Qualitative Decision Theory:
from Savage's Axioms to Non-Monotonic Reasoning

Didier Dubois, Hélène Fargier, Patrice Perny[‡], Henri Prade

Institut de Recherche en Informatique de Toulouse
Université Paul Sabatier – 118 route de Narbonne
31062 Toulouse Cedex 4 – France
Email: {dubois, fargier, prade}@irit.fr

[‡]LIP6 - POLE IA
University Paris VI
8 rue du Capitaine Scott
75015 Paris, France
Email: perny@poleia.lip6.fr

Abstract: This paper investigates to what extent a purely symbolic approach to decision making under uncertainty is possible, in the scope of Artificial Intelligence. Contrary to classical approaches to decision theory, we try to rank acts without resorting to any numerical representation of utility nor uncertainty, and without using any scale on which *both* uncertainty and preference could be mapped. Our approach is a variant of Savage's where the setting is finite, and the strict preference on acts is a partial order. It is shown that although many axioms of Savage theory are preserved and despite the intuitive appeal of the ordinal method for constructing a preference over acts, the approach is inconsistent with a probabilistic representation of uncertainty. The latter leads to the kind of paradoxes encountered in the theory of voting. It is shown that the assumption of ordinal invariance enforces a qualitative decision procedure that presupposes a comparative possibility representation of uncertainty, originally due to Lewis, and usual in nonmonotonic reasoning. Our axiomatic investigation thus provides decision-theoretic foundations to preferential inference of Lehmann and colleagues. However, the obtained decision rules are sometimes either not very decisive or may lead to overconfident decisions, although their basic principles look sound. This paper points out some limitations of purely ordinal approaches to Savage-like decision making under uncertainty, in perfect analogy with similar difficulties in voting theory.

Keywords: Decision theory; preference relations; comparative uncertainty; possibility theory; nonmonotonic reasoning.

1 - Introduction

Traditionally, decision making under uncertainty (DMU) relies on a probabilistic framework. When modeling a decision maker's rational choice between acts, it is assumed that the uncertainty about the state of the world is described by a probability distribution, and that the ranking of acts is done according to the expected utility of the consequences of these acts. This proposal was made by economists in the 1950's, and justified on an axiomatic basis by Savage (1972) and colleagues. More recently, in Artificial Intelligence, this setting has been applied to problems of planning under uncertainty, and is at the root of the influence diagram methodology for multiple stage decision problems.

However, in parallel to these developments, Artificial Intelligence has witnessed the emergence of a new decision paradigm called qualitative decision theory, where the rationale for choosing among decisions no longer relies on probability theory nor numerical utility functions. Papers by Brafman and Tennenholtz (1997) and Doyle and Thomason (1999) stress the main issues in this area. Motivations for this new proposal are twofold. On the one hand there exists a tradition of symbolic processing of information in Artificial Intelligence, and it is not surprising this tradition should try and stick to symbolic approaches when dealing with decision problems. On the other hand, the emergence of new information technologies has generated many new decision problems involving intelligent agents like information systems or autonomous robots.

An information system is supposed to help an end-user retrieve information and choose among courses of action, based on a limited knowledge of the user needs. It is not clear that numerical approaches to DMU, that were developed in the framework of economics, are fully adapted to these new problems. Expected utility theory might sound too demanding a tool for handling queries of end-users. Numerical utility functions and subjective probabilities presuppose a rather elaborate elicitation process that is worth launching for making strategic decisions that need to be carefully analyzed. Users of information systems are not necessarily capable of describing their state of uncertainty by means of a probability distribution, nor may they be willing to quantify their preferences (Boutilier, 1994). This is typical of electronic commerce, or recommender systems 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. Moreover the expected utility criterion makes full sense for repeated decisions whose successive results accumulate (for instance money in gambling decisions). Some decisions made by end-users are rather one-shot, in the sense that getting a wrong advice one day cannot always be compensated by a good advice the next day. Note that this kind of application often needs multiple-criteria decision-making rather than DMU. However there is a strong similarity between the two problems. Besides, there is no need for a precise modelling of information in the handling of many queries to recommender systems, and qualitative information is sometimes good enough, because little is at stake with such queries.

In the case of autonomous robots, conditional plans are often used to monitor the robot behavior, and the environment of the robots are sometimes only partially observable. The theory of partially observable Markov decision processes leads to highly complex methods, because handling infinite state spaces. A qualitative, finitistic, description of the goals of the robot, and of its knowledge of the environment might lead to more tractable methods (Sabbadin, 2000, for instance). Besides, the expected utility criterion is often adopted because of its mathematical properties (it enables dynamic programming principles to be used). However it is not clear that this criterion is always the most cogent one, for instance in risky environment, where cautious policies should be followed.

It is thus tempting to try and define decision rules that would be simple to implement and do not require from the user numerical information he/she is not able to provide. Formulating decision problems in a symbolic way may also be more compatible with a declarative expression of uncertainty and preferences in the setting of some logic-based language (Boutilier, 1994; Thomason 2000).

So, there is a need for qualitative decision rules. However there is no real agreement on what "qualitative" means. Some authors assume incomplete knowledge about classical additive utility models, whereby the utility function is specified via symbolic constraints (Lang, 1996; Bacchus and Grove, 1996 for instance). Others use sets of integers and the like to describe rough probabilities or utilities (Tan and Pearl, 1994). Lehmann (1996) injects some qualitative concepts of negligibility in the classical expected utility framework. However some approaches are genuinely qualitative in the sense that they do not involve any form of quantification. Boutilier (1994) exploits preferential inference of nonmonotonic reasoning and makes decisions on the basis of most plausible states of nature. Brafman & Tenenholtz (1996) adopt a pessimistic attitude to decision making, and cast the maximin criterion in a logical setting. Dubois and Prade (1995c) generalize the maximin criterion in the setting of possibility theory and propose a Von Neumann and Morgenstern style axiomatisation. However, none of these approaches are in agreement with the most classical of decision theories, i.e. Savage's. The natural question is then whether it is possible to found rational decision making in a purely qualitative setting, without moving too far away from Savage.

We take it for granted that a qualitative decision theory is one that does not resort to the full expressive power of numbers for the modeling of uncertainty not for the representation of utility. Two lines can be followed in agreement with this definition: the relational approach and the absolute approach. Following the relational approach, the decision maker uncertainty is represented by a partial ordering relation among events (expressing relative likelihood), and the utility function is just encoded as another ordering relation between potential consequences of decisions. This approach is close to the framework of voting theories after Arrow (1951), Sen (1986) etc. The advantage is that it is faithful to the kind of elementary information users can directly provide. The other approach (Dubois et al. 2001) presupposes the existence of a totally ordered scale (typically a finite one) for grading both likelihood and utility. Decision rules generalizing maximin and maximax criteria can be defined on this ordinal scale. This approach is very simple but has some shortcomings. First, the obtained ranking of decisions is bound to be coarse since there cannot be more classes of preference-equivalent decisions than levels in the finite scale used. Moreover, the assumption of a common scale for grading uncertainty and preference (we call the commensurability assumption) is strong. It can be questioned, although it is already taken for granted in classical decision theory (via the notion of certainty equivalent of an uncertain event). When only relations are used, one does not make such an assumption, and the expressive power of purely relational approaches is thus worth investigating.

This investigation is carried out in this paper, in the setting of decision theory after Savage. Namely, we address the following question: If an uncertainty relation on events is given, and a preference relation on consequences is independently given as well, how can a preference relation on acts be rationally constructed with a *purely symbolic approach*, without assuming that the uncertainty relation and the preference relation can both be mapped to some unique scale be it numerical or ordinal. The aim of the paper is to lay bare the formal consequences of adopting such a purely ordinal point of view on DMU, while retaining as much as possible from Savage's axioms, and especially the sure thing principle which is the cornerstone of the theory.

Suppose only a relative likelihood relation on events that describes a decision maker's knowledge about the state of the world, and a preference ordering on the consequences of feasible acts are available. A natural, intuitively appealing technique that solves the act-ranking problem is as follows: suppose an act f is preferred to another g if the (relative) likelihood that f outperforms g is greater than the likelihood that g outperforms f. It is called the *likely dominance* rule. In our investigation, Savage's framework is slightly relaxed so as to allow for intransitive indifference between acts. This is necessary in order to avoid trivialization. Moreover we assume that both the set of states and the set of consequences are finite. An axiom, originally due to Fishburn (1975) in another context, is then added. It expresses the purely ordinal nature of the decision criterion, and it is proved to enforce the likely dominance rule.

More precisely, the results of this paper show that:

1) There is no other way of building a preference on acts from an uncertainty relation on events and a preference relation of consequences, in the sense of reasonable assumptions coming from Savage's framework and the requirement of a purely qualitative (ordinal) approach.

2) The decision rule generated by the axiomatic framework is generally incompatible with a probabilistic representation of uncertainty. Indeed, applying it to an ordering on events induced by a probability measure may lead to a non-transitive strict preference relation on acts. The ordinal Savagean framework actually leads to a representation of uncertainty that is closely related to the ordering of formulas in the nonmonotonic logic ("System P") of Kraus, Lehmann and Magidor (1990), also exhibited by Friedman and Halpern (1996).

3) The enforced simple decision criterion is a mixture of lexicographic priority and unanimity. Preferred acts are selected by restricting to the most plausible states of the world, and using a unanimity rule on these maximally plausible states. Ties are broken by lower level oligarchies. However cognitively appealing it may look, it has a limited expressive and decisive power.

The difficulty met by the purely relational approach to the decision problem under uncertainty is very similar to the one faced in voting theories (see for instance

Moulin(1988)) where a priori natural procedures, like the Condorcet pairwise majority rule, exhibit counterintuitive intransitivities. These results question the very possibility of a purely ordinal solution to this problem, in the framework of transitive and complete preference relations on acts. This kind of impediment was already pointed out by Doyle and Wellman (1991) for preference-based default theories.

This paper extends preliminary results that were presented in (Dubois et al., 1997). It is organized as follows. Section 2 recalls some classical decision rules and Savage's framework for DMU. Section 3 presents the likely dominance rule that can rank acts from a purely ordinal knowledge of uncertainty on states, and utility on consequences. This method is similar to the pairwise majority rule in social choice but is much more general. It is shown that this decision rule is basically inconsistent with Savage axioms. An axiomatic framework for ordinal DMU is presented in Section 4. This framework is the one of Savage but the indifference between acts is allowed not to be transitive. An ordinal invariance axiom is introduced and it is proved that it enforces the likely dominance rule. Section 5 studies the type of subjectivist uncertainty theory induced by purely ordinal Savage-like frameworks with or without the ordinal invariance assumption. It shows that a connection exists between Savage's framework adapted to ordinal decision criteria, and nonmonotonic reasoning after Kraus, Lehmann and Magidor (1990). Namely the subjective likelihood relation induced from the decision-making axioms is equivalently represented by a preferential consequence relation. Section 6 exhibits a model of the set of axioms, where uncertainty is represented by a single weak ordering of states, in the framework of comparative possibility theory (Lewis, 1973; Dubois and Prade, 1998). The relation on events induced by the likely dominance expresses relative likelihood and plays the role of a probability measure, but only its strict part is transitive. This type of relation, refines both possibility and necessity relations. A representation theorem for ordinal DMU is then provided in terms of a family of possibility relations. Section 7 studies the type of decision rules generated by the theory. Proofs of the results appear in Appendix 2.

2 – Background on decision theory

A decision problem is cast in the framework of a set S of states (of the world) and a set X of potential consequences of decisions. States encode possible situations, states of affairs, etc. An act is viewed as a mapping f from the state space to the consequence set, namely, in each state $s \in S$, an act f produces a well-defined result $f(s) \in X$. The decision maker must rank acts without knowing what is the current state of the world in a precise way. The consequences of an act can often be ranked in terms of their relative appeal: some consequences are judged better than others. This is often modeled by means of a numerical utility function u which assigns to each consequence $x \in X$ a utility value $u(x) \in \mathbb{R}$. Classically there are two approaches to modeling the lack of knowledge of the decision maker about the state. The most widely found assumption is that there is a probability distribution p on S. It is either obtained from statistics (this is called decision under risk, Von Neumann and Morgenstern, 1953) or it is a subjective probability supplied by the agent via suitable elicitation methods. Then the most usual decision rule is based on the expected utility criterion:

$$EU(f) = \sum_{s \in S} p(s)u(f(s))$$

where by act f is preferred to act g if and only if EU(f) > EU(g). Another proposal is the maximin criterion (Arrow and Hurwicz, 1972). It applies when no information about the current state is available, and it ranks acts according to its worst consequence:

$$W(f) = \min_{s \in S} u(f(s)).$$

This criterion, first proposed by Wald (1950), has the major defect of being extremely pessimistic. In practice it is never used for this reason. An optimistic counterpart of W(f) is obtained by turning minimum into maximum. Hurwicz has proposed to use a weighted average of W(f) and its optimistic counterpart, where the weight bearing on W(f) is viewed as a degree of pessimism of the decision maker. Other decision rules have been proposed, especially some that generalize both EU(f) and W(f); see (Jaffray, 1989) and (Schmeidler 1989), among others. Dubois et al. (1995c, 2001) propose a non-numerical generalization of W(f) on an ordinal scale. However the expected utility criterion is by far the most commonly used one.

Savage (1972) proposed an axiomatic framework for justifying the use of expected utility, that we shall briefly recall here as our ordinal approach to DMU heavily relies on it. Savage is interested in proving that rational decision makers should use this criterion when ranking acts according to their preferences. From a complete and transitive preference relation over acts, Savage derives a likelihood relation representing uncertainty on events in the state space, and a preference relation on the consequences. If the preference on acts satisfies suitable properties, according to which the decision maker is considered to be "rational", then the uncertainty on events can be represented by a unique probability distribution, the preference relation on consequences by a numerical utility function on an interval scale, and acts are ranked according to their expected utility.

Let us recall this framework. It is adapted to our purpose of devising a purely ordinal approach because the starting point of Savage theory is indeed based on relations and their representation on an interval scale. Suppose a decision maker supplies a preference relation ≥ over acts. An act is modeled as a mapping $f: S \to X$ and X^S usually denotes the set of all such mappings. In Savage's approach, any mapping in the set X^S is considered as a possible act (even if it is an imaginary one rather than a feasible one). The first requirement stated by Savage is:

Axiom S1: $(X^S, \geq)$ is a weak order.

Although Savage had his own defense of this axiom, it is unavoidable in the scope of utility theory: If acts are ranked according to expected utility, then the preference over acts will be transitive, reflexive, and complete (f ≥ g or g ≥ f for any f, g).

What this axiom also implies, if X and S are finite, is that there exists a totally ordered scale, say $\mathcal{L}$, that can serve to evaluate the worth of acts. Indeed the indifference relation (f ~ g if and only if f ≥ g and g ≥ f) is an equivalence relation, and the set of equivalence classes, denoted $X^S/\sim$ is totally ordered via the strict preference >. If [f] and [g] denote the equivalence classes of f and g, [f] > [g] holds if and only if f > g holds for any representatives of each class. So it is possible to rate acts on $\mathcal{L}$ =

$X^S/\sim$ and [f] is the qualitative utility level of f.

The set of acts is closed under the following combination involving acts and events. An event is modeled by a subset of states, understood disjunctively. To say that an event A occurs means that one of the states in A is the real state. Events are the same as propositions of the form "$s \in A$". Let $A \subseteq S$ be an event, f and g two acts, and denote by fAg the act such that:

$$fAg(s) = f(s) \text{ if } s \in A, \text{ and } g(s) \text{ if } s \notin A.$$

For instance, f may mean "bypass the city", g mean "cross the city", and A represents the presence of a traffic jam. Then S represents descriptions of the state of the road network, and X represents a time scale for the time spent by an agent who drives to his/her working place. fAg then means: bypass the city if there is a traffic jam, and cross the city otherwise. More generally the notation $f_1A_1f_2A_2, \ldots, A_{n-1}f_nA_n$, where $A_1, \ldots, A_{n-1}A_n$ is a partition of S, denotes the act whose result is $f_i(s)$ if $s \in A_i$, $\forall i = 1, \ldots n$. fAg is actually short for $fAgA^c$ where A^c is the complement of A.

Savage proposed an axiom that he called the sure-thing principle. It requires that the relative preference between two acts does not depend on states where the acts have the same consequences. In other words, the preference between fAh and gAh does not depend on the choice of h:

Axiom S2 (Sure-thing principle): $\forall A, f, g, h, h'$ $\quad fAh \geq gAh$ iff $fAh' \geq gAh'$.

For instance, if you bypass the city (f) rather than cross it (g) in case of a traffic jam (A), this preference does not depend on what you would do in case of fluid traffic (A^c), say, cross the city (h = g), bypass it anyway (h = f) or make a strange decision such as staying at home. Grant et al. (1997) pointed out that the name "sure-thing principle" for this postulate was not fully justified since it is hard to grasp where the sure thing is. Grant et al. propose several expressions of a genuine sure-thing principle (he calls weak sure-thing principle) one version being as follows:

Axiom WSTP: $\quad fAg \geq g$ and $gAf \geq g$ imply $f \geq g$.

The above property really means that the weak preference of f over g does not depend on whether A occurs or not. It is obvious that WSTP is implied by S1 and S2, since from $fAg \geq g = gAg$ and S2 we derive $f = fAf \geq gAf$ and using transitivity of $\geq$ due to S1, $f \geq g$ follows.

The sure-thing principle enables two notions to be simply defined, namely conditional preference and null events.

Definition 1: f is said to weakly preferred to g, conditioned on A if and only if $\forall h, fAh \geq gAh$. This is denoted by $(f \geq g)_A$.

Clearly, the sure-thing principle enables $(f \geq g)_A$ to hold as soon as $fAh \geq gAh$ for some h. Conditional preference $(f \geq g)_A$ means that f is weakly preferred to g when the state space is restricted to A, regardless of the decision made when A does not occur.

Note that $f \geq g$ is short for $(f \geq g)_S$. Moreover $(f \geq g)_\emptyset$ always holds, for any f and g, since it is equivalent to the reflexivity of $\geq$ (i.e., $h \geq h$).

Definition 2: An event A is said to be null if and only if $\forall f, \forall g, (f \geq g)_A$ holds.

Any non-empty set A on which all acts make no difference is then like the empty set: the reason why all acts make no difference is because this event is considered impossible by the decision maker.

Conditional preference enables the weak sure-thing principle to be expressed like a unanimity principle in the terminology of voting theory, provided that the sure-thing principle holds.

Axiom U: $(f \geq g)_A$ and $(f \geq g)_{A^c}$ implies $f \geq g$ (unanimity).

Proposition 1: Under axiom S2, WSTP is equivalent to unanimity.

Note that in the absence of S2, (U) implies (WSTP) but not the converse. The unanimity postulate has been formulated by Lehmann (1996).

Among acts in X^S are *constant acts* such that: $\forall x \in X, \forall s \in S, f(s) = x$. They are denoted fx. It seems reasonable to identify the set of constant acts $\{fx, x \in X\}$ and X. The preference on X can be induced from $(X^S, \geq)$ as follows:

Definition 3: Given $(X^S, \geq)$, the preference relation $\geq_P$ on X is of the form

$\forall x, y \in X$, $x \geq_P y$ if and only if $fx \geq fy$.

This definition is self-consistent provided that the preference between constant acts is not altered by conditioning. This is the third Savage's postulate:

Axiom S3: $\forall A \subseteq S$, A not null, $(fx \geq fy)_A$ if and only if $x \geq_P y$.

An act f is said to *Pareto-dominate* another act g if and only if $\forall s \in S$, $f(s) >_P g(s)$ or $f(s) = g(s)$, which is denoted $f \geq_P g$. Clearly, Pareto-dominance should imply weak preference for acts. And indeed under S1, S2, and S3, $f \geq_P g$ implies $f \geq g$.

The preference on acts also induces a likelihood relation among events. For this purpose, it is enough to consider the set of binary acts, of the form $f_x A f_y$, which due to (S3) can be denoted xAy, where $x \in X$, $y \in X$, and $x >_P y$. Clearly for fixed $x >_P y$, the set of acts $\{x,y\}^S$ is isomorphic to the set of events 2^S. However the restrictions of $(X^S, \geq)$ to $\{x, y\}^S$ may be inconsistent with the restriction to $\{x', y'\}^S$ for other choices of consequences $x' >_P y'$. A partial ordering among events can however be recovered, as suggested by Lehmann (1996):

Definition 4: Relative likelihood $\geq_L$:

$\forall A, B \subseteq S$, $A \geq_L B$ if and only if $xAy \geq xBy$, $\forall x, y \in X$ such that $x >_P y$.

In order to get a weak order of events, Savage introduced a new postulate:

Axiom S4: $\forall x, y, x', y' \in X$ s.t. $x >_P y$, $x' >_P y'$, $xAy \geq xBy \Rightarrow x'Ay' \geq x'By'$.

Under this property, the choice of $x, y \in X$ with $x >_P y$ does not affect the ordering between events in terms of binary acts, namely: $A \geq_L B$ is short for $\forall x >_P y$, $xAy \geq xBy$.

Lastly, Savage assumed that the set X is not trivial:

Axiom S5: X contains at least two elements x, y such that fx > fy (or $x >_P y$).

Under S1-S5, the likelihood relation on events is a comparative probability ordering. Savage introduces another postulate that enables him to derive the existence (and uniqueness) of a numerical probability measure on S that can represent the likelihood relation $\geq_L$. This axiom reads:

Axiom S6: For any f, g with f > g in X^S and any $x \in X$, there is a partition $\{S_1, \ldots, S_n\}$ of S such that $\forall i = 1, \ldots n$, $xS_if > g$ and $f > xS_ig$.

Under the postulates S1-S6, not only can $\geq_L$ be represented by a numerical probability function but $(X^S, \geq)$ can be represented by the expected utility of acts $u(f) = \sum_{s \in S} p(s)u(f(s))$ where the utility function u represents the relation $\geq_P$ on X uniquely, up to a linear transformation. However S6 presupposes that the state space S is infinite so that the probability of S_i can be made arbitrary small, thus not altering the relation f > g when x is very bad (so that $xS_if > g$) or very good (so that $f > xS_ig$). In contrast, we assume that both S and X are finite in this paper, and S6 is trivially violated in such a finite setting. In that case, it is well-known that the axioms of comparative probability do not ensure that the likelihood relation can always be represented by a probability (see Fishburn, 1986). This feature does not matter here since our aim is to remain at the ordinal level. However, there is to our knowledge no representation of subjective expected utility that would assume a purely finite setting for states and consequences.

3 The Likely Dominance Rule

Savage approach to the justification of expected utility starts from a ranking of acts. However the real decision problem is to build this relation from information regarding the likelihood of states and the decision maker preference on consequences. This Section introduces a natural decision rule that computes such a preference relation on acts from a purely symbolic perspective, no longer assuming a probability function on S is available nor a numerical utility function on X. Moreover we assume the set of states is finite, contrary to the framework of Savage, and we do not use axiom S6.

A "likelihood" relation on the set of events is an irreflexive and transitive relation $>_L$ on 2^S, and a non-trivial one ($S >_L \emptyset$,), faithful to deductive inference. $A >_L B$ means that event A is more likely than B. Given that an event A stands for a proposition "$s \in A$", a deductive inference "A implies B" is modeled by the inclusion $A \subseteq B$. So, if $A \subseteq B$ then $A >_L B$ should not hold (inclusion-monotony). The inclusion-monotony property states that if A implies B, then A cannot be more likely than B. Let us define the weak likelihood relation $\geq_L$ induced from $>_L$ via complementation, and the indifference relation $\sim_L$ as usual:

$A \geq_L B$ if and only if not $(B >_L A)$; $A \sim_L B$ iff $A \geq_L B$ and $B \geq_L A$.

The preference relation on the set of consequences X is supposed to be a *weak order* (a complete preordering, e.g. Roubens and Vincke, 1985). Namely, $\geq_P$ is a reflexive and transitive relation, and completeness means $x \geq_P y$ or $y \geq_P x$, $\forall x, y \in X$. $x \geq_P y$ means that consequence x is not worse than y. The induced strict preference relation is derived as usual: $x >_P y$ if and only if $x \geq_P y$ and not $y \geq_P x$. It is assumed that X has at least two elements x and y s.t. $x >_P y$. The assumptions pertaining to $\geq_P$ are natural in the scope of numerical representations of utility, however we do not require that the weak likelihood relation be a weak order too.

If the likelihood relation on events and the preference relation on consequences are not comparable, a natural way of lifting the pair $(>_L, \geq_P)$ to X^S is as follows: an act f is more promising than an act g if and only if the event formed by the disjunction of states in which f gives better results than g, is more likely than the event formed by the disjunction of states in which g gives results better than f. A state s is more promising for act f than for act g if and only if $f(s) >_P g(s)$. The following notation is adopted in the paper: while $f >_P g$ stands for: $\forall s, f(s) >_P g(s)$, $[f >_P g]$ is an event made of all states where f outperforms g, that is $[f >_P g] = \{s \in S, f(s) >_P g(s)\}$.

Accordingly, we define the preference between acts ($\geq$), the corresponding indifference ($\sim$) and strict preference ($>$) relations as follows:

Definition 5 (Likely dominance rule) :
$f > g$ if and only if $[f >_P g] >_L [g >_P f]$;
$f \geq g$ if and only if not$(g > f)$ i.e. if and only if $[f >_P g] \geq_L [g >_P f]$;
$f \sim g$ if and only if $f \geq g$ and $g \geq f$.

This approach looks very natural, and is the first one that comes to mind when information is only available under the form of an ordering of events and an ordering of consequences and *when the preference and uncertainty scales are not comparable*. Events are only compared to events, and consequences to consequences. Another definition, that may look as natural, uses the weak preference ordering and is obtained by changing $[f >_P g]$ into $[f \geq_P g]$. However, in the scope of Savage approach, this alternative definition should be equivalent to Definition 5, because of the sure thing principle: the event $[f \sim_P g]$ should not affect the likelihood ordering between two events that both contain it.

The properties of the relations $\geq$, $\sim$, and $>$ on X^S will depend on the properties of $\geq_L$ with respect to Boolean connectives. The most obvious choice for $\geq_L$ is a comparative probability enforced by S1-S5 (e.g., Fishburn, 1986):

Definition 6: $\geq_L$ is a comparative probability if and only if

i) $\geq_L$ is a weak order
ii) $S >_L \emptyset$ (non-triviality),
iii) $\forall A$, $A \geq_L \emptyset$ (consistency)
iv) $A \cap (B \cup C) = \emptyset \Rightarrow (B \geq_L C \Leftrightarrow A \cup B \geq_L A \cup C)$ (preadditivity).

Note that these properties do not ensure a numerical additive representation of $\geq_L$. A first negative result is that if $\geq_L$ is a comparative probability ordering then the strict preference relation $>$ in X^S is not necessarily transitive.

Example 1:

A very classical and simple example of undesirable lack of transitivity is when $S = \{s_1, s_2, s_3\}$ and $X = \{x_1, x_2, x_3\}$ with $x_1 >_P x_2 >_P x_3$, and the comparative probability ordering is generated by a uniform probability on S. Suppose three acts f, g, h such that
$f(s_1) = x_1 >_P f(s_2) = x_2 >_P f(s_3) = x_3$,
$g(s_3) = x_1 >_P g(s_1) = x_2 >_P g(s_2) = x_3$,
$h(s_2) = x_1 >_P h(s_3) = x_2 >_P h(s_1) = x_3$.
Then $[f >_P g] = \{ s_1, s_2 \}$; $[g >_P f] = \{ s_3\}$; $[g >_P h] = \{ s_1, s_3\}$; $[h >_P g] = \{ s_2\}$;
$[f >_P h] = \{ s_1\}$; $[h >_P f] = \{ s_2, s_3\}$.
The likely dominance rule yields $f > g$, $g > h$, $h > f$. Note that the presence of this cycle does not depend on figures of utility that could be attached to consequences insofar as the ordering of utility values is respected for each state. In contrast the ranking of acts induced by expected utility completely depends on the choice of utility values, even if we keep the constraint $u(x_1) > u(x_2) > u(x_3)$. The reader can check that, by symmetry, any of the three linear orders $f > g > h$, $g > h > f$, $h > f > g$ can be obtained by suitably quantifying the utility values of states without changing their preference ranking. Note that the undesirable cycle remains if probabilities $p(s_1) > p(s_2) > p(s_3)$ are attached to states, and the degrees of probability remain close to each other (so that $p(s_2) + p(s_3) > p(s_1)$).

This situation can be viewed as an analog of the Condorcet paradox in social choice, here in the setting of DMU. Indeed the problem of ranking acts on the basis of the likelihood relation $>_L$ and the preference on consequences $\geq_P$ can be cast in the setting of a voting problem (See Moulin, 1988, for an introduction to voting methods). Let V be a set of voters, C be a set of candidates and let $\geq_v$ be a relation on C that represents the preference of voter v on the set of candidates. $\geq_v$ is a weak order, by assumption. The decision method consists in constructing a relation R on C that aggregates $\{\geq_v, v \in V\}$ as follows. Let $V(c_1,c_2) = \{v \in V, c_1 >_v c_2\}$ be the set of voters who find c_1 more valuable than c_2, and $|V(c_1,c_2)|$ the cardinality of that set. Then the social preference relation R on C is defined as follows by Condorcet :

$$c_1 \; R \; c_2 \text{ if and only if } |V(c_1,c_2)| > |V(c_2,c_1)|.$$

This is the so-called pairwise majority rule. It is well-known that such a relation is often not transitive and may contain cycles. More generally, Arrow (1951) proved that the transitivity of R is impossible with natural requirements of the voting procedure such as independence of irrelevant alternatives, unanimity, and non-dictatorship (i.e., $R \neq \geq_v$ for the same voter v systematically).

The probabilistic version of the likely dominance rule is very similar to Condorcet procedure, taking $V = S$, $C = X^S$, and considering for each $s \in S$ the relation R_s on acts such that

$$\forall f, g \in X^S: f\ R_s\ g \text{ if and only if } f(s) >_P g(s).$$

Computing the probability Prob([f $>_P$ g]) is a weighted version of $|V(c_1,c_2)|$ with $V = S$, $c_1 = f$, $c_2 = g$, which explains the intransitivity phenomenon. However, the likely dominance rule makes sense for any inclusion-monotonic likelihood relation between events and is then much more general than the Condorcet pairwise majority rule.

The above counter-example suggests that the likely dominance rule is generally not coherent with axioms (S1-S5) of Savage's setting for DMU since they enforce comparative probability relations as representing subjective uncertainty, and S1 assumes the transitivity of the preference relation over acts. Here we shall precisely describe the reason for this incompatibility.

More precisely, the following result explains why probabilistic version of likely dominance may violate the transitivity of the strict preference over acts:

Proposition 2: Suppose that

i)(X^S, $\geq$) is a weak order satisfying axioms S4, S2,

ii) S and X have at least three elements,

iii) (X^S, $\geq$) projects on 2^X as (2^X, $\geq_L$) using axiom S4,

iv) (X^S, $\geq$) projects on X as (X, $\geq_P$) where $x >_P y >_P z$ for some $x, y, z \in X$,

v) The likely dominance rule generates (X^S, $\geq$).

Then, there cannot exist three pairwise disjoint non-null events A, B, C such that

$C \cup B >_L A$, $C \cup A >_L B$ and $A \cup B \geq_L C$.

Proposition 2 means that, as soon as some acts are neither constant nor binary, the likely dominance rule is compatible with axioms S1, S2 and S4 only for uncertainty orderings such that :

Axiom N: $\forall$ A, B, C disjoint, $B \cup C >_L A$ and $A \cup C >_L B$ imply $C >_L A \cup B$.

The proof precisely shows that the conditions in proposition 2 may create cycles in preferences. Axiom (N) has been introduced by Friedman and Halpern (1996), and Dubois and Prade (1995b) in connection with the nonmonotonic logic framework of Kraus, Lehmann and Magidor (1990). It means that if $C >_L A$, the likelihood of C is much higher than the likelihood of A. Indeed the likelihood of events can never be attained by lumping events of lower plausibility. So $C >_L A$ means that the likelihood of A is negligible in front of the likelihood of C. This result considerably restricts the applicability of the likely dominance rule in Savage's framework. Indeed, probability measures almost always violate property (N). More precisely, we can prove the following result:

Proposition 3: If a likelihood relation $\geq_L$ is a comparative probability on S such that (N) holds, then if S has more than two non-impossible elements□

i) there is a permutation of the *non-null* elements of S, such that $s_1 >_L s_2 >_L \dots >_L s_{n-1} \geq_L s_n >_L \emptyset$ and

ii) $\forall\, i = 1,\dots n-2$, $s_i >_L \{s_{i+1}, \dots, s_n\}$.

Propositions 2 and 3 trivialize the likely dominance rule in the Savage's framework since it enforces likelihood relation to be one of comparative probability property satisfying (N), when there are at least three states and three consequences. Proposition 3 claims that the likely dominance rule forbids a Savagean decision maker to believe that there are two equally likely states of the world, each of which is more likely than a third state. This is clearly not acceptable in practice. If we analyze the reason why this phenomenon occurs, it is easy to see that axiom S1 plays the crucial role, in so far as we wish to keep the sure-thing principle. S1 assumes the full transitivity of the likelihood relation $\geq_L$. Giving up the transitivity of $\geq_L$ may relax the unnatural restriction of an almost total ordering of states.

When |X| < 3 or if there are no consequences such that fx > fy > fz the likely dominance rule cannot generate intransitivities on X^S. If X has only two elements $x >_P y$, then the set of acts is truly isomorphic to the set of events and the likely dominance rule then follows from S1, S2, S3, S4, S5. Indeed, for any act f consider the event $A_f \subseteq S$, where A_f is defined as follows: $s \in A_f$ if f(s) = x and $s \notin A_f$ if f(s) = y. Now:

$f \geq g$ if and only if $xA_fy \geq xA_gy$

if and only if $x(A_f \cap (A_g)^c)y \geq x(A_g \cap (A_f)^c)y$ (due to S2)

if and only if $[f >_P g] \geq_L [g >_P f]$.

In particular for binary acts the likely dominance rule yields the same preference ordering as expected utility. Indeed, if u is a utility function on X and Prob a probability measure on S, it is obvious that, given f = xAy and g = xBy,

$$u(x)\cdot Prob(A) + u(y)\cdot(1-Prob(A)) > u(x)\cdot Prob(B) + u(y)\cdot(1-Prob(B))$$

if and only if Prob(A) > Prob(B) if and only if $Prob(B^c \cap A) > Prob(B \cap A^c)$

if and only if $[xAy >_P xBy] >_L [xBy >_P xAy]$.

When |X| ≥ 3, the trivialization result in Proposition 3 does not rule out all probabilistic representations. The search for probability functions satisfying axiom (N) has been carried out by Snow (1994, 1999) and Benferhat et al. (1997b). They are very special probability functions such that $\forall\, A \subseteq S$, $\forall\, s \in A$, $Prob(\{s\}) > Prob(A \setminus \{s\})$ so that $Prob(A) \geq Prob(B)$ if and only if $\max_{s \in A} Prob(s) \geq \max_{s \in B} Prob(s)$ (this is due to Proposition 3): the magnitude of p(s) can never be attained by summing the probabilities of states that are individually less probable than s. Moreover $p(s_1) > p(s_2) > \dots > p(s_{n-1}) \geq p(s_n)$ for states $s_1,\dots, s_n$ of nonzero probability. Benferhat et al. (1999) call these probability functions "big-stepped probabilities". They form what Snow (1994) calls an atomic-bound system. Clearly a lot of probability measures are ruled out by this condition.

4 - An ordinal axiomatic framework for decision making under uncertainty

We now use a relaxed Savage-like framework that tries to overcome the above trivialization result. In view of this result, the only way of accommodating the likely dominance rule while retaining most of Savage's framework is to weaken postulate S1. The lack of transitivity of the strict preference is not acceptable. But it is possible to modify S1 and drop the transitivity of indifference between acts. Although we shall have to give up probabilistic representations of uncertainty for the most part, such a relaxation improves the consistency between the sure-thing principle and the likely dominance rule. However, is this decision rule the only one that ensures a separate treatment of uncertainty and preference and avoids making them comparable? To address this question, we shall add an axiom precisely expressing this non-comparability requirement. We show that this axiom indeed enforces the likely dominance rule.

4.1 A relaxed Savage framework

We shall start from a strict preference relation > on the set of acts X^S and assume the following property, a weak form of S1:

WS1: (X^S, >) is a transitive, irreflexive, partially ordered set.

Let ≥ denote the transposed complement of >. Because the strict preference > among acts is irreflexive and transitive, it is obvious that the weak preference ≥ is reflexive and complete (since it cannot hold that f > g and g > f). Moreover denote f ∼ g when f ≥ g and g ≥ f, and call it indifference relation. It is obvious that ∼ is reflexive and symmetric. The following weak transitivity properties hold: f > g and g ∼ h $\Rightarrow$ f ≥ h and f ∼ g and g > h $\Rightarrow$ f ≥ h. Indeed, for instance, f > g and h > f imply h > g (transitivity of >) and this is inconsistent with g ∼ h. However none of ≥, ∼ is supposed to be transitive. This framework differs from Lehmann's (1996) who assumes that the weak preference ≥ is not complete but the indifference is transitive.

The lack of transitivity of the indifference relation is not counter-intuitive. If f ∼ g is interpreted as "f is close to g" in the sense that their utilities are close ($|u(f) – u(g)| \leq c$), then the lack of transitivity is expected. This is usually the case in preference modeling for multiple criteria analysis (Roubens and Vincke, 1985). One could alternatively interpret f ∼ g as "f and g are not comparable". However since relation ≥ is complete, the present framework does not allow to distinguish between genuine indifference and genuine incomparability. The latter more naturally means that neither f ≥ g nor g ≥ f holds, a case ruled out by our assumptions here. In our framework, indifference is equivalent to when neither f > g nor g > f holds. while the latter Treating indifference and non-comparability as distinct situations requires a weak preference ≥ that is not complete.

Let us derive the preference relation $\geq_P$ from ≥ according to definition 5. Note that definition 5 applies to whatever preference relation among acts, so that given such a relation, the set X is automatically equipped with a preference relation having at least the same properties as the preference relation on acts it is induced from. The strict part of $\geq_P$ (resp. $\geq_L$) is denoted $>_P$ (resp. $>_L$) and the associated indifference $\sim_P$ (resp. $\sim_L$). Due to WS1, it is obvious that $>_P$ is irreflexive and transitive and that $\geq_P$ is reflexive and complete.

Dropping the transitivity of $\geq$ cancels some useful consequence of the sure-thing principle under S1, which are nevertheless consistent with the likely dominance rule. Typically, WSTP (or equivalently, the unanimity axiom U) will not derive from the relaxed framework. We shall have to add it so as to keep it**.** Another very natural property that is implied by Savage's framework, is that f $\geq_P$ g and g > h implies f > h (where f $\geq_P$ g means that whatever s, f(s) $\geq_P$ g(s)). Similarly, f > g and g $\geq_P$ h implies f > h. This form of transitivity which could be called "Pareto stability" sounds also very natural, but is not required in the above setting. However the additional axiom of the next section obviates the need for assuming Pareto stability, up to a pathological case, as seen further on.

At this point, and in order to stick as close as possible to the classical framework, it is natural to assume WS1, WSTP, S2, S3, S4, S5 hold.

4.2 The Ordinal Invariance axiom

Basically, an ordinal description of preference presupposes that preference values are taken from an ordinal scale L and are attached to each consequence $x \in X$. It means that the inference drawn from such a representation should be invariant if L is changed into $\varphi(L) = \{\varphi(\lambda): \lambda \in L\}$ for any non-decreasing bijection φ of L to itself. Here, we shall not use a preference scale explicitly. All that is needed is the finite set X of consequences, equipped with a preference relation $\geq_P$. Intuitively we shall require the following: if, in each state, the preference pattern between the consequence of act f and the consequence of act g is not altered when changing g and f into f' and g' then the preference pattern between f and g should remain the same as the one between f' and g'.

More rigorously, two pairs of acts (f, g) and (f', g') will be called *statewise order-equivalent* if and only if $\forall s \in S$, f(s) $\geq_P$ g(s) if and only if f'(s) $\geq_P$ g'(s). This is denoted (f, g) ≡ (f', g'). The following *Ordinal Invariance* axiom (Fargier and Perny, 1999) fully embodies our requirement of a purely ordinal approach to DMU.

OI: $\forall$f, f' g, g' $\in X^S$, if (f, g) ≡ (f', g') then (f $\geq$ g $\Leftrightarrow$ f' $\geq$ g').

This axiom does express that what matters for preference between two acts is the relative positions of consequences of acts for each state, not the consequences themselves, nor the positions of these acts relative to other acts. If (f, g) ≡ (f', g'), then for each state s there is an order-preserving mapping φ_s from ({f(s), g(s)}, $\geq_P$) to ({f'(s), g' (s)}, $\geq_P$) such that φ_s(f(s)) = f'(s) and φ_s(g(s)) = g'(s).

It is easy to check that the likely dominance rule obeys axiom OI. This is obvious noticing that if (f, g) ≡ (f', g') then by definition, [f $>_P$ g] = {s, f(s) $>_P$ g(s)} = [f' $>_P$ g']

In the following, the consequence of adding OI to the relaxed Savage-like framework is laid bare. As it turns out the presence of OI makes the explicit statement of the sure thing principle and of axiom S4 superfluous.

Proposition 4: Under assumptions that OI holds, and that relation $\geq$ on acts is reflexive, it follows that S2, and S4 hold too.

For S2, one has to see that (fAh, gAh) $\equiv$ (fAh', gAh'). For S4, if $x >_P y$ and $x' >_P y'$ then (xAy, xBy) $\equiv$ (x'Ay', x'By'). The role of S3 is to ensure that $x \geq_P y$ if and only if xAf $\geq$ yAf. It must be kept since it is not true that (fx, fy) $\equiv$ (xAh, yAh), for any act h, consequences x, y and event A.

Note that axiom OI always holds when $|X| = 2$ since in that case, (f, g) $\equiv$ (f', g') and f≠g imply f' = f and g' = g. Hence the ordinal invariance assumption makes a difference only if $|X| \geq 3$ and $\exists$x, y, z, $x >_P y >_P z$. Moreover , if $|X| = 2$, the transitivity of $\geq_P$ is trivially satisfied. In fact, if $|X| = 2$, all non-constant acts are binary, and the ordinal setting coincides with the Savage setting, since the likely dominance rule produces the same ordering of acts ad the expected utility, in that case, as pointed out in the previous section. Hence we may have to strengthen the non triviality axiom S5 into:

- S'5: $\exists$fx, fy, fz three constant acts such that fx > fy > fz,

in order to focus on the specific features of the decision rule induced by axiom OI. In view of the above discussion, motivated by a purely ordinal approach, we propose to adopt the following axiomatic framework, that is faithful to Savage's as much as possible: WS1, WSTP (or U), S3, S5 and OI. In the following, this is called the *ordinal decision setting*.

This framework (in fact, not all axioms are needed) actually enforces the likely dominance rule as the only rational decision rule (Fargier and Perny, 1999):

Proposition 5: Under the assumption of ordinal invariance, if the weak preference on acts is reflexive and the induced weak preference on consequences is complete, the only possible decision rule is likely dominance.

The proof relies on the statewise order-equivalence between the pairs of acts (f, g) and ($x[f \geq_P g]y$, $x[g \geq_P f]y$) for $x >_P y$. The assumptions of proposition 5 are clearly verified under axiom WS1. As a consequence, if one insists on sticking to a purely ordinal view of DMU, and one adopts the very weakly constrained framework defined by WS1, WSTP (or U), S3, S'5, then ordinal invariance, embodied by axiom OI enforces the likely dominance rule. There is thus no alternative to this decision rule in a purely ordinal setting.

It remains to check the type of uncertainty representation this framework leads to, as it was proved to be almost contradictory with a standard probabilistic representation. This is the topic of the next section.

5 - Subjective comparative likelihood induced by the ordinal setting

In this section we first consider the weak Savage setting defined by postulates WS1, S2, S3, S4, S5, WSTP (or unanimity). In subsection 5.1, we first study the properties of the likelihood relations $>_L$, $\geq_L$ and $\sim_L$ induced via S4 without assuming OI. In that case, the setting is more general than Savage's (but it is finite), since we do not assume the transitivity of the weak preference. In subsection 5.2, we add the OI axiom, and the framework then moves away from comparative probability. In subsection 5.3, it is shown that we get much closer to non-monotonic reasoning based on preferential inference.

5.1 - Without Ordinal Invariance

Suppose we keep the five basic axioms of Savage framework, just relaxing S1 into WS1. We call this framework the *quasi-transitive Savage setting*. We first prove properties that hold *without even assuming unanimity* (U or WSTP).

Proposition 6: In the quasi-transitive Savage setting, the subjective likelihood relation $>_L$ induced by S4 verifies the following properties

- $S >_L \emptyset$ (Non-triviality)
- $\square\ A \ \square\ S, A \geq_L \emptyset$ (Non-Contradiction)
- $A \ \square\ (B \ \square\ C) = \emptyset\ \square\ (B \geq_L C \ \square\ A \ \square\ B \geq_L A \ \square\ C)$ (Preadditivity)
- A null $\square$ $A \sim_L \emptyset$ (Compatibility with null events)
- $A \ \square\ B, A >_L \emptyset\ \square\ \square >_L \emptyset$ (Stability of non-null events)

Corollaries :

- $A >_L B \ \square\ B^c >_L A^c$ $A \geq_L B \ \square\ B^c \geq_L A^c$ (Auto-duality)
- $A \ \square\ B \ \square\ \square \geq_L A$ (Inclusion-monotony)
- $A \sim_L \emptyset\ \square\ \square\ B \ \square\ A,\ B \sim_L \emptyset$ (Characterization of null events)

These properties are satisfied by comparative probability, of course. The preadditivity property is driven by the sure thing principle and it implies that $A \geq_L B$ $\square$ $A \ \square\ B^c \geq_L A^c \ \square\ B$ and $A >_L B \ \square\ A \ \square\ B^c >_L A^c \ \square\ B$. Moreover autoduality and inclusion monotony are derived from the other properties (non-contradiction, non-triviality, preadditivity of$\square_L$. However, the weak comparative likelihood $\geq_L$ is not transitive and transitivity is not even preserved when $\geq_L$ is composed with its strict part$\square$$A \geq_L B$ and $B >_L C$ do not imply $A \geq_L C$. Moreover :

- $\{s\} \sim_L \emptyset$ and $\{s'\} \sim_L \emptyset$ do not imply $\{s, s'\} \sim_L \emptyset$, nor $\{s\} \sim_L \{s'\}$;
- $\{s\} \sim_L \emptyset$ and $\{s'\} >_L \emptyset$ do not imply $\{s, s'\} >_L \emptyset$;
- $E >_L F$ does not imply $E >_L \emptyset$.

In order to recover these properties, what is needed is the weak sure-thing principle WSTP or the unanimity rule (U) equivalently. Remember they do hold in Savage's framework due to S1, and have to be added here since the weak preference is not transitive. This new setting, called the *quasi-transitive Savage setting with unanimity,* remains weaker than Savage's, but the situation then slightly improves due to the following property:

Proposition 7: In the quasi-transitive Savage setting with unanimity: if $(A \cup B) \cap (C \cup D) = \emptyset$, then $A \geq_L B$ and $C \geq_L D \Rightarrow A \cup C \geq_L B \cup D$

Corollaries:

- $\{s\} \sim_L \emptyset$ and $\{s'\} \sim_L \emptyset \Rightarrow \{s, s'\} \sim_L \emptyset$ (7.1)
- $\{s\} \sim_L \emptyset$ and $\{s'\} \sim_L \emptyset \Rightarrow \{s\} \sim_L \{s'\}$ (7.2)
- $E >_L F \Rightarrow E >_L \emptyset$ (7.3)
- $E \cap F = \emptyset$ and $G \subseteq E$ and $G \geq_L F \Rightarrow E \geq_L F$ (7.4)
- There is at least one non-null state. (7.5)

Hence, the Unanimity axiom added to WS1, S2, S3, S4, S5, ensures that the null events form a class closed under union (proposition 7.1) and intersection (last item of proposition 6) since $\sim_L$ becomes transitive for null events. So there exists a largest null set, say NULL, such that any set of states A is null if and only if it is a subset of NULL. It also means that under this weakened Savage framework it is always possible to assume that no state is null without any loss of information: if we delete all null states from S as being impossible, a state space where no non empty set is null is obtained.

Remark: $A >_L C$ and $B \sim_L \emptyset$ do not imply $A \cup B >_L C$, except if $C = \emptyset$. This is because $\{s\} \sim_L \emptyset$ and $\{s\} \sim_L \{s'\}$ do not yet imply $\{s'\} \sim_L \emptyset$. (A simple counterexample is $S = \{s, s'\}$ with $\{s'\} >_L \emptyset$ being the only requirement). It is not surprising if $\sim_L$ is interpreted in terms of incomparability. If this transitivity property were valid, and $A \cup B \sim_L C$, then $A \cup B \sim_L \emptyset$ and $A \sim_L \emptyset$, consequently, which contradicts $A >_L C$.

5.2 - Consequences of Ordinal Invariance

Let us now add OI to our set of axioms characterizing the preference on acts. This is called the *ordinal decision setting*, defined by axioms WS1, S3, S'5, OI, WSTP (or U). We know that axioms S2 and S4 need no longer be stated explicitly. Under these conditions we can show that :

Proposition 8: In the ordinal decision setting, if S contains at least two non-null states, the preference on X is a weak order (but the ordering of states in S may be partial).

The proof only uses WS1 and OI. The above proposition leads us to consider the quotient set $X/\sim_P$ using the equivalence relation associated to the consequence preordering $\geq_P$. It provides a utility scale. It makes sense to cluster equivalent consequences and replace X by $X/\sim_P$. For instance if $X/\sim_P$ has only two elements, then the set of acts is truly isomorphic to the set of events and the likely dominance rule then follows from the quasi-transitive Savage setting as pointed out in Section 3.2.

If there is only one non-null state $s^* \in S$ such that $s^* >_L \emptyset$, then $f > g$ if and only if $(f > g)_{\{s^*\}}$. Hence we can then assume that S reduces to a singleton, that is, the state is known. This is the case when the agent acts as if the knowledge about the world were deterministic. Then the likelihood relation $\geq_L$ is transitive, since for any event $A \sim_L S$ or $A \sim_L Ø$. Axiom S3 leads to $f > g$ if and only if $f(s^*) >_P g(s^*)$. Hence the set of acts and the set of consequences coincide. If there is only one non-null state $s^* \in S$, axioms S2, S4, U are trivially satisfied, as well as the OI axiom since $f > g$ if and only if $s^* \in [f >_P g]$. However, Proposition 8 may fail to hold when there is only one non-null state and X has more than three non-indifferent elements because the properties of the ordering on X reflect the properties of the ordering on acts.

Another important consequence of ordinal invariance OI is a fully coherent behavior of the strict likelihood relation $>_L$ with respect to set inclusion, because it respect the orderly property of Halpern (1996):

Proposition 9: In the ordinal decision setting, the likelihood relation is orderly, that is:

$$A >_L B \cup C \Rightarrow A >_L B \quad (9.1)$$
$$A >_L B \Rightarrow A \cup C >_L B \quad (9.2)$$

However the presence of the Ordinal Invariance Axiom leads us away from comparative probability, since the following very strong, non probabilistic property, already encountered in Section 3, and denoted (N), holds for the comparative likelihood:

Proposition 10: In the ordinal decision setting, if A, B, C are disjoint, then

$A \cup C >_L B$ and $B \cup C >_L A$ imply $C >_L A \cup B$

This is a variant of Proposition 2, whose proof is based on OI and does not require from axiom S1 more than the transitivity of the strict part of the preordering $\geq$ on events and thus applies here. Unanimity is not needed either.

However Propositions 9 and 10 generally do not hold when X has only two non indifferent elements $x >_P y$, (i.e., when S'5 does not hold), and there are at least two non-null states[1]. In that case, consequences belong to two different equivalence classes, because of Proposition 8 (indeed, WS1, S3, OI and S5 imply that the preference relation on consequences is a weak order). Because of S4, acts and events coincide, and the likely dominance rule is a consequence of S2 and S4. In this case, the preference on acts can be a weak order, as in Savage setting. We thus get a theory of uncertainty, that satisfies Proposition 6 and, if Unanimity is required, Propositions 7 and 8. But Propositions 9 and 10 can be violated.

---

[1] Indeed, if there is only one non null state, say s^*, $A >_L B \cup C$ implies that $A >_L Ø$ and thus $s^* \in A$ and $s^* \notin B$. It is impossible to have $B \geq_L A$, which would lead to $xBy \geq xAy$, i.e. $xBy(s^*) \geq_P xAy(s^*)$: $y \geq_P x$. Similarily, $A >_L B$ implies that $A >_L Ø$ and thus $s^* \in A$ and $s^* \notin B$. It is impossible to have $B \geq_L A \cup C$, which would lead to $xBy \geq x(A \cup C)y$, i.e. $xBy(s^*) \geq_P x(A \cup C)y(s^*)$: $y \geq_P x$. Hence, in this case, proposition 9 still holds, provided that the system obeys S5. Concerning proposition 10 : $A \cup C >_L B$ and $B \cup C >_L A$ imply that s^* belongs to $A \cup C$ and $B \cup C$, i.e. that s^* belongs to C. $A \cup B \geq_L C$ would lead to $x(A \cup B)y \geq xCy$, i.e. $x(A \cup B)y(s^*) \geq_P xCy(s^*)$: $y \geq_P x$

Violations of proposition 9 when X = {x, y} and $x >_P y$, are easy to find: if S = {a, b, c}, a likelihood relation generated by $\{a\} >_L \{b, c\} >_L \emptyset$ only, is such that $\{a\} \sim_L \{b\}$ and violates no property induced by the ordinal decision setting, except (9.1). Similarly, a likelihood relation generated by $\{a\} >_L \{b\} >_L \emptyset$ only (hence $\{a, c\} >_L \{b, c\}$), is such that $\{a\} \sim_L \{b, c\}$ and violates no property induced by the ordinal decision setting, except (9.1). So the case of binary consequence sets is pathological. Proposition 9 directly follows from Pareto-stability of the strict preference over acts, that is " $f \geq_P g$ and $g > h$ implies $f > h$", and $f > g$ and " $g \geq_P h$ implies $f > h$". Hence Proposition 9 is recovered even if X = {x, y}, when Pareto stability is added to WS1, S2, S3, S4, S5, U. In fact, Pareto stability is consequence of Savage framework and it must be added in its non-transitive relaxation using WS1.

Proposition 10 cannot be deduced in Savage's framework : counterexamples are easy to find when $>_L$ is supposed to come from a probability distribution.

5.3. - Savage axioms, ordinal invariance and preferential inference

Property (N) and the orderly property of proposition 9 are closely related to one of the characteristic properties for uncertainty relations that (Friedman and Halpern, 1996) call "plausibility" relations, which are closely connected to non-monotonic consequence relations. Given a subjective likelihood relation $>_L$, define an inference relation as follows:

$$A \mid\!\sim B \text{ if and only if } A \cap B >_L A \cap B^c .$$

$A \mid\!\sim B$ intuitively means that B is more likely than not, when A is true. This Section gives a self-contained proof that the uncertainty theory induced by the relaxed Savage framework augmented with the OI axiom generates a full-fledged nonmonotonic inference, thus giving decision-theoretic foundations to nonmonotonic reasoning. Of course, this is no longer surprising due to Proposition 10 and Friedman and Halpern (1996)'s results. Our presentation makes it clear which properties of non-monotonic consequences are related to which Savage-like axioms.

A conditional assertion is an expression of the form $\alpha \mid\!\sim \beta$, where α and β are propositional formulae, and $\mid\!\sim$ is an inference relation. In the following we shall never use formulae, but only sets of models thereof, and consider nonmonotonic inferences of the form $A \mid\!\sim B$. It means that if all that is known by an agent is that event A obtains, then the agent will reason as if B obtained as well. In other words, in context A, the agent accepts B. Basic properties of nonmonotonic inference have been advocated by Kraus, Lehmann and Magidor (1990), and form what is known as “system P”:

R:	$A \mid\!\sim A$ (reflexivity)
RW:	$A \mid\!\sim B$ and $B \subseteq C$ imply $A \mid\!\sim C$ (right weakening)
AND:	$A \mid\!\sim B$ and $A \mid\!\sim C$ imply $A \mid\!\sim B \cap C$ (closure by conjunction)
OR:	$A \mid\!\sim C$ and $B \mid\!\sim C$ imply $A \cup B \mid\!\sim C$ (reasoning by cases)
CM:	$A \mid\!\sim B$ and $A \mid\!\sim C$ imply $A \cap B \mid\!\sim C$ (cautious monotony)
CUT:	$A \mid\!\sim B$ and $A \cap B \mid\!\sim C$ imply $A \mid\!\sim C$ (modified transitivity)

The above rules of inference embody the notion of plausible inference in the presence of incomplete information. Namely, they describe the properties of deduction under the assumption that the state of the world is as normal as can be. The crucial rules are Cautious Monotony and the Cut. Cautious Monotony claims that if A holds, and if the normal course of things is that B and C hold in this situation, then knowing that B and A hold should not lead us to situations that are exceptional for A: C should still normally hold. The Cut is the converse rule: If C usually holds in the presence of A and B then, if situations where A and B hold are normal ones among those where A holds (so that A normally entails B), one should take it for granted that A normally entails C as well. The other above properties are not specific to plausible inference: OR enables disjunctive information to be handled without resorting to cases. The Right Weakening rule, when combined with the Right AND, just ensures that the set of nonmonotonic consequences $\{B, A \cap B >_L A \cap B^c\}$ (the set of propositions "accepted" by $>_L$ when A is true) is deductively closed in every context (see Dubois and Prade, 1995b; Dubois, Fargier and Prade, 1998). Reflexivity sounds natural but can be challenged for null events as seen below.

These basic properties can be used to form the syntactic inference rules of a logic of plausible inference. There exist six semantics that have been proposed for such a logic: preferential semantics (Kraus, Lehmann and Magidor, 1990), based on a partial ordering relation on an abstract set, projected to the set S via a mapping (it has little intuitive appeal); three-valued logic semantics based on an abstraction of conditional probabilities (Dubois and Prade, 1994); infinitesimal probabilistic semantics (Adams, 1975; Pearl, 1988; Lehmann and Magidor, 1992); possibilistic semantics (Dubois and Prade, 1995a) whereby a set of conditional assertions is equivalent to a family of possibility orderings; finitistic probabilistic semantics by means of a very special class of probability measures, the one encountered in Section 3 (Benferhat et al., 1997b); and a last semantics (Friedman and Halpern, 1996) based on partial orderings of events satisfying Property (N).

The above natural properties of these inference relations, for the purpose of reasoning tolerant to exceptions proposed by Kraus, Lehmann and Magidor (1990) can be translated as follows:

R: Whatever A, $A >_L Ø$.

AND: $A \cap C >_L A \cap C^c$ and $A \cap B >_L A \cap B^c \Rightarrow A \cap B \cap C >_L A \cap (C^c \cup B^c)$

OR: $A \cap C >_L A \cap C^c$, $B \cap C >_L B \cap C^c \Rightarrow (A \cup B) \cap C >_L (A \cup B) \cap C^c$

RW: $B \subseteq C$, $A \cap B >_L A \cap B^c \Rightarrow A \cap C >_L A \cap C^c$

CM: $A \cap B >_L A \cap B^c$ and $A \cap C >_L A \cap C^c \Rightarrow A \cap B \cap C >_L A \cap B \cap C^c$.

CUT: $A \cap B >_L A \cap B^c$ and $A \cap B \cap C >_L A \cap B \cap C^c \Rightarrow A \cap C >_L A \cap C^c$.

Property R does not hold in the setting of likelihood relations. In system P, the reflexivity property $A \mid\!\sim A$ reads $\forall A, A >_L Ø$ in our setting. This axiom fails due to the presence of null events. Moreover it implies $Ø >_L Ø$ which is also wrong since $>_L$ is irreflexive. This is not so important at the technical level for getting representation results, as shown in Benferhat et al. (1997a). Moreover, it is not clear that $A \mid\!\sim A$ makes sense when reasoning under exceptions, when $A = Ø$: one may admit that a contradiction preferentially infers nothing. Reflexivity can be replaced by two conditions conjointly (Benferhat et al. 1997):

– restricted reflexivity: $\forall A \neq \emptyset$, $A \mid\sim A$, that is, $A >_L \emptyset$

– consistency preservation: $\neg(A \mid\sim \emptyset)$, which expresses that $\emptyset >_L A$ holds for no event A, a property that relation $>_L$ satisfies.

Since the set of null events has a maximal element NULL for inclusion and every subset of NULL is null, system P with restricted reflexivity and consistency preservation is satisfied on S \ NULL. Null events can be defined directly in system P by just dropping the restricted reflexivity axiom and calling an event A null whenever $A \mid\sim A$ does not hold. Using right weakening and OR, it is clear that $A \mid\sim A$ and $B \mid\sim B$ imply $A \cup B \mid\sim A \cup B$. Hence if $A \mid\sim A$ does not hold then $\forall B \subseteq A$, $B \mid\sim B$ does not hold either. So there will always be a maximal event NULL such that reflexivity holds except for subsets of NULL.

The other properties of system P are induced by the ordinal decision setting on the inference relation based on subjective likelihood:

Proposition 11: OR and AND follow from (WS1, S3, S'5, OI).

At this point of our axiomatisation, we cannot get all the properties of system P, but only some of them. In order to prove RW, CUT and CM, we need the properties:

$B \cap C = \emptyset$ and $A \subseteq B$ and $A \geq_L C \Rightarrow B \geq_L C$ (Proposition 7)

$B \cap C = \emptyset$ and $A \subseteq B$ and $A >_L C \Rightarrow B >_L C$ (Proposition 9)

which cannot be derived without the unanimity axiom[2]. The next proposition holds if this axiom is valid.

Proposition 12: If the relation on X^S satisfies the axioms of the ordinal decision setting, then the nonmonotonic inference built from the relation projected from X^S to 2^S satisfies the properties following of system P: OR, AND, RW, CM, CUT.

This result establishes a bridge between decision theory and non-monotonic reasoning. It also show the similarity of spirit between the axiomatic approaches of Savage and Lehmann.

6 - Ordinal invariance enforces qualitative possibility theory.

A concrete non-trivial example of a subjective likelihood relation that fits decision theory based on ordinal invariance can be provided in the setting of qualitative possibility theory (Lewis, 1973; Dubois, 1986; Dubois and Prade, 1998).

---

2 Except if there are no null events. Null events generally never appear in nonmonotonic reasoning nor belief revision theories. But supposing that there is no null event is not appropriate in the context of a Savage-like theory of decision.

6.1 Qualitative possibility theory

Assume that a decision maker supplies a weak order of states $\geq_\pi$ on S, describing their relative plausibility. This relation is called a comparative possibility distribution. Intuitively s $\geq_\pi$ s' means that state s is at least as plausible, normal, expected as state s'. The simplicity of possibility theory stems from the fact that it uses weak orderings of events completely characterized by the comparative possibility distribution. Namely, a pair of dual possibility and necessity orderings on events can be generated on events, from a weak order $\geq_\pi$ on states only. Namely denoting max(A) any (most plausible) state s $\in$ A such that s $\geq_\pi$ s' for any s' $\in$ A.

A $\geq_\Pi$ B if and only if max(A) $\geq_\pi$ max(B)

and A $\geq_N$ B if and only if max{B^c} $\geq_\pi$ max{B^c}.

This lifting rule must be completed by the property {s} $\geq_\Pi$ Ø for all s $\in$ S. If we enforce the property {s} $>_\Pi$ Ø, for all states, the obtained possibility ordering on events is called non-dogmatic as it does not rule out any state in an absolute manner. Of course there are other ways of lifting weak ordering relations from elements to subsets, that are less intuitive here, for instance changing max into min (see Dubois and Prade, 1998). Halpern (1997) has studied in detail the process of lifting a plausibility relation from elements of a set to its subsets, assuming only a partial ordering of states. See Appendix 1 for the axiomatic definitions of possibility and necessity orderings and their set-function representations.

The main idea behind this modeling is that the state of the world is always believed to be as normal as possible, neglecting less normal states. Thus B $\geq_\Pi$ C means that there is a normal state where B holds that is at least as normal as any normal state where C holds. B $\geq_\Pi$ C also means that B is at least as plausible as C where plausibility refers to consistency with an agent's knowledge: B is plausible when it does not contradict the agent's beliefs described by the comparative plausibility distribution. The dual ordering B $\geq_N$ C is intuitively understood as "B is at least as certain as C", in the sense that all situations where B fails to hold are less normal than the most normal situation where C does not hold. In particular the events accepted as true are those which hold in all the most plausible situations.

Note that as soon as A $\cap$ B $\geq_\Pi$ A^c $\cap$ B and A $\cap$ B $\geq_\Pi$ A $\cap$ B^c events A and B are of equal possibility and/or certainty. To overcome this lack of discrimination, consider the following comparative likelihood relation between events:

Definition 7: A *possibilistic likelihood* relation $>_{\Pi L}$ is defined from a weak order $\geq_\pi$ on S by

$$A \geq_{\Pi L} B \text{ if and only if } A \cap B^c \geq_\Pi A^c \cap B \qquad (*)$$

where $\geq_\Pi$ is the possibility ordering on events induced by $\geq_\pi$.

This relation has been proposed independently by various authors like Brewka (1989), Brass (1991), Cayrol et al. (1992), Geffner (1992) and Delgrande (1994), mainly for the purpose of ordering subsets of formulas from a prioritized knowledge base, in the scope of non-monotonic reasoning and inconsistency management. Here it is used as a tool to define a comparative likelihood relation among events.

Note that $>_{\Pi}$ and $>_{\Pi L}$ coincide when restricting to mutually exclusive events A and B since then $B^c \cap A = A$ and $B \cap A^c = B$. Similarly $>_{N}$ and $>_{\Pi L}$ coincide when restricting to events A and B whose union is the whole state space, since then $B^c \cup A = B^c$ and $B \cup A^c = A^c$. The possibilistic likelihood relation $>_{\Pi L}$ is a refinement of both possibility and necessity orderings since it may discriminate among events of equal possibility and/or certainty, when this equality is due to the common parts of A and B or their complements.

The possibilistic likelihood relation is only a partial ordering. In particular, the relation $A \sim_{\Pi L} B$ if and only if neither $A >_{\Pi L} B$ nor $B >_{\Pi L} A$ hold on 2^S is not transitive. Namely, in terms of possibility measures (Appendix 1), the assumptions $\Pi(A \cap B^c) = \Pi(A^c \cap B)$ and $\Pi(B \cap C^c) = \Pi(B \cap C^c)$ do not imply $\Pi(A \cap C^c) = \Pi(A^c \cap C)$. This is because, even if Π is two-valued, $\max(\pi_5, \pi_3) = \max(\pi_4, \pi_6)$ and $\max(\pi_2, \pi_6) = \max(\pi_3, \pi_7)$ do not imply $\max(\pi_4, \pi_7) = \max(\pi_2, \pi_5)$ denoting $\pi_i = \pi(s_i)$, $A = \{s_2, s_3, s_5\}$, $B = \{s_2, s_4, s_6\}$, $C = \{s_3, s_4, s_7\}$ (just take $\pi_2 = \pi_3 = \pi_5 = \pi_6 = 1$, $\pi_4 = \pi_7 = 0$).

Finally, the preadditivity condition (iv) of comparative probabilities holds, and the relation $\geq_{\Pi L}$ turns out to be self dual:

Proposition 13:

- $A \cap (B \cup C) = \emptyset \Rightarrow (B \geq_{\Pi L} C \Leftrightarrow A \cup B \geq_{\Pi L} A \cup C)$ (preadditivity).
- $A \geq_{\Pi L} B \Leftrightarrow B^c \geq_{\Pi L} A^c$ (self duality)
- $A >_{\Pi L} B \cup C \Rightarrow A >_{\Pi L} B$ (orderly)
- $A >_{\Pi L} B \Rightarrow A \cup C >_{\Pi L} B$ (orderly)
- $A \cup C >_{\Pi L} B$ and $B \cup C >_{\Pi L} A$ imply $C >_{\Pi L} A \cup B$ (N)

So, the relation $\geq_{\Pi L}$ possesses all properties of a comparative likelihood induced by the ordinal decision setting. It is similar to a comparative probability, except for the transitivity of indifference. Because of the lack of transitivity of the indifference relation, we cannot generally represent such relations by means of probability measures, nor can we assume properties that are usually derived from comparative probabilities. For instance, the following property of comparative probability is NOT satisfied in general by possibilistic likelihood:

If $A \cap B = \emptyset$ and $C \cap D = \emptyset$, then $A \sim_{\Pi L} C$ and $B \sim_{\Pi L} D \Rightarrow A \cup B \sim_{\Pi L} C \cup D$

Counter-example:

$S = \{a, b, c, d, e, f, g, h\}$

$A = \{a, b, c\}$ $B = \{e, f, g\}$ $C = \{b, g, h\}$ $D = \{c, d, e\}$

$\pi(a) = \pi(f) = \pi(c) = \pi(g) = 1$

$\pi(b) = \pi(d) = \pi(e) = \pi(h) = 0.$

$\Pi(A \cap C^c) = \Pi(\{a, c\}) = 1$ and $\Pi(C \cap A^c) = \Pi(\{g, h\}) = 1$: $A \sim_{\Pi} C$
$\Pi(B \cap D^c) = \Pi(\{f, g\}) = 1$ and $\Pi(D \cap B^c) = \Pi(\{c, d\}) = 1$: $B \sim_{\Pi L} D$
$\Pi((A \cup B) \cap C^c \cap D^c) = \Pi(\{a, f\}) = 1 > \Pi((C \cup D) \cap A^c \cap B^c) = \Pi(\{d, h\}) = 0$.
Thus : $A \cup B >_{\Pi L} C \cup D$.

6.2 Possibilistic likelihood as a model of subjective uncertainty in the ordinal decision setting

It is easy to verify that possibilistic likelihood is a model of subjective likelihood consistent with our ordinal decision setting. Assume that a decision maker supplies a weak order of states in the form of a possibility distribution π on S and a weak order of consequences $\geq_P$ on X. Let $>_{\Pi}$ be the induced possibilistic ordering of events. Let $>_{\Pi L}$ denote the associated possibilistic likelihood relation (*). The two relations coincide for disjoint sets. Define the preference on acts in accordance with the likely dominance rule, that is , for any two acts f and g:

$$f > g \Leftrightarrow [f >_P g] >_{\Pi L} [g >_P f] \Leftrightarrow [f >_P g] >_{\Pi} [g >_P f]$$
$$f \geq g \Leftrightarrow \neg(g > f).$$

The OI axiom holds for this preference relation, by construction. In the following we prove that the obtained preference on acts also satisfies the other axioms of the ordinal decision setting (WS1, S3, S'5, U, OI). To do it, we shall use the properties of the possibility ordering of events (definition A1 in the appendix). First we show that the notions of constant acts, null events, and conditional preference are well-defined.

Proposition 14:
- $\forall x, y \in X \quad fx \geq fy \Leftrightarrow x \geq_P y$ (14.1)
- $(f \geq g)_A \Leftrightarrow [f >_P g] \cap A \geq_{\Pi} [g >_P f] \cap A$ (14.2)
- A null $\Leftrightarrow A \sim_{\Pi} Ø$ (14.3)

Proposition 15: If the preference on acts derives from a possibilistic likelihood relation on events and a weak order of consequences via the likely dominance rule, then $(X^S, >)$ satisfies WS1, S3, OI and WSPT.

Hence all properties of the ordinal decision setting are satisfied when the preference on acts derives from a possibility distribution on S and a weak ordering on X.

Example 1 (continued)
Consider again the 3-state/3-consequence example of Section 3. If a uniform probability is changed into a uniform possibility distribution, then it is easy to check that the likely dominance rule yields
$f \sim g \sim h$. However, if $s_1 >_{\Pi} s_2 >_{\Pi} s_3$ then
$[f >_P g] = \{s_1, s_2\} >_{\Pi L} [g >_P f] = \{s_3\}$;
$[g >_P h] = \{s_1, s_3\} >_{\Pi L} [h >_P g] = \{s_2\}$;
$[f >_P h] = \{s_1\} >_{\Pi L} [h >_P f] = \{s_2, s_3\}$.
So f > g > h follows. It contrasts with the cycles obtained with a probabilistic approach.

The lack of transitivity of the indifference relation can be observed in very simple cases, where only 3 states and 2 consequences are used:

Counterexample for the transitivity of $\sim$ for acts

Consider again the 3-state example of Section 3, with a uniform possibility distribution. Suppose $X = \{x, y\}$ and $x >_P y$.

f: $f(s_1) = x$ $f(s_2) = x$ $f(s_3) = y$

g: $g(s_1) = y$ $g(s_2) = y$ $g(s_3) = x$

h: $h(s_1) = y$ $h(s_2) = x$ $h(s_3) = y$

One can verify that:

$[f \geq_P g] = \{s_1, s_2\} \sim$ □$_L$ $[g \geq_P f] = \{s_3\}$ thus: $f \sim g$;

$[g \geq_P h] = \{s_1, s_3\} \sim$ □$_L$ $[h \geq_P g] = \{s_1, s_2\}) = 0$ thus: $g \sim h$;

$[f \geq_P h] = \{s_1, s_2, s_3\} >$□$_L$ $[h \geq_P f] = \{s_2, s_3\}$ thus: $f > h$.

6.3 Representation results

Possibility theory is closely related to preferential inference in nonmonotonic reasoning, as defined by Kraus, Lehmann and Magidor (1990). Indeed, they call rankings what we call comparative possibility distributions. Assume no null events. Then, Dubois and Prade (1995a) have shown that an inference relation □ is preferential if and only if there exists a set of *non-dogmatic* possibility orderings F such that if A ≠ Ø:

A □ B if and only if □ $>_□$ □ F, A □ B $>_□$ A □ B^c

Non-dogmatic possibility orderings are such that A $>_□$ Ø if and only if A ≠ Ø. According to non-dogmatic possibility orderings, no event is impossible, but Ø. Since the likelihood relation coming from our decision-theoretic setting $>_L$ satisfies OR, AND, RW, CM, CUT, this result applies when restricting $>_L$ to disjoint events :

Proposition 16: If the relation on X^S satisfies the axioms of the ordinal decision setting, there is a family F of possibility orderings such that □X,Y disjoint,

$X >_L Y$ if and only if □ $>_□$ □ F, X $>_□$ Y

Sketch of proof

Consider non null events : all the properties of system P are satisfied, including reflexivity. So, there is a family of possibility orderings such that

A □ B □ □ $>_□$ □ F, A □ B $>_□$ A □ B^c □ A □ B $>_L$ A □ B^c

with X = A □ B and Y = A □ B^c, we get □ $>_□$ □ F, X $>_□$ Y □ X $>_L$ Y.

To encompass null events, it is enough to extend these possibility orderings so that whatever Y null, Y$\sim$ □ 0. Indeed, the set of null events is the family subsets of NULL. In terms of possibility distributions, a possibility distribution assigns degree 0 to each null state.

Since $>_L$ is an autodual additive relation, the ordering between non disjoint events is completely defined by the ordering between disjoint events. Hence Proposition 16 holds for any pair of events X, Y without restricting to disjoint events. Overall, putting together the results obtained so-far and the ones of Dubois and Prade (1995a) on the possibilistic representation of the non-monotonic system P, one gets the bridge between the ordinal decision making framework and possibility theory:

Theorem 1: If a partially ordered set of acts $(X^S, \geq)$ satisfies the axioms of the ordinal decision setting, then there is a family F of possibility orderings on S and a relation $\geq_P$ on X such that f > g □ □ $>_□$ □ F, $[f >_P g] >_□ [g >_P f]$. Generally, relation $\geq_P$ is a weak order unless there is only one non null state.

Proof:
Suppose 2 non null states at least. A consequence of Proposition 8 we know that the projection of $(X^S, \geq)$ on X is a weak order. Let us denote $\geq_P$ this weak order. If holds that f > g □ $[f \geq_P g] >_L$ and that $[g \geq_P f]$□ $[f >_P g] >_L [g >_P f]$.
From proposition 15 we also know that the projection $(X^S, \geq)$ on the disjoint events of S defines a family of possibility orderings such that $X >_L Y$ if and only if □ >□ □ F, X >□ Y where X and Y are disjoint. With X = $[f >_P g]$ and Y = $[g >_P f]$, f > g □ □ >□ □ F, $[f >_P g]$ >□ $[g >_P f]$.
When there is only one non null state the theorem obviously holds since the likelihood relation reduces to a single, very particular possibility ordering.

So, starting from a general framework on acts, that respects Savage's approach to a large extent, we find a likelihood relation which can be represented by a family of possibility orderings, since the preferential entailment of system P can always be represented in terms of such a family. Notice in particular that the relation on states obtained from our axiomatic framework is not necessarily a weak order, so that it cannot always be represented by a single possibility ordering. However, given any family F of possibility orderings, it is routine to check that the induced likely dominance decision rule such that:

f > g □ □ >□ □ F, $[f >_P g]$ >□ $[g >_P f]$

does satisfy propositions 14 and 15. All above proofs for a single possibility distribution straightforwardly extend to families thereof. Hence Theorem 1 can be strengthened into a genuine representation theorem.

Theorem 2: A partially ordered set of acts $(X^S, \geq)$ satisfies the axioms of the ordinal decision setting, *if and only if* there is a family F of possibility orderings on S and a relation $\geq_P$ on X such that: f > g □ □ >□ □ F, $[f >_P g]$ >□ $[g >_P f]$

In order to get a *unique* possibilistic likelihood relation as the representation of uncertainty, first recall that the nonmonotonic inference which expresses that A □ B >□ A □ B^c for one possibility ordering >□ coincides with the so-called "rational" inference of Lehmann (Lehmann and Magidor, 1992), that is an inference relation which satisfies rational monotony (Benferhat et al., 1997) .

Rational monotony: A □ C and not A □ B^c imply A □ B □ C

In the above, "not A □ B^c " means that it is not the case that B generally does not hold, in situations where A holds. Indeed if B^c is expected then it might well be an exceptional A-situation, where C is no longer normal. In terms of the subjective likelihood relation this axiom reads:

Axiom RM: A □ C $>_L$ A □ C^c and A □ B $\geq_L$ A □ B^c □ A □ B □ C $>_L$ A □ B □ C^c.

Adding the RM axiom to system P forces the relation on states to be a weak order, thus corresponding to a single possibilistic ordering as proved in (Benferhat et al., 1997). However, it is not easy to express RM in terms of acts in a non-trivial way. The proof of CM in Proposition 12 could be mimicked assuming A □ C $>_L$ A □ C^c and A □

$B \geq_L A \cap B^c$ (instead of $A \cap B >_L A \cap B^c$). Then we get respectively $f \geq g$ and $g > h$ for the corresponding acts. RM implies $A \cap C \cap B >_L A \cap C^c \cap B$ which means $f > h$. It yields : "$f \geq g$ and $g > h$ implies $f > h$". However this axiom is equivalent to requiring the transitivity of $\geq$ on acts. It makes the system collapse to the usual Savage setting and then leads to special orderings on S such as described in Proposition 3. It thus captures only very special possibility orderings. Decision axioms for deriving unique general comparative possibility orderings are under study (Dubois et al, 2002).

To summarize this Section, we have proved that in the presence of the ordinal invariance assumption, the only underlying "rational" subjective theory of uncertainty can be expressed in terms of possibilistic likelihood relations. Then the uncertainty on states must be represented by one or several possibility distributions instead of a probability distribution. By adopting an ordinal invariance assumption, we rule out additive numerical representations. Moreover, it leads to a unique decision rule that reconstructs the ordering on acts from the uncertainty on events and the preference relation on consequences: the likely dominance rule, whose expressiveness turns out to be limited as shown now.

7 - What is the Behavior of the Ordinal Decision Maker under Uncertainty?

Let us study the likely dominance rule induced by a single possibility distribution (i.e., the possibilistic likelihood relation it induces). It is clear that when the subjective likelihood relation derives from several possibility distributions, it only produces more reasons for incomparability between acts. So the case of a single possibility distribution is the most favorable for the likely dominance rule to be decisive.

If the decision maker is ignorant about the state of the world, all states are equipossible, and all events but Ø are equally possible as well. Possibilistic likelihood then coincides with set-inclusion. Namely, $A >_{\Pi L} B$ if and only if $B \subset A$. So, if f and g are such that $[g >_P f] \neq Ø$ and $[f >_P g] \neq Ø$ hold, then none of $f > g$ and $g > f$ hold as per the likely dominance rule, since such events are always disjoint. Then X^S is partially ordered only via Pareto dominance in the sense of $\geq_P$ that is:

$$f > g \text{ if and only if } \forall s, f(s) \geq_P g(s) \text{ and } \exists s, f(s) >_P g(s).$$

This method, although totally sound, is not decisive at all (it corresponds to the unanimity rule in voting theory).

Conversely, if there is an ordering $s_1, \ldots, s_n$ of S such that $\pi(s_1) > \pi(s_2) > \ldots > \pi(s_n)$, then for any A, B, $A \cap B = Ø$, either $A >_{\Pi L} B$ or $B >_{\Pi L} A$. Hence $\forall f \neq g$, either $f > g$ or $g < f$. Moreover this is a lexicographic ranking:

$$f > g \text{ if and only if } \exists k \text{ such that } f(s_k) >_P g(s_k) \text{ and } f(s_i) \sim_P g(s_i), \forall i < k.$$

It corresponds to the procedure: check if f is better than g in the most normal state; if yes prefer f; if f and g give equally preferred results in s_1, check in the second most normal state, and so on recursively. This comes down to a lexicographic ranking of vectors ($f(s_1)$, …, $f(s_n)$) and ($g(s_1)$, …, $g(s_n)$). It is a form of dictatorship by most plausible states, in voting theory terms.

More generally any weak order splits S into a well ordered partition $S_1 \cup S_2 \cup \ldots \cup S_n = S$, $S_i \cap S_j = \emptyset$ ($i \neq j$), such that states in each S_i are equally plausible and states in S_i are more plausible than states in S_j, $\forall j > i$. In that case the ordering of events is defined as follows: $A >_{L} B$ if and only if $\min\{i: S_i \cap A \cap B^c\} < \min\{i: S_i \cap B \cap A^c\}$, and the ordering of acts is defined by:

- $f > g$ if and only if $\exists k \geq 1$ such that: $\forall s \in S_1 \cup S_2 \cup \ldots \cup S_{k-1}$, $f(s) \sim_P g(s)$, and $\forall s \in S_k$, $f(s) \geq_P g(s)$ and $\exists s \in S_k$, $f(s) >_P g(s)$

- $f \sim g$ if and only if $\exists k \geq 1$ such that: $\forall s \in S_1 \cup S_2 \cup \ldots \cup S_{k-1}$, $f(s) \sim_P g(s)$, and $\exists s \neq s' \in S_k$ $f(s) >_P g(s)$ and $g(s') >_P f(s')$.

Informally, the decision maker proceeds as follows: f and g are compared on the set of most normal states (S_1): if f performs better than g in each $s \in S_1$, then f is preferred to g; if there is a disagreement in S_1 about the relative performance of f and g then f and g are indifferent. If f and g have equally preferred consequences in each most normal state then the decision maker considers the set of second most normal states S_2, etc. In other words, the decision maker applies a unanimity rule on the most plausible states, and a prioritized (lexicographic) procedure when a situation of indifference between acts is met; the next most plausible states are used to break ties. In a nutshell it is a prioritized Pareto-dominance relation.

This behavior contrasts, although a clear similarity exists, with Lehmann (1996)'s approach which keeps the prioritized procedure across unequally plausible states (implausible states are neglected) and recommends a standard expected utility rule among the equally plausible ones. Moreover it is striking to notice the similarity between the decision rules displayed above, that require Pareto-dominance of best acts over the most plausible states, and voting procedures found in social choice (Mas-Colell and Sonnenschein, 1972, Weymark 1984; see also Sen,1986) by relaxing the transitivity of the social preference and assuming the transitivity of its strict part. They do find that decisions should be made unanimously by an oligarchy of voters (here, the most plausible states), but they do not insist on the prioritized sequence of oligarchies.

Example 2:

Consider the omelet example of Savage (1972, pages 13 to 15) recalled on Table 1. The problem is how to make a six-egg omelet from a five-egg one. The new egg can be fresh or rotten. There are three feasible acts : break the egg in the omelet (BIO); break it apart in a cup (BAC); or throw it away (TA). The set of 6 consequences is given in Table 1. Integers between parentheses indicate the intuitive ordering of consequences, and the greater the number the higher the preference. Since only two states are present, the ordering of events can always be represented by only one possibility ordering. Let us apply the likely dominance rule. If fresh egg is more likely than rotten egg, then

$[BIO >_P BAC] = [BIO >_P TA] = [BAC >_P TA] = \{fresh\}$

$>_L [BAC >_P BIO] = [TA >_P BIO] = [TA >_P BAC] = \{rotten\}$

and the best act is clearly BIO. If the decision maker thinks the egg is rotten, then the best act is TA. In case of total ignorance, the three acts are indifferent. So the decision making attitude induced by the approach is: break the egg in the omelet if you think the egg is fresh, throw it away if you think it is rotten, and do anything you like if you have no opinion (all acts equally preferred then). Clearly, this may result in many starving days, and garbage cans filled with lots of spoiled fresh eggs, in the case when the state of the egg is usually hard to tell.

ACTS STATES	fresh egg	rotten egg
break the egg in the omelet BIO	a 6 egg omelet (6)	nothing to eat (1)
break it apart in a cup BAC	a 6 egg omelet, a cup to wash (5)	a 5 egg omelet, a cup to wash (3)
throw it away TA	a 5 egg omelet, one spoiled egg (2)	a 5 egg omelet (4)

Table 1

Of course the extreme peculiarity of this example masks the richness of possible blends between unanimity and dictatorship occurring in the case of more than two states, when a single possibility ordering is involved.

It is interesting to compare the ranking of acts obtained above on the egg example with the one obtained using the pessimistic possibilistic utility proposed in (Dubois, and Prade, 1995c). This criterion presupposes that utilities and degrees of necessity belong to the same totally ordered scale, here L = {0, 1, 2, 3, 4, 5, 6}, equipped with its order-reversing map n. The pessimistic utility of an act f when the knowledge of the state is given by a L-valued possibility distribution π: S $\to$ L and the utility of consequence given by a L-valued utility u: X $\to$ L, is expressed as:

$$u_*(f) = \min_{s \in S} \max(n(\pi(s)), u(f(s))) = \min_{s \in S} \max(N(S\setminus\{s\}), u(f(s))).$$

This utility is a mild extension of the pessimistic maximin decision rule (that rates an act on the basis of its least preferred consequence) advocated and axiomatized by Brafman and Tennenholtz (2000) in terms of conditional policies (rather an acts). See Dubois Prade and Sabbadin (2001) for the act-driven axiomatization of $u_*(f)$.

The utilities of the three acts in the egg example are then given as a function of states F and R:

$$u_*(BIO) = \min(\max(N(R), 6), \max(N(F), 1)) = \max(N(F), 1)$$

$$u_*(BAC) = \min(\max(N(R), 5), \max(N(F), 3))$$

$$u_*(TA) = \min(\max(N(R), 2), \max(N(F), 4))$$

Table 2 exhibits the best acts as a function of the decision-maker's knowledge about the egg. The model recommends act BAC in case of relative ignorance on the egg state, that is when max(N(F), N(R)) is not high enough (less than 3). In practice, it is indeed advisable to act cautiously, and to break the egg in a spare cup in case of serious doubt. The pessimistic possibilistic criterion seems to deliver a more realistic advice than the one prescribed by the likely dominance rule which does not presupposes commensurability between uncertainty and utility.

N(F)	N(R)	u_*(BIO)	u_*(BAC)	u_*(TA)	Best acts
6	0	6	5	2	BIO
5, 4, 3	0	N(F)	N(F)	2	BIO or BAC
2	0	2	3	2	BAC
1, 0	0, 1, 2	1	3	2	BAC
0	3	1	3	3	BAC or TA
0	≥ 4	1	3	4	TA

Table 2

Other decision rules proposed in the literature could benefit from the above study. For instance Boutilier (1994) compares acts on the basis of the worst (or the best) consequences in the set of most plausible states. This decision rule lacks discrimination power. It can be made more discriminating by allowing ties to be broken by checking the second most plausible states, and so on like for the likely dominance rule. The above possibilistic decision rule and the maximin rule can consider two acts as indifferent while one Pareto-dominates the other. Of course, the acts are then totally preordered. On the contrary, the likely dominance rule respects the Pareto dominance, but it may lead to incomparabilities that can sometimes be solved using a maximin rule. This complementarity suggests a combination of ideas of prioritization and respect of Pareto-dominance stemming from our study and the maximin rule. The maximin rule can be forced to respect Pareto-dominance by comparing acts f and g on the basis of their worst *discriminating* consequences (in $[f >_P g] \square [g >_P f]$ only). This decision rule was proposed and axiomatized by Cohen and Jaffray (1980). Clearly, acts will be only partially ordered using this improved maximin decision rule, but this partial ordering refines the maximin rule. Again one may apply this rule in a prioritized way: the improved maximin decision rule can be substituted to unanimity within the likely dominance rule inside the oligarchies of states. The Savagean-like framework proposed by Lehmann (1996) also leads to a sequence of oligarchies of states, with a comparative probability inside the oligarchies. No decision rule is propose but it is easy to imagine a counterpart of the likely dominance rule where expected utility applies inside the oligarchies of states. However reasonable these refined decision rules may look, they need further work to be formally justified.

8 - Conclusion

This paper suggests that purely symbolic approaches to both rational and practically useful criteria for decision making under uncertainty have serious limitations in terms of expressive power. We have tried to adapt Savage's decision theory to a purely ordinal setting. The addition of the OI axiom to the relaxed Savage turns out to drastically restrict the type of uncertainty that can represent the decision maker's knowledge about the world, as well as the possible decision rules that can be rationally applied. Despite the presence of well-known Savage axioms, including the sure-thing principle, the admissible uncertainty functions contain only very special kinds of probability functions ("big-stepped probabilities" pointed out in Section 3.2), due to a Condorcet-like intransitivity effect. By relaxing the weak order assumptions for comparing acts, decision-theoretic foundations of non-monotonic reasoning based on preferential entailment have been laid bare as a by-product.

The decision rule obtained from first principles in this way is likely dominance. It is either overconfident by overfocusing on most plausible states, or indecisive because requiring Pareto-dominance of acts on these most plausible states. Since preferential inference is very cautious, the relations on acts which do not correspond to a total likelihood ordering on states will not be very discriminating. On the contrary, if the set of states is totally ordered in terms of plausibility (when the weak preference on acts is transitive), the decisions will be very risky: as usual with nonmonotonic inference, the decision maker must assume that the world is in the most normal state. Cautious decisions will never be preferred. The decision rules obtained by means of a purely ordinal representation of uncertainty and utility have common features with other ones proposed on intuitive grounds in the nonmonotonic reasoning literature (especially Boutilier, 1994). They logically recommend to stick to the most plausible states when discriminating between acts, in the tradition of preferential entailment. While being natural when modeling acceptance of statements for the purpose of making inferences about the world, this view may be questioned in the case of making decisions whose results affect the world, and where focusing on a few plausible states may be harmful if very unlikely states have very undesirable consequences.

It is thus difficult to maintain that the decision guidelines offered by the theory developed in this paper are completely reasonable since they reject the most usual decision rules, like expected utility and the maximin criterion. In contrast, the possibilistic extension of the maximin criterion, based on an absolute finite scale for utility and likelihood looks more intuitively satisfactory as shown in the example of the previous Section. However the possibilistic criterion relies on a commensurability assumption between uncertainty and preference, which may be questioned by tenants of a purely symbolic approach, (although it is part of Savage's framework proper).

The results of this paper are rather negative for decision theory when only ordinal information about uncertainty and preference on consequences is available and no common evaluation scale is assumed for uncertainty and preference. This is not really a surprise for anybody familiar with voting theory and aware of its similarity with DMU (Fargier and Perny, 1999) emphasize this point). In voting theory, the ordinal setting leads to either complex and rationally debatable procedures, or simple but morally dubious ones, or yet indecisive ones. This paper seems to conclude that there is no alternative to the likely dominance rule that would enable a reasonably cautious and albeit decisive ordering on acts to be computed on the single basis of ordering relations on events and consequences, under the constraints induced by the sure thing principle. In other words, we cannot do away with comparability assumptions between preference and uncertainty, nor avoid numerical approaches if human-like decision attitudes under uncertainty, that remain decisive and rational are to be captured.

This view is probably too pessimistic. Actually, there might be some unexplored possibilities by weakening axiom OI while avoiding the notion of certainty equivalent of an uncertain act. It must be stressed that OI requires more than the simple ordinal nature of preference and uncertainty (i.e. more than separate ordinal scales for each of them). Condition OI also involves a condition of *independence with respect to irrelevant alternatives* (in the sense of Arrow 1969, Sen 1986). It says that the preference f ≥ g only depends on the relative positions of quantities f(s) and g(s) on the preference scale. This unnecessary part of the condition could be cancelled within the proposed framework, thus leaving room for a new family of rules not considered in this paper, for instance involving a third act or some prescribed consequence considered as an aspiration level.

Besides, it is noticeable that neither the set of axioms proposed by Savage nor the ones suggested in this paper contain one that refers to the decision maker attitude towards uncertainty (like in the maximin rule). This may explain why cautious decisions seem to be ignored. Sometimes, this is due to overfocusing on plausible states, which tend to assume the decision maker thinks the state of the world is indeed what he/she considers most plausible, while other states may not be so implausible as to justify their being neglected. Sometimes it is due to indecisiveness, when the decision maker has little knowledge of the state, and cautious acts that might be advisable never dominate other ones. In order to explicitly handle the decision maker 's level of confidence in front of incomplete information, one might think of reverting the likely dominance rule by exchanging the role of states and consequences. It would lead to express the preference on acts in terms of a comparison of sets of consequences instead of sets of states as done here. Since only single consequences are compared by the utility relation, this method presupposes that the preference relation on consequences be itself lifted to subset of consequences. Then pessimistic and optimistic attitudes can be captured in the sets of axioms, by specifying whether a set of consequences is appraised in terms of best or worst consequence in the set. This is explicitly done when axiomatizing the maximin criterion, or its possibilistic extensions, already mentioned. Future works should also strive towards exploiting the complementarities of prioritized Pareto-efficient decision methods as the likely dominance rule, and of the pessimistic decision rules related to the maximin criterion. Lastly, our results also apply, with minor adaptation, to multicriteria decision making, where the various objectives play the role of the states of nature in DMU (Dubois et al., 2001).

Appendix 1 Comparative Possibility Relations

The most natural framework for representing uncertainty without numbers is to define a comparative likelihood relation among events, viewed as subsets A, B, C... of a state space S. Comparative probabilities, introduced in Section 2, are examples of such relations. Perhaps the simplest types of other comparative likelihood relations are those induced by possibility and necessity measures:

Definition A1: $\geq_\Pi$ is a possibility ordering if and only if

i) $\geq_\Pi$ is complete and transitive

ii) S $>_\Pi$ Ø,

iii) $\forall A$, $A \geq_\Pi$ Ø

iv) B $\geq_\Pi$ C implies $A \cup B \geq_\Pi A \cup C$

$B \geq_\Pi C$ reads B is at least as possible (= plausible) as C.

Definition A2: $\geq_N$ is a necessity ordering if and only if

i) $\geq_N$ is complete and transitive

ii) $S >_N \emptyset$,

iii) $\forall A$, $S \geq_N A$

iv) $B \geq_N C \Rightarrow A \cap B \geq_N A \cap C$

$B \geq_N C$ reads B is at least as necessary (= certain) as C.

Necessity orderings, introduced by Dubois (1986), coincide with epistemic entrenchment relations in the sense of Gärdenfors (1988) (see Dubois and Prade, 1991), and possibility orderings have been introduced by Lewis (1973), then rediscovered by Dubois (1986), and also studied by Grove (1988) in the scope of belief revision. Possibility and necessity orderings are dual in the sense that for a given possibility ordering $\geq_\Pi$, the relation defined by $A \geq_N B$ if and only if $B^c \geq_\Pi A^c$ is a necessity ordering, and conversely. A pair $(\geq, \geq_\Pi)$ such that $A \geq B$ if and only if $B^c \geq_\Pi A^c$ is called a pair of dual uncertainty relations (while comparative probability in the sense of definition 2 is autodual).

A pair of dual necessity and possibility orderings is a particularly simple on finite sets since that they both derive from a weak order $\geq_\pi$ on states only.

Any necessity ordering on a finite set can be represented by a necessity function on a linearly ordered scale L:

Proposition A3 (Dubois, 1986): $\geq_N$ is a necessity ordering if and only if there is a finite totally ordered set L and $\exists N: 2^S \rightarrow L$ such that: $B \geq_N C \Leftrightarrow N(B) \geq N(C)$, where $N(F) \in L$ and $N(F \cap G) = \min(N(F), N(G))$ $\forall$ F and G.

L is isomorphic to the quotient set $2^S / \sim_N$ when $\sim_N$ is the indifference part of $\geq_N$. $0 \in L$ denotes the bottom element of L (equivalence class of Ø) and $1 \in L$ denotes the top element of L (equivalence class of S), and $N(S) = 1 > N(\emptyset) = 0$.

The same kind of property holds for possibility orderings with respect to possibility functions:

Proposition A4 (Dubois, 1986): $\geq_\Pi$ is a possibility ordering if and only if $\exists \Pi$: $2^S \rightarrow L$ such that $B \geq_\Pi C \Leftrightarrow \Pi(B) \geq \Pi(C)$ where $\Pi(B) \in L$ and $\Pi(B \cup C) = \max(\Pi(B), \Pi(C))$ $\forall$ B and C.

The comparative possibility distribution $\geq_\pi$ can be encoded as a possibility distribution π, which is a mapping $S \rightarrow L$. Namely $\Pi(A) = \max_{s \in A} \pi(s)$ and $N(A) = n(\Pi(A^c)) = \min_{s \notin A} n(\pi(s))$ where n is the order reversing function $L \rightarrow L$. In terms of possibility measures (Zadeh, 1978), $\Pi(A) = n(N(A^c)) = 1 - N(A^c)$ if $L = [0,1]$.

Appendix 2: Proofs

Proof of Proposition 1

Assume (U) and $fAg \geq g$, $gAf \geq g$. Then using S2, $fAh \geq gAh$ for any h, and $h'Af \geq h'Ag$ for any h'. The first preference statement reads $(f \geq g)_A$, and the second $(f \geq g)_{A^c}$ noticing that h'Af is the same act as fA^ch'. Hence $f \geq g$ by (U). Conversely assume $(f \geq g)_A$, $(f \geq g)_{A^c}$ and WSTP. The two first assumptions mean $\forall h$, $fAh \geq gAh$ and $\forall h'$, $h'Af \geq h'Ag$. Now just specialize these with h = g and h' = g and apply WSTP to get $f \geq g$.

Proof of Proposition 2

Let $x, y, z \in X$ be such that $x >_P y >_P z$. Suppose $\exists A, B, C$, pairwise disjoint, such that $B \cup C >_L A >_L \emptyset$, $A \cup C >_L B >_L \emptyset$, $A \cup B \geq_L C$. Consider the three acts f, g, h as follows

	f	g	h
A	x	y	z
B	y	z	x
C	z	x	y

and $\forall s \notin A \cup B \cup C$, $f(s) = g(s) = h(s)$. Let $D = \overline{A \cup B \cup C}$. Note that

$[f \geq_P g] = A \cup B \cup D$, $[g \geq_P f] = C \cup D$

$[g \geq_P h] = A \cup C \cup D$, $[h \geq_P g] = B \cup D$

$[h \geq_P f] = B \cup C \cup D$, $[f \geq_P h] = A \cup D$

$B \cup C >_L A \Rightarrow B \cup C \cup D >_L A \cup D$ (using S2) $\Rightarrow h > f$ (using the likely dominance rule). Similarly $g > h$ derives from $A \cup C >_L B$. Lastly, using again the likely dominance rule we find that $A \cup B \geq_L C$ implies $f \geq g$. However since $\geq$ is a weak order, the strict part of $\geq$ is transitive so that $g > h$ follows. Contradiction.

Proof of proposition 3:

Assume $\exists i, j, k$, $s_i \sim_L s_j >_L s_k >_L \emptyset$ then the preadditivity axiom leads to accept $\{s_i, s_k\} >_L s_j$ and $\{s_j, s_k\} >_L s_i$. Indeed $s_k >_L \emptyset$ implies $\{s_k, s_i\} >_L s_i$ using the preadditivity axiom (that derives from S2)) and since $s_i \sim_L s_j$, using transitivity of $\geq_L$ we find $\{s_k, s_i\} >_L s_j$. Now using (N) we conclude $s_k >_L \{s_i, s_j\}$ contrary to the assumption. Hence if $s_i >_L s_k > \emptyset$ and $s_j >_L s_k > \emptyset$ we have that either $s_i >_L s_j$ or $s_j >_L s_i$. This proves (i). As for (ii) just notice that $s_i >_L s_j$ and $s_i >_L s_k$ imply $\{s_i, s_k\} >_L s_j$ and $\{s_i, s_j\} >_L s_k$ and apply (N). It is noticeable that the transitivity of the likelihood relation plays no role in the latter result.

Proof of proposition 4

First let us prove that S2 holds, that is, $\forall A, f, g, h, h'$, $fAh \geq gAh$ if and only if $fAh' \geq gAh'$. It is enough to notice that $(fAh, gAh) \equiv (fAh', gAh')$. Indeed if $s \in A$ then the consequences remain the same; if $s \notin A$, then since relation $\geq$ on acts is reflexive, it holds that $h(s) \geq_P h(s)$ and $h'(s) \geq_P h'(s)$.

Now for S4, a similar reasoning applies. S4 reads $\forall x, y, x', y' \in X$ s.t. $x >_P y$, $x' >_P y'$, $xAy \geq xBy \Leftrightarrow x'Ay' \geq x'By'$. We shall show that $(xAy, xBy) \equiv (x'Ay', x'By')$. There are four cases:

-) $s \in A \cap B$: $xAy(s) = x \geq_P xBy(s) = x$ and $x'Ay'(s) = x' \geq_P x'By'(s) = x'$ using reflexivity.

-) $s \in A^c \cap B$: $xBy(s) = x >_P xAy(s) = y$ and $x'By'(s) = x' >_P x'By'(s) = y'$

-) $s \in A \cap B^c$: $xAy(s) = x >_P xBy(s) = y$ and $x'By'(s) = x' >_P x'By'(s) = y'$

-) $s \in A^c \cap B^c$: $xAy(s) = y \geq_P xBy(s) = y$ and $x'By'(s) = y' \geq_P x'By'(s) = y'$ using reflexivity.

Hence S4 holds under OI.

Proof of Proposition 5

Note that since OI holds S2 and S4 hold too. Let $\geq_P$ is the restriction of $\geq$ on X (due to definition 5) and $\geq_L$ be the projection of $\geq$ on events (due to S4). If all consequences on X are indifferent then all acts are indifferent and the problem is trivial since $[f \geq_P g] = [g \geq_P f] = X$ and $\geq_P$ is complete. So, for all acts, $f \sim g$ and $[f \geq_P g] \sim_L [g \geq_P f]$.

Due to S5, there are $x, y \in X$ such that $x >_P y$. Under S4 the likely dominance rule also reads $f \geq g \Leftrightarrow x[f \geq_P g]y \geq_L x[g \geq_P f]y$, in terms of binary acts. Now we just have to show that $(f, g) \equiv (x[f \geq_P$

g]y, x[g $\geq_P$ f] y). Indeed,

-) if s is such that f(s) $\sim_P$ g(s), then s $\in$ [f $\geq_P$ g] $\cap$ [g $\geq_P$ f]. Hence

$$x[f \geq_P g]y(s) = x[g \geq_P f]\, y(s) = x$$

and x[f $\geq_P$ g]y(s) $\sim_P$ x[g $\geq_P$ f] y(s) using reflexivity of $\geq_P$

-) if s is such that f(s) $>_P$ g(s), then s $\in$ [f $\geq_P$ g] and s $\notin$ [g $\geq_P$ f]. Hence

x[f $\geq_P$ g]y(s) = x and x[g $\geq_P$ f] y(s) = y and x[f $\geq_P$ g]y(s) $>_P$ x[g $\geq_P$ f] y(s)

-) if s is such that g(s) $>_P$ f(s), then x[g $\geq_P$ f]y(s) $>_P$ x[f $\geq_P$ g] y(s) using the same reasoning.

There are no other cases since $\geq_P$ is complete. Hence under the assumptions of completeness and reflexivity of the preference relations, OI is strictly equivalent to the likely dominance rule.

Proof of Proposition 6

By S5 it holds that: $\exists$x, y s.t. fx > fy.

Non-triviality: Since xSy = fx and xØy = fy, we get xSy > xØy i.e., by S4: S $>_L$ Ø.

Non-Contradiction: By S3, if A is not null, it holds that $(fx > fy)_A$, that is xAh > yAh for all acts h. Take h = fy, we find xAy > yAy = xØy, hence using S4: A $>_L$ Ø. If A is null, fAh ≥ gAh, for all f, g, h. Choosing f = fx; h = fy; g = fy leads to xAy ≥ yAy = xØy, hence using S4: A $\geq_L$ Ø.

Compatibility with null events: Continuing the proof with a null event A, we now let f = h = fy; g := fx and get yAy = xØy ≥ xAy, i.e., Ø $\geq_L$ A. Hence if A is null then A $\sim_L$ Ø. And it is proved above that if A is not null then A $>_L$ Ø.

Preadditivity: By definition of $\geq_L$ between events and S4, B $\geq_L$ C $\Leftrightarrow$ xBy ≥ xCy.

Write xBy as (xBy)(B $\cup$ C)y, and xCy as (xCy)(B $\cup$ C)y.

Then using (S2) (xBy)(B $\cup$ C)y ≥(xCy)(B $\cup$ C)y is equivalent to (xBy)(B $\cup$ C)h ≥ (xCy)(B $\cup$ C)h for all h. Choose h = xAy. Then xBy ≥ xCy if and only if (xBy)(B $\cup$ C)(xAy) ≥ (xCy)B $\cup$ C(xAy). Now A $\cap$ (B $\cup$ C) = Ø, then act (xBy)(B $\cup$ C)(xAy) has value x on B $\cup$ A, and y otherwise and is equal to x(B $\cup$ A)y, similarly (xCy)(B $\cup$ C)(xAy) = x(C $\cup$ A)y.

Hence B $\geq_L$ C is equivalent to B $\cup$ A $\geq_L$ C $\cup$ A.

Stability of non null events: From S5, $\exists$ x $>_P$ y, fx > fy. A $>_L$ Ø writes xAy > xØy; suppose A $\subseteq$ B and B $\sim_L$ Ø, then B is null. Then $\forall$f, g, $(f \sim g)_B$, i.e., fBh $\sim$ gBh, $\forall$f, g, h. Assume f = xAy, g = fy = h then (xAy)By = xAy, yBy = xØy. Hence B $\sim_L$ Ø implies xAy $\sim$ xØy, i.e., A $\sim_L$ Ø contradiction.

Autoduality: A $\geq_L$ B $\Leftrightarrow$ (A $\cap$ B) $\cup$ (A $\cap$ B^c) $\geq_L$ (A $\cap$ B) $\cup$ (A^c $\cap$ B)

$\Leftrightarrow$ A $\cap$ B^c $\geq_L$ A^c $\cap$ B (from preadditivity)

$\Leftrightarrow$ (A $\cap$ B^c) $\cup$ (A^c $\cap$ B^c) $\geq_L$ (A^c $\cap$ B) $\cup$ (A^c $\cap$ B^c)

$\Leftrightarrow$ B^c $\geq_L$ A^c

A $>_L$ B $\Leftrightarrow$ B^c $>_L$ A^c is proved by contraposition of the previous property.

Inclusion-monotony: By non-contradiction, we have: B \ A = A^c $\cap$ B $\geq_L$ Ø. Since A $\subseteq$ B, it also reads A^c $\cap$ B $\geq_L$ A $\cap$ B^c = Ø, i.e., using preadditivity: B $\geq_L$ A.

Characterization of null events comes from stability of non null events (contraposition)

Proof of Proposition 7

Let x and y be two consequences such that x$>_P$y. Let f = fx(A $\cup$ C)fy and g= fx(B $\cup$ D)fy.

A $\geq_L$ B implies that $(f \geq g)_{A \cup B}$. C $\geq_L$ D implies that $(f \geq g)_{C \cup D}$. Since (A $\cup$ B) $\cap$ (C $\cup$ D) = Ø, i.e (C $\cup$ D) $\subseteq$ (A^c $\cap$ B^c) and f = g = fy on A^c $\cap$ B^c $\cap$ C^c $\cap$ D^c, we get $(f \geq g)_{A^c \cap B^c}$. By (U), we get f ≥ g, i.e. A $\cup$ C $\geq_L$ B $\cup$ D.

- 7.1 : direct from Prop. 7 with A = C = Ø, B = {s} and D = {s'}.
- 7.2 : direct from Prop. 7 with A = D = Ø, B = {s} and C = {s'}.
- 7.3 By contraposition: E $\sim_L$ Ø $\Rightarrow$ F $\geq_L$ E. Suppose that E $\sim_L$ Ø. Since for any F, F $\geq_L$ Ø : Prop. 7 yields: E $\sim_L$ Ø and F $\geq_L$ Ø imply F $\geq_L$ E as for (7.2)
- 7.4 This property is verified without using WSP nor U if E \ D $>_L$ Ø or G $>_L$ F (propositions 7.2 and 7.3). Suppose that G $\sim_L$ F and E \ G $\sim_L$ Ø. From Prop. 7, we get E $\geq_L$ F (with A = G, B = F, C = E \ G, D := Ø.)
- 7.5 : if every state is null, then by 7.1, S is null : contradiction with Prop. 6 (non triviality).

Proof of Proposition 8

By (WS1), we know that ≥ is complete and > transitive. We now have to show that the indifference relation $\sim$ that one can define from $\geq$ is also transitive on constant acts. If X has only 2 elements the problem is trivial. Suppose more than 2 elements.

Suppose that x,y, and z are three elements of X s.t. $x \sim_P y$, $y \sim_P z$ and $x >_P z$.

Suppose that at least two states of S are not null states. Let s_1 and s_2 be two of the non null states and compare the decisions g, h, k:

g: $g(s_1) = x$, $g(s_2) = y$, $g(s) = x$ if $s \notin \{s_1, s_2\}$

h: $h(s_1) = z$, $h(s_2) = x$, $h(s) = x$ if $s \notin \{s_1, s_2\}$

k: $k(s_1) = y$, $k(s_2) = z$, $k(s) = x$ if $s \notin \{s_1, s_2\}$

Thus, $[g \geq_P h] = S$ and $[h \geq_P g] = S \setminus \{s_1\}$. Since $\{s_1\} >_L \emptyset$, applying auto-duality and OI, the likely dominance rule leads to obtain: $g > h$. Similarly, $[h \geq_P k] = S$ and $[k \geq_P h] = S - \{s_2\}$.Since $\{s_2\} >_L \emptyset$, we get $h > k$. Similarly, $[g \geq_P k] = S$ and $[k \geq_P g] = S$. Hence $g \sim k$.

Hence, assuming $x \sim_P y$, $y \sim_P z$ and $x >_P z$ leads to get: $g > h$, $h > k$, $g \sim k$, which is in contradiction with the transitivity of $>$. Hence, $\sim_P$ is transitive on X.

Proof of Proposition 9

Let $x, y, z \in X$ be such that $x >_P y >_P z$. Suppose first A,B,C disjoint. Let $D = \overline{A \cup B \cup C}$.

• 9.1, Suppose $A >_L B \cup C$. Define three acts f, g, h as follows:

	f	g	h
A	x	y	z
B	y	x	x
C	y	x	y

and $\forall s \notin A \cup B \cup C$, $f(s) = g(s) = h(s)$.

Note that $[f \geq_P g] = A \cup D$, $[g \geq_P f] = B \cup C \cup D$

$[g \geq_P h] = A \cup B \cup C \cup D$, $[h \geq_P g] = B \cup D$

$[h \geq_P f] = B \cup C \cup D$, $[f \geq_P h] = A \cup C \cup D$

$A >_L B \cup C \Rightarrow A \cup D >_L B \cup C \cup D$ (S2) $\Rightarrow f > g$ (likely dominance).

Since $A >_L B \cup C$, also $A >_L \emptyset$ (7.3) and $A \cup C >_L \emptyset$ (Proposition 6). Thus (S2, OI) imply $g > h$.

By transitivity : $f > h$. Thus, by likely dominance : $A \cup C \cup D >_L B \cup C \cup D$: and by S2, we get $A >_L B$

• 9.2, Suppose $A >_L B$. Define three acts f, g, h as follows: $\forall s \in D$, $f(s) = g(s) = h(s)$

	f	g	h
A	x	y	z
B	z	x	y
C	x	x	y

Note that $[f \geq_P g] = A \cup C \cup D$, $[g \geq_P f] = B \cup C \cup D$

$[g \geq_P h] = A \cup B \cup C \cup D$, $[h \geq_P g] = D$

$[h \geq_P f] = B \cup D$, $[f \geq_P h] = A \cup C \cup D$

$A >_L B \Rightarrow A \cup C \cup D >_L B \cup C \cup D$(S2) $\Rightarrow f > g$ (likely dominance)

Since $A >_L B$, $A >_L \emptyset$ (7.3) and $A \cup B \cup C >_L \emptyset$ (7.4). Thus (S2, OI) imply $g > h$..

By transitivity : $f > h$. Thus, by likely dominance : $A \cup C \cup D >_L B \cup D$, by S2, $A \cup C >_L B$.

Now we do not assume A, B, C, disjoint.

9.1 general case : $A >_L B \cup C \Rightarrow A \cap B^c \cap C^c >_L A^c \cap (B \cup C)$ (preadditivity).

By 9.1 disjoint case : $A \cap B^c \cap C^c >_L A^c \cap B$.

By 9.2, disjoint case $(A \cap B^c \cap C^c) \cup (A \cap B^c \cap C) >_L A^c \cap B$ i.e. $A \cap B^c >_L A^c \cap B$.

Thus, by preadditivity, $A >_L B$.

9.2 general case : From $A >_L B$, we get $A \cap B^c >_L A^c \cap B$. Thus, by 9.1, disjoint case, $A \cap B^c >_L A^c \cap B \cap C^c$; by 9.2, disjoint case, $(A^c \cap B^c \cap C) \cup (A \cap B^c) >_L A^c \cap B \cap C^c$.

Thus adding $A \cap B$ and $A^c \cap B \cap C$ on both sides, we get by S2 : $A \cup C >_L B$

Proof of Proposition 11

AND: Consider the following acts where $x >_P y >_P z$

	ABC	ABC^c	AB^cC	AB^cC^c
f	x	y	z	z
g	y	z	x	y
h	z	x	y	x

[f $>_P$ g] = A □ B and [g $>_P$ f] = A □ B^c: hence f > g
[g $>_P$ h] = A □ C, [h $>_P$ g] = A □ C^c: hence g > h
Since [f $>_P$ h] =A □ B □ C and [h $>_P$ f] = A □ (C^c □ B^c),by transitivity of > we get:
A □ B □ C $>_L$ A □ (C^c □ B^c)
OR: Consider the following acts:

	ABC	ABC^c	AB^cC	AB^cC^c	A^cBC	A^cBC^c
f	x	z	x	y	x	y
g	y	y	y	x	x	y
h	z	x	y	x	y	x

Since [f $>_P$ g] = A □ C $>_L$ A □ C^c = [g $>_P$ f], we have f > g.
Since [g $>_P$ h] = A □ B > A □ B^c = [h $>_P$ g], we have g > h.
Thus f > h: hence [f > h] $>_L$ [h > f].
Since [f $>_P$ h] = (A □ B) □ C and [h $>_P$ f] = (A □ B) □ C^c: (A □ B) □ C $>_L$ (A □ B) □ C^c.

Sketches of proofs for proposition 12

RW can be easily proved from Propositions 8 and 10 : if B □ C, A □ B $>_L$ A □ B^c writes
A □ B □ C $>_L$ (A □ B^c □ C) □ (A □ B^c □ C^c).
By Proposition 9, we get : A □ B □ C $>_L$ A □ B^c □ C^c
and then (A □ B □ C)□(A □ B^c □ C) $>_L$ A □ B^c □ C^c, i.e. A □ C $>_L$ A □ C^c
CM: Consider the following acts where x $>_P$ y $>_P$ z:

	A □ B □ C	A □ B □ C^c	A □ B^c □ C	A □ B^c □ C^c
f	x	y	y	z
g	y	z	x	y
h	z	x	y	x

From A □ B $>_L$ A □ B^c, we get f > g. From A □ C $>_L$ A □ C^c we get g > h.
By transitivity of >: f > h, i.e., A □ B □ C $>_L$ (A □ B □ C^c) □ (A □ B^c □ C^c).
From proposition 9, A □ B □ C $>_L$ A □ B □ C^c.
We can also obtain the CUT in the same way.

Proof of Proposition 13

• A □ B ≥□L A □ C □ (A □ B) □ A^c □ C^c ≥□ (A □ C) □ A^c □ B^c
□ B □ A^c □ C^c ≥□ C □ A^c □ B^c
Since A □ (B □ C) = Ø, B □ A^c □ C^c = B □ C^c and C □ A^c □ B^c = C □ B^c.
Thus: A □ B ≥□L A □ C □ B □ C^c ≥□ C □ B^c □ B ≥□L C
• A ≥□L B □ (A □ B) □ (A □ B^c) ≥□L (A □ B) □ (A^c □ B)
□ (A □ B^c) ≥□L (A^c □ B) (preadditivity)
□ (A^c □ B^c) □ (A □ B^c) ≥□L (A^c □ B^c) □ (A^c □ B) (preadditivity again)
□ B^c ≥□L A^c.
The other properties are very easy to check.

Proof of proposition 14

14.1: y >P x □ [fx >P fy] = Ø and [fy >P fx] = S □ [fx >P fy] >□ [fy >P fx] (due to def. A1 ii). Hence fy > fx. And x ≥P y □ [fy >P fx] = Ø □ [fx>P fy] ≥□ [fy >P fx] (def. A1 (iii)) □ fx ≥ fy

14.2: (f ≥ g)A □ fAh ≥ gAh, □h. Let us denote B = [fAh >P gAh] and C = [gAh >P fAh]. Note that, by definition, whatever h: B = [f >P g] □ A and C = [g>P f] □ A. Hence: B ≥□ C □ (f ≥ g)A.

14.3: Consider f, g s.t. f(s) >P g(s) if s □ A, f(s) = g(s) if s □ A: then, [f >P g] = A and [g >P f] = Ø.
□ Suppose A is null: □f, g, (f ∼ g)A, i.e.: [f >P g] □ A∼□ [g >P f] □ A, thus A ∼□ Ø.
□ Suppose A ∼□ Ø: □B, B □ A ≤ □ A∼□Ø. Hence: □f, g, [f >P g] □ A∼□ [g >P f] □ A ∼□ A∼□ Ø i.e., □f,g (f ∼ g)A that is, A is null.y

Proof of Proposition 15

• WS1: [f >P f]∼□[f >P f]: there is no act f s.t. f > f: the relation > is obviously irreflexive.
• > is transitive. Let f, g, h be three acts s.t. f > g, g > h, h ≥ f, i.e.,
[f >P g] >□ [g >P f], [g >P h]) >□ [h >P g] and [h >P f]) ≥□ [f >P h] (i)

Let us partition S in 13 subsets:
A = {s / f(s) = g(s) = h(s)}; B = {s / h(s) $>_P$ f(s) = g(s)}; C = {s / f(s) = g (s) $>_P$ h(s)};
D = {s / g(s) $>_P$ f(s) = h(s)}; E = {s / f(s) = h(s) $>_P$ g(s)}; F = {s / f(s) $>_P$ g(s) = h(s)};
G = {s / g(s) = h(s) $>_P$ f(s)}; H = {s / h(s) $>_P$ g(s) $>_P$ f(s)}; I = {s / g(s) $>_P$ h(s) $>_P$ f(s)};
J = {s / h(s) $>_P$ f(s) $>_P$ g(s)}; K = {s / g(s) $>_P$ f(s) $>_P$ h(s)}; L = {s l f(s) $>_P$ h(s) $>_P$ g(s)};
M = {s / f(s) $>_P$ g(s) $>_P$ h(s)}.
The system of inequations (i) can be rewritten as (where c is the possibility degree of C, etc.):
max(e, f, j, l, m) > max(d, g, h, i, k) (ii-1)
max(c, d, i, k, m) > max(b, e, h, j, l) (ii-2)
max(b, g, h, i, j) ≥ max(c, f, k, l, m) (ii-3)
It is inconsistent. Indeed, from ii-1 and ii-2, one must have max(f, c, m) greater than all of b, d, e, g, h, i, j, k, l, and this is in contradiction with ii-3.

WSP: [fAg $>_P$ f] = [g $>_P$ f] ∩ A^c and [gAf $>_P$ f] = [g $>_P$ f] ∩ A
[f $>_P$ fAg] = [f $>_P$ g] ∩ A^c and [f $>_P$ gAf] = [f $>_P$ g] ∩ A
By construction fAg ≥ f ⇔ [g $>_P$ f] ∩ A^c $\geq_\Pi$ [f $>_P$ g] ∩ A^c
gAf ≥ f ⇔ [g $>_P$ f] ∩ A $\geq_\Pi$ [f $>_P$ g] ∩ A
Since: $s_1 \geq_\pi s_2$ and $s_3 \geq_\pi s_4$ ⇒ {s_1 , s_2} $\geq_\Pi$ {s_3 , s_4} , this implies [g $>_P$ f] $\geq_\Pi$ [f $>_P$ g] ⇔ g ≥ f
Clearly, the unanimity property U also holds using a similar proof.

S3: fx ~ fy ⇔ x $\sim_P$ y ⇔ [fx $>_P$ fy] = [fy $>_P$ fx] = Ø thus:
∀A, [fx $>_P$ fy] ∩ A $\sim_\Pi$ [fy $\geq_P$ fx] ∩ A ⇔ ∀A, $(fx \sim fy)_A$ (i)
• fx > fy ⇔ x $>_P$ y ⇔ [fx $>_P$ fy] = S and [fy $>_P$ fx] = Ø thus:
∀A, [fx $>_P$ fy] ∩ A = A and [fy $>_P$ fx] ∩ A = Ø;
since A is not null: A $>_\Pi$ Ø and [fx $>_P$ fy] ∩ A $\sim_\Pi$ A $>_\Pi$ [fy $>_P$ fx] ∩ A.
Hence: $(fx > fy)_A$ (ii)
• By (i) and (ii): x $\geq_P$ y ⇒ $(fx \geq fy)_A$.
Moreover, by (ii) contraposed: $(fy \geq fx)_A$ ⇒ y $\geq_P$ x.

OI: The result is trivial due to the use of the likely dominance rule.

References

E.W. Adams (1975) The Logic of Conditionals. Dordrecht: D. Reidel.

K. Arrow (1951) Social Choice and Individual values. Wikey, New York.

K. Arrow, L. Hurwicz (1972) An optimality criterion for decision-making under ignorance. In Uncertainty and Expectations in Economics (C.F. Carter, J.L. Ford, Eds) Basil Blacwell & Mott Ltd, Oxford.

F. Bacchus and A. Grove (1996) Utility independence in a qualitative decision theory. Proc. 5rd Inter. Conf. on Principles of Knowledge Repres. and Reasoning (KR'96), Cambridge, Mass., 542-552.

S. Benferhat, D. Dubois, H. Prade (1997a) Nonmonotonic reasoning, conditional objects and possibility theory, Artificial Intelligence, 92, 259-276

S. Benferhat, D. Dubois, H. Prade (1997b) Possibilistic and standard probabilistic semantics of conditional knowledge. Proc. 14th National Conf. on Artificial Intelligence (AAAI-97), Providence RI, pp. 70-75. Revised version to appear in the J. Logic & Computation.

C. Boutilier (1994) Towards a logic for qualitative decision theory. Proc. 4rd Inter. Conf. on Principles of Knowledge Representation and Reasoning (KR'94), Bonn, Germany, May. 24-27, 75-86.

R.I Brafman, M. Tennenholtz (1996) On the foundations of qualitative utility theory. Proc. 13th National Conf. on Artificial Intelligence (AAAI-96), Portland, OR, pp. 1291-1296

R.I Brafman, M. Tennenholtz(1997) Modeling agents as qualitative decision makers. Artificial Intelligence, 94, 217-268.

R.I Brafman, M. Tennenholtz (2000) On the Axiomatization of Qualitative Decision Criteria, J. ACM, 47, 452-482.

S. Brass (1991) Deduction with supernormal defaults. Proc. 2d Int. Worshop on Nonmonotonic and Inductive logics. LNAI 659, Springer Verlag, Berlin, 153-174

G. Brewka (1989) Preferred subtheories: an extended logical framework for default reasoning. Proc. of the 11th Inter. Joint Conf. on Artificial Intelligence (IJCAI'89), Detroit, Aug. 20-25, 1043-1048.

C. Cayrol, V. Royer, C. Saurel (1992) Management of preferences in assumption-based reasoning. In Proc. 4th Int. Conf. on Information Processing, and Management of Uncertainty in Kwowledge-based Systems (IPMU'92), LNCS 682, Springer Verlag, Berlin, 13-22

M. Cohen and J.-Y. Jaffray (1980) Rational behavior under complete ignorance. Econometrica, 48, 1280-1299.

J. P. Delgrande (1994) A preference-based approach to default reasoning. Proc. 12th National Conf. on Artificial Intelligence (AAAI-94), 902-908.

J. Doyle, R. Thomason (1999) Background to qualitative decision theory. The AI Magazine, 20 (2), Summer 1999, 55-68

J. Doyle, M.P. Wellman (1991) Impediments to universal preference-based default theories. Artificial Intelligence, 49, 97-□28.

D. Dubois (1986) Belief structures, possibility theory and decomposable confidence measures on finite sets. Computers and Artificial Intelligence (Bratislava), 5(5), 403-416.

D. Dubois, H. Fargier, H. Prade (1997) Decision making under ordinal preferences and uncertainty. Proc. 13th Conf. on Uncertainty in Artificial Intelligence (D. Geiger, P.P. Shenoy, eds.), Providence, RI, Morgan & Kaufmann, San Francisco, CA, 157-164. Full paper Report IRIT/9803, Toulouse, France.

D. Dubois, H. Fargier, H. Prade (1998) Comparative uncertainty, belief functions and accepted beliefs. Proc. of the 14th Conf. on Uncertainty in Artificial Intelligence (G. Cooper, S. Moral, eds.), Madison, WI, Morgan & Kaufmann, San Francisco, CA, 113-120.

D. Dubois, H. Fargier, P. Perny H. Prade (2001) A characterization of generalized concordance rules in multicriteria decision-making, Proceedings of EUROFUSE Workshop on Preference Modelling and Applications, Granada, Spain , April 2001, p. 121-129 , 25 -27. To appear in Int. J. Intell. Syst.

D. Dubois, H. Fargier, P. Perny (2002) Qualitative models for decision under uncertainty□an axiomatic approach. In preparation.

D. Dubois, H. Prade (1988) Possibility Theory — An Approach to the Computerized Processing of Uncertainty. Plenum Press, New York

D. Dubois, H. Prade (1991) Epistemic entrechment and possibilistic logic. Artificial Intelligence, 50, 223-239.

D. Dubois, and H. Prade (1994). Conditional objects as nonmonotonic consequence relations. IEEE Trans. on Systems, Man and Cybernetics, 24, 1724-1740.

D. Dubois, H. Prade (1995a) Conditional objects, possibility theory and default rules. In Conditional: From Philosophy to Computer Science (G. Crocco, L. Farinas del Cerro, A. Herzig eds.), Oxford University Press, 311-346.

D. Dubois, H. Prade (1995b) Numerical representation of acceptance. Proc. of the 11th Conf. on Uncertainty in Articicial Intelligence, Montréal, August, 149-156.

D. Dubois, H. Prade (1995c) Possibility theory as a basis for qualitative decision theory. Proc. of the Inter. Joint Conf. on Artificial Intelligence (IJCAI'95), Montréal, August, 1924-1930.

D. Dubois, H. Prade (1998) Possibility theory: qualitative and quantitative aspects. P. Smets, editor, *Handbook on Defeasible Reasoning and Uncertainty Management Systems — Volume 1: Quantified Representation of Uncertainty and Imprecision*. Kluwer Academic Publ., Dordrecht, The Netherlands, 169-226

D. Dubois, H. Prade, R. Sabbadin (2001) Decision-theoretic foundations of possibilty theory. European Journal of Operational Research , 128, 459-478.

H. Fargier, P. Perny (1999) Qualitative models for decision under uncertainty without the commensurability assumption. Proc. of the 15th Conf. on Uncertainty in Artificial Intelligence (K. Laskey, H. Prade, eds.), Providence, RI, Morgan & Kaufmann, San Francisco, CA, 157-164.

P. Fishburn (1975) Axioms for lexicographic preferences. Review of Economical Studies, 42, 415-419

P. Fishburn (1986) The axioms of subjective probabilities. Statistical Science 1, 335-358.

N. Friedman, J. Halpern (1996) Plausibility measures and default reasoning. Proc of the 13th National Conf. on Artificial Intelligence (AAAI'96), Portland, 1297-1304. To appear in J. Assoc. for Comp. Mach.

J. Halpern (1997) Defining relative likelihood in partially-ordered preferential structures. J. AI Research, 7, 1-24

P. Gärdenfors (1988) Knowledge in Flux. MIT Press, Cambridge.

H. Geffner (1992) High probabilities, model preference and default arguments, Mind and Machines 2, 51-70.

S. Grant , A. Kajii, B. Polak (1997) Weakening the sure-thing principle: decomposable choice under uncertainty. Workshop on Decision Theory, Chantilly, France, June 12-14. A revised version, "Decomposable Choice under Uncertainty", is to appear in the J. Economic Theory.

A. Grove (1988) Two modellings for theory change. J. Philos. Logic, 17, 157-170.

J.-Y. Jaffray (1989) Linear utility theory for belief functions. Operations research Letters, 8, 107-112.

K. Kraus, D. Lehmann, M. Magidor (1990) Nonmonotonic reasoning, preferential models and cumulative logics. Articicial Intelligence, 44, 167-207.

J. Lang (1996) Conditional desires and utilities: an alternative logical approach to qualitative decision theory. Proc. of the 12th European Conf. on Artificial Intelligence (ECAI'96), Budapest, 318-322.

D. Lehmann (1996) Generalized qualitative probability: Savage revisited. Proc. of the 12th Conf. on Uncertainty in Artificial Intelligence, Portland, August, Morgan & Kaufman, San Mateo, CA, 381-388.

D. Lehmann, M. Magidor (1992) What does a conditional knowledge base entail? Artificial Intelligence, 55(1), 1-60.

D. Lewis (1973) Counterfactuals. Basil Blackwell, London.

A. Mas-Colell, H.F. Sonnenschein (1972) General possibility theorems for group decisions. Review of Economic Studies, 39, 185-192

H. Moulin (1988). Axioms of Cooperative Decision Making. Cambridge University Press, Cambridge, MA.

J. Pearl (1988). Probabilistic Reasoning in Intelligent Systems. San Mateo, CA: Morgan Kaufmann.

M. Roubens, P. Vincke (1985) Preference Modelling. Lecture Notes in Economics and Mathematical Systems, Vol. 250, Springer Verlag, Berlin.

R. Sabbadin, 2000, Empirical comparison of probabilistic and possibilistic Markov decision processes algorithms. Proc. 14th Europ. Conf. on Artificial Intelligence (ECAI'00), Berlin, Germany, 586-590.

L.J. Savage (1972) The Foundations of Statistics. Dover, New York.

D. Schmeidler (1989) Subjective probability and expected utility without additivity, Econometrica, 57, 571-587.

A.K. Sen (1986) Social choice theory. In K. Arrow, M.D. Intrilligator (Eds) Handbook of Mathematical Economics, Chap. 22, Elsevier, Amtsterdam, 1173-1181.

P. Snow (1994) The emergence of ordered belief from initial ignorance. Proc. 11th National Conf. on Artificial Intelligence (AAAI-94), Seattle, WA, pp. 281-286.

P. Snow (1999) Diverse confidence levels in a probabilistic semantics for conditional logics. Artificial Intelligence 113, 269-279.

S.W. Tan, J. Pearl. (1994) Qualitative decision theory. Proc. 11th National Conf. on Artificial Intelligence (AAAI-94), Seattle, WA, pp. 70-75.

R. Thomason (2000), Desires and defaults□a framework for planning with inferred goals. Proc. of the Inter. Conf. on Principles of Knowledge Representation and Reasoning (KR'00), Breckenridge, Col.., Morgan & Kaufmann, San Francisco, 702-713.

J. Von Neumann and O. Morgenstern (1944): *Theory of Games and Economic Behaviour* (Princeton Univ. Press, Princeton, NJ).

A. Wald (1950), "Statistical Decision Functions." Wiley, New York.

J.A. Weymark (1984), "Arrow's theorem with social quasi-orderings, *Public Choice*, 42, 235-246.

L.A. Zadeh (1978) Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1, 3-28.