Gaines (Eds.), North Holland, Amsterdam, 1977.

[39] S. Tan and J. Pearl. Specification and evaluation of preferences under uncertainty. In Pietro Torasso Jon Doyle, Erik Sandewall, editor, *Proceedings of KR'94*, pages 530–539. Morgan Kaufmann, May 1994.

[40] R. Thomason. Desires and defaults: A framework for planning with inferred goals. In *Proceedings of KR'2000*, pages 702–713, 2000.

[41] Z. Wang. The autocontinuity of set function and the fuzzy integral. *Journal of Mathematical Analysis and Applications*, 99:195–218, 1984.

[42] Z. Wang. Asymptotic structural characteristics of fuzzy measures and their applications. *Fuzzy Sets and Systems*, 16:277–290, 1985.

[43] T. Whalen. Decision making under uncertainty with various assumptions about available information. *IEEE Transactions on Systems, Man and Cybernetics*, 14:888–900, 1984.

[44] S. K. M. Wong, P. Bollmann Y. Y. Yao, and H. C. Burger. Axiomatization of qualitative belief structure. *IEEE transactions on SMC*, 21(34):726–734, 1991.

[45] R. Yager. Possibilistic decision making. *IEEE Transactions on Systems, Man and Cybernetics*, 9:388–392, 1979.

[46] L.A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.