ELSEVIER

European Journal of Operational Research 128 (2001) 459–478

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/dsw

Invited Review

Decision-theoretic foundations of qualitative possibility theory

Didier Dubois [a,*], Henri Prade [a], Régis Sabbadin [b,☆]

[a] IRIT, CNRS, Université Paul Sabatier, Lab. Languages/Systemes Informatiques, 118 route de Narbonne, F-31062 Toulouse Cedex 4, France

[b] INRA de Toulouse, Unité de Biométrie et Intelligence Artificielle, BP 27 – 31326 Castanet-Tolosan Cedex, France

Received 7 January 1999; accepted 25 November 1999

Abstract

This paper presents a justification of two qualitative counterparts of the expected utility criterion for decision under uncertainty, which only require bounded, linearly ordered, valuation sets for expressing uncertainty and preferences. This is carried out in the style of Savage, starting with a set of acts equipped with a complete preordering relation. Conditions on acts are given that imply a possibilistic representation of the decision-maker uncertainty. In this framework, pessimistic (i.e., uncertainty-averse) as well as optimistic attitudes can be explicitly captured. The approach thus proposes an operationally testable description of possibility theory. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Decision theory; Uncertainty; Possibility theory

1. Introduction

The expected utility criterion for decision under uncertainty was the first to receive axiomatic justifications both in terms of probabilistic lotteries [36] and in terms of preference between acts [30]. These axiomatic frameworks have been questioned later, challenging some of the postulates leading to the expected utility criterion, on the basis of systematic violations of these postulates (e.g., [1,17]). For instance Gilboa [19] and Schmeidler [31] have advocated lower and upper expectations expressed by Choquet integrals attached to non-additive measures, sometimes corresponding to a family of probability measures (see also [20,29]). In this paper, we propose axiomatic justifications for two qualitative criteria, an optimistic and a pessimistic one whose definitions only require finite linearly ordered scales. The pessimistic criterion can be viewed as a refinement of the Wald criterion, where uncertainty is expressed in a qualitative way and is captured in the framework of possibility theory [13,15,44].

2. Background on qualitative possibility theory

A possibility distribution π on a set of possible worlds or states S is a mapping from S to a

☆ This work was done while the author was preparing a PhD at IRIT.

* Corresponding author. Tel.: +33-561-556-331; fax: +33-561-556-239.

E-mail addresses: dubois@irit.fr (D. Dubois), prade@irit.fr (H. Prade), sabbadin@toulouse.inra.fr (R. Sabbadin).

PII: S 0 3 7 7 - 2 2 1 7 (9 9) 0 0 4 7 3 - 7

bounded, linearly ordered valuation set $(L, >)$. This ordered set is supposed to be equipped with an order-reversing map denoted by n_L, that is, a bijection of L on itself such that if $\alpha > \beta \in L$, then $n_L(\beta) > n_L(\alpha)$. Let 1 and 0 denote the top and the bottom of L, respectively. Then $n_L(0) = 1$ and $n_L(1) = 0$. In the numerical setting, $L = [0, 1]$, and function n_L is generally taken as $1 - \cdot$. Here, it is only assumed that L is a finite chain, and n_L just puts L upside down. [1] The referential set S represents a set of "states of affairs" or possible worlds, each state being an unambiguous description of a cluster of situations, at a certain level of granularity.

A possibility distribution describes knowledge about the unknown value taken by one or several attributes used to describe states of affairs. For instance it may refer to the age of a man, the size of a building, the temperature of a room, etc. Here it will refer to the ill-known consequence of a decision. A possibility distribution can represent a state of knowledge (about the state of affairs) distinguishing what is plausible from what is less plausible, what is the normal course of things from what is not, what is surprising from what is expected. The function $\pi : S \to L$ represents a flexible restriction on the actual state of affairs, with the following conventions: $\pi(s) = 0$ means that state s is rejected as impossible; $\pi(s) = 1$ means that s is totally possible (plausible). Distinct states may simultaneously have a degree of possibility equal to 1. Flexibility in this description is modeled by letting $\pi(s)$ vary between 0 and 1 for some states s. The quantity $\pi(s)$ thus represents the degree of possibility of the state s, some states being more possible than others. Clearly, if S is the complete range of states, at least one of the elements of S should be fully possible, so that $\exists s, \pi(s) = 1$ (normalization). In this paper we consider only normalized possibility distributions. Strictly speaking a possibility distribution can be viewed as the generalized characteristic function of a fuzzy set [44]. The fundamental point made by Zadeh [44] is the following: as set-characteristic functions can be used to express equipossibility, fuzzy set membership functions are the basis of gradual possibility.

A possibility distribution π is said to be at least as specific as another π' if and only if for each state of affairs $s : \pi(s) \leqslant \pi'(s)$ [43]. Then, π is at least as restrictive and informative as π'. In the possibilistic framework extreme forms of partial knowledge can be captured, namely:

- *complete knowledge*: for some $s_0, \pi(s_0) = 1$ and $\pi(s) = 0\ \forall s \neq s_0$ (only state s_0 is possible);
- *complete ignorance*: $\pi(s) = 1\ \forall s$ (all states in S are possible).

In the following, subsets are denoted $A, B, C, \ldots$. $\overline{A}$ denotes the complement of A. Given a simple query of the form "does the actual state belong to A?", where A is a prescribed subset of situations, the response to the query can be obtained by computing the partial belief induced on A by the knowledge encoded by the possibility distribution π, noticeably to what extent:

- A is consistent with π, with degree
$$\Pi(A) = \sup_{s \in A} \pi(s),$$
- A is certainly implied by π, with degree
$$N(A) = n_L(\Pi(\overline{A})) = \inf_{s \in \overline{A}} n_L(\pi(s)).$$

$\Pi(A)$ is called the degree of possibility of A, and is defined by assuming that, if it is only known that A occurs, then the most plausible situation compatible with A is the one that takes place. It expresses a level of unsurprisingness. The basic axiom of possibility measures in the finite case is $\Pi(A \cup B) = \max(\Pi(A), \Pi(B))$. It is justified by the assumption of jumping to the most plausible situation. By convention, $\Pi(\emptyset) = 0$. A systematic assumption in possibility theory is that the actual situation is normal, i.e., it is any s such that $\pi(s)$ is maximal given other known constraints. It justifies the evaluation $\Pi(A)$, and contrasts with the probabilistic evaluation of the likelihood of events. $N(A)$ is called degree of necessity of A.

When $N(A) \geqslant \alpha > 0$, it means that the most plausible situation where A is false is rather impossible, i.e., not possible to a level greater than

[1] As kindly pointed out by a referee, in the infinite case, not any bounded, totally ordered set can be equipped with an order-reversing map. For instance, $L = [0, 0.5] \cup \{1\}$ cannot. So, L should be everywhere dense, in order to be on the safe side. For a similar reason, n_L should be continuous. However, since we stick to a finite setting here, these problems do not occur.

$n_L(\alpha)$. Moreover $N(A) > 0$ also means that A holds in all the most normal situations. Since the assumption of normality is always made, $N(A) > 0$ thus means that A is an accepted belief, i.e., one may act as if A were true. This assumption is always a default one and can be revised if further pieces of evidence contradict it. Necessity measures satisfy an axiom dual of the one of possibility measures, namely $N(A \cap B) = \min(N(A), N(B))$. This decomposability axiom, as well as the above maxitivity axiom, presupposes a finite setting in order to be characteristic. Otherwise the axiom must hold for infinite families of sets.

Set-functions Π and N are, respectively, called *possibility* and *necessity measures* [15], and can provide simple ordinal representations of graded belief that are fully compatible with preferential representations of uncertainty very common in non-monotonic reasoning [18]. Their particular character lies in their ordinal nature, i.e., the valuation set L is used only to rank-order the various possible situations in S, in terms of their compatibility with the normal course of things as encoded by the possibility distribution π. To each possibility distribution π, we can associate its comparative counterpart, a complete preorder denoted by $\geqslant_\pi$, defined by $s \geqslant_\pi s'$ if and only if $\pi(s) \geqslant \pi(s')$, which induces a well-ordered partition [34] $\{E_1, \ldots, E_n\}$ of S, that is, $\{E_1, \ldots, E_n\}$ is a partition of S such that $\forall s \in E_i\ \forall s' \in E_j, \pi(s) \geqslant \pi(s')$ iff $i \leqslant j$ (for $1 \leqslant i, j \leqslant n$). By convention E_1 represents the most normal states of fact. Thus, a possibility distribution partitions S into classes of equally possible states. Dubois [5] defined *comparative possibility* as a relation on events, denoted $\geqslant_\Pi$, satisfying:

A1. $\geqslant_\Pi$ is complete and transitive.
A2. $S >_\Pi \emptyset$ (non-triviality).
A3. $A \geqslant_\Pi \emptyset$.
Pos. $\forall B, C, D, B \geqslant_\Pi C$ implies $B \cup D \geqslant_\Pi C \cup D$.

Qualitative necessity relations are defined by duality, i.e., $A \geqslant_N B$ if and only if $\overline{B} \geqslant_\Pi \overline{A}$. Their characteristic property, on top of **A1**, **A2**, and a dual property of **A3**,

A3′. $S \geqslant_N A$,

is

N. $\forall B, C, D, B \geqslant_N C$ implies $B \cap D \geqslant_N C \cap D$.

In the finite case, Dubois [5] has shown that the only numerical counterparts to comparative necessity (resp. possibility) relations are necessity (resp. possibility) measures. Qualitative necessity relations are closely related to the epistemic entrenchment relation underlying any revision of a belief set in the sense of Gärdenfors [18]. Possibility orderings are an optimistic view on the relative likelihood of events since they focus on their most plausible realization. Conversely, necessity orderings are cautious since they focus on the most plausible realization of the converse event.

In the above lines, a possibility distribution encodes imprecise knowledge about a situation; in that case, no choice is at stake, that is, the actual situation is what it is, and π encodes plausible guesses about it. However, there exists a different understanding of a possibility distribution: possibility distributions can also express what are the states in which an agent *would like to be*, under the form of a flexible constraint on the state space. In this case possibility is interpreted in terms of *graded preference* or subjective *feasibility* and necessity degrees are interpreted as *priority levels*. A possibility distribution is then similar to a utility function, or, better, a value function, but it may range on a qualitative valuation set (see also [7] for a detailed discussion of the preference view of possibility theory in the setting of constraint satisfaction). Using the two types of possibility distributions conjointly leads to qualitative utility theory.

3. Qualitative counterparts of expected utility

Generally, decisions are made in an uncertain environment. In the Savage framework [30], the consequence of a decision depends on the state of the world in which it takes place. If S is a set of states and X a set of possible consequences, the decision-maker has some knowledge of the actual state and some preference on the consequences of his decision. Here, a belief state about which situation in S is the actual one, is supposed to be represented by a normalized possibility distribution π from S to a plausibility scale L. $\pi(s) \in L$ estimates the plausibility level of being in situation s. As already said, possibility theory, in its qualitative version, represents uncertainty by means of a complete pre-ordering on S, that can be mapped to the totally ordered scale L.

3.1. Possibilistic criteria

It makes sense, if information is qualitative, to represent not only the incomplete knowledge on the state by a possibility distribution π on S with values in a plausibility scale L but also the decision-maker's preference on X by means of another possibility distribution μ with values on a preference scale U. Let x^* and x_* be the best and worst consequences in X, with $\mu(x^*) = 1$ and $\mu(x_*) = 0$. A decision is represented by a function, called an act, from S to X. The utility of a decision $\mathbf{f}$ whose consequence in state s is $x = \mathbf{f}(s) \in X$ for all states s, can be evaluated by combining the plausibilities $\pi(s)$ and the utilities $\mu(x)$ in a suitable way. Two qualitative criteria that evaluate the worth of decision $\mathbf{f}$ have been put forward in the literature of fuzzy sets, provided that a commensurability assumption between plausibility and preference is made:

- *A pessimistic criterion*

  $$v_*(\mathbf{f}) = \inf_{s \in S} \max(n(\pi(s)),\ \mu(\mathbf{f}(s))),$$

  which generalizes the max–min Wald criterion in the absence of probabilistic knowledge. Mapping n is order-reversing from L to U.
- *An optimistic criterion*

  $$v^*(\mathbf{f}) = \sup_{s \in S} \min(m(\pi(s)),\ \mu(\mathbf{f}(s))),$$

  which generalizes the maximax optimistic criterion. Mapping m is order-preserving from L to U.

The optimistic criterion has been first proposed by Yager [42] and the pessimistic criterion by Whalen [40], and also used in [24]. These criteria are clearly based on the possibility and necessity of the fuzzy event with membership function $\mu(\mathbf{f}(\cdot))$. They are special cases of Sugeno integrals [35,38] as proved by Dubois and Prade [11] for the optimistic criterion, and Inuiguchi et al. [24] for the pessimistic criterion; see also [21].

3.2. Axiomatization on possibilistic lotteries

The pessimistic criterion has been axiomatically justified by Dubois and Prade [14] in the style of von Neumann and Morgenstern utility theory [36]. Expected utility theory of von Neumann and Morgenstern relies on the principle that the decision maker's behavior in the face of risk is entirely determined by his/her preferences on the probability distributions about the consequences of his/her actions. Preferences about probabilistic lotteries should fulfill a set of axioms describing the attitude of a "rational" decision maker in the face of risk. Expected utility provides a simple criterion to rank-order the lotteries, and thus the acts, since each lottery is associated with the uncertain consequences of an act. The idea of possibilistic decision theory is that if the uncertainty on the state is represented by a possibility distribution π, each decision induces on the set of consequences X a possibility distribution such that $\pi_{\mathbf{f}}(x) = \Pi(\mathbf{f}^{-1}(x))$. So ranking decisions comes down to ranking possibility distributions on X. Assume the decision-maker supplies an ordering between possibility distributions on X, thus expressing his attitude in front of uncertainty, that is, in front of various possibilities of happy and unhappy consequences in X.

Let $\pi_{\mathbf{f}}(x)$ be the plausibility of getting x under decision $\mathbf{f}$. The question is to know what kind of axioms on the ordering between possibility distributions on X make it representable by the ranking of decisions according to the above pessimistic or optimistic criteria. Let x and y be two elements of X, the possibility distribution $\pi_{\mathbf{f}}$ defined by $\pi_{\mathbf{f}}(x) = \lambda, \pi_{\mathbf{f}}(y) = v, \pi_{\mathbf{f}}(z) = 0$ for $z \neq x, z \neq y$ with $\max(\lambda, v) = 1$ (in order to have $\pi_{\mathbf{f}}$ normalized), is called a *qualitative binary lottery* and will be denoted by $(\lambda/x, v/y)$, which means that we get either consequence x or consequence y, with the respective levels of possibility λ and v. A subset $A = \{x_1, \ldots, x_k\}$ corresponds to the lottery $(1/x_1, \ldots, 1/x_k)$. Here $\mathbf{f}$ is a binary act. More generally, any possibility distribution π can be viewed as a multiple consequence lottery $(\lambda_1/x_1, \ldots, \lambda_m/x_m)$ where $X = \{x_1, \ldots, x_m\}$ and $\lambda_i = \pi_{\mathbf{f}}(x_i)$. For simplicity we drop subscript $\mathbf{f}$ in the following. The notation $(\lambda/\pi, v/\pi')$ denotes the higher-order qualitative lottery yielding the uncertainty distribution π with possibility λ and π' with possibility v. Of course, $\max(\lambda, v) = 1$. A singleton $\{x_0\}$ corresponds to the possibility distribution which is zero everywhere except in x_0, where $\pi(x_0) = 1$.

Let $\succeq$ denote the preference relation between possibility distributions ("possibilistic lotteries") given by the decision maker, which extends the preference ordering over X, to normalized possibility distributions in L^X. Relation $\succeq$ is supposed to satisfy the following axioms, where $\pi \sim \pi'$ means that both $\pi \succeq \pi'$ and $\pi' \succeq \pi$ hold:

Axiom 1. $\succeq$ is a complete pre-ordering.

Axiom 2 (*independence*).

$$\pi_1 \sim \pi_2 \Rightarrow (\lambda/\pi_1, v/\pi') \sim (\lambda/\pi_2, v/\pi').$$

Axiom 3 (*continuity*).

$$\pi \succeq \pi' \Rightarrow \exists \lambda \in L, \pi' \sim (1/\pi, \lambda/X).$$

Axiom 4 (*reduction of lotteries*).

$$(\lambda/s, v/(\alpha/s, \beta/t)) \sim (\max(\lambda, \min(v, \alpha))/s, \min(v, \beta)/t).$$

Axiom 5 (*uncertainty aversion or "precision is safer"*).

$$\pi \leqslant \pi' \Rightarrow \pi \succeq \pi'.$$

Axiom 1 makes it possible to represent utility on a totally ordered scale. Axioms 2–4 are counterparts of axioms proposed by von Neumann and Morgenstern. Axiom 4 reduces higher-order lotteries to standard ones. The resulting possibility distribution is here the qualitative counterpart of a probabilistic mixture $\lambda p_1 + (1 - \lambda) p_2$. Axiom 4 is motivated by the particular form of mixtures in possibility theory (see [9]). The risk aversion axiom states that the less informative π is, the more risky the situation is: the worst epistemic state is total ignorance (here represented by X). So this axiom expresses an aversion for a lack of information. Continuity says that the utility of π goes down without jump if the uncertainty about π raises. Due to continuity and uncertainty aversion, it can be proved that if the lottery is represented by a subset A of possible consequences, then $\exists x \in A$, $x \sim A$ (see [10]). This property, violated by expected utility, suggests that contrary to it, the pessimistic utility is not based on the idea of average and repeated decisions, but makes sense for one-shot decisions. It is based on the idea that when the decision is made and put to work, then the consequence will be some $x \in A$, and the benefit of the decision will indeed be the one of consequence x. It comes down to rejecting the notion of mean value. In fact lottery A is then equivalent to the worst consequence in A.

The possibilistic pessimistic criterion is thus an extension of Wald [37] pessimistic criterion, which evaluates decisions on the basis of their worst consequences, however unlikely they are. But the possibilistic criterion is less pessimistic. It focuses on the idea of usuality and relies on the worst *plausible* consequences induced by the decision. Some unlikely states are neglected by a variable thresholding and the threshold is determined by comparing the possibility distributions valued on L and U via the mapping n. A decision will be rated low if there is a plausible consequence of the decision that has low utility.

A dual set of axioms can be devised for the optimistic criterion (see [10]). The latter can be used as a secondary criterion, for breaking ties between decisions which are equivalent w.r.t. the pessimistic criterion. Clearly the optimistic criterion is very optimistic since $v^*(\pi)$ is high as soon as there exists a situation with a high plausibility and a high utility.

This approach sounds realistic in settings where information about plausible states and preferred consequences is poor and linguistically expressed, and where decisions will not be repeated, and also for repeated decisions whose results do not accumulate. These qualitative counterparts of the expected utility theory nicely fit the setting of flexible constraint propagation [7] illustrating the difference between a fuzzy set modeling preference (in terms of fuzzy constraints) and a fuzzy set modeling uncertainty on ill-controlled parameters, for making decisions. See [6] for an application of the pessimistic possibilistic utility to scheduling.

Example *(The omelette* [30, pp. 13–15]). The problem is about deciding whether or not to add an egg to a 5-egg omelette. The possible states of the world are: *The egg is good* (denoted by g), and *The egg is rotten* (denoted by r). The uncertain part of the knowledge base consists only in our opinion about the state of freshness of the egg. The available acts are: *Break the egg in the omelette* (BIO), *Break it apart in a cup* (BAC), and *Throw it away* (TA). The possible consequences are:

- $6e$ (meaning that we obtain a 6-egg omelette) if g holds and we choose BIO;
- $6c$ (we obtain a 6-egg omelette and we have a cup to wash) if g holds and we choose BAC;
- $5e$ (we obtain a 5-egg omelette) if r holds and we choose TA;
- $5c$ (we obtain a 5-egg omelette and we have a cup to wash) if r holds and we choose BAC;
- $5w$ (we obtain a 5-egg omelette and an egg is wasted) if g holds and we choose TA; and
- wo (the omelette is wasted) if r holds and we choose BIO.

Concerning the preferences: first of all, we do not want to waste the omelette, then if possible, we prefer not to waste an egg. Then, if possible, we prefer to avoid having a cup to wash if the egg is rotten (that is, it would have been better to throw it away directly). Finally, if all these preferences are satisfied, then we prefer to have a 6-egg omelette, and the best situation would be to have, in addition, no cup to wash.

Let us use the scale $\{0, 1, 2, 3, 4, 5\}$ for assessing the certainty levels and preferences. Just notice that we could have used linguistic values instead of numbers: only comparison and order-reversing are meaningful operations here. The preferences can be expressed by means of a *symbolic* utility function μ.

According to the above discussion, the utilities assigned to the consequences are:

$$\mu(6e) = 5, \quad \mu(6c) = 4, \quad \mu(5e) = 3,$$
$$\mu(5c) = 2, \quad \mu(5w) = 1, \quad \mu(wo) = 0.$$

In this example, the possibility distribution π_d restricting the more or less plausible consequences of a decision d, depends only on the possibility distribution on the two possible states g and r, namely, on $\Pi(g)$ and $\Pi(r)$. Let $N(g) = n(\Pi(r))$ and $N(r) = n(\Pi(g))$ (the certainty or necessity of an event is the impossibility of the opposite event). Note that $\min(N(g), N(r)) = 0$, where 0 is here the bottom element of our scale (since the possibility distribution over $\{g, r\}$ should be normalized whatever decision d).

The pessimistic utilities of the possible decisions, given by v_* are the following, according to the levels of certainty of g and r:

- $$v_*(\text{BIO}) = \min(\max(n(\Pi(r)),\ \mu(wo)), \max(n(\Pi(g)), \mu(6e))),$$

  which simplifies into

  $$v_*(\text{BIO}) = N(g).$$

- $$v_*(\text{BAC}) = \min(\max(n(\Pi(r)),\ \mu(5c)), \max(n(\Pi(g)), \mu(6c))).$$

  Thus, $v_*(\text{BAC}) = \min(\max(N(g), 2), 4)$.

- $$v_*(\text{TA}) = \min(\max(n(\Pi(r)),\ \mu(5e)), \max(n(\Pi(g)), \mu(5w))).$$

  Thus,

  $$v_*(\text{TA}) = 1 \text{ if } N(g) > 0 \text{ and } \min(3, \max(N(r), 1)) \text{ if not.}$$

The best decisions are therefore:

- BIO if $N(g) = 5$ (we are sure that the egg is good).
- BIO or BAC if $N(g) \in \{2, 3, 4\}$ (we are rather sure that the egg is good).
- BAC if $N(g) < 2$ and $N(r) < 2$ (we are rather ignorant on the quality of the egg).
- TA or BAC if $N(r) = 2$ (we have a little doubt on its quality).
- TA if $N(r) > 2$ (we do not think that the egg is good).

Notice the importance of the commensurability assumption in the computation of v_* where both degrees of certainty and preferences are involved. Note also the qualitative nature of the approach, since the results depend only on the ordering between the levels in the scale.

4. The axiomatics of Savage for expected utility

The weak point of the above axiomatic justification of qualitative utility theory is that the uncertainty theory (here possibility theory) is part of the set of assumptions. While this approach is natural when uncertainty is captured by objective probabilities, as done by von Neumann and Morgenstern, it is more debatable for subjective uncertainty. On the contrary Savage has proposed a framework for axiomatizing decision rules under uncertainty where both the uncertainty function and the utility function are derived from first principles on acts. The proposed axioms can be operationally verified by checking how the decision-maker ranks acts. This section recalls Savage's setting and his axioms for justifying expected utility and probability functions. Eventually, we propose a Savagean justification of the two above mentioned possibilistic utilities.

In Savage's approach a preference relation $\succeq$ between acts (or decisions) is assumed to be given by a decision-maker. Such a preference relation is observable from the decision-maker's behavior. Acts are defined as functions **f** from an *infinite* state space S to a set X of consequences. Indeed the result of an act depends on the state of the world in which it is performed: the effect of braking a car depends on the state of the brake. Let us denote $\mathscr{F} = X^S$ the set of potential acts. The set of actually feasible acts is generally only a subset of $\mathscr{F}$.

The first assumption of Savage is that the preference relation on $\mathscr{F}$ is transitive and complete ($\mathbf{g} \succeq \mathbf{f}$ or $\mathbf{f} \succeq \mathbf{g}$):

Sav 1 *(Ranking)*. $(\mathscr{F}, \succeq)$ is a complete preorder.

Two particular families of acts are crucial to recover the preference information on consequences and the uncertainty information on the state space S: constant acts and binary acts respectively. A *constant act*, denoted **x** for $x \in X$ is such that $\forall s \in S$, $\mathbf{x}(s) = x$. Since $\succeq$ is a complete preorder on $\mathscr{F}$, the set of acts, it is also a complete preorder on the set of constant acts (which can be identified with X). Therefore, we can define the following complete preorder $\geqslant_P$ on X:

Definition 1 *(Preference on consequences induced by the ranking of acts)*. $\forall x, y \in X$ if $\mathbf{f}(s) = x \ \forall s \in S$, and $\mathbf{g}(s) = y \ \forall s \in S$, then $x \geqslant_P y \iff \mathbf{f} \succeq \mathbf{g}$.

In order to avoid the trivial case when there is only one consequence, or all consequences are equally preferred, Savage has enforced the following condition: [2]

Sav 5 *(Non-triviality)*. There exist $x, x' \in X$ such that $x >_P x'$, where $>_P$ is the strict part of the complete preordering on X.

The ranking of acts also induces a ranking of events, i.e. subsets of the state space: this is based on the use of binary acts. A *binary act* is an act **f** such that there is a set $A \subseteq S$ and two consequences $x >_P x' \in X$, where $\mathbf{f}(s) = x$ if $s \in A$, $\mathbf{f}(s) = x'$ if $s \in \overline{A}$ and $\overline{A}$ is the complement of A. Such a binary act is denoted xAx'. A partial ordering $\geqslant_L$ of events can be defined by restricting the complete preordering on acts to binary acts:

Definition 2 *(Relative likelihood of events)*. Let $A, B \subseteq S$. Event A is at least as likely as event B, denoted $A \geqslant_L B$, if and only if $\forall x, y \in X$, $x >_P y$, $xAy \succeq xBy$.

Of course relation $\geqslant_L$ is only a partial preordering. In order to turn it into a complete preordering, Savage proposed the following axiom:

Sav 4 *(Projection from acts over events)*. Let $x, y, x', y' \in X$, $x >_P x'$, $y >_P y'$. Let $A, B \subseteq S$. Then $xAx' \succeq xBx' \iff yAy' \succeq yBy'$.

This axiom ensures that for any choice of consequences $x >_P y$, the restriction of the preordering on acts to binary acts xAy defines a complete preordering of events in a unique way. The preference ordering on events expresses the uncertainty of the decision-maker about the state of the world, implicit in the way acts are ranked.

The notion of binary act is a particular case of a compound act:

[2] For the sake of clarity we use Savage's original numbering of axioms.

Definition 3 *(Compound act)*. $\forall A \subseteq S$, $\mathbf{f}A\mathbf{g}$ is the act defined by: $\mathbf{f}A\mathbf{g}(s) = \mathbf{f}(s)$ for all $s \in A$, and $\mathbf{f}A\mathbf{g}(s) = \mathbf{g}(s)$ for all $s \in \overline{A}$.

A binary act is thus a compound constant act. Any act can be viewed as a compound act. Combining acts and events and forming compound acts enables any act to be generated by a suitable finite sequence of combinations of events and constant acts, if the state space is finite. Savage has introduced a cancelation property, that boils down to the following assumption: if two acts give the same results on a subset of states, their relative preference does not depend on what these results are. This is called the sure thing principle and is modeled as follows:

Sav 2 *(Sure thing principle)*. Let $\mathbf{f}, \mathbf{g}, \mathbf{h}, \mathbf{h}' \in \mathscr{F}$, let $A \subseteq S$. $\mathbf{f}A\mathbf{h} \succeq \mathbf{g}A\mathbf{h} \Rightarrow \mathbf{f}A\mathbf{h}' \succeq \mathbf{g}A\mathbf{h}'$.

This principle says that the ordering between two acts does not depend on their common consequences.

If two acts **f** and **g** are such that for any third act **h**, $\mathbf{f}A\mathbf{h} \succeq \mathbf{g}A\mathbf{h}$ holds, then **g** is said to be *conditionally preferred* to act **f** on event (a set of states) A, denoted $(\mathbf{f} \succeq \mathbf{g})_A$. Clearly, due to the sure thing principle, conditional preference requires only that $\mathbf{f}A\mathbf{h} \succeq \mathbf{g}A\mathbf{h}$ holds for a single act **h**, since the property $(\mathbf{f} \succeq \mathbf{g})_A$ does not depend on the choice of act **h**. Moreover it is a complete preordering of acts. There is a type of event such that conditioning on them blurs all preferences: null events. An event A is said to be null if and only if $\mathbf{f}A\mathbf{h} \succeq \mathbf{g}A\mathbf{h}$ for any **f**, **h** and **g**. It can be proved that null events are impossible in the sense that $A \sim_L \emptyset$ if and only if A is null.

The restriction of conditional preference to constant acts must coincide with the preference ordering on consequences (except for null events). This is achieved by the following axiom:

Sav 3 *(Conditioning over constant acts)*. Let $x, y \in X, A \subseteq S, A$ not null. Let **x**, **y** be the constant acts: $\mathbf{x}(s) = x$ and $\mathbf{y}(s) = y\ \forall s \in S$. Then, $(\mathbf{x} \succeq \mathbf{y})_A \iff x \geqslant_P y$.

Under the above five conditions the likelihood relation on events induced by the preference on acts is a comparative probability relation, namely it obeys the following characteristic properties:

A1. $\geqslant_L$ is complete and transitive.
A2. $S >_L \emptyset$ (non-triviality).
A3. $\forall A\ A \geqslant_L \emptyset$ (consistency).
P. If $A \cap (B \cup C) = \emptyset$ then: $B \geqslant_L C$ if and only if $A \cup B \geqslant_L A \cup C$ (additivity).

If S is finite, the above four axioms are not enough to ensure the existence of a numerical probability function representing $\geqslant_L$ (see [25]). The setting proposed by Savage presupposes that the set of states is infinite. This assumption is necessary for the introduction of the following axiom:

Sav 6 *(Quantitative probability)*. Let $\mathbf{f}, \mathbf{g} \in \mathscr{F}$, such that $\mathbf{f} \succ \mathbf{g}$ and let $x \in X$. There exists a partition $\{B_1, \ldots, B_n\}$ of S such that $\forall i\ \mathbf{x}B_i\mathbf{f} \succ \mathbf{g}$ and $\mathbf{f} \succ \mathbf{x}B_i\mathbf{g}$.

This condition which allows to partition S into tiny parts with arbitrarily low probability values is necessary in order to obtain a quantitative representation of the comparative probability ordering. Savage proved that a preference relation satisfying **Sav 1**–**Sav 6** can be represented by a utility function u from the set of acts to the reals. For any act **f**, $u(\mathbf{f})$ is the expected utility of the consequences of **f** in the sense of a probability distribution on S. Lastly Savage introduced an axiom that copes with infinite consequence sets:

Sav 7 *(Extension to an infinite number of consequences)*. Let $\mathbf{f}, \mathbf{g} \in \mathscr{F}$ and $A \subseteq S$. $((\mathbf{f} \preceq \mathbf{g}(s))_A\ \forall s \in A) \Rightarrow (\mathbf{f} \preceq \mathbf{g})_A$.

Sav 7 expresses that if every possible consequence of **g** on A is preferred or indifferent to act **f** (conditionally on A) then act **g** shall be preferred or indifferent to act **f** conditionally on A.

The two axioms **Sav 6** and **Sav 7** are clearly technical, not so natural as the other ones, and not so essential to the framework.

5. Properties of possibilistic utility

One of the key postulates of Savage is the sure thing principle which expresses, roughly speaking, that if **f** is preferred or is equivalent to **g** and these two acts result in identical consequences on a subset $B \subseteq S$, then if **f** and **g** are modified in the same way on B, the two modified acts remain ordered as **f** and **g**.

However, two acts may be found equivalent just because they have identical and extreme (very good or very bad) likely consequences on $B \subseteq S$, while one act would be strictly preferred to the other in case their consequences on B were not so dramatic. In other words, extreme and likely consequences may be allowed to blur minor differences on $\overline{B}$. Of course changing the identical parts of two acts should not lead to a preference reversal. This rationale suggests the following weakening of Savage **Sav 2** postulate:

WI *(Weak independence)*. Let $\mathbf{f}, \mathbf{g}, \mathbf{h}, \mathbf{h}' \in \mathscr{F}$, let $A \subseteq S$. $\mathbf{f}A\mathbf{h} \succ \mathbf{g}A\mathbf{h} \Rightarrow \mathbf{f}A\mathbf{h}' \succeq \mathbf{g}A\mathbf{h}'$.

Proposition 1. *The possibilistic utilities v^* and v_* introduced in Section 3.1 satisfy the weak independence property, but not the sure thing principle.*

Proof. $v_*(\mathbf{f}A\mathbf{h}) = \min(\inf_{s \in A} \max(n(\pi(s)), \mu(\mathbf{f}(s))), \inf_{s \in \overline{A}} \max(n(\pi(s)), \mu(\mathbf{h}(s))))$. Let us write $v_*^A\mathbf{f} = \inf_{s \in A} \max(n(\pi(s)), \mu(\mathbf{f}(s)))$. Hence if the term $v_*^{\overline{A}}\mathbf{h}$ in the above term is smaller than both $v_*^A\mathbf{f}$ and $v_*^A\mathbf{g}$ then $\mathbf{f}A\mathbf{h} \sim \mathbf{g}A\mathbf{h}$. However, changing act **h** into the best constant act $\mathbf{x}^*$, the preference between $\mathbf{f}A\mathbf{x}^*$ and $\mathbf{g}A\mathbf{x}^*$ only reflects the ordering between $v_*^A\mathbf{f}$ and $v_*^A\mathbf{g}$ since $v_*^{\overline{A}}\mathbf{x}^* = 1$. The same reasoning holds for the optimistic utility, with some adaptation. The possibilistic utilities violate the sure thing principle because min and max fail to be cancelative. □

As a consequence of Proposition 1, the notion of conditional preference defined in Section 4 is no longer valid for possibilistic utilities. Especially, $\mathbf{f}A\mathbf{h} \succ \mathbf{g}A\mathbf{h}$ for any **h** means, for the pessimistic utility, $\min(v_*^A\mathbf{f}, \alpha) > \min(v_*^A\mathbf{g}, \alpha)$ for any $\alpha \in L$, which is impossible. A similar conclusion holds for the optimistic utility.

A weaker notion of conditional preference could be adopted, such that

- $(\mathbf{f} \sim \mathbf{g})_A$ iff $\mathbf{f}A\mathbf{h} \sim \mathbf{g}A\mathbf{h}\ \forall \mathbf{h}$,
- $(\mathbf{f} \succ \mathbf{g})_A$ iff $\mathbf{f}A\mathbf{h} \succeq \mathbf{g}A\mathbf{h}\ \forall \mathbf{h}$ and $\mathbf{f}A\mathbf{h} \succ \mathbf{g}A\mathbf{h}$ for some **h**.

In the following we avoid the notion of conditional preference and stick to representing preference on X^S. We use preference between compound acts instead of conditional preference.

The failure of the sure thing principle also suggests that axioms **Sav 3** and **Sav 4** will not hold. Possibilistic utility only obeys weak versions of these axioms:

WS3 *(Weak coherence with constant acts)*. If **x** and **y** are constant acts then $x \geqslant_P y \Rightarrow \mathbf{x}A\mathbf{h} \succeq \mathbf{y}A\mathbf{h}\ \forall A \subseteq S$ and all acts **h**.

It is obvious that **WS3** is satisfied by both pessimistic and optimistic utilities. However, these utilities fail to satisfy **Sav 3**, for the same reason as they fail to satisfy **Sav 2**, namely the blurring effect of act **h** in compound acts $\mathbf{x}A\mathbf{h}$. Fortunately our possibilistic utilities satisfy a more general property of consistency with a dominance relation between acts that is similar to Pareto-dominance, in the sense that it is a pointwise preference property.

Definition 4 *(Pointwise preference)*. An act **f** is said to *dominate* another act **g**, which is denoted $\mathbf{f} \geqslant_P \mathbf{g}$ if and only if $\forall s \in S,\ \mathbf{f}(s) \geqslant_P \mathbf{g}(s)$ (the ordering on X induced by constant acts). We also say that **f** is *pointwisely preferred* to **g**.

In the terminology of fuzzy sets, pointwise preference corresponds to fuzzy set inclusion. It is easy to check that for the pessimistic and optimistic utilities, pointwise preference implies weak preference. The monotonicity of the pessimistic and optimistic utility is obvious from their definitions: increasing $\mu(\mathbf{f}(s))$ in

$$v_*(\mathbf{f}) = \inf_{s \in S} \max(n(\pi(s)), \mu(\mathbf{f}(s))) \text{ and}$$

$$v^*(\mathbf{f}) = \sup_{s \in S} \min(m(\pi(s)), \mu(\mathbf{f}(s)))$$

cannot decrease the utilities.

More specifically the pessimistic utility satisfies the following axiom which Grant et al. [22] claim to be one form of the genuine sure thing principle.

WSP *(Weak sure thing principle)*. If $\mathbf{f}A\mathbf{g} \succ \mathbf{f}$ and $\mathbf{g}A\mathbf{f} \succ \mathbf{f}$ then $\mathbf{g} \succ \mathbf{f}$.

This principle means that if by changing act **f** into act **g** one improves the expectations both when A occurs and when its opposite occurs, then **g** should be better than **f** regardless of event A. The pessimistic and optimistic utilities satisfy two different principles, respectively:

- **PES.** *Pessimism*: $\forall \mathbf{f}, \mathbf{g}, \forall A \subseteq S$,

  $\mathbf{f}A\mathbf{g} \succ \mathbf{f} \Rightarrow \mathbf{f} \succeq \mathbf{g}A\mathbf{f}$.

- **OPS.** *Optimism*: $\forall \mathbf{f}, \mathbf{g}, \forall A \subseteq S$,

  $\mathbf{f} \succ \mathbf{f}A\mathbf{g} \Rightarrow \mathbf{g}A\mathbf{f} \succeq \mathbf{f}$.

PES implies **WSP** because they state that the "if" part of **WSP** is ever false. **OPT** violates the "if" part of the following dual expression of Grant's axiom: **WSP′** If $\mathbf{f} \succ \mathbf{f}A\mathbf{g}$ and $\mathbf{f} \succ \mathbf{g}A\mathbf{f}$ then $\mathbf{f} \succ \mathbf{g}$. (**WSP′** is equivalent to **WSP** if **Sav 1**, **Sav 3** and **Sav 6** hold [22]). The pessimism axiom means the following: given some act **f**, if by changing act **f** into act **g** one improves the expectations of the act **f** when $\overline{A}$ occurs, then there is no way of forming an act better than **f** by turning **f** into **g** when the opposite event A occurs. The reason is that the decision-maker considers it as plausible that $\overline{A}$ occurs as its opposite, and he pays no attention to good consequences that may occur if A occurs, due to pessimism.

For instance suppose a game of chance according to which a coin is tossed that makes you win 10,000 Euros if head, and lose 10,000 Euros if tail (Game 1). Usually, you will prefer another game, whereby you win 10,000 Euros if head, and nothing otherwise (Game 2). Now, you are proposed yet another game, whereby you win 20,000 Euros if head, and lose 10,000 Euros if tail (Game 3). If, preferring Game 2 to Game 1, you are nevertheless indifferent between Games 1 and 3, then you are a pessimist. Indeed, it holds that $\mathbf{f}A\mathbf{g} \succ \mathbf{f}$ and $\mathbf{f} \sim \mathbf{g}A\mathbf{f}$, where $A=$ head, $\mathbf{f} =$ playing Game 1, **g** is a game where you win 20,000 Euros if head and nothing otherwise, so that $\mathbf{f}A\mathbf{g} =$ Game 2 and $\mathbf{g}A\mathbf{f} =$ Game 3. It indeed reveals that you consider the outcome "tail" as not unlikely, and that you focus on the worst possible consequences. Standard expected utility cannot model this behavior.

The pessimistic utility satisfies the pessimism axiom and the optimistic utility satisfies the optimism axiom. To see it let us first introduce a notion of conjunction and disjunction of acts. First given two acts **f** and **g**, define the act $\mathbf{f} \wedge \mathbf{g}$ (resp. $\mathbf{f} \vee \mathbf{g}$) which in each state s gives the worst (resp. the best) of the results $\mathbf{f}(s)$ and $\mathbf{g}(s)$, following the ordering on X (induced by the ordering of constant acts). In terms of fuzzy sets this is the fuzzy union and intersection of fuzzy sets viewed as acts. Then, it is easy to check, due to elementary properties of min and max, that the following properties, violated by expected utility, hold for qualitative utility.

Lemma 1.

$$v_*(\mathbf{f} \wedge \mathbf{g}) = \min(v_*(\mathbf{f}), v_*(\mathbf{g})),$$

and

$$v^*(\mathbf{f} \vee \mathbf{g}) = \max(v^*(\mathbf{f}), v^*(\mathbf{g})).$$

The two other decomposability properties do not hold except if we consider disjunctions and conjunctions of acts $\mathbf{f} \vee \mathbf{g}$ and $\mathbf{f} \wedge \mathbf{g}$ one of **f** or **g** being a constant act. Namely, if **x** is a constant act

$$v^*(\mathbf{f} \wedge \mathbf{x}) = \min(v^*(\mathbf{f}), \mu(\mathbf{x})),$$
$$v_*(\mathbf{f} \vee \mathbf{x}) = \max(v_*(\mathbf{f}), \mu(\mathbf{x})).$$

This is again obvious to check due to properties of min and max. Let us call the latter property *semi-decomposability*. It leads to introduce a property that is respected by the possibilistic utilities and, again, not generally by the expected utility:

The following lemma holds.

Lemma 2. *Under the pointwise preference monotonicity assumption, the two following properties are equivalent*:
(i) $\mathbf{g} \succ \mathbf{f}$ *and* $\mathbf{h} \succ \mathbf{f}$ *imply* $\mathbf{g} \wedge \mathbf{h} \succ \mathbf{f}$,
(ii) $\mathbf{f} \sim \mathbf{f} \wedge \mathbf{g}$ *or* $\mathbf{g} \sim \mathbf{f} \wedge \mathbf{g}$.

Proof. Suppose (i) and both $\mathbf{f} \succ \mathbf{f} \wedge \mathbf{g}$, $\mathbf{g} \succ \mathbf{f} \wedge \mathbf{g}$. Then $\mathbf{f} \wedge \mathbf{g} \succ \mathbf{f} \wedge \mathbf{g}$, which is impossible, hence $\mathbf{f} \preceq \mathbf{f} \wedge \mathbf{g}$ or $\mathbf{g} \preceq \mathbf{f} \wedge \mathbf{g}$. However, the pointwise preference assumption implies both $\mathbf{f} \succeq \mathbf{f} \wedge \mathbf{g}$ and $\mathbf{g} \succeq \mathbf{f} \wedge \mathbf{g}$. Hence (ii) holds.

Conversely, suppose (ii) and $\mathbf{g} \succ \mathbf{f}$ and $\mathbf{h} \succ \mathbf{f}$. Then, one of $\mathbf{g}$ and $\mathbf{h}$ can be changed into $\mathbf{g} \wedge \mathbf{h}$, which means (i). □

Similarly, and under the same assumptions as in Lemma 2, $\mathbf{f} \succ \mathbf{g}$ and $\mathbf{f} \succ \mathbf{h}$ imply $\mathbf{f} \succ \mathbf{g} \vee \mathbf{h}$ if and only if $\mathbf{f} \sim \mathbf{f} \vee \mathbf{g}$ or $\mathbf{g} \sim \mathbf{f} \vee \mathbf{g}$.

Due to Lemmas 1 and 2, it is clear that the pessimistic utility satisfies the following properties:

CD *(Conjunctive-dominance).*

$\mathbf{g} \succ \mathbf{f} \quad \text{and} \quad \mathbf{h} \succ \mathbf{f} \Rightarrow \mathbf{g} \wedge \mathbf{h} \succ \mathbf{f}.$

RDD *(Restricted disjunctive-dominance).*

$\mathbf{f} \succ \mathbf{g} \quad \text{and} \quad \mathbf{f} \succ \mathbf{x} \Rightarrow \mathbf{f} \succ \mathbf{g} \vee \mathbf{x},$

where $\mathbf{x}$ is the constant act that always yields consequence x.

Dually, the optimistic utility satisfies:

DD *(Disjunctive-dominance).*

$\mathbf{f} \succ \mathbf{g} \quad \text{and} \quad \mathbf{f} \succ \mathbf{h} \Rightarrow \mathbf{f} \succ \mathbf{g} \vee \mathbf{h}.$

RCD *(Restricted conjunctive-dominance).*

$\mathbf{g} \succ \mathbf{f} \quad \text{and} \quad \mathbf{x} \succ \mathbf{f} \Rightarrow \mathbf{g} \wedge \mathbf{x} \succ \mathbf{f}.$

To see that expected utility violates **RCD**, for instance, it is enough to find real values a, b, a', b', c and a number α in the unit interval such that

$a \cdot \alpha + b \cdot (1 - \alpha) > a' \cdot \alpha + b' \cdot (1 - \alpha),$

$c > a' \cdot \alpha + b' \cdot (1 - \alpha),$ and

$\min(a,c) \cdot \alpha + \min(b,c) \cdot (1 - \alpha) \leqslant a' \cdot \alpha + b' \cdot (1 - \alpha).$

The reader can check that the values $a = 1000$, $b = 2$, $a' = 3$, $b' = 100$, $c = 10$, and $\alpha = 0.93$ yield such a counterexample.

The decomposability with respect to the disjunction of acts, or maxitivity property, of the optimistic utility, v^* is a clear counterpart of the additivity of expected utility for the sum of acts. Similarly, the semi-decomposability of v^* for the conjunction of an act and a constant act is the counterpart of the linearity of expected utility with respect to the multiplication of an act by a constant. These properties were used by de Campos and Bolaños [3] when characterizing the possibility of a fuzzy event. However, they do not consider the necessity of fuzzy events. Now we can relate these decomposability properties to pessimism and optimism axioms:

Proposition 2. *The pessimistic utility v_* satisfies* **PES** *and the optimistic utility v^* satisfies* **OPT**.

Proof. Now assume $v_*(\mathbf{f}A\mathbf{g}) > v_*(\mathbf{f})$ and $v_*(\mathbf{g}A\mathbf{f}) > v_*(\mathbf{f})$. Then, $\min(v_*(\mathbf{f}A\mathbf{g}), v_*(\mathbf{g}A\mathbf{f})) > v_*(\mathbf{f})$. But using the above min decomposability of the pessimistic utility, this also reads: $v_*(\mathbf{f}A\mathbf{g} \wedge \mathbf{g}A\mathbf{f}) > v_*(\mathbf{f})$ and since $(\mathbf{f}A\mathbf{g}) \wedge (\mathbf{g}A\mathbf{f}) = (\mathbf{f} \wedge \mathbf{g})$, we find $v_*(\mathbf{f} \wedge \mathbf{g}) > v_*(\mathbf{f})$ which is impossible since $\mathbf{f}$ is pointwisely better than $\mathbf{f} \wedge \mathbf{g}$ and the pessimistic utility respects pointwise preference. The negation of $v_*(\mathbf{f}A\mathbf{g}) > v_*(\mathbf{f})$ and $v_*(\mathbf{g}A\mathbf{f}) > v_*(\mathbf{f})$ is precisely the pessimism axiom. A similar proof can be proposed for showing that the optimistic utility satisfies the optimism axiom. □

Let us now consider binary acts of the form xAy $(x >_P y)$. Note that

$$\begin{aligned} v_*(xAy) &= \max(\mu(y),\ \min(N(A),\ \mu(x))) \\ &= \min(\mu(x),\ \max(N(A),\ \mu(y))). \end{aligned}$$

This form of the pessimistic utility is easy to understand: if the agent is sure enough that A occurs $(N(A) > \mu(x))$ then the utility of the act xAy is $\mu(x)$. If the agent has too little knowledge $(\max(N(A), N(\overline{A})) < \mu(y))$ he is cautious and the utility is $\mu(y)$, the worst case. Of course the same happens if the agent is at least somewhat certain that $\overline{A}$ occurs. If the agent's certainty that A occurs is positive but not extreme, the utility reflects the certainty level and is equal to $N(A)$. Note that the pessimistic utility of the binary qualitative lottery is the median of $\{\mu(x), N(A), \mu(y)\}$, thus contrasting with expected utility, which is a mean. Similarly, the optimistic utility of the binary act takes the simplified form

$$v^*(xAy) = \max(\min(\Pi(A), \mu(x)), \mu(y)),$$

and can be interpreted similarly as the median of $\{\mu(x), \Pi(A), \mu(y)\}$, but here the utility is $\mu(x)$ as soon as the agent believes that obtaining x is possible enough ($\Pi(A) > \mu(x)$).

Both pessimistic and optimistic utilities violate axiom **Sav 4**, because of the blurring effects of almost sure events with drastic consequences. Indeed, considering binary acts xAx', xBx', yAy', and yBy', one may have $v_*(xAx') = N(A) > v_*(xBx') = N(B)$ and $v_*(yAy') = v_*(yBy') = \mu(y)$, for instance, when $\mu(y) \leqslant \min(N(A), N(B))$.

It is easy to verify that $\forall x >_P y$, the set $\mathscr{F}_{xy}$ of binary acts of the form xAy is isomorphic to $\mathscr{P}(S)$ (the set of all subsets of S). Let $\trianglerighteq^{xy}$ be the total preorder on events, induced by the possibilistic utilities, restricted to $\mathscr{F}_{xy}$: $A \trianglerighteq^{xy} B \iff xAy \succeq xBy$.

Via **Sav 4**, Savage [30] required that the induced weak ordering on events should not depend on the values of the outcomes x, y. Here, $\trianglerighteq^{xy}$ depends on the values of x and y. In fact the possibilistic utilities satisfy a weak version of **Sav 4**, whereby the preference ordering of binary acts remains weakly coherent when changing the consequences x and y.

WS4. Let $x >_P x', y >_P y'; A, B \subseteq S$:

$xAx' \succ xBx' \Rightarrow yAy' \succeq yBy'$.

If furthermore, we have: $x' \leqslant_P y' <_P y \leqslant_P x$, then:

$xAx' \succeq xBx' \Rightarrow yAy' \succeq yBy'$.

This means that if two binary acts with the same outcomes are equivalent, changing the outcomes into outcomes that are less *extreme* keeps both acts equivalent. However, changing the consequences of binary acts while preserving their respective preference orderings will not create a preference reversal: we cannot have $v_*(xAx') > v_*(xBx')$ and $v_*(yAy') < v_*(yBy')$.

Proposition 3. *Axiom* **WS4** *holds for the possibilistic utilities.*

Proof. If $v_*(xAy) > v_*(xBy)$ several cases occur:

- $v_*(xAy) = N(A) > v_*(xBy) = \mu(y)$. Then it means that $\mu(x) \geqslant N(A) > \mu(y) \geqslant N(B)$, hence $N(A) > N(B)$,
- the same holds if $v_*(xAy) = N(A)$ and $v_*(xBy) = N(B)$ of course,
- $v_*(xAy) = \mu(x)$ and $v_*(xBy) = \mu(y)$ or $N(B)$. Then again $N(A) > \mu(x) \geqslant N(B)$.

Then $v_*(xAy) > v_*(xBy)$ implies $N(A) > N(B)$ hence $v_*(x'Ay') \geqslant v_*(x'By')$ since the function $\min(a, \max(b, c))$ is non-decreasing. We may have that $v_*(x'Ay') = v_*(x'By')$, if $N(A) \geqslant N(B) \geqslant \mu(x') > \mu(y')$ for instance. But no preference reversal is possible. Moreover choosing $\mu(x') > \mu(x) > \mu(y) > \mu(y')$ increases the chance for $N(A)$ and $N(B)$ to be the values of the utilities $v_*(x'AY')$ and $v_*(x'By')$ namely checking the above three cases shows that $v_*(xAy)$ cannot but increase, and $v_*(xBy)$ cannot but decrease. One becomes convinced that $v_*(xAy) > v_*(xBy)$ implies $v_*(x'Ay') > v_*(x'By')$. The same reasoning works for the optimistic utility. □

The above analysis shows that possibilistic utility functions have properties that noticeably differ from those of expected utility. Especially, axiomatizing possibilistic utilities cannot rely on the sure-thing principle, nor on Savage definition of the uncertainty relation induced from preference on acts via **Sav 4**.

6. Act-driven axiomatization of possibility theory and qualitative utility

In this section it is shown that the pessimistic and optimistic possibilistic utilities can be axiomatized in the style of Savage, just like expected utility. The main difference is that a finite setting is enough to prove the results. In a first step, we point out a general framework for describing many families of monotonic set-functions in terms of acts, thus providing a practically testable framework for many non-probabilistic uncertainty theories. Namely, by asking a decision-maker to rank acts in an uncertain environment, one may "guess" the kind of uncertainty measure he is implicitly working with. In particular, possibility theory thus receives some operational foundations.

6.1. Various uncertainty measures induced by the preference on acts

Generally uncertainty is represented by set-functions $\sigma : S \to L$ which are Sugeno measures [35,38], that is:

$$\sigma(\emptyset) = 0_L, \quad \sigma(S) = 1_L, \quad \text{and}$$
$$A \subseteq B \Rightarrow \sigma(A) \leqslant \sigma(B).$$

This kind of set-function is very general and represents the minimal requirement for the representation of partial belief. Especially the last condition is called *monotonicity*, and is verified by probability measures and most other well-known representations of partial belief.

6.1.1. Representation of monotonic set-functions

In terms of acts Sugeno measures can be recovered as follows, if we consider the restrictions of a preference relation on acts to binary acts of the form $\mathscr{F}_{xy} = \{xAy, A \subseteq S\}$ with $x, y \in X$ whose corresponding constant acts $\mathbf{x}$ and $\mathbf{y}$ satisfy $\mathbf{x} \succ \mathbf{y}$. Only very few axioms are needed. However, we need the notion of conditional preference on acts defined in Section 4:

Lemma 3. *If an act* $\mathbf{f}$ *is conditionally preferred to an act* $\mathbf{g}$ *both on a set A and its complement then axiom* **Sav1** *implies that* $\mathbf{f}$ *is preferred to* $\mathbf{g}$.

Proof. Assume $(\mathbf{f} \succeq \mathbf{g})_A$ that is, $\mathbf{f}A\mathbf{h} \succeq \mathbf{g}A\mathbf{h}$ holds for all $\mathbf{h}$, and $(\mathbf{f} \succeq \mathbf{g})_{\overline{A}}$ as well. Then due to the transitivity of **Sav 1**, $\mathbf{f} \succeq \mathbf{g}A\mathbf{f}$ (using $\mathbf{h} = \mathbf{f}$), $\mathbf{g}A\mathbf{f} = \mathbf{f}\overline{A}\mathbf{g} \succeq \mathbf{g}$ (using $\mathbf{h} = \mathbf{g}$ and conditional preference on $\overline{A}$.) Hence $\mathbf{f} \succeq \mathbf{g}$. □

Lemma 4 (Monotonicity). *If the set of acts* $\mathscr{F} = X^S$ *is equipped with a preference relation* $\succeq$ *that satisfies* **Sav 1**, *and* **WS3**, *then pointwise preference implies preference:* $\mathbf{f} \geqslant_P \mathbf{g} \Rightarrow \mathbf{f} \succeq \mathbf{g}$.

Proof. We recall that relation $\geqslant_P$ is a complete preordering on X obtained by restricting $\succeq$ to constant acts. Assume $\mathbf{f} \geqslant_P \mathbf{g}$, in such a way that $\mathbf{f}(s) = \mathbf{g}(s)$ except for some state s_0 where $\mathbf{f}(s_0) >_P \mathbf{g}(s_0)$. Such two acts exist, otherwise X is an equivalence class for $\geqslant_P$ and the result trivially holds. Clearly $\mathbf{f}$ is pointwisely preferred to $\mathbf{g}$. In this particular situation we say that $\mathbf{f}$ is *simply* pointwisely preferred to $\mathbf{g}$. Now due to **WS3**, $\mathbf{f}(s_0) >_P \mathbf{g}(s_0)$ implies $\mathbf{f} \succeq \mathbf{g}$. More generally, if $\mathbf{f}$ is pointwisely preferred to $\mathbf{g}$, then it is possible to build a finite sequence of acts $\mathbf{f}_0, \ldots, \mathbf{f}_n$ such that $\mathbf{f}_0 = \mathbf{f}, \mathbf{f_n} = \mathbf{g}$ where $\mathbf{f}_i$ is simply pointwisely preferred to $\mathbf{f}_{i+1}$. Then, by transitivity, $\mathbf{f} \succeq \mathbf{g}$. □

Due to the above lemmas the following theorem is obvious.

Theorem 1 (Representation of Sugeno measures). *If the set of acts* $\mathscr{F} = X^S$ *is equipped with a preference relation* $\succeq$ *that satisfies* **Sav 1**, **WS3**, **Sav 5** *then the uncertainty relation induced by restricting to binary acts with fixed consequences can be represented by a Sugeno measure.*

Proof. Let $x >_P y$, due to **Sav 5**. Consider the relation $\trianglerighteq^{xy}$ among events defined by $A \trianglerighteq^{xy} B$ if and only if $xAy \succeq xBy$. This relation is a complete preordering and can be mapped to a finite linear scale L_{xy} whose elements are the equivalence classes of $\mathscr{F}_{xy}$. Let $[xAy]$ denote the equivalence class of xAy. Let σ denote the set function such that $\sigma(A) = [xAy]$. Note that $\forall A \subseteq B, xBy \geqslant_P xAy$ and due to **Sav 1**, **WS3**, via the preceding monotonicity lemma, we get $xBy \succeq xAy$ and $\sigma(B) \geqslant \sigma(A)$. Therefore σ is monotonic with respect to set inclusion. **Sav 5** ensures that $\sigma(\Omega) > \sigma(\emptyset)$. □

Clearly the problem at this point is to ensure that the sets $\mathscr{F}_{xy}$ of binary acts remain coherent with one another in the sense that the orderings of events induced by $\mathscr{F}_{xy}$ and $\mathscr{F}_{x'y'}$ for two pairs of consequences (x, y) and (x', y') do not contradict each other. A minimal coherence is ensured by the axiom **WS4**. Then we are sure that the relation $\trianglerighteq^{xy}$ among events is a refinement of another one $\trianglerighteq^{x'y'}$ if $x \geqslant_P x' >_P y' \geqslant_P y$. Moreover the case of outright contradiction $xAy \succ xBy$ and $x'By' \succ x'Ay'$ should not be observed. It is already ruled out if $A \subseteq B$ by the above theorem. Axiom **WS4** ensures that it does not occur in other cases. Let x_* and x^* be the least and greatest elements of $(X, \geqslant_P)$. Clearly the constant acts $\mathbf{x}_*$ (resp. $\mathbf{x}^*$) they induce are not preferred to (resp. are better than) any other act, due to the pointwise preference theorem, under **Sav 1**, **WS3**, **Sav 5**. If L denotes a linearly ordered scale isomorphic to the set of equivalence classes of $(\mathscr{F}, \succeq)$ then $\mathbf{x}_*$ and $\mathbf{x}^*$ correspond to the bottom 0 and the top 1 of L. The most refined uncertainty relation on events that can be obtained from $(\mathscr{F}, \succeq)$ is thus via $\mathscr{F}_{x^*x_*}$ which for simplicity we shall denote $\mathscr{F}_{10}$ and whose elements will be denoted $1A0$, binary acts where the best consequence obtains if A occurs and the worst otherwise. We shall denote the uncertainty relation between events $A \trianglerighteq B$ if and only if $1A0 \succeq 1B0$ as *the* uncertainty relation induced from $(\mathscr{F}, \succeq)$, representing the implicit epistemic state of a decision-maker rank-ordering acts and respecting **Sav 1**, **WS3**, **WS4**, **Sav 5**.

6.1.2. Pseudo-additive set functions

It is interesting to see which kind of uncertainty measures can be captured in terms of acts apart from Sugeno measures and probability measures. To see it we shall consider relaxations of the sure-thing principle (that leads to comparative probability), and first of all the weak independence axiom **WI** that just prevents preference reversals of the form $\mathbf{f}A\mathbf{h} \succ \mathbf{g}A\mathbf{h}$ while $\mathbf{f}A\mathbf{h}' \prec \mathbf{g}A\mathbf{h}'$, still coping with a blurring effect of strict preferences when moving from $\mathbf{h}$ to $\mathbf{h}'$ when $\overline{A}$ occurs.

Dubois [5] proposed a relaxation of the comparative probability axiom **P** that, in conjunction with the other basic axioms **A1**–**A3** subsumes both qualitative probability and qualitative possibility:

DM. $\forall A, B, C, A \cap (B \cup C) = \emptyset$,

$B \trianglerighteq C \Rightarrow B \cup A \trianglerighteq C \cup A$,

and a dual axiom to **DM**, which is satisfied by qualitative probability and qualitative necessity:

DDM. $\forall A, B, C, A \cup (B \cap C) = S$,

$B \trianglerighteq C \Rightarrow B \cap A \trianglerighteq C \cap A$.

Chateauneuf [2], improving results in [5], has proved that any uncertainty ordering that obeys **A1**–**A3** and **DM** can be represented by a *pseudo-additive measure*, that is, a set-function σ mapping on $L = \sigma(2^S)$ such that there exists an operation $\oplus$ in L that verifies the following properties:

- $1 \oplus \lambda = 1$;
- $0 \oplus \lambda = \lambda$;
- $\oplus$ is commutative and associative;
- moreover $\sigma(A \cup B) = \sigma(A) \oplus \sigma(B)$ for any disjoint events A and B.

Such pseudo-additive measures have been introduced by Dubois and Prade [12] and Weber [39] when $\oplus$ is a triangular conorm in the sense of Schweizer and Sklar [32]. Clearly adequate candidates for $\oplus$ are maximum and the bounded sum (if L is numerical), so that decomposable measures include possibility and probability measures. Axiom **DM** can be called decomposable monotonicity.

By duality, any uncertainty ordering that obeys **A1**, **A2** and **A3**$'$ and **DDM** can be represented by a *dual pseudo-additive measure*, that is, a set-function ρ with range $L = \rho(2^S)$ such that there exists an operation $\odot$ in L that verifies the following properties:

- $1 \odot \lambda = \lambda$;
- $0 \odot \lambda = 0$;
- $\odot$ is commutative and associative;
- moreover $\rho(A \cap B) = \rho(A) \odot \rho(B)$ for any events A and B such that $A \cup B = S$.

Such dual pseudo-additive measures are of the form $\rho(A) = n_L(\sigma(\overline{A}))$ where n_L is an involutive order-reversing map of L. Operation $\odot$ can be taken as a triangular norm in the sense of Schweizer and Sklar [32]. Clearly adequate candidates for $\odot$ are minimum and the Lukasiewicz conjunction ($\max(0, a + b - 1)$ if L is numerical), so that dual pseudo-additive measures include necessity and probability measures.

However, the above relaxation of the probabilistic framework is still too restrictive to result from the weak independence axiom. In order to find the proper class of set functions that is captured by the latter, we consider a relaxed version of **DM** that we call weak decomposable monotonicity:

WDM. $\forall A, B, C, A \cap (B \cup C) = \emptyset$,

$B \triangleright C \Rightarrow B \cup A \trianglerighteq C \cup A$.

It must be pointed out that **WDM** can be stated differently in an equivalent way:

$\forall A, B, C, A \cup (B \cap C) = S,$

$B \rhd C \Rightarrow B \cap A \unrhd C \cap A.$

To prove this point, just let $E = \overline{B \cap A}$, $F = \overline{C \cap A}, G = A$ and consider the contraposed form of **WDM**.

The following theorem can be shown.

Theorem 2. *Let $\succeq$ be an order relation on acts, satisfying* **Sav1**, **WI**, **WS3**, **Sav5**. *Then the order relation on events $\unrhd$ induced by the preference relation on acts via binary acts $1A0$ satisfies* **A1**, **A2**, **A3**, **A3′** *and* **WDM**.

Proof. Only **WDM** needs to be established. Just write **WI**($\mathbf{f}D\mathbf{h} \succ \mathbf{g}D\mathbf{h} \Rightarrow \mathbf{f}D\mathbf{h}' \succeq \mathbf{g}D\mathbf{h}'$) for $\mathbf{f} = 1B0$ and $\mathbf{g} = 1C0, D = B \cup C$ $\mathbf{h} = \mathbf{x}_*$. $\mathbf{f}D\mathbf{h} \succ \mathbf{g}D\mathbf{h}$ then writes $1B0 \succ 1C0$, i.e. $B \rhd C$. Now let $\mathbf{h}' = 1A0$ where A is disjoint from D. Then $\mathbf{f}D\mathbf{h}' \succeq \mathbf{g}D\mathbf{h}'$ reads $B \cup A \unrhd C \cup A$. □

Note that **WS4** is not used to prove the result, which holds for all the relations induced from $\mathscr{F}_{xy}$.

However weak **WDM** may look, it is satisfied neither by belief functions, nor by plausibility functions of Shafer [33]. To see it, first recall that a belief function Bel is defined from a non-negative mass function $m : 2^S \to [0, 1]$, such that $\sum_{E \subseteq S} m(E) = 1$ and $m(\emptyset) = 0$, as follows:

$$\mathrm{Bel}(A) = \sum_{E \subseteq A} m(E).$$

Then, let A, B, C be such that $A \cap (B \cup C) = \emptyset$, suppose $m(B) > m(C) > 0$, $E_1 \subset (A \cup B) \cap \overline{C}$, $E_1 \cap \overline{B} \neq \emptyset$, $m(E_1) > 0$, and $E_2 \subset (A \cup C) \cap \overline{B}$, $E_2 \cap \overline{C} \neq \emptyset$, $m(E_2) > 0$, but $m(E) = 0$ for $E \notin \{B, C, E_1, E_2\}$ (see Fig. 1).

Then

$$Bel(B) = m(B) > Bel(C) = m(C) \quad \text{and}$$

$$Bel(A \cup B) = m(E_1) + m(B),$$

$$Bel(A \cup C) = m(E_2) + m(C).$$

It is easy to choose $m(E_1)$ and $m(E_2)$ such that $m(E_2) + m(C) > m(E_1) + m(B)$, and then $Bel(A \cup B) < Bel(A \cup C)$.

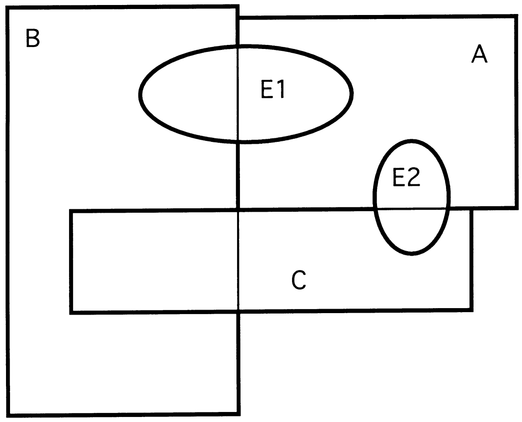

Fig. 1. Belief functions do not satisfy **WDM**.

Plausibility functions Pl do not satisfy **WDM** either. Indeed, $Pl(A) = 1 - Bel(\overline{A}) = \sum_{E \cap A \neq \emptyset} m(E)$. Let

$$m(B \cap \overline{C}) = m_1 > 0, \ m(C \cap \overline{B}) = m_2 > 0,$$

$$m(A \cup (B \cap \overline{C})) = m_3 > 0, m(A \cup (C \cap \overline{B})) = m_4 > 0.$$

Then, $Pl(B) = m_1 + m_3, Pl(C) = m_2 + m_4$. Now,

$$Pl(A \cup B) = m_1 + m_3 + m_4,$$

$$Pl(A \cup C) = m_2 + m_3 + m_4.$$

Clearly it is easy to have $Pl(B) > Pl(C)$ while $Pl(A \cup B) < Pl(A \cup C)$.

In fact belief functions and plausibility functions can represent all orderings that are such that [41]:

- $A \subseteq B \Rightarrow B \unrhd A$ (monotonicity);
- if $C \subseteq B$ and $A \cap B = \emptyset$ then

  $B \rhd C \Rightarrow A \cup B \rhd A \cup C$ (**Bel**);
- if $C \subseteq B$ and $A \cup B = S$ then

  $B \rhd C \Rightarrow A \cap B \rhd A \cap C$ (**Pl**).

Clearly axiom **Bel** implies **WDM** only if $C \subseteq B$, and the same holds for **Pl**, considering the alternative form of **WDM**.

If we want the uncertainty relation $\unrhd$ to be a pseudo-additive measure, we may strengthen axiom **WI** in the following way:

PA *(Pseudo-additivity)*. Let $\mathbf{f}, \mathbf{g}, \mathbf{h}, \mathbf{h}' \in \mathscr{F}$, let $A \subseteq S$. If $\mathbf{f}A\mathbf{h} \succeq \mathbf{g}A\mathbf{h}$ and $\mathbf{h}'$ is pointwisely better than $\mathbf{h}$ then $\mathbf{f}A\mathbf{h}' \succeq \mathbf{g}A\mathbf{h}'$.

Theorem 3. *Let $\succeq$ be an order relation satisfying* **Sav 1**, **PA**, **WS3**, **Sav 5**. *Then, the uncertainty relation $\trianglerighteq$ based on $\succeq$ is a pseudo-additive measure.*

Proof. We have to prove *DM*. Let B, C, D be such that $D \cap (B \cup C) = \emptyset, \mathbf{f} = 1B0, \mathbf{g} = 1C0, \mathbf{f}' = \mathbf{f}$ and $\mathbf{g}' = \mathbf{g}$ on $B \cup C$, and $\mathbf{f}' = \mathbf{g}' = 1D0$ on $\overline{B \cup C}$, so, $\mathbf{f}' \succeq \mathbf{f}$ on $\overline{B \cup C}$, and by **PA**, we get

$$\begin{aligned} \mathbf{f} \succeq \mathbf{g} &\Longleftrightarrow \mathbf{f}(B \cup C)0 \succeq \mathbf{g}(B \cup C)0 \\ &\Rightarrow \mathbf{f}(B \cup C)\mathbf{f}' \succeq \mathbf{g}(B \cup C)\mathbf{g}' \\ &\Longleftrightarrow 1(D \cup B)0 \succeq 1(D \cup C)0 \end{aligned}$$

(i.e. $B \trianglerighteq^{10} C \Rightarrow D \cup B \trianglerighteq^{10} D \cup C$). $\square$

If we change **PA** into its dual axiom **DPA**, we now ensure the satisfaction of **DDM** for $\trianglerighteq$.

DPA *(Dual pseudo-additivity)*. Let $\mathbf{f}, \mathbf{g}, \mathbf{h}, \mathbf{h}' \in \mathscr{F}$, let $A \subseteq S$. If $\mathbf{f}A\mathbf{h} \succeq \mathbf{g}A\mathbf{h}$ and $\mathbf{h}$ is pointwisely better than $\mathbf{h}'$ then $\mathbf{f}A\mathbf{h}' \succeq \mathbf{g}A\mathbf{h}'$.

6.1.3. Qualitative possibility theory

We could base our decision theory on these axioms, choosing to represent uncertainty by pseudo-additive measures, or their dual, or even weaker measures such as Sugeno measures. See [16,23] for decision-theoretic foundations of Sugeno integrals in the style of von Neumann and Morgenstern or Savage, respectively.

In this paper, stronger axioms than **PA** or **DPA** are used so as to recover the "possibilistic" qualitative utilities. First let us recover possibility and necessity measures. To do so, we prove that the characteristic act-based axiom of the former is the optimism axiom, and the pessimism axiom for the latter.

Lemma 5. *Under* **Sav 1**, **WS3** *and* **Sav 5**, *the pessimism Axiom* **PES** *implies that if an act* $\mathbf{f}$ *can be improved by a suitable modification when A occurs, then there is no way of improving* $\mathbf{f}$ *by any other modification when its contrary occurs, namely* $\forall \mathbf{f}, \mathbf{g},\ \forall A \subseteq S,\ \mathbf{g}A\mathbf{f} \succ \mathbf{f} \Rightarrow \mathbf{f} \succeq \mathbf{f}A\mathbf{h}$, *for any act* $\mathbf{h}$.

Proof. The pessimism axiom reads $\forall \mathbf{f}, \mathbf{g},\ \forall A \subseteq S$, $\mathbf{g}A\mathbf{f} \succ \mathbf{f} \Rightarrow \mathbf{f} \succeq \mathbf{f}A\mathbf{g}$. Now $\mathbf{g}A\mathbf{f} \succ \mathbf{f} \Rightarrow 1A\mathbf{f} \succ \mathbf{f}$ using pointwise preference. Now suppose $\mathbf{f}A\mathbf{h} \succ \mathbf{f}$ for some act $\mathbf{h}$. Again, $\mathbf{f}A1 \succ \mathbf{f}$, for the same reason. However, the pessimism axiom forbids that both $\mathbf{f}A1 \succ \mathbf{f}$, and $1A\mathbf{f} \succ \mathbf{f}$ hold. $\square$

As recalled above in Section 2, necessity and possibility measures satisfy, respectively, the two following axioms, also stronger than **WDM** [5]:

N. $B \trianglerighteq C \Rightarrow B \cap A \trianglerighteq C \cap A$.

Pos. $B \trianglerighteq C \Rightarrow B \cup A \trianglerighteq C \cup A$.

Lemma 6. *Under Axioms* **A1**, **A2**, **A3**$'$, *Axiom* **N** *is equivalent to the conjunction of the two following properties:*
Monotonicity: $A \subseteq B \Rightarrow B \trianglerighteq A$,
$B \cap C \sim C$ *or* $B \cap C \sim B$.

Proof. To see that **N** implies monotonicity, just use $\mathbf{A3}' = S \trianglerighteq A$ and assume $B = A \cup C$. Then, **N** implies $B \trianglerighteq A$. Now, let $A = C$ in **N**. It then reads $B \trianglerighteq A \Rightarrow B \cap A \trianglerighteq A$. Since $\trianglerighteq$ is monotonic, $B \cap A \sim A$. Then $B \cap A \sim A$ or $B \cap A \sim B$ follows from the fact that either $B \trianglerighteq A$ or $A \trianglerighteq B$.

Conversely, if $B \cap C \sim B$ or $B \cap C \sim C$ and monotonicity holds then suppose $B \trianglerighteq C$. By assumption, $B \cap A \sim A$ or $B \cap A \sim B$, and $C \cap A \sim A$ or $C \cap A \sim C$. If $B \cap A \sim A$ and $C \cap A \sim A$ or if $B \cap A \sim B$ and $C \cap A \sim C$ then $B \cap A \trianglerighteq A \cap C$ trivially. If $B \cap A \sim A$ and $C \cap A \sim C$ then $B \cap A \sim A \trianglerighteq A \cap C$. If $B \cap A \sim B$ and $C \cap A \sim C$ then $B \cap A \sim B \trianglerighteq C \trianglerighteq A \cap C$. $\square$

A similar lemma holds for Axiom **Pos**, which, under **A1**–**A3**, is equivalent to the conjunction of the two properties: monotonicity and the disjunction $A \sim A \cup B$ or $B \sim A \cup B$.

The above lemma makes it clear that a set-function that satisfies **N** is a necessity measure since $\sigma(A \cap B) \leqslant \min(\sigma(A), \sigma(B))$ due to monotonicity and $\sigma(A \cap B) = \sigma(A)$ or $\sigma(B)$. Similarly, a set function that satisfies **Pos** is a possibility measure, since $\sigma(A \cup B) \geqslant \max(\sigma(A), \sigma(B))$ due to monotonicity and $\sigma(A \cup B) = \sigma(A)$ or $\sigma(B)$. Now, thanks to the above lemmas we prove:

Theorem 4 (Representation of necessity measures). *If the set of acts $\mathscr{F} = X^S$ is equipped with a preference relation $\succeq$ that satisfies* **Sav1**, **WS3**, **Sav5** *and the pessimism Axiom* **PES** *then the uncertainty relation induced by restricting to binary acts with fixed consequences is representable by and only by a necessity measure.*

Proof. Let $\mathbf{f} = 1B0$, $\mathbf{g} = 1C0$ and $\mathbf{h} = 1D0$. Then $\mathbf{g}A\mathbf{f} \succ \mathbf{f}$ reads $(A \cap C) \cup (\overline{A} \cap B) \rhd B$, and $\mathbf{f} \succeq \mathbf{f}A\mathbf{h}$ for any act $\mathbf{h}$, reads $B \unrhd (A \cap B) \cup (\overline{A} \cap D)$, for any event D. In particular, letting $C = A$ and $D = \overline{A}$, the former reads $A \cup B \rhd B$ and the latter reads $B \unrhd \overline{A} \cup B$. Using Lemma 5, the pessimism axiom induces the following property for the uncertainty relation:

$$A \cup B \rhd B \Rightarrow B \unrhd \overline{A} \cup B.$$

Let $E = A \cup B$ and $F = \overline{A} \cup B$. Then $B = E \cap F$ and the property reads: $E \rhd E \cap F$ implies $E \cap F \unrhd F$. But since $\unrhd$ is monotonic, $F \unrhd E \cap F$ and we find that either $F \sim E \cap F$ or $E \cap F \sim F$. But, due to Lemma 6, this axiom, along with monotonicity, is equivalent to the one of comparative necessity measures **N** which are characteristic of necessity measures only. □

Of course, a similar theorem holds for representing possibility measures which are the way uncertainty on events is captured in terms of preference between acts, under **Sav 1**, **WS3**, **Sav 5** and the optimism Axiom **OPT**.

6.2. A representation theorem for qualitative utility

Finally we can propose representation theorems for the qualitative possibilistic utilities introduced in Section 3 of this paper. As shown below the key axioms to be added now are **RDD** and **RCD**, which ensure the semi-decomposability of the utilities, and lead to the maxmin or minmax structure.

Theorem 5 (Representation of the qualitative pessimistic utility). *Let $\succeq$ be a preference relation over the set $\mathscr{F}$ of all acts* **f** *from S to X, satisfying* **Sav1**, **WS3**, **Sav5**, **PES** *and* **RDD**. *Then there exists a finite qualitative scale L, a utility function μ from X to L, a possibility distribution π on S, also taking its values on L, and a utility function v_* with values in L such that: $\mathbf{f} \succeq \mathbf{f}' \iff v_*(\mathbf{f}) \geqslant v_*(\mathbf{f}')$. Moreover v_* can be chosen of the form*

$$v_*(\mathbf{f}) = \min_{s \in S} \max(n(\pi(s)), \mu(\mathbf{f}(s)))$$

on X, where n is an order-reversing map on L.

In order to prove the theorem more easily we need the following lemmas, which use the conjunction $\mathbf{f} \wedge \mathbf{g}$ and the disjunction $\mathbf{f} \vee \mathbf{g}$ of two acts, introduced in Section 5 before Lemma 1.

Lemma 7. *Assume* **Sav1**, **WS3**, **Sav5**, *and* **PES**. *If $\mathbf{h} = \mathbf{f} \wedge \mathbf{g}$, then $\mathbf{h} \sim \mathbf{f}$ or $\mathbf{h} \sim \mathbf{g}$.*

Proof. Let **f** and **g** be any two acts and $\mathbf{h} = \mathbf{f} \wedge \mathbf{g}$. Let $A = \{s, \mathbf{f}(s) >_P \mathbf{g}(s)\}$. Then $\mathbf{f} = \mathbf{f}A\mathbf{h}$ and $\mathbf{g} = \mathbf{h}A\mathbf{g}$. It holds that $\mathbf{g} \geqslant_P \mathbf{h}$ and $\mathbf{f} \geqslant_P \mathbf{h}$. Hence $\mathbf{f} \succeq \mathbf{h}$ and $\mathbf{g} \succeq \mathbf{h}$ by Lemma 3. Assume both $\mathbf{f} \succ \mathbf{h}$ and $\mathbf{g} \succ \mathbf{h}$ hold. It reads $\mathbf{f}A\mathbf{h} \succ \mathbf{h}$ and $\mathbf{h}A\mathbf{g} \succ \mathbf{h}$, which is impossible due to **PES**. □

Lemma 8. *Assume* **Sav1**, **WS3**, **Sav5**, *and* **RDD**. *If $\mathbf{h} = \mathbf{f} \vee \mathbf{x}$, where* **x** *is a constant act with value x, then $\mathbf{h} \sim \mathbf{f}$ or $\mathbf{h} \sim \mathbf{x}$.*

Proof. Axiom **RDD** says that $\mathbf{f} \succ \mathbf{g}$ and $\mathbf{f} \succ \mathbf{x}$ imply $\mathbf{f} \succ \mathbf{g} \vee \mathbf{x}$. But due to the other axioms, $\mathbf{h} = \mathbf{f} \vee \mathbf{x} \succeq \mathbf{f}$ and $\mathbf{h} \succeq \mathbf{x}$ (pointwise dominance). Suppose both $\mathbf{h} \succ \mathbf{f}$ and $\mathbf{h} \succ \mathbf{x}$ hold. Then, by **RDD**, $\mathbf{h} \succ \mathbf{h}$, which is impossible. Hence $\mathbf{h} \sim \mathbf{f}$ or $\mathbf{h} \sim \mathbf{x}$. □

Lemma 8 is also a consequence of Lemma 3. Now we can prove the representation theorem. In Section 3, we proved that the pessimistic utility does satisfy the axioms **Sav 1**, **WS3**, **Sav 5**, **RDD** and **PES**. The other direction of the proof, that is, any preference on acts obeying these axioms can be represented by a pessimistic qualitative utility, is done in four steps:

1. *Building a utility scale.* From **Sav 1**, we know that the set of acts $(\mathscr{F} = X^S, \succeq)$ is a complete preorder. Since X and S are finite, it can be

structured into a linearly order set of equivalence classes $\mathscr{F}/\sim$, that can be bijectively mapped in a finite linearly ordered scale L with a least element denoted 0 and a greatest one, denoted 1, called a utility scale. To each act **f** the image of the equivalence class [**f**] in L is called the utility of **f** and is denoted $v_*(\mathbf{f})$. Considering a constant act **x** with value x, we define the value function μ over X by $\mu(x) = v_*(\mathbf{x})$. Due to pointwise preference $\mu(x_*) = 0$ and $\mu(x^*) = 1$.
2. *Building a qualitative possibility distribution on states.* We then consider the uncertainty relation $\trianglerighteq$ induced by the restriction of $\succeq$ on $\mathscr{F}_{10}$ the set of binary acts of the form $1A0$. From Theorem 4 we know that this is a necessity relation. Hence the utility $v_*(1A0)$ of such acts when A varies defines a necessity measure N, such that $N(A) = v_*(1A0)$. Let n be the order-reversing map in L. Then the function π from S to L defined by $\pi(s) = n(v_*(1(S \setminus \{s\})0))$ is the possibility distribution associated to N (such that $N(A) = \inf_{s\in\overline{A}} n(\pi(s)))$.
3. *Computation of the utilities of binary acts of the form xAy.* Consider an act $1Ax$. It can be written as a disjunction $1A0 \vee \mathbf{x}$. From Lemma 8, $v_*(1Ax) = N(A)$ or $\mu(x)$. But pointwise preferences $1Ax \geqslant_P \mathbf{x}$ and $1Ax \geqslant_P 1A0$ imply $v_*(1Ax) \geqslant \max(N(A), \mu(x))$. Hence $v_*(1Ax) = \max(N(A), \mu(x))$. Now any binary act of the form xAy with $x \geqslant_P y$ is of the form $1Ay \wedge \mathbf{x}$. Using Lemma 7, and a similar reasoning as above, it is obvious that the utility is conjunctively decomposable, and that

$$v_*(xAy) = \min(v_*(1Ay), \mu(x)) = \min(\max(N(A), \mu(y)), \mu(x)),$$

and more generally $v_*(\mathbf{f} \wedge \mathbf{g}) = \min(v_*(\mathbf{f}), v_*(\mathbf{g}))$.
4. *Computing the utility of any act.* We finally extend the computation of the utility function v_* to the whole set of acts X^S, and prove that $v_*(\mathbf{f}) = \min_{s\in S} \max(n(\pi(s)), \mu(\mathbf{f}(s)))$. Any act can be written as a conjunction

$$\mathbf{f} = \wedge_{s\in S} 1(S \setminus \{s\})\mathbf{f}(s).$$

From the above calculation,

$$v_*(1(S \setminus \{s\})\mathbf{f}(s)) = \max(N(S \setminus \{s\}), \mu(\mathbf{f}(s))) = \max(n(\pi(s)), \mu(\mathbf{f}(s))).$$

Then just apply conjunctive decomposability to get the result.

In a similar way one can easily prove the dual result pertaining to the optimistic utility:

Theorem 6 (Representation of the qualitative optimistic utility). *Let $\succeq$ be a preference relation over the set $\mathscr{F}$ of all acts* **f** *from* S *to* X*, satisfying* **Sav1**, **WS3**, **Sav5**, **OPT** *and* **RCD**. *Then there exists a finite qualitative scale L, a utility function μ from X to L and a possibility distribution π on S, also taking its values on L, and a utility function v^* with values in L such that:* $\mathbf{f} \succeq \mathbf{f}' \iff v^*(\mathbf{f}) \geqslant v^*(\mathbf{f}')$. *Moreover v^* can be chosen of the form* $v^*(\mathbf{f}) = \max_{s\in S} \min(\pi(s), \mu(\mathbf{f}(s)))$.

Proof. The only differences with the previous proof are as follows:
- *Building a qualitative possibility distribution on states.* The uncertainty relation $\trianglerighteq$ induced by the restriction of $\succeq$ on $\mathscr{F}_{10}$ the set of binary acts of the form $1A0$ is a possibility relation due to **OPT**. Hence the utility $v^*(1A0)$ of such acts when A varies defines a possibility measure Π, such that $\Pi(A) = v^*(1A0)$. Then the function π from S to L defined by $\pi(s) = v^*(1\{s\}0)$ is the possibility distribution associated to Π (such that $\Pi(A) = \max_{s\in A} \pi(s)$).
- *Computation of the utilities of binary acts of the form xA0.* Consider an act $xA0$. It can be written as a conjunction $1A0 \wedge \mathbf{x}$. From axiom RCD, $v^*(xA0) = \Pi(A)$ or $\mu(x)$. But pointwise preference $xA0 \leqslant_P \mathbf{x}$ and $xA0 \leqslant_P 1A0$ implies $v^*(xA0) \leqslant \min(\Pi(A), \mu(x))$. Hence $v^*(xA0) = \min(\Pi(A), \mu(x))$. Using the axiom **OPT**, and a similar reasoning as above, it is obvious that the optimistic utility is disjunctively decomposable, and that $v^*(\mathbf{f} \vee \mathbf{g}) = \max(v^*(\mathbf{f}), v^*(\mathbf{g}))$.
- *Any act can be written as a disjunction* $\mathbf{f} = \vee_{s\in S}\mathbf{f}(s)\{s\}0$. From the above calculation, $v^*(\mathbf{f}(s)\{s\}0) = \min(\pi(s), \mu(\mathbf{f}(s)))$. Then just apply disjunctive decomposability to get the result. □

This theorem appears in a different, not act-driven form, in [3] with a mathematical justification of the possibility of a fuzzy event as a special case of Sugeno integral. The above act-driven construction can indeed be generalized so as to show that general Sugeno integrals also qualify as utility functions [16]. See also [23] for an alternative construct based on fuzzy lotteries. The results heavily rely on semi-decomposability of Sugeno integrals with respect to conjunction and disjunction of acts one of which being a constant one, or alternatively on comonotonic acts. Such representation results come close to already existing characterizations of Sugeno integrals [3,4,27] that they put in a decision-theoretic perspective. They also have counterparts in the study of aggregation techniques in multicriteria decision making [28].

7. Concluding remarks

One strong assumption has been made in this paper, which is that uncertainty levels and utility levels are commensurate. This is already a consequence of the first axiom of Savage. An attempt to relax this assumption has been made in [8]. These authors point out that working without the commensurability assumption leads to a decision method based on uncertainty representations connected to non-monotonic reasoning. Unfortunately, that method also proves to be either very little decisive or to lead to very risky decisions. On the contrary decisions made on the basis of possibilistic utilities, especially the pessimistic one, sound very reasonable. The latter is a mild extension of the Wald criterion, that recommends cautiousness over the most plausible consequences of an act. By providing an act-driven axiomatization of possibility and necessity measures, possibility theory ceases to be a purely intuitively plausible construct based on introspection. It becomes an observable assumption that can be checked from the actual behavior of a decision-maker choosing among acts, just like subjective probabilities, after Savage axiomatics. This is why the result of this paper is significant from the point of view of Artificial Intelligence, as laying some foundations for qualitative decision theory.

The failure of the Sure-thing principle in the possibilistic setting implies that the notion of conditional preference of Savage no longer makes sense in such a setting. One may distinguish between hypothetical conditioning and preference revision, which may no longer coincide outside the Savage approach. Hypothetical conditioning means that the preference between acts is studied on a subset A of states, regardless of the plausibilities of non-A states. Handling such a conditioning at the axiomatic level means studying a family of preference relations $\succeq_A$ on X^A for all $A \subseteq S$, and directly represent them in terms of conditional utilities [26]. Preference revision means that some states become impossible, that is $\Pi(\bar{A}) = 0$ is enforced in the preference patterns. It comes down to define conditional preference $(\mathbf{f} \succeq \mathbf{g})_A$ as $\mathbf{f}Ax^* \succeq \mathbf{g}Ax^*$ for the pessimistic criterion. Decision-theoretic justifications of qualitative possibilistic conditioning are a topic for further research.

References

[1] M. Allais, Le comportement de l'homme rationnel devante risque: Critique des postulats et axiomes de l'école américaine, Econometrica 21 (1953) 503–546.

[2] A. Chateauneuf, Decomposable measures, distorted probabilities and concave capacities, in: Proceedings of the Conference on the Foundations of Utility and Risk Theories (FUR-IV), Budapest, 1988, Published in Mathematical and Social Sciences 31 (1996) 19–37.

[3] L.L. de Campos, M.J. Bolaños, Characterization and comparison of Sugeno and Choquet integrals, Fuzzy Sets and Systems 52 (1992) 61–67.

[4] L.M. de Campos, M.T. Lamata, S. Moral, A unified approach to define fuzzy integrals, Fuzzy Sets and Systems 39 (1991) 75–90.

[5] D. Dubois, Belief structures, possibility theory and decomposable confidence measures on finite sets, Computers and Artificial Intelligence 5 (1986) 404–416.

[6] D. Dubois, H. Fargier, H. Prade, Fuzzy constraints in job-shop scheduling, Journal of Intelligent Manufacturing 64 (1995) 215–234.

[7] D. Dubois, H. Fargier, H. Prade, Possibility theory in constraint satisfaction problems: Handling priority, preference and uncertainty, Applied Intelligence 6 (1996) 287–309.

[8] D. Dubois, H. Fargier, H. Prade, Decision-making under ordinal preferences and comparative uncertainty, in: Proceedings of the Uncertainty in AI Conference (UAI97) Providence, RI, 1997, pp. 157–164.

[9] D. Dubois, J.C. Fodor, H. Prade, M. Roubens, Aggregation of decomposable measures with application to utility theory, Theory and Decision 41 (1996) 59–95.
[10] D. Dubois, L. Godo, H. Prade, A. Zapico, Possibilistic representation of qualitative utility: An improved characterization, in: Proceedings of the International Conference on Information Processing and Management of Uncertainty (IPMU'98), Paris, July 1998, pp. 180–187.
[11] D. Dubois, H. Prade, Fuzzy Sets and Systems: Theory and Applications, Academic Press, New York, 1980.
[12] D. Dubois, H. Prade, A class of fuzzy measures based on triangular norms, International Journal of General Systems 8 (1982) 225–233.
[13] D. Dubois, H. Prade, Possibility Theory, Plenum Press, New York, 1988.
[14] D. Dubois, H. Prade, Possibility theory as a basis for qualitative decision theory, in: Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI'95), Montreal, 20–25 August 1995, pp. 1925–1930.
[15] D. Dubois, H. Prade, Possibility theory: Qualitative and quantitative aspects. in: D.M. Gabbay, Ph. Smets (Eds.), Handbook of Defeasible Reasoning and Uncertainty Management Systems, vol. I, Kluwer Academic Publishers, Netherlands, pp. 169–226.
[16] D. Dubois, H. Prade, R. Sabbadin, Qualitative decision theory with Sugeno integrals, in: Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI'98), Madison, WI, USA, 1998, pp. 121–128.
[17] D. Ellsberg, Risk, ambiguity and the Savage axioms, Quarterly Journal of Economics 75 (1961) 643–669.
[18] P. Gärdenfors, Knowledge in Flux Modeling the Dynamics of Epistemic States, MIT Press, Cambridge, MA, 1988.
[19] I. Gilboa, Expected utility with purely subjective non-additive probabilities, Journal of Mathematical Economics 16 (1987) 65–88.
[20] M.M. Grabisch, H.T. Nguyen, E.A. Walker, Fundamentals of Uncertainty Calculi with Applications to Fuzzy Inference, Kluwer Academic Publishers, Dordrecht, 1995.
[21] M. Grabisch, T. Murofushi, M. Sugeno, Fuzzy measure of fuzzy events defined by fuzzy integrals, Fuzzy Sets and Systems 50 (1992) 293–313.
[22] S. Grant, A. Kajii, B. Polak, Weakening the sure-thing principle: Decomposable choice under uncertainty, Workshop on Decision Theory, Chantilly, France, Department of Economics, Australian National University, June 1997.
[23] J.L. Hougaard, H. Keiding, Representation of preferences on fuzzy measures by a fuzzy integral, Mathematical and Social Sciences 31 (1996) 1–17.
[24] M. Inuiguichi, H. Ichihashi, H. Tanaka, Possibilistic linear programming with measurable multiattribute value functions, ORSA Journal on Computing 1 (1989) 146–158.
[25] C.H. Kraft, J.W. Pratt, A. Seidenberg, Intuitive probability on finite sets, Annals of Mathematical Statistics 30 (1959) 408–419.
[26] D. Lehmann, Generalized qualitative probability : Savage revisited, in: Proceedings of the 12th Uncertainty in Artificial Intelligence (UAI 96), Portland, 1996, Morgan Kaufmann, pp. 381–388.
[27] D. Ralescu, M. Sugeno, Fuzzy integral representation, Fuzzy Sets and Systems 84 (1996) 127–133.
[28] J.-L. Marichal, Aggregation operators for multicriteria decision aid, Ph.D. Thesis, Université de Liège, Belgium, December 1998.
[29] R. Sarin, P.P. Wakker, A simple axiomatization of nonadditive expected utility, Econometrica 60 (6) (1992) 1255–1272.
[30] L.J. Savage, The Foundations of Statistics, Wiley, New York, 1954.
[31] D. Schmeidler, Subjective probability and expected utility without additivity, Econometrica 57 (1989) 571–587.
[32] B. Schweizer, A. Sklar, Probabilistic Metric Spaces, North-Holland, Amsterdam, 1983.
[33] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, NJ, 1976.
[34] W. Spohn, Ordinal conditional functions: A dynamic theory of epistemic states, in: W. Harper, B. Skyrms (Eds.), Causation in Decision, Belief Change and Statistics, Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 105–134.
[35] M. Sugeno, Fuzzy measures and fuzzy integrals-A survey, in: M.M. Gupta, G.N. Saridis, B.R. Gaines (Eds.), Fuzzy Automata and Decision Processes, North-Holland, Amsterdam, 1977, pp. 89–102.
[36] J. von Neumann, O. Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, Princeton, NJ, 1944.
[37] A. Wald, Statistical Decision Functions, Wiley, New York, 1950.
[38] Z. Wang, G.J. Klir, Fuzzy Measure Theory, Plenum Press, New York, 1992.
[39] S. Weber, $\perp$-Decomposable measures and integrals for Archimedean t-conorms $\perp$, Journal of Mathematical Analytical Applications 101 (1984) 114–138.
[40] T. Whalen, Decision making under uncertainty with various assumptions about available information, IEEE Transactions on Systems, Man and Cybernetics 14 (1984) 888–900.
[41] S.K.M. Wong, Y.Y. Yao, P. Bollmann, H.C. Burger, Axiomatization of qualitative belief structure, IEEE Transactions on Systems, Man and Cybernetics 21 (4) (1991) 726–734.
[42] R.R. Yager, Possibilistic decision making, IEEE Transactions on Systems, Man and Cybernetics 9 (1979) 388–392.
[43] R.R. Yager, On the specificity of a possibility distribution, Fuzzy Sets and Systems 50 (1992) 279–292.
[44] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1 (1978) 3–28.