A Survey of Qualitative Decision Rules Under Uncertainty*

Didier Dubois, Hélène Fargier, Henri Prade, Régis Sabbadin

October 16, 2009

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 [67] and his school. 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 (see chapters by Jaffray and Sabbadin in this volume).

However, in parallel to these developments, Artificial Intelligence has witnessed the emergence of a new decision paradigm called *qualitative decision theory* [14], where the rationale for choosing among decisions no longer relies on probability theory nor on numerical utility functions. 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. Formulating decision problems in a symbolic way may be more compatible with a declarative expression of uncertainty and preferences in the setting of some logic-based language [6]; [75].

On the other hand, the emergence of new information technologies like information systems or autonomous robots has generated many new decision problems involving intelligent agents [8]. 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 sophisticated 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 complex decisions that need to be carefully

*This paper is chapter 11 in D. Bouyssou, D. Dubois, M. Pirlot, H. Prade, eds, Decision-Making Process, ISTE, London, UK & Wiley, Hoboken, N.J. USA, 2009

analyzed (see [7]). 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 [6]. This is typical of electronic commerce, or recommender systems (that provide advice or suggestions) for instance. In many cases, it sounds more satisfactory to implement a choice method that is fast, and based on rough information about the user preferences and knowledge. Moreover the expected utility criterion makes full sense for repeated decisions whose successive results accumulate (for instance money in gambling decisions). In contrast, 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 [33], and some notions and results in this chapter can be expressed in the setting of multiple-criteria-decision-making.

In the case of autonomous robots, conditional plans are often used to monitor the robot behaviour, and the environment of the robots is sometimes only partially observable. The theory of partially observable Markov decision processes leads to highly complex methods, because handling infinite state spaces. A qualitative, finitist, description of the goals of the robot, and of its knowledge of the environment might lead to more tractable methods ([66], 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 environments, where cautious policies should be followed. Anyway, dynamic programming techniques are also compatible with qualitative settings such as possibility theory ([42] [30] and the chapter by R. Sabbadin in this book).

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 ([58]; [4] for instance). Others use sets of integers and the like to describe rough probabilities or utilities [74, 49]. Lehmann [60] 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. 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, nor for the representation of utility.

This chapter proposes an overview of qualitative decision theory, focused on discussing the rationale of the various possible decision rules and their properties. It is stressed that two kinds of approach exist, according to whether degrees of uncertainty and degrees of utility are commensurate (i.e. belong to a unique measurement scale) or not. A Savage-like axiomatics is proposed for each of these two approaches. In this setting, acts are modellled as functions from the set of states to the set of consequences, and decision rules stem from properties the preference relation over acts is requested to satisfy. Generally it turns out that the natural uncertainty theory at work in qualitative frameworks is possibility theory rather than probability theory. However qualitative decision rules are often either little decisive (due to incomparability incurred, or ties) or too adventurous (because concentrating on most plausible states of nature and neglecting other ones). Some rules lack discrimination simply because the measurement scale is too coarse. Recent results show how to refine some criteria, using a specific form of standard expected utility encoding lexicographic refinements.

The chapter is organized as follows. First, a survey of qualitative decision rules is proposed, first those that assume commensurability between utility and uncertainty (Section 2). Section 3 then motivates a decision rule that does not presuppose it, and shows its potential limitations. Section 4 presents Savage-like representation results for such qualitative decision rules, on the basis of axioms expressing properties of the preference relation between acts. Finally, Section 5, acknowledging the lack of discrimination power of qualitative decision rules, explains how to refine them so as to tackle this weakness.

2 Quantitative versus qualitative decision rules

A decision problem can be cast in the following framework: consider 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 when modeling the lack of knowledge of the decision maker about the state of affairs. 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 [64]) or it is a subjective probability [67] supplied by the agent via suitable elicitation methods. Then the most usual decision rule is based on the expected utility criterion:

$$EU_{p,u}(f) = \sum_{s \in S} p(s)u(f(s)). \tag{1}$$

An act f is strictly preferred to act g if and only if $EU_{p,u}(f) > EU_{p,u}(g)$. The expected utility criterion is by far the most commonly used one. This criterion makes sense especially for repeated decisions whose results accumulate. It also clearly presupposes subjective notions like belief and preference to be precisely quantified. In particular, in the expected utility model, the way in which the preference on consequences is numerically encoded will affect the induced preference relation on acts. The model exploits some extra information not contained solely in preference relations on X, namely, the absolute order of magnitude of utility grades. Moreover the same numerical scale is used for utilities and degrees of probability. This is based on the notion of *certainty equivalent*, i.e. the idea of that a lottery (involving uncertainty) can be compared to a sure gain or a sure loss (involving utility only) in terms of preference.

Another proposal is the maximin criterion often credited to Wald (1950) [77]. It applies when no information about the current state is available, and it ranks acts according to its worst consequence:

$$W_u^-(f) = \min_{s \in S} u(f(s)). \tag{2}$$

It is a (very) pessimistic criterion. An optimistic counterpart $W_u^+(f)$ to W_u^- is obtained by turning minimum into maximum in (2). Clearly, the maximin and maximax criteria do not need numerical utility values. Only a total ordering on consequences is needed. No knowledge about the state of the world is necessary. However, these criterion have the major defect of being respectively extremely pessimistic and overoptimistic. In practice, it is never used for this reason. Hurwicz has proposed to use a weighted average of $W_u^-(f)$ and its optimistic counterpart, where the weight bearing on $W_u^-(f)$ is viewed as a degree of pessimism of the decision maker. Other decision rules have been proposed, especially some that generalize both $EU_{p,u}(f)$ and W_u^-, see [55] and [68], among others (see the chapter by Chateauneuf and Cohen in this book). However, all these extensions again require the quantification of preferences and/or uncertainty.

Qualitative extensions of the maximin criterion, that account for some knowledge about the state of affairs, nevertheless exist. Boutilier [6] is inspired by preferential inference of nonmonotonic reasoning whereby a proposition A entails another one B by default if B is true in the most normal situations where A is true. He assumes that states of nature are ordered in terms of their relative plausibility using a weak order relation $\trianglerighteq$ on S. He proposes to choose decisions on the basis of the most plausible states of nature in accordance with the available information, neglecting other states. If the available information is that $s \in A$, a subset of states, and if A^* is the set of maximal elements in A according to the plausibility ordering $\trianglerighteq$, then the criterion is defined by

$$W^-_{\trianglerighteq,u}(f) = \min_{s \in A^\star} u(f(s)). \tag{3}$$

Another refinement of Wald criterion is the possibilistic qualitative criterion. It is based on a utility function u on X and a possibility distribution π on S [35] [37], both mapping on the same totally ordered value scale V, with top 1 and bottom 0. The ordinal value $\pi(s)$ represents the relative plausibility of state $s \in S$. A pessimistic criterion $W^-_{\pi,u}(f)$ is proposed of the form [36]:

$$W^-_{\pi,u}(f) = \min_{s \in S} \max(\nu(\pi(s)), u(f(s))) \tag{4}$$

Here, V is equipped with its involutive order-reversing map ν; in particular $\nu(1) = 0$, $\nu(0) = 1$. So, $\nu(\pi(s))$ represents the degree of potential surprise [71] caused if that the state of the world were s. In particular, $\nu(\pi(s)) = 1$ for impossible states. The value of $W^-_{\pi,u}(f)$ is small as soon as there exists a highly plausible state ($\nu(\pi(s)) = 0$) with low utility value. This criterion is actually a prioritized extension of the Wald maximin criterion $W_u^-(f)$. The latter is recovered if $\pi(s) = 1$ for all $s \in S$. The decisions are again made according to the merits of acts in their worst consequences, now restricted to the most plausible states, like (3). But the set of most plausible states ($S^* = \{s, \pi(s) \geq \nu(W^-_{\pi,u}(f))\}$)) now depends on the act itself. It is defined by the compromise between belief and utility expressed in the min-max expression. However, contrary to the other qualitative criteria, the possibilistic qualitative criterion presupposes that degrees of utility $u(f(s))$ and possibility $\pi(s)$ share the same scale and can be compared.

The optimistic counterpart to this criterion [36] is:

$$W^+_{\pi,u}(f) = \max_{s \in S} \min(\pi(s), u(f(s))). \tag{5}$$

This expression is due to Zadeh [83] but its interpretation as a decison criterion has been first proposed by Yager [82]. The pessimistic criterion was first proposed as such by Whalen [81]. Similar ideas actually appeared already in the works of Shackle [70], a forerunner of possibility theory.

These two criteria have been used for a long time in fuzzy information processing for the purpose of triggering fuzzy rules in expert systems [10] and flexible querying of an incomplete information database [40]. They were used in scheduling under flexible constraints and uncertain task durations, when minimizing the risk of delayed jobs [23].

These optimistic and pessimistic possibilistic criteria are actually particular cases of a more general criterion based on the Sugeno integral [73], a qualitative counterpart to Choquet integral used in extensions of expected utility approaches; see [53] and the chapter by J.-L. Marichal in this book, for details and references. One expression of this criterion can be written as follows:

$$S_{\gamma,u}(f) = \max_{\lambda \in V} \min(\lambda, \gamma(F_\lambda)) \tag{6}$$

where $F_\lambda = \{s \in S, u(f(s)) \geq \lambda\}$ is a set of preferred states for act f, $\gamma(A)$ is the degree of likelihood of event A. The set-function γ reflects the decision-maker attitude in front of uncertainty. This expression achieves a trade-off between the degrees of likelihood of preferred events and the figures of merit of the worst consequences when they occur. If the set of states is rearranged in decreasing order of merit via f in such a way that $u(f(s_1)) \geq \cdots \geq u(f(s_n))$, then denoting $A_i = \{s_1 \ldots, s_i\}$, it turns out that $S_{\gamma,u}(f)$ is the median of the set

$$\{u(f(s_1)), \ldots, u(f(s_n))\} \cup \{\gamma(A_1), \ldots, \gamma(A_{n-1})\}.$$

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

- if A is likely enough, $\gamma(A) \geq u(x)$ and $S_{\gamma,u}(f) = u(x)$, then the decision-maker thinks she can get x;
- if A unlikely to a sufficient degree, $\gamma(A) \leq u(y)$ and $S_{\gamma,u}(f) = u(y)$, then the decision-maker thinks she can get only y;
- else ($u(x) > \gamma(A) > u(y)$), the value of act f exactly reflects the likelihood of the successful event ($S_{\gamma,u}(f) = \gamma(A)$).

In this approach, the attitude of the decision-maker in front of uncertainty is encoded by the choice of a likelihood function γ. If optimistic, she selects a possibility measure $\gamma = \Pi$, i.e. a function $\Pi(A) = \max_{s \in A} \pi(s)$, where $\pi(s) \in V$ is the grade of plausibility of event A. $\Pi(A) = 1$ as soon as the decision-maker thinks the opposite event A^c of A has no certainty of occurring. She then bets on the occurrence of A, hence $S_{\gamma,u}(f) = u(x)$ (more generally $S_{\Pi,u} = W^{+}_{\pi,u}$). If pessimistic, she selects a necessity mesure $\gamma = N$, adjoint to a possibility measure: $N(A) = 1 - \Pi(A^c)$, which evaluates the grade of certainty of A. $N(A) = 0$ as soon as the decision-maker is uncertain about A. She then bets on the non-occurrence of A. Then $S_{\gamma,u}(f) = u(y)$ (more generally $S_{N,u} = W^{-}_{\pi,u}$).

3 An ordinal decision rule without commensurateness

Several among the above decision rules presuppose that utility functions and uncertainty functions share the same range, so that it makes sense to write $\min(\pi(s)), u(f(s)))$, for instance. In contrast, one may look for a natural decision rule that computes a preference relation on acts from a purely symbolic perspective, no longer assuming that utility and partial belief are commensurate, that is, share the same totally ordered scale [25]. The decision-maker then only supplies a likelihood relation $\succeq_L$ between events and a preference relation $\succeq_P$ on consequences. The strict part $\succ_L$ of the likelihood relation $\succeq_L$ is defined by $A \succ_L B$ if and only if $A \succeq_L B$, but not $B \succeq_L A$, and the indifference relation $\sim_L$ induced by $\succeq_L$ is defined in the usual way: $A \sim_L B$ if and only if $A \succeq_L B$ and $B \succeq_L A$. $A \succeq_L B$ means that event A is at least as likely as B.

In the most realistic model, the strict part $\succ_L$ of a likelihood relation on the set of events is irreflexive, transitive, and non-trivial ($S \succ_L \emptyset$). Moreover, it should be faithful to deductive inference, which means for the strict part of the likelihood relation:

$$\forall A, B, C, D, A \succ_L B \text{ implies } A \cup C \succ_L B \cap D.$$

Finally, if $A \subseteq B$ then it should hold that $B \succeq_L A$ (inclusion-monotony). The inclusion-monotony property states that if A implies B, then A cannot be more likely than B. Let $s_i \trianglerighteq s_j$ denote the plausibility relation between states induced by $\succeq_L$ on elements of S.

The preference relation on the set of consequences X is supposed to be a weak order (a complete preordering, e.g. see Chapter 2 of this book). Namely, $\succeq_P$ is a reflexive and transitive relation, and completeness means $x \succeq_P y$ or $y \succeq_P x$. So, $x \succeq_P y$ means that consequence x is not worse than y. The induced strict preference relation is derived as usual: $x \succ_P y$ if and only if $x \succeq_P y$ and not $y \succeq_P x$. It is assumed that X has at least two elements x and y s.t. $x \succ_P y$. The assumptions pertaining to $\succeq_P$ are natural in the scope of numerical representations of utility, however we do not require that the 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 $(\succ_L, \succeq_P)$ 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 $f(s) \succ_P g(s)$. Let $[f \succ_P g]$ be an event made of all states where f outperforms g, that is $[f \succ_P g] = \{s \in S, f(s) \succ_P g(s)\}$. Accordingly, we define the preference between acts ($\succeq$)), the corresponding indifference ($\sim$) and strict preference ($\succ$) relations as follows:

- $f \succ g$ if and only if $[f \succ_P g] \succ_L [g \succ_P f]$;
- $f \succeq g$ if and only if $\neg(g \succ f)$, i.e. if and only if $[f \succ_P g] \succeq_L [g \succ_P f]$;
- $f \sim g$ if and only if $f \succeq g$ and $g \succeq f$.

This is the Likely Dominance Rule [25]. It 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. The properties of the relations $\succeq, \sim$ and $\succ$ on X^S will depend on the properties of $\succeq_L$ with respect to Boolean connectives. An interesting remark is that if $\succeq_L$ is a comparative probability ordering then the strict preference relation $\succ$ in X^S is not necessarily transitive, nor acyclic.

Exemple 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 $x_1 \succ_P x_2 \succ_P x_3$, and the comparative probability ordering is generated by a uniform probability on S. Suppose three acts f, g, h such that

- $f(s_1) = x_1 \succ_P f(s_2) = x_2 \succ_P f(s_3) = x_3$,
- $g(s_3) = x_1 \succ_P g(s_1) = x_2 \succ_P g(s_2) = x_3$,
- $h(s_2) = x_1 \succ_P h(s_3) = x_2 \succ_P h(s_1) = x_3$.

Then $[f \succ_P g] = \{s_1, s_2\}$; $[g \succ_P f] = \{s_3\}$; $[g \succ_P h] = \{s_1, s_3\}$; $[h \succ_P g] = \{s_2\}$; $[f \succ_P h] = [h \succ_P f] = \{s_2, s_3\}$. The likely dominance rule yields $f \succ g, g \succ h, h \succ f$ (while, of course $f \succ f$ does not hold). 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. Moreover, the undesirable cycle remains as long as probabilities $p(s_1) > p(s_2) > p(s_3)$ of states remain close to each other, so that $p(s_i) + p(s_j) > p(s_k), \forall i, j, k$ distinct. 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 $f \succ g \succ h, g \succ h \succ f, h \succ f \succ g$ can be obtained by the expected utility criterion, by suitably quantifying the utility values of states without changing their preference ranking.

This situation can be viewed as a counterpart to the Condorcet paradox in social choice, here in the setting of DMU. Indeed the problem of ranking acts can be cast in the setting of a voting problem (See [63], for an introduction to voting methods, and the chapter by Bouyssou, Marchant and Perny in this book). Let $\mathcal{V}$ be a set of voters, C be a set of candidates and let $\succeq_i$ be a relation on C that represents the preference of voter i on the set of candidates. $\succeq_i$ is a weak order, by assumption. The decision method consists in constructing a relation R on C that aggregates the relations $\{\succeq_i, i \in V\}$ as follows. Let $\mathcal{V}(c_1, c_2) = \{i \in \mathcal{V}, c_1 \succ_i c_2\}$ be the set of voters who find c_1 more valuable than c_2, and $|\mathcal{V}(c_1, c_2)|$ the cardinality of that set. Then the social preference relation R on C is defined as follows by Condorcet: $c_1 R c_2$ if and only if $|\mathcal{V}(c_1, c_2)| > |\mathcal{V}(c_2, c_1)|$. 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 [3] proved that the transitivity of R is impossible under natural requirements on the voting procedure such as independence of irrelevant alternatives, unanimity, and non-dictatorship (i.e., there should be no voter i enforcing her preference relation: $R \neq \succeq_i, \forall i \in \mathcal{V}$).

Condorcet procedure is thus a special case of the likely dominance rule based on a uniform probability distribution, letting $\mathcal{V} = S, C = X^S$, and considering for each $s \in S$ the relation R on acts such that $\forall f, g \in X^S : f \succeq_s g \iff f(s) \succ_P g(s)$. Computing the probability $Prob([f \succ_P g])$ is a weighted version of $|V(c_1, c_2)|$ with $\mathcal{V} = S, c_1 = f, c_2 = g$, which explains the intransitivity phenomenon. Such weighted extensions of Condorcet procedure are commonly found in multicriteria decision-making [76]. 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 even in its weighted version.

Assume now that a decision maker supplies a weak order of states $\trianglerighteq$ and a weak order of consequences $\succeq_P$ on X. Let $\succeq_\Pi$ be the induced possibilistic ordering of events ([61] ; [16]). Namely, denote $\max(A)$ any (most plausible) state $s \in A$ such that $s \trianglerighteq s', \forall s \in A$. Then, define $A \succeq_\Pi B$ if and only if $\max(A) \trianglerighteq \max(B)$. The preference on acts in accordance with the likely dominance rule, for any two acts f and g, is : $f \succ g$ if and only if $[f \succ_P g] \succ_\Pi [g \succ_P f]$; $f \succeq g$ if and only if$\neg(g \succ f)$. Then, the undesirable intransitivity of the strict preference vanishes.

Example 1 (continued) Consider again the 3-state + 3-consequence example. 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 $s_1 \triangleright s_2 \triangleright s_3$ then

- $[f \succ_P g] = \{s_1, s_2\} \succ_\Pi [g \succ_P f] = \{s_3\}$;
- $[g \succ_P h] = \{s_1, s_3\} \succ_\Pi [h \succ_P g] = \{s_2\}$;
- $[f \succ_P h] = \{s_1\} \succ_\Pi [h \succ_P f] = \{s_2, s_3\}$.

So $f \succ g \succ h$ follows. It contrasts with the cycles obtained with a probabilistic approach. However the indifference relation between acts is generally not transitive.

Let us study the likely dominance rule induced by a single possibility distribution (and the possibilistic likelihood relation it induces).

1. If the decision maker is ignorant about the state of the world, all states are equipossible, and all events as well but for $\emptyset$. So, if f and g are such that $[g \succ_P f] \neq \emptyset$ and $[f \succ_P g] \neq \emptyset$, then neither $f \succ g$ nor $g \succ f$ holds, following the likely dominance rule. The case when $[f \succ_P g] \neq \emptyset$ and $[g \succ_P f] = \emptyset$ corresponds to when f Pareto-dominates g, i.e. $f \succeq_P g$ and $\exists s \in S, f(s) \succ_P g(s)$. The preference relation induced on acts by the likely dominance rule then reduces to Pareto-dominance. This method, although natural, is not at all decisive (it corresponds to a unanimity rule in voting theories).

2. Conversely, if there is a total ordering $s_1 \triangleright s_2 \triangleright \cdots \triangleright s_n$ of S, then for any A, B such that $A \cap B = \emptyset$, it holds that $A >_\Pi B$ or $B >_\Pi A$, then $\forall f \neq g$, either $f \succ g$ or $g \succ f$. Moreover this is a lexicographic ranking:

   $$f \succ g \text{ if and only if } \exists k \text{ s.t. } f(s_k) \succ_P g(s_k) \text{ and } f(s_i) \sim_P g(s_i), \forall i < k.$$

   It corresponds to the following 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, do the same test in the second most normal state, and so on recursively. This comes down to a lexicographic ranking of vectors $(f(s_1), \ldots, f(s_n))$ and $(g(s_1), \ldots, g(s_n))$. It is a form of dictatorship by most plausible states, in voting theory terms. It also coincides with Boutilier's criterion, except that ties can be broken by less normal states.

3. More generally any weak order splits S into a well-ordered partition $E_1 \cup E_2 \cup \cdots \cup E_n = S$, $E_i \cap E_j = \emptyset (\forall i \neq j)$ such that states in each E_i are equally plausible and any state in E_i is more plausible than all states in $E_j, \forall j > i$. Then, the ordering of events is defined as follows:

   - $f \succ g$ if and only if $\exists k \geq 1$ such that: $\forall s \in E_1 \cup E_2 \cup \cdots \cup E_{k-1}, f(s) \sim_P g(s)$, and $\forall s \in E_k, f(s) \succeq_P g(s)$ and $\exists s \in E_k, f(s) \succ_P g(s)$.

- $f \sim g$ if and only if either $\forall s \in S, f(s) \sim_P g(s)$, or $\exists k \geq 1$ such that: $\forall s \in E_1 \cup E_2 \cup \cdots \cup E_{k-1}, f(s) \sim_P g(s)$, and $\exists s \neq s' \in E_k, f(s) \succ_P g(s)$ and $g(s') \succ_P f(s')$.

This decision criterion is a blending of lexicographic priority and unanimity among states. Informally, the decision maker proceeds as follows: f and g are compared on the set of most normal states E_1: if f Pareto-dominates g in E_1, then f is preferred to g; if there is a disagreement in E_1 about the relative performance of f and g then f and g are not comparable. 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 E_2, etc. In a nutshell, it is a prioritized Pareto-dominance relation. Preferred acts are selected by restricting to the most plausible states of the world, and a unanimity rule is used on these maximally plausible states. Ties are broken by lower level oligarchies. So this procedure is similar to Boutilier's decision rule in that it focuses on the most plausible states, but Pareto-dominance is required instead of the maximin rule on them, and ties can be broken by subsets of lower plausibility. This decision rule is cognitively appealing, but it has a limited expressive and decisive power.

One may also apply the maximin rule in a prioritized way: the maximin decision rule can be substituted to unanimity within the likely dominance rule inside the oligarchies of states. It is also easy to imagine a counterpart to the likely dominance rule where expected utility applies inside the oligarchies of states [59]. However reasonable these refined decision rules may look, they need to be formally justified.

4 Axiomatics of qualitative decision theory

A natural question is then whether it is possible to found rational decision-making in a purely qualitative setting, under an act-driven framework à la Savage. The idea of the approach is to extract the decision maker's likelihood relation and the decision maker's preference on consequences from the decision maker's preference pattern on acts, as only the latter is observable from human behavior. Enforcing "rationality" conditions on the way the decision maker should rank acts then determines the kind of uncertainty theory implicitly "used" by the decision maker for representing the available knowledge on states. It also prescribes a decision rule. Moreover, this framework is operationally testable, since choices made by individuals can be observed, and the uncertainty theory at work is determined by these choices. However, it is right away assumed that states of nature are imperfectly perceived : it comes down to considering a finite set S representing states in a granular way, each element in S clustering indiscernible states of nature according to the decision-maker's language and perception.

As seen in sections 2 and 3, two research lines can be followed in agreement with this definition: the relational approach and the absolute approach. Following the relational approach [21, 19], 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. The advantage is that it is faithful to the kind of elementary information users can directly provide. The other approach, which can be dubbed the absolute approach [38, 39], presupposes the existence of a totally ordered scale (typically a finite one) for grading both likelihood and utility. Both approaches lead to an act-driven axiomatization of the qualitative variant of possibility theory [83, 37].

4.1 Savage theory: a refresher

The Savage framework (already described in chapter 9 of this volume), is adapted to our purpose of devising a purely ordinal approach because its starting point is indeed based on relations even if its representation is eventually made on an interval scale. Suppose a decision maker supplies a preference relation $\succeq$ over acts $f : S \to X$. X^S usually denotes the set of all such mappings. In Savage's approach, any mapping in the set X^S is considered to be a possible act (even if it is an imaginary one rather than a feasible one). The first requirement stated by Savage is:

Axiom P1: $(X^S, \succeq)$ is a weak order.

This axiom is unavoidable in the scope of expected utility theory: If acts are ranked according to expected utility, then the preference over acts will be transitive, reflexive, and complete ($f \succeq g$ or $g \succeq f$ for any f, g). What this axiom also implies, if X and S are finite, is that there exists a totally ordered scale, say V, that can serve to evaluate the worth of acts. Indeed the indifference relation ($f \sim g$ if and only if $f \succeq g$ and $g \succeq f$) is an equivalence relation, and the set of equivalence classes, denoted $X^S/\sim$ is totally ordered via the strict preference $\succ$. If $[f]$ and $[g]$ denote the equivalence classes of f and g, $[f] \succ [g]$ holds if and only if $f \succ g$ holds for any pair of representatives of each class. So it is possible to rate acts on $V = X^S/\sim$ and $[f]$ is interpreted as the qualitative utility level of f.

An event is modeled by a subset of states, understood as a disjunction thereof. The set of acts is closed under the following combination involving acts and events. Let $A \subseteq S$ be an event, f and g two acts, and denote by fAg the act such that:

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

For instance, f may mean "bypass the city", g means "cross the city", and A represents the presence of a traffic jam in the city. 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. Act 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 act fAh and act gAh does not depend on the choice of act h:

Axiom P2 : $\forall A, f, g, h, h', fAh \succeq gAh$ if and only if $fAh' \succeq gAh'$, 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.* [54] 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, one version they called the weak sure-thing principle being as follows:

Axiom WSTP : $fAg \succeq g$ and $gAf \succeq g$ implies $f \succeq g$.

The above property really means that the weak preference of act f over act g does not depend on whether A occurs or not. It is obvious that **WSTP** is implied by **P1** and **P2**, since from $fAg \succeq g = gAg$ and **P2** we derive $f = fAf \succeq gAf$ and using transitivity of $\succeq$ due to **P1**, $f \succeq g$ follows.

The sure-thing principle enables two notions to be simply defined, namely conditional preference and null events. An act f is said to be weakly preferred to another act g, *conditioned* on event A if and only if $\forall h, fAh \succeq gAh$. This is denoted by $(f \succeq g)_A$. Conditional preference $(f \succeq g)_A$ 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 $f \succeq g$ is short for $(f \succeq g)_S$. Moreover $(f \succeq g)_\emptyset$ always holds, for any f and g, since it is equivalent to the reflexivity of $\succeq$ (i.e. $h \succeq h$). Clearly, the sure-thing principle enables $(f \succeq g)_A$ to hold as soon as $fAh \succeq gAh$ for *some* act h.

An event A is said to be *null* if and only if $\forall f, \forall g, (f \succeq g)_A$ holds. Any non-empty set of states A on which no act makes a difference behaves then like the empty set in the perspective of choosing a best decision.

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

Note that in the absence of **P2**, axiom **U** implies **WSTP** but not the converse. The unanimity postulate has been formulated by Lehmann [59].

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

$\forall x, y \in X, x \succeq_P y$ if and only if $f_x \succeq f_y$.

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 P3 : $\forall A \subseteq S, A$ not null, $(f_x \succeq f_y)_A$ if and only if $x \succeq_P y$.

Clearly, Pareto-dominance should imply weak preference for acts. And indeed under **P1, P2**, and **P3**, $f \succeq_P g$ (that is, $\forall s \in S, f(s) \succeq_P g(s)$) implies $f \succeq g$.

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_xAf_y, which due to **P3** can be denoted xAy, where $x \in X, y \in X$, and $x \succ_P y$. Clearly for fixed $x \succ_P y$, the set of binary acts $\{x,y\}^S$ is isomorphic to the set of events 2^S. However the restriction of $(X^S, \succeq)$ to $\{x,y\}^S$ may be inconsistent with the restriction to $\{x',y'\}^S$ for other choices of consequences $x' \succ_P y'$. A relative likelihood $\succeq_L$ among events can anyway however be recovered, as suggested by Lehmann :[59] :

$\forall A, B \subseteq S, A \succeq_L B$ if and only if $xAy \succeq xBy, \forall x, y \in X$ s.t. $x \succ_P y$.

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

Axiom P4 : $\forall x, y, x', y' \in X$ s.t. $x \succ_P y, x' \succ_P y'$, it holds that $xAy \succeq xBy$ if and only if $x'Ay' \succeq x'By'$.

Under this property, the choice of $x \in X, y \in X$, with $x \succ_P y$ does not affect the ordering between events in terms of binary acts, namely: $A \succeq_L B$ is short for $\exists x \succ_P y, xAy \succeq xBy$.

Lastly, Savage assumed that the ordering $\succ$ is not trivial:

Axiom P5: X contains at least two elements x, y such that $f_x \succ f_y$ (or $x \succ_P y$).

Under **P1-P5**, the likelihood relation on events is a comparative probability ordering (see Fishburn[47]), i.e. satisfies the preadditivity property

If $A \cap (B \cup C) = \emptyset$, then $B \succeq_L C$ if and only if $A \cup B \succeq_L A \cup C$.

Such relations are induced by probability measures, but the converse is not true[56].

Savage introduces yet another postulate that enables him to derive the existence (and uniqueness) of a numerical probability measure on S that can represent the likelihood relation $\succeq_L$. This axiom reads:

Axiom P6: For any f, g with $f \succ g$ in X^S and any $x \in X$, there is a partition $\{E_1, \ldots, E_n\}$ of S such that $\forall i = 1, \ldots n, xE_if \succ g$ and $f \succ xE_ig$.

Under the postulates **P1-P6**, not only can $\succeq_L$ be represented by a unique numerical probability function but $(X^S, \succeq)$ can be represented by the expected utility of acts:

$$u(f) = \int_{s \in S} u(f(s))\, dP(s),$$

where the numerical utility function u represents the relation $\succeq_P$ on X uniquely, up to an affine transformation. According to postulate **P6**, the probability of subset E_i can be made arbitrary small, thus not altering the relation $f \succ g$ when x is very bad (so that $xE_if \succ g$) or very good (so that $f \succ xE_ig$). Clearly, such a postulate makes sense only if the state space S is infinite, which goes against our assumptions. In contrast, we assume that both S and X are finite in this chapter, and **P6** is trivially violated in such a finite setting. There is to our knowledge no joint representation of subjective probability and expected utility that would assume a purely finite setting for both states and consequences.

4.2 The relational approach to decision theory

The relational approach introduced in [25], and further developed in [21, 19], tries to lay bare the formal consequences of adopting 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 theory. To this end, an axiom of ordinal invariance, originally due to [46] in another context, is then added [44]. This axiom says that what matters for determining the preference between two acts is the relative merit of consequences of acts for each state, not the figures of merit of these consequences, nor the relative positions of these acts relative to other acts. More rigorously, two pairs of acts (f, g) and (f', g') such that

$$\begin{aligned} \forall s \in S, f(s) \succeq_P g(s) &\iff f'(s) \succeq_P g'(s) \\ \text{and } g(s) \succeq_P f(s) &\iff g'(s) \succeq_P f'(s) \end{aligned}$$

are called statewise order-equivalent[1]. This is denoted $(f, g) \equiv (f', g')$. It means that, in each state, consequences of f, g, and of f', g', are rank-ordered likewise. The Ordinal Invariance axiom reads:

Axiom OI: $\forall f, f'g, g' \in X^S$, si $(f, g) \equiv (f', g')$ then ($f \succeq g$ if and only if $f' \succeq g'$).

It expresses the purely ordinal nature of the decision criterion. It is easy to check that the likely dominance rule obeys axiom **OI**. This is obvious noticing that if $(f, g) \equiv (f', g')$ then by definition, $[f \succ_P g] = \{s, f(s) \succ_P g(s)\} = [f' \succ_P g']$. More specifically, under **OI**, 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, and **OI** implies the validity of Savage **P2**, and **P4** axioms. Adopting axiom **OI** and sticking to a transitive weak preference on acts lead to problems exemplified in the previous section by the probabilistic variant of the likely dominance rule. Indeed the following result was proved in [21]:

Theorem 1 *If $(X^S, \succeq)$ is a weak order on acts satisfying axiom* **OI**, *and S and X have at least three elements, let $\succeq_L$ be the likelihood relation induced by axiom* **P4** *(implied by* **OI***). Then there is a permutation of the non-null elements of S, such that $\{s_1\} \succ_L \{s_2\} \succ_L \ \succ_L \{s_{n-1}\} \succ_L \{s_n\} \succ_L \emptyset$ and $\forall i = 1, \ldots, n-2, \{s_i\} \succ_v \{s_{i+1}, \ldots, s_n\}$.*

In the general case where X has more than two elements, the Ordinal Invariance axiom forbids a Savagean decision maker to believe that there are two equally likely states of the world, each of which being more likely than a third state. This is clearly not acceptable in practice. Nevertheless, if X only has two consequences of distinct values, then such a trivialization is avoided.

If we analyze the reason why this phenomenon occurs, it is found that axiom **P1** plays the crucial role. **P1** assumes the full transitivity of the likelihood relation $\succeq_L$. Giving up the transitivity of

[1] As pointed out in [20] only one of these two equivalence conditions is explicitly stated in [21] and [19], even though the intent was clearly to have both of them, and the proofs of subsequent results presuppose it. However stating only one of these equivalence conditions is not sufficient for proving the representation theorem.

$\succeq_L$ suppresses the unnatural restriction of an almost total plausibility ordering of states, in so far as we wish to keep the sure-thing principle. We are led to formulate a weak form of **P1** [21]:

Axiom WP1: $(X^S, \succ)$ is a transitive, irreflexive, partially ordered set.

Dropping the transitivity of $\succeq$ cancels some useful consequences of the sure-thing principle under **P1**, which are nevertheless consistent with the likely dominance rule. For instance, axiom **WSTP** (or equivalently, the unanimity axiom **U**) will not follow from the relaxed framework. We must add it to get it. As a consequence, if one insists on sticking to a purely ordinal view of DMU, we come up to the framework defined by axioms **WP1, WSTP** (or **U**), **P3, P5**, and **OI**. The likelihood relation induced by **P4** is provably in agreement with classical deduction:

$$\text{If } B \succ_L A \text{ then } B \cup C \succ_L A \text{ and } B \succ_L A \cap C. \tag{7}$$

The null events are then all subsets of a subset N of null states. Moreover, if X has more than two elements, the likelihood relation satisfies the following strongly non-probabilistic property [41]: for any three pairwise disjoint non-null events A, B, C,

$$B \cup C \succ_L A \text{ and } A \cup C \succ_L B \text{ imply } C \succ_L A \cup B. \tag{8}$$

The statement $B \succ_L A$ then really means that event B is *much* more likely than A, because B will ever be more likely than any disjunction of events nor more likely than A. This likelihood relation can always be represented by a family of *possibility relations*. Namely, there is a family $\mathcal{F}$ of possibility relations $\succeq_\Pi$ on S and a weak order relation $\succeq_P$ on X such that the preference relation on acts is defined by the likely dominance rule restricted to possibility relations [27] :

$$f \succ g \text{ if and only if } \forall \succeq_P \in \mathcal{F}, [f \succ_P g] \succ_\Pi [g \succ_P f].$$

In [27], it is shown that this ordinal Savagean framework actually leads to a representation of uncertainty at work in the nonmonotonic logic system of Kraus, Lehmann and Magidor [57] (see Friedman and Halpern [48] who also study property (8).

A more general setting starting from a reflexive weak preference relation on acts is used in Dubois *et al.* [19]. In this framework **P3** is replaced by a monotonicity axiom on both sides, that is implied by Savage's framework, namely for any event A:

Monotonicity : If $h \succ_P f$ and $f \succeq g$ then $fAh \succeq g$;
if $g \succ_P h$ and $f \succeq g$ then $f \succeq gAh$.

Two additional axioms (also valid in Savage's framework) enable a unique possibility relation to be enforced as the resulting likelihood relation between disjoint events [19] :

Axiom EUN: $\forall A, B \subseteq S, (f \succeq g)_A$ and $(f \succeq g)_B$ jointly imply $(f \succeq g)_{A \cup B}$.

Axiom ANO If $s_1 \sim_L s_2$ then: $\forall f, g, f \succeq g \iff f(s_1)\{s_2\}f(s_2)\{s_1\}f \succeq g(s_1)\{s_2\}g(s_2)\{s_1\}g$.

The first axiom, EUN, extends unanimity U to any disjunction of events. The second one is an anonymity property ensuring that exchanging consequences of two equally plausible states does not alter the preferences between acts. Note that $f(s_1)\{s_2\}f(s_2)\{s_1\}f$ represents act f where states s_1 and s_2 have been exchanged. The following result obtains:

Theorem 2 *The two following properties are equivalent:*

- *Relation* $\succeq$ *on* X^S *is reflexive and complete, has a transitive strict part with at least two non-null states, and it satisfies axioms* **OI, A1, A5, LM, RM, EUN, ANO***;*
- *there exists a complete preorder* $\succeq_P$ *on* X *and a unique non-trivial possibility relation* $\succeq_\Pi$ *on events, such that* $\forall f, g \in X^S$, $f \succeq g \iff [f \succ g] \succ_\Pi [g \succ f]$.

While these results do characterise the possibilistic approach to uncertainty in qualitiative decision theory, the family of "rational" decision rules in the purely relational approach to decision under uncertainty is very restrictive. This restricted family only reflects the situation faced in voting theories where natural axioms lead to impossibility theorems (see for instance, Arrow [3], and Sen [69]). This kind of impediment was already pointed out by Doyle and Wellman [15] for preference-based default theories. 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.

The likely dominance rule lacks discrimination, not because of indifference between acts, but because of incomparability. Actually, it may be possible to weaken 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 [3]). It says that the preference $f \succ g$ only depends on f and g. 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 [65].

4.3 Qualitative decision rules under commensurateness

Let us now consider the axiomatization of absolute qualitative criteria (4), (5), (6), based on Sugeno integral in the scope of Savage theory. Before, note that the maximin and maximax criteria, to which possibilistic decision rules reduce when total ignorance prevails ($\forall s \in S, \pi(s) = 1$), had been axiomatized very early by Chernoff [11]. Arrow and Hurwicz[2] characterized the pair of criteria ($W_u^-(f)$, $W_u^+(f)$) for decision-making under total ignorance. Interestingly they carefully distinguish it from equiprobable states, and credit Shackle [70] for the invention of these criteria. More recently, Brafman and Tennenholtz [9] axiomatically characterise the refinement $W^-_{\trianglerighteq,u}$ of the maximin rule to unequally plausible states, due to Boutilier, in terms of conditional policies (rather than acts). Finally the first axiomatization of possibilistic decision rules was proposed by Dubois and Prade [36] in the style of decision under risk, assuming a possibility distribution is known, and adapting Von Neumann and Morgenstern axioms [64] (see [31] for the complete study).

Clearly, pessimistic, optimistic possibilistic criteria and Sugeno Integral satisfy **P1**. However the Sure Thing Principle can be severely violated by Sugeno integral. It is easy to show that there may exist f, g, h, h such that $fAh \succ gAh$ while $gAh \succ fAh$. It is enough to consider binary acts (events) and notice that, generally if A is disjoint from $B \cup C$, nothing prevents a fuzzy measure γ

from satisfying $\gamma(B) > \gamma(C)$ along with $\gamma(A \cup C) > \gamma(A \cup B)$ (for instance, Shafer's belief functions are such). The possibilistic criteria (4), (5) violate the sure-thing principle to a lesser extent since:

$\forall A \subseteq S, \forall f, g, h, h', W^{-}_{\pi,u}(fAh) > W^{-}_{\pi,u}(gAh)$ implies $W^{-}_{\pi,u}(fAh') \geq W^{-}_{\pi,u}(gAh')$

And likewise for $W^{+}_{\pi,u}$. Moreover, only one part of **P3** holds, for Sugeno integrals. The obtained ranking of acts satisfies the following

Axiom WP3: $\forall A \subseteq S, \forall x, y \in X, \forall f, x \succeq_P y$ implies $xAf \succeq yAf$.

Besides, axiom **P4** is violated by Sugeno integrals, but to some extent only. Namely, the preference relation between binary acts satisfies a weak form of it:

Axiom WP4: $\forall x \succ_P y, x' \succ_P y' \in X : xAy \succ xBy$ implies $x'Ay' \succeq x'By'$

which forbids preference reversals when changing the pair of consequences used to model events A and B. Moreover the strict preference is maintained if the pair of consequences is changed into more extreme ones:

$$\text{If } u(x') > u(x) > u(y) > u(y') \text{ then } S_{\gamma,u}(xAy) > S_{\gamma,u}(xBy) \Rightarrow S_{\gamma,u}(x'Ay') > S_{\gamma,u}(x'By'). \quad (9)$$

Sugeno integral and its possibilistic specializations are weakly Pareto-monotonic since $f \succeq_P g$ implies $S_{\gamma,u}(f) \geq S_{\gamma,u}(g)$, but one may have $f(s) \succ_P g(s)$ for some state s, while $S_{\gamma,u}(f) = S_{\gamma,u}(g)$. This is the so-called drowning effect, which also appears in the violations of **P4**. It is because some states are neglected when comparing acts.

The basic properties of Sugeno integrals exploit disjunctive and conjunctive combinations of acts. Namely, given a preference relation $(X^S, \succeq)$, and two acts f and g, define $f \wedge g$ and $f \vee g$ as follows

$$f \wedge g(s) = f(s) \text{ if } g(s) \succeq_P f(s) \text{ and } g(s) \text{ otherwise} \quad (10)$$

$$f \vee g(s) = f(s) \text{ if } f(s) \succeq_P g(s) \text{ and } g(s) \text{ otherwise.} \quad (11)$$

Act $f \wedge g$ always produces the worst consequences of f and g in each state, while $f \vee g$ always makes the best of them. They are union and intersection of fuzzy sets viewed as acts. Obviously

$$S_{\gamma,u}(f \wedge g) \leq \min(S_{\gamma,u}(f), S_{\gamma,u}(g)) \text{ and } S_{\gamma,u}(f \vee g) \geq \max(S_{\gamma,u}(f), S_{\gamma,u}(g))$$

from weak Pareto monotonicity. These properties hold with equality whenever f or g is a constant act. These properties are in fact characteristic of Sugeno integrals for monotonic aggregation operators [62]. Actually, these properties can be expressed by means of axioms, called restricted conjunctive and disjunctive dominance (**RCD** and **RDD**) on the preference structure $(X^S, \succeq)$ [38]:

- **Axiom RCD**: if f is a constant act, $f \succ h$ and $g \succ h$ jointly imply $f \wedge g \succ h$
- **Axiom RDD**: if f is a constant act,, $h \succ f$ and $h \succ g$ jointly imply $h \succ f \vee g$.

The area of significance of qualitative decision theory and more precisely the one of axioms **RCD** and **RDD**, is restricted to the case where X and S are finite and where the value scale is coarse. For instance, **RCD** means that limiting from above the potential utility values of an act g, that is better than another one h, to a constant value that is better than the utility of act h, still yields an act better than h. This is in contradiction with expected utility theory and debatable in that setting. Indeed, suppose g is a lottery where you win 1000 euros against nothing with equal chances. Suppose the certainty equivalent of this lottery is 400 euros, received for sure, and h is the fact of receiving 390 euros for sure. Now, it is likely that, if f represents the certainty-equivalent of g, $f \wedge g$ will be felt strictly less attractive than h, as the former means you win 400 euros against nothing with equal chances. Axiom **RCD** implies that such a lottery should ever be preferred to receiving $400 - \epsilon$ euros for sure, for arbitrary small values of ϵ. This axiom is thus strongly counterintuitive in the context of economic theory, with a continuous consequence set X. However the area of significance of qualitative decision theory is precisely when both X and S are finite.

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

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

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

Sugeno integral can be axiomatized in the style of Savage [38]. Namely:

Theorem 3 *If the preference structure $(X^S, \succeq)$ satisfies* **P1, WP3, P5, RCD** *and* **RDD**, *then there a finite chain V of preference levels, a $V-$valued capacity function γ, and a $V-$valued utility function u on the set of consequences X, such that the preference relation on acts is defined by $f \succeq g$ if and only if $S_{\gamma,u}(f) \geq S_{\gamma,u}(g)$.*

The proof goes as follows. Clearly, **P1, WP3**, and **P5** jointly imply Pareto-monotonicity between acts. In the representation method, V is the quotient set $X^S / \sim$, the utility value $u(x)$ is the equivalence class of the constant act f_x. Because the sure-thing principle is lacking, the degree of likelihood $\gamma(A)$ is the equivalence class of the binary act $1A0$, having extreme consequences. It yields the most refined likelihood relation between events due to (9). Then the equality $u(xA0) = \min(u(x), u(1A0))$ can be proved using **RCD**. Finally, due to **RDD**, $u(xA0 \vee yB0) = \max(u(xA0), u(yB0))$ can be obtained. The result is the easy to obtain by expressing any act f in a canonical form $\vee_{x \in X} x F_x 0$, with $F_x = \{s, u(f(s)) \geq u(x)\}$.

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

$$\textbf{Axiom NC}: \begin{cases} 1_L A y \sim y \quad \text{or} \quad 1_L A y \sim 1_L A 0 \\ \text{and} \\ x A 0_L \sim x \quad \text{or} \quad x A 0_L \sim 1_L A 0_L \end{cases}$$

Non-compensation formalizes the following intuition: in order to evaluate act $1_L A y$, there is no middle term between values $u(y)$ and $\gamma(1_L A 0)$.

Theorem 3 still holds if in the expression of **RCD** and **RDD** one considers any two comonotonic acts f and g (i.e. $f(s) >_P f(s') \Rightarrow g(s) \succeq_P g(s'), \forall s, s' \in S$). Indeed Sugeno integrals are "linear" for operations maximum and minimum with respect to disjunctions and conjunctions of comonotonic acts f, g: $S_{\gamma,u}(f \wedge g) = \min(S_{\gamma,u}(f), S_{\gamma,u}(g))$ and $S_{\gamma,u}(f \vee g) = \max(S_{\gamma,u}(f), S_{\gamma,u}(g))$. In this sense, Sugeno integral is a qualitative counterpart to Choquet integral.

It is easy to check that these equalities hold with any two acts f and g, for the pessimistic and the optimistic possibilistic preference functionals respectively $W^-_{\pi,u}(f \wedge g) = \min(W^-_{\pi,u}(f), W^-_{\pi,u}(g))$ and $W^+_{\pi,u}(f \vee g) = \max(W^+_{\pi,u}(f), W^+_{\pi,u}(g))$. The criterion $W^-_{\pi,u}(f)$ can thus be axiomatized by strengthening axiom **RCD** as follows:

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

This axiom means that if two acts f, g are individually better than a third one, the act $f \wedge g$ which yields the worse result of both acts still remains better than the third one. It makes sense in the scope of a one-shot-decision. Together with **P1, WP3, RDD** and **P5, CD** implies that the set-function γ is a necessity measure and so, $S_{\gamma,u}(f) = W^-_{\pi,u}(f)$, for some possibility distribution π.

In order to figure out why axiom CD leads to a pessimistic criterion, Dubois, Prade and Sabbadin [39] have noticed that it can be equivalently replaced by the following property:

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

This property can be explained as follows: if changing g into f when A occurs results in a better act, the decision maker has enough confidence in event A to consider that improving the results on A is worth trying. But, in this case, there is less confidence on the complement A^c than in A, and any possible improvement of g when A^c occurs is neglected. Alternatively, the reason why $fAg \succ g$ holds may be that the consequences of g when A occurs are very bad and the occurrence of A is not unlikely enough to neglect them, while the consequences of g when A^c occurs are acceptable. Then suppose that consequences of f when A occurs are acceptable as well. Then $fAg \succ g$. But act gAf remains undesirable because, regardless of whether the consequences of f, when A^c occurs, are acceptable, act gAf still possesses plausibly bad consequences when A occurs. So, $g \succeq gAf$. For instance, g means losing ($= A$) or winning ($= A^c$) 10,000 euros with equal chances according to whether A occurs or not, and f means winning either nothing ($= A$) or 20,000 euros ($= A^c$) conditioned on the same event. Then fAg is clearly safer than g as there is no risk of losing money. However, if axiom **PESS** holds, then the chance of winning much more money (20,000 euros) by choosing act $gAf$is neglected because there is still a good chance to lose 10,000 euros with this lottery. Such a behaviour is clearly cautious.

Similarly, the optimistic criterion $W^+_{\pi,u}(f)$ can be axiomatized by strengthening the axioms **RDD** as follows:

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

Together with P1, WP3, RCD, P5, DD implies that the set-function γ is a possibility measure and so, $S_{\gamma,u}(f) = W^+_{\pi,u}(f)$ for some possibility distribution π. The optimistic counterpart to property axiom **PESS** that can serve as a substitute to axiom DD for the representation of criterion $W^+_{\pi,u}$ is:

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

5 Toward more efficient qualitative decision rules

The absolute approach to qualitative decision criteria is simple (especially in the case of possibility theory). Naturally, a complete preorder on acts is obtained. The restriction of the pessimistic approach to the most plausible states, at work in possibilistic criteria, makes them more realistic than the maximin criterion, and more flexible than purely ordinal approaches based on the likely dominance rule.

However, approaches based on an absolute qualitative value scale have their own shortcomings. First, one has to accept the commensurability assumption between utility and degrees of likelihood. It assumes the existence of a common scale for grading uncertainty and preference. 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). It is already implicit in Savage approach, and looks acceptable for decision under uncertainty (but more debatable in social choice). And as a consequence, the acts are then totally preordered. This is not really a drawback from a normative point of view.

More importantly, absolute qualitative criteria lack discrimination due to many indifferent acts. 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. The above possibilistic decision rule and the maximin rule are consistent with Pareto dominance only in the wide sense: they can consider two acts as indifferent even if one Pareto-dominates the other. The Sure Thing principle may be violated (even if not drastically for possibilistic criteria). This section tries to survey approaches that try to remedy this problem.

5.1 Refining qualitative criteria

The main reason for the lack of discrimination power of absolute qualitative criteria is the fact that they do not use all the available information for discriminating among acts, since an act f can be considered indifferent to another act g, even if f is at least as good as g in all states and strictly in some states (including some of the most plausible ones). This defect is absent from the expected utility model.

One may consider refining the optimistic possibilistic criterion by the pessimistic one or vice-versa [32]. Along this line, Giang and Shenoi [49, 50] have tried to obviate the need for making assumptions on the pessimistic or optimistic attitude of the decision-maker and thus, improve the discrimination power in the absolute qualitative setting, by using, as a utility scale, a totally ordered set of possibility measures on a two element set $\{0, 1\}$ containing the values of the best and the worst consequences. Each such possibility distribution represents a qualitative lottery. Let $V_\Pi = \{(a, b), \max(a, b) = 1, a, b \in V\}$. Coefficient a represents the degree of possibility of obtaining the worst consequence, and coefficient b the degree of possibility of obtaining the best. This set can be viewed as a bipolar value scale ordered by the following complete preordering relation:

$$(a, b) \geq_V (c, d) \text{ if and only if } (a \leq c \text{ and } b \geq d)$$

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

$$A \succeq_{L\Pi} B \text{ if and only if } (\Pi(A^c), \Pi(A)) \geq_V (\Pi(B^c), \Pi(B))$$

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

Each consequence x is supposed to have a utility value (α, β) in V_Π. The proposed preference functional maps acts, viewed as $n-$tuples $f = ((\alpha_1, \beta_1), \ldots, (\alpha_n, \beta_n))$ of values in V_Π, to V_Π itself. The uncertainty is described by possibility weights $(\pi_1, \ldots, \pi_n)$ with $\max_{i=1,\ldots,n} \pi_i = 1$. The utility of an act f, called *binary possibilistic utility* is computed as the pair

$$W_{GS}(f) = (\max_{i=1,\ldots,n} \min(\pi_i, \alpha_i), \max_{i=1,\ldots,n} \min(\pi_i, \beta_i)) \in V_\Pi.$$

This form results from simple and very natural axioms on possibilistic lotteries, which are counterparts to the Von Neumann and Morgenstern axioms [64]: a complete preorder of acts, increasingness in the wide sense according to the ordering in V_Π, substitutability of indifferent lotteries, and the assumption that any consequence of an act is valued on V_Π. More recently, Weng [79] proposed a Savage-style axiomatization of binary possibilistic utility. It puts together the axiomatizations of the optimistic and the pessimistic possibilistic criteria by Dubois *et al.* [39], adding, to the axioms justifying Sugeno integral, two conditions: (i) the self-adjointness of the preference relation on binary acts, and (ii) an axiom enforcing axiom **OPT** on the subset of acts $\{f : f \succeq h, f \in X^S\}$ weakly preferred to an act h that plays the role of a neutral point separating favorable from unfavorable acts in X^S.

Pessimistic and optimistic possibilistic criteria turn out to be special case of this bipolar criterion. They respectively correspond to either using the negative part of V_Π only (not telling $(1, 1)$ from $(0, 1)$ in case of pessimism) or using the positive part of V_Π only (not telling $(1, 0)$ from $(1, 1)$ in case of optimism). The decision rule of Giang and Shenoy can capture the lexicographic use of possibilistic criteria $W^-_{\pi,n(u)}$ and $W^+_{\pi,u}$, where the optimist one is used when the pessimistic one cannot discriminate (or conversely)[51, 78]. Yet, this criterion has a major drawback: Whenever two states s_i and s_j are such that $\alpha_i = 1$ and $\beta_j = 1$ (respectively a bad or neutral, and a good or neutral state) and these states have maximal possibility $\pi_i = \pi_j = 1$, then $W_{GS}(f) = (1, 1)$ results, expressing indifference. This limited expressiveness seems to be unavoidable when using finite bipolar scales [52].

Lehmann [60] axiomatizes a refinement of the maximin criterion whereby ties between equivalent worst states are broken by considering their respective likelihoods. This decision rule takes the form of an expected utility criterion with qualitative (infinitesimal) utility levels. An axiomatization is carried out in the Von Neumann-Morgenstern style [64].

The lack of discrimination of the maximin rule itself was actually addressed a long time ago by Cohen and Jaffray [13] who improve it by comparing acts on the basis of their worst consequences of *distinct* merits, i.e. one considers only the set $D(f, g) = \{s, u(f(s)) \neq u(g(s))\}$ when performing a minimization. Denoting by $f \succ_D g$ the strict preference between acts,

$$f \succ_D g \iff \min_{s \in D(f,g)} f(s) > \min_{s \in D(f,g)} g(s) \tag{12}$$

and the weak preference is $f \succeq_D g \iff \neg(g \succ_D f)$. This refined rule always rates an act f better than another act g whenever f is at least as good as g in all states and better in some states (strict compatibility with Pareto-dominance). However, only a partial ordering of acts is then obtained. This last decision rule is actually no longer based on a preference functional (i.e. it cannot be encoded by a numerical function, like expected utility). This decision rule has been independently proposed by Fargier and Schiex [43] and used in fuzzy constraint satisfaction problems [29] under the name *discrimin ordering*.

This criterion can be further refined by the so-called *Leximin* ordering [63]: The idea is to reorder utility vectors $\vec{f} = (u(f(s_1)), \ldots u(f(s_n)))$ by non-decreasing values as $(f_{(1)}, \ldots, f_{(n)})$, where $f_{(k)}$ is the kth smallest component of $\vec{f}$ (i.e., $f_{(1)} \leq \ldots \leq f_{(n)}$). Similarly, a *Leximax* preorder can be envisaged as a refinement of the one induced by the maximum. Let $\vec{f}, \vec{g} \in L^N$. Define the *Leximin* ($\succeq_{lmin}$) and *Leximax* ($\succeq_{lmax}$) rules as:

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

The two possible decisions f and g are indifferent if and only if the corresponding reordered vectors are the same. The *Leximin*-ordering is a refinement of the discrimin ordering, hence of both the Pareto-ordering and the maximin-ordering [24]: $f \succ_D g$ implies $f \succ_{lmin} g$. Leximin optimal decisions are always discrimin maximal decisions, and thus indeed min-optimal and Pareto-maximal: $\succ_{lmin}$ is the most selective among these preference relations. The *Leximin* ordering can discriminate more than any symmetric aggregation function, since when, e.g. in the numerical setting, the sum of the $u(f(s_i)$'s equals the sum of the $u(f(s_i)$'s, it does not mean that the reordered vectors are the same.

5.2 A bridge between generalized maxmin criteria and expected utility.

Criteria of the discrimin and *Leximin* type unfortunately never take into account the available information on the state of affairs, contrary to possibilistic criteria. One idea is to refine the latter by changing minimum into *Leximin* and maximum into *Leximax* in each of them so as to integrate the plausibility of states within the lexicographic approach. It turns out that this can be encoded by means of an extreme form of expected utility([45]).

First note that, in a finite setting, the qualitative *Leximin* and *Leximax* rules can be simulated by means of a sum of utilities provided that the levels in the qualitative (finite) utility scale V are mapped to values sufficiently far away from one another on a numerical scale. Consider an increasing mapping ϕ from V to the reals. It is possible to define this mapping in such a way as to refine the maximax ordering:

$$\max_{i=1,\ldots n} f_i > \max_{i=1,\ldots n} g_i \text{ implies } \sum_{i=1,\ldots n} \phi(f_i) > \sum_{i=1,\ldots n} \phi(g_i) \tag{13}$$

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

It can actually be checked that the *Leximax* ordering is equivalent to applying the Bernoulli criterion with respect to such a convex utility function $\phi(.)$:

$$f >_{Leximax} g \text{ if and only if } \sum_{i=1,\ldots n} \phi(f_i) > \sum_{i=1,\ldots n} \phi(g_i). \tag{14}$$

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

$$\min_{i=1,\ldots n} f_i > \min_{i=1,\ldots n} g_i \text{ implies } \sum_{i=1,\ldots n} \psi(f_i) > \sum_{i=1,\ldots n} \psi(g_i) \tag{15}$$

The *Leximin* ordering comes down to applying the Bernoulli criterion with respect to such a concave utility function $\psi(.)$. Notice that these transformations are not possible when V is not finite [63] although the *Leximin* and *Leximax* procedures still make mathematical sense even in this case. The qualitative pessimistic and optimistic criteria under total ignorance are thus refined by means of a classical criterion with respect to a risk-averse and risk-prone utility function respectively, as can be seen by plotting V against numerical values in $\phi(V)$ and $\psi(V)$.

These refinement principles have been extended to possibilistic criteria [45] using weighted averages. Consider first the optimistic possibilistic criterion $W^+_{\pi,\mu}$ under a given possibility distribution π. We can again define an increasing mapping χ from V to the reals such that $\chi(\lambda_0) = 0$ and especially:

$$\begin{array}{l} \max_i \min(\pi(s_i), \mu(f(s_i))) > \max_i \min(\pi(s_i), \mu(g(s_i))) \\ \text{ implies} \\ \sum_{i=1,\ldots n} \chi(\pi(s_i)) \cdot \chi(\mu(f(s_i))) > \sum_{i=1,\ldots n} \chi(\pi(s_i)) \cdot \chi(\mu(g(s_i))) \end{array} \tag{16}$$

A sufficient condition is that:

$$\forall i \in \{1, \ldots, m\}, \chi(\lambda_i)^2 \geq N\chi(\lambda_{i-1}) \cdot \chi(1) \tag{17}$$

for $N > n$. The increasing mapping $\chi(\lambda_i) = \frac{N}{N^{2^{m-i}}}, i = 1, \ldots, m$, and $\chi(\lambda_0) = 0$, with $N = n + 1$ can be chosen, with $n = |S|; m = |V|$. It is such that $\chi(\lambda_m) = 1$

Moreover, let $\{E_0, \ldots, E_k\}$ be the well-ordered partition of S induced by π, E_k containing the most plausible states, and E_0 the null states. Let $K = \frac{1}{\sum_{i=1,k} |E_i| \cdot \chi(\pi(s_i))}$. Define $\chi^*(\lambda_i) = K\chi(\lambda_i)$, it holds that:

- $p = \chi^*(\pi(\cdot))$ is a probability assignment respectful of the possibilistic ordering of states. In particular, p is uniform on equi-possible states (the sets E_i). Moreover, if $s \in E_i$ then $p(s)$ is greater than the sum of the probabilities of all less probable states, that is, $p(s) > P(E_{i-1} \cup \cdots \cup E_0)$. Such probabilities, here said to be *big-stepped*, generalize the linear big-stepped probabilities that form a super-increasing sequence of numbers assigned to singletons. They are introduced by Snow [72] and also studied in Benferhat *et al.* [5] in connection with non-monotonic reasoning. Linear big-stepped probabilities are recovered when the E_i's are singletons.
- $\chi(u(\cdot))$ is a big-stepped numerical utility function (a super-increasing sequence of reals $u_l > \cdots > u_1$ such that $\forall l \geq i > 1, u_i > n \cdot u_{i-1}$) that can be encoded by a convex real mapping.
- The preference functional

$$EU_+(f) = \sum_{i=1,\ldots n} \chi^*(\pi(s_i)) \cdot \chi(u(f(s_i))) \quad (18)$$

is an expected (big-stepped) utility criterion for a risk-seeking decision-maker, and $W^+_{\pi,\mu}(f) > W^+_{\pi,\mu}(g)$ implies $EU_+(f) > EU_+(g)$. Namely this is precisely equation (16) up to the multiplicative constant K, i.e., the expected utility criterion so-obtained refines the possibilistic optimistic criterion. As a refinement, it is perfectly compatible with, but more decisive than, the optimistic utility. Since it is based on expected utility, it obviously satisfies the Sure Thing Principle as well as Strict Pareto Dominance, actually recovering Savage's five first axioms. Moreover, it does not use any other information but the original ordinal one. It can be shown that it is not the only criterion in this family of sound "unbiased" refinements, but it is the most efficient among them (up to an equivalence relation), since it refines any unbiased refinement of the possibilistic optimistic criterion (see [45] for more details).

The pessimistic criterion can be similarly refined. Notice that $W^-_{\pi,\mu}(f) = \nu(W^+_{\pi,\nu(u)}(f))$ using the order-reversing map ν of V. Then, choosing the same mapping χ^* as above, one may have that

$$\begin{array}{l} \min_i \max(\pi(s_i), \mu(f(s_i))) > \min_i \max(\pi(s_i), u(g(s_i))) \\ \text{implies} \\ \sum_{i=1,\ldots n} \chi^*(\pi(s_i)) \cdot \phi(u(f(s_i))) > \sum_{i=1,\ldots n} \chi^*(\pi(s_i)) \cdot \phi(u(g(s_i))) \end{array} \quad (19)$$

where $\phi(\lambda_i) = 1 - \chi(\nu(\lambda_i))$ (it is equal to $1 - \frac{n+1}{(n+1)^{2^i}}$, here). Function $\phi(u(\cdot))$ is a super-increasing numerical utility function that can be encoded by a concave real mapping, and the expected utility criterion

$$EU_-(f) = \sum_{i=1,\ldots n} \chi^*(\pi((s_i))) \cdot \phi(u(f(s_i))) \quad (20)$$

is a risk-averse one, that refines $W_{\pi,\mu}$ in the sense that $W^{-}_{\pi,\mu}(f) > W^{-}_{\pi,\mu}(g)$ implies $EU_{-}(f) > EU_{-}(g)$.

These results point out the deep agreement between qualitative possibilistic criteria and expected utility. The former is just coarser than the latter, and as such cannot account for compensative effects. Actually both types of criteria are subsumed within a more abstract algebraic approach using operations on a semi-ring by Chu and Halpern [12] and Weng [80].

5.3 Weighted *Leximax* / *Leximin* criteria

The orderings induced by $EU_{+}(f)$ and $EU_{-}(f)$ actually correspond to generalizations of *Leximin* and *Leximax* to prioritized minimum and maximum aggregations, thus bridging the gap between possibilistic criteria and classical decision theory. To make this generalization clear, let us simply consider that *Leximin* and *Leximax* orderings are defined on sets of tuples whose components belong to a totally ordered set $(\Omega, \trianglerighteq)$, say $Leximin(\trianglerighteq)$ and $Leximax(\trianglerighteq)$. Now, suppose $(\Omega, \trianglerighteq) = (V^l, Leximin)$ or $(\Omega, \trianglerighteq) = (V^l, Leximax)$, with any $l \in \mathbb{N}$. Then, nested lexicographic ordering relations can be recursively defined by nesting procedures such as $(Leximin(Leximin(\geq))$, $Leximax(Leximin(\geq))$, $Leximin(Leximax(\geq))$, $Leximax(Leximax(\geq)))$, in order to compare V-valued matrices.

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

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

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

$$\max_i \min_j a_{ij} > \max_i \min_j b_{ij} \text{ implies } [a] \succ_{lmax(\succeq lmin)} [b].$$

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

Comparing acts f and g in the context of a possibility distribution π can be done using relations $\succeq_{lmax(\succeq lmin)}$ applied to $n \times 2$ matrices on $(V, \leq)$, n being the number of states in S, namely on the matrix $[f]_{\pi,u}$ and $[g]_{\pi,u}$ with coefficients $f_{i1} = \pi(s_i)$ and $f_{i2} = \mu(f(s_i))$, $g_{i1} = \pi(s_i)$ and $g_{i2}\mu(g(s_i))$.

The big-stepped expected utility $EU_{+}(f)$ defined in the previous subsection precisely encodes the relation $\succeq_{lmax(\succeq lmin)}$:

Theorem 4 *[45]: $EU_{+}(f) \geq EU_{+}(f)$ if and only if $[f]_{\pi,u} \succeq_{lmax(\succeq lmin)} [g]_{\pi,u}$.*

In other terms, EU_+ applies a *Leximax* procedure to utility degrees weighted by possibility degrees. Similarly, EU_- applies a *Leximin* procedure to utility degrees weighted by "impossibility degrees":

Theorem 5 *[45]: $EU_-(f) \geq EU_-(g)$ if and only if $[f]_{n(\pi),u} \succeq_{lmin(\succeq lmax)} [g]_{n(\pi),u}$.*

i.e., $EU_-(f)$ just encodes the application of a procedure *Leximin(Leximax)* not directly on $[f]_{\pi,u}$ and $[g]_{\pi,u}$ but on the corresponding π-reverse matrix $[f]_{n(\pi),u}$ and $[g]_{n(\pi),u}$ with coefficients $f_{i1} = \nu(\pi(s_i))$ and $f_{i2} = \mu(f(s_i))$, $g_{i1} = \nu(\pi(s_i))$ and $g_{i2} = u(g(s_i))$.

As a consequence, the additive preference functionals $EU_+(f)$ and $EU_-(f)$ refining the possibilistic criteria are qualitative despite their numerical encoding. Moreover, the two orderings $\succeq_{lmax(\succeq lmin)}$ and $\succeq_{lmin(\succeq lmax)}$ of acts are defined even on coarse ordinal scales V while *obeying Savage five first axioms of rational decision.* Weng [78] has extended this approach to the binary possibilistic utility of Giang and Shenoy, recalled in subsection 5.1.

5.4 The representation of uncertainty underlying *Leximax(Leximin)* and *Leximin(Leximax)* criteria

The two relations $\succeq_{lmax(\succeq lmin)}$ and $\succeq_{lmin(\succeq lmax)}$ coincide if the utility functions are Boolean, and then compare events by their likelihood. The corresponding uncertainty representation is precisely the lexi-refinement of possibility orderings $\succeq_\Pi$ ($A \succeq_\Pi B \iff \Pi(A) \geq \Pi(B)$) already identified by [26]:

$$A \succeq_{l\Pi} B \text{ if and only if } \vec{\pi_A} \succeq_{lmax} \vec{\pi_B} \tag{21}$$

where $\vec{\pi_A}$ is the vector $(a_1, \ldots, a_n)$ such that $a_i = \pi(s_i)$ if $s_i \in A$ and $a_i = \bot$ otherwise. This relation among events is called "*Leximax*" likelihood [26, 17]. It is a complete preordering whose strict part refines the possibilistic ordering of events together its adjoint necessity ordering ($A \succeq_N B \iff \Pi(B^c) \geq \Pi(A^c)$). This is not surprising since $\succeq_{lmin}$ and $\succeq_{lmax}$ are adjoint : $\vec{u} \succeq_{lmin} \vec{v}$ if and only if $(\nu(v_1), \ldots, \nu(v_k)) \succeq_{lmax} (\nu(u_1), \ldots, \nu(u_k))$. If $\vec{u}$ and $\vec{v}$ are Boolean and encode events A and B, it comes down to $A \succeq_{l\Pi} B$ if and only if $B^c \succeq_{l\Pi} A^c$. Relation $\succeq_{l\Pi}$ is thus self-adjoint.

Another natural way of refining a possibility relation is to delete, in the spirit of Cohen-Jaffray's decision rule $\succ_D$, states common to two events A and B [26]:

$$A >_{D\Pi} B \iff \Pi(A \setminus B) > \Pi(B \setminus A)$$

This partial ordering relation, called *possibilistic likelihood* also refines the weak order induced by the necessity measure. It is preadditive like probability relations, and also self-adjoint.

The relation $\succeq_{l\Pi}$ refines the above possibilistic likelihood relation and coincides with it for linear plausibility rankings of states (by which S is totally ordered). In the case of a uniform distribution, the *Leximax* likelihood relation coincides with a qualitative probability relation induced by a uniform probability distribution (comparing the cardinality of sets, $A \succ_{l\Pi} B$ if and only if $|B| > |A|$).

The uncertainty representation underlying the $\succeq_{lmax(\succeq lmin)}$ and $\succeq_{lmin(\succeq lmax)}$ decision rules is thus probabilistic, although qualitative. The *Leximax* likelihood relation is a special qualitative probability relation representable by means of the big-stepped probability P, involved in functionals

$EU_+(f)$ and $EU_-(f)$, i.e. $A \succeq_{l\Pi} B \iff P(A) \geq P(B)$. Another formulation of the reason why the *Leximax* likelihood relation is self-adjoint thus consists in noticing that $EU_+(f)$ and $EU_-(f)$ share the same big-stepped probability function P.

6 Conclusion

This chapter is an overview of qualitative criteria for decision under uncertainty. These results also apply, to some extent, to multicriteria decision, where the various objectives play the role of states and the likelihood relation is used to compare the relative importance of groups of objectives [22, 34]. Indeed, the commensurability assumptions of the absolute possibilistic approach are often more difficult to advocate between objectives having different natures in multicriteria evaluation than for states of the world in decision under uncertainty. To reconcile the two frameworks, evaluation methods in this chapter should be articulated with conjoint measurement methods described by Bouyssou and Pirlot in this book.

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

- when the problem is located in a dynamic environment involving a large state space, a non-quantifiable goal to be reached, and a partial information on the current state. This case can be found in robotic planification problems [66];
- when only a very high level description of a decision problem is available, where states and consequences of decisions are coarsely defined (for instance in some kinds of strategic decision-making);
- or yet when there is no time to quantify utilities and probabilities because a fast advice is requested (like in recommender systems).

Possibilistic criteria were also used in scheduling problems, so as to produce robust sequencings of tasks, ensuring moderate and balanced violations of due-dates in case of unexpected events [23]. These criteria compatible with dynamic programing in multi-step decision problems [42].

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

1. Faithfulness to available information supplied by decision-makers, as poor as it be : an ordinal declarative approach sounds closer to human capabilities.
2. Cognitive Relevance : The number of levels in the value scale must be small enough (according to well-known psychological studies, man does not make a difference beyond seven value levels).
3. Good Discrimination : especially respecting the strict Pareto-dominance.

4. Decisive Power : avoiding incomparability and favor linear rankings.
5. Taking into account the decision-maker's attitude in the face of risk and uncertainty.
6. Taking into account the available information on the current state of affairs.

These requirements are often conflicting. Expected utility is information-demanding, and hardly compatible with the limited perception capabilities of human decision-makers. The maximin criterion of Wald and its refinements neglect available information on the state of affairs. In the present overview, two kinds of qualitative decision rules, compatible with the two first requirements, have been laid bare. Approaches based on ordinal preference relations, including the likely dominance rule, are in full agreement with Pareto-dominance, satisfy the sure-thing principle, but either leave room to incomparability to a large extent, or focus too much on the most plausible states. Approaches based on an absolute value scale improve the expressivity of maximin and maximax criteria by injecting the respective plausibility of states. They provide rankings of decisions but lack discrimination power. There is some inconsistency between the requirement of a fine-grained discrimination (respecting Pareto-dominance) and the requirement of a total (especially transitive) ranking of acts in the qualitative framework. Enforcing both conditions seems to bring us back to a special case of expected utility as explained in Section 5.2. In a purely qualitative setting, weighted lexicographic criteria seem to maximize the number of satisfied natural requirements listed above.

Many problems remain open in this area. Let us cite some of them:

- How to refine the coarse ranking induced by Sugeno integral? Recent results [18] suggest it can be refined by means of Choquet integral.
- The likely dominance rule compares two acts independently of others. More expressive decision rules involving the comparison with a third reference act can be envisaged by weakening the ordinal invariance axiom [65].
- How do qualitative criteria behave in a dynamic environment where new information is acquired, in the absence of the Sure Thing Principle to ensure dynamic consistency? It requires the study of conditional acts and qualitative conditional preference functionals. See [28] for a preliminary study in the case of possibilistic criteria.
- Qualitative decision criteria make it easier to explain why some decisions are better than others, laying bare arguments for or against them. See [1] for a discussion of the role of argumentation in decision under uncertainty.

References

[1] L. Amgoud and H. Prade. Explaining qualitative decision under uncertainty by argumentation. In *Proc. 21st National Conference on Artificial Intelligence, Boston, Massachusetts, USA*. AAAI Press, 2006.

[2] K. Arrow and L. Hurwicz. An optimality criterion for decision-making under ignorance. In C.F. Carter and J.L. Ford, editors, *Uncertainty and Expectations in Economics*, pages 1–12. Basil Blackwell and Mott Ltd, 1972.

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

[4] F. Bacchus and A. J. Grove. Utility independence in a qualitative decision theory. In *Proc. of the Fifth International Conference on Principles of Knowledge Representation and Reasoning (KR'96), Cambridge, Mass., USA*, pages 542–552, 1996.

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

[6] C. Boutilier. Toward a logic for qualitative decision theory. In *Proc. of the 4th International Conference on Principles of Knowledge Representation and Reasoning (KR'94), Bonn, Germany, USA*, pages 75–86, 1994.

[7] D. Bouyssou, Th. Marchant, M. Pirlot, P. Perny, A. Tsoukiàs, and Ph. Vincke. *Evaluation and Decision Models: A critical perspective.* Kluwer Academic Publishers, 2000.

[8] R. Brafman and M. Tennenholtz. Modeling agents as qualitative decision makers. *Artificial Intelligence*, 94:217–268, 1997.

[9] R. Brafman and M. Tennenholtz. On the axiomatization of qualitative decision criteria. *Journal of the ACM*, 47:452–482, 2000.

[10] M. Cayrol, H. Farreny, and H. Prade. Fuzzy pattern matching. *Kybernetes*, 11:103–116, 1982.

[11] H. Chernoff. Rational selection of decision functions. *Econometrica*, 22:422–444, 1954.

[12] F. C. Chu and J. Y. Halpern. Great expectations. part ii: generalized expected utility as a universal decision rule. *Artif. Intell.*, 159(1-2):207–229, 2004.

[13] M. Cohen and J.Y. Jaffray. Rational behavior under complete ignorance. *Econometrica*, 48(5):1281–99, July 1980.

[14] J. Doyle and R. H. Thomason. Background to qualitative decision theory. *AI Magazine*, 20(2):55–68, 1999.

[15] J. Doyle and M. P. Wellman. Impediments to universal preference-based default theories. *Artificial Intelligence*, 49:97–128, 1991.

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

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

[18] D. Dubois and H. Fargier. Lexicographic refiniments of Sugeno integrals. In K. Mellouli, editor, *European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU), Hammamet, Tunisie*, number 4724 in Lecture Notes in Artificial Intelligence, LNAI, pages 611–622, http://www.springerlink.com, 2007. Springer.

[19] D. Dubois, H. Fargier, and P. Perny. Qualitative models for decision under uncertainty: an axiomatic approach. *Artificial Intelligence*, 148:219–260, 2003.

[20] D. Dubois, H. Fargier, and P. Perny. Corrigendum to qualitative decision theory with preference relations and comparative uncertainty: an axiomatic approach. *Artificial Intelligence*, 171:361–362, 2007.

[21] D. Dubois, H. Fargier, P. Perny, and H. Prade. Qualitative decision theory: from Savage's axioms to non-monotonic reasoning. *J. of the ACM*, 49:455–495, 2002.

[22] D. Dubois, H. Fargier, P. Perny, and H. Prade. A characterization of generalized concordance rules in multicriteria decision making. *Int. J. of Intelligent Systems*, 18:751–774, 2003.

[23] D. Dubois, H. Fargier, and H. Prade. Fuzzy constraints in job-shop scheduling. *J. of Intelligent Manufacturing*, 6:215–234, 1995.

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

[25] D. Dubois, H. Fargier, and H. Prade. Decision making under ordinal preferences and comparative uncertainty. In *Proc. of Int. Conf. on Uncertainty in Artificial Intelligence (UAI'97)*, pages 157–164, 1997.

[26] D. Dubois, H. Fargier, and H. Prade. Possibilistic likelihood relations. In *Proc. of the International Conference on Information Processing and Management of Uncertainty (IPMU'98), Paris, France*, pages 1196–1203. EDK, 1998.

[27] D. Dubois, H. Fargier, and H. Prade. Ordinal and probabilistic representations of acceptance. *J. Artificial Intelligence Research*, 22:23–55, 2004.

[28] D. Dubois, H. Fargier, and B. Vantaggi. An axiomatization of conditional possibilistic preference functionals. In K. Mellouli, editor, *European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU), Hammamet, Tunisie*, volume 4724 of *Lecture Notes in Artificial Intelligence, LNAI*, pages 803–815, http://www.springerlink.com, 2007. Springer.

[29] D. Dubois and P. Fortemps. Computing improved optimal solutions to max-min flexible constraint satisfaction problems. *European Journal of Operational Research*, 118:95–126, 1999.

[30] D. Dubois and P. Fortemps. Selecting preferred solutions in the minimax approach to dynamic programming problems under flexible constraints. *European Journal of Operational Research*, 19:441–471, 2005.

[31] D. Dubois, L. Godo, H. Prade, and A. Zapico. On the possibilistic decision model: from decision under uncertainty to case-based decision. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 7:631–670, 1999.

[32] D. Dubois, L. Godo, H. Prade, and A Zapico. Advances in qualitative decision theory: Refined rankings. In *7th Ibero-American Conference on AI and 15 th Brazilian Symposium on AI, IBERAMIA-SBIA'2000, LNAI 1952,*, pages 427–436, Berlin, 2000. Springer.

[33] D. Dubois, M. Grabisch, F. Modave, and H. Prade. Relating decision under uncertainty and multicriteria decision making models. *International Journal of Intelligent Systems*, 15(10):967–979, 2000.

[34] D. Dubois, J.-L. Marichal, H. Prade, M. Roubens, and R. Sabbadin. The use of the discrete Sugeno integral in decision-making: a survey. *Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9:539–561, 2001.

[35] D. Dubois and H. Prade. *Possibility Theory.* Plenum Press, New York (NY), 1988.

[36] D. Dubois and H. Prade. Possibility theory as a basis for qualitative decision theory. In *Proc. of the Int. Joint Conference on Artificial Intelligence (IJCAI'95)*, pages 1925–1930, 1995.

[37] D. Dubois and H. Prade. Possibility theory: qualitative and quantitative aspects. In P. Smets, editor, *Handbook on Defeasible Reasoning and Uncertainty Management Systems Volume 1: Quantified Representation of Uncertainty and Imprecision*, pages 169–226. Kluwer Academic Publ., 1998.

[38] D. Dubois, H. Prade, and R. Sabbadin. Qualitative decision theory with Sugeno integrals. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integrals - Theory and Applications*, pages 314–322. Physica Verlag, 2000.

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

[40] D. Dubois, H. Prade, and C. Testemale. Weighted fuzzy pattern matching. *Fuzzy Sets and Systems*, 28:313–331, 1988.

[41] Didier Dubois and Henri Prade. Numerical representations of acceptance. In *Proc. of Int. Conf. on Uncertainty in Artificial Intelligence (UAI'95)*, pages 149–156, 1995.

[42] H. Fargier, J. Lang, and R. Sabbadin. Towards qualitative approaches to multi-stage decision making. *International Journal of Approximate Reasoning*, 19:441–471, 1998.

[43] H. Fargier, J. Lang, and T. Schiex. Selecting preferred solutions in fuzzy constraint satisfaction problems. In *Proc. 1st European Congress on Fuzzy and Intelligent Technologies (EUFIT'93)*, pages 1128–1134, Aachen, Germany, 1993.

[44] H. Fargier and P. Perny. Qualitative models for decision under uncertainty without the commensurability assumption. In *Proc. of the 15th Conf. on Uncertainty in Artificial Intelligence (K. Laskey, H. Prade, eds.),*, pages 157–164, Stockholm, Suede, 1999.

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

[46] P. C. Fishburn. Axioms for lexicographic preferences. *Review of Economic Studies*, 42:4.5–419, 1975.

[47] P. C. Fishburn. The axioms of subjective probabilities. *Statistical Science*, 1:335–358, 1986.

[48] N. Friedman and J. Y. Halpern. Plausibility measures and default reasoning. In *Proceedings of AAAI'96*, pages 1297–1304, 1996.

[49] P. H. Giang and P. P. Shenoy. A qualitative utility theory for spohn's theory of epistemic beliefs. In *Proc. of the 16th Conf. on Uncertainty in Artificial Intelligence*, pages 220–227, 2000.

[50] P. H. Giang and P. P. Shenoy. Two axiomatic approaches to decision-making using possibility theory. *European Journal of Operational Research*, 162:450–467, 2005.

[51] L. Godo and A Zapico. Lexicographic refinements in the context of possibilistic decision theory. In *Proc. of International Conference in Fuzzy Logic and Technology (EUSFLAT05)*, pages 569–575, Barcelona, 2005. Universitat Politecnica.

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

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

[54] S. Grant, A. Kaji, and B. Polak. Decomposable choice under uncertainty. *J. Economic Theory*, 92:169–197, 2000.

[55] J.-Y. Jaffray. Linear utility theory for belief functions. *Operations Research Letters*, 8:107–112, 1989.

[56] C. H. Kraft, J. W. Pratt, and A. Seidenberg. Intuitive probability on finite sets. *Annals of Mathematical Statistics*, 30:408–419, 1958.

[57] S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44(1-2):167–207, 1990.

[58] J. Lang. Conditional desires and utilities: an alternative logical approach to qualitative decision theory. In *Proceedings of the 12th European Conf. on Artificial Intelligence (ECAI'96)*, pages 318–322, Budapest, 1996.

[59] D. Lehmann. Generalized qualitative probability : Savage revisited. In *Proc. of Int. Conf. on Uncertainty in Artificial Intelligence (UAI'96)*, pages 381–388, 1996.

[60] D. Lehmann. Expected qualitative utility maximization. *Games and Economic Behavior*, 35(1-2):54–79, 2001.

[61] D. Lewis. *Counterfactuals.* Basil Blackwell, London, 1973.

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

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

[64] J. Von Neumann and O. Morgenstern. *Theory of games and economic behavior.* Princeton University Press, 1947.

[65] P. Perny and A. Rolland. Reference-dependent qualitative models for decision making under uncertainty. In Gerhard Brewka, Silvia Coradeschi, Anna Perini, and Paolo Traverso, editors, *ECAI*, pages 422–426. IOS Press, 2006.

[66] R. Sabbadin. Possibilistic markov decision processes. *Engineering Appl. of Artificial Intelligence*, 14:287–300, 2001.

[67] L.J. Savage. *The Foundations of Statistics.* Wiley, New York, 1954.

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

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

[70] G.L.S.. Shackle. *Expectation in Economics.* Cambridge University Press, U.K, 1949.

[71] G.L.S.. Shackle. *Decision Order and Time In Human Affairs.* Cambridge University Press, U.K, 1961.

[72] P. Snow. Diverse confidence levels in a probabilistic semantics for conditional logics. *Artificial Intelligence*, 113(1–2):269–279, 1999.

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

[74] S. Tan and J. Pearl. Specification and evaluation of preferences under uncertainty. In Pietro Torasso Jon Doyle, Erik Sandewall, editor, *Proceedings of Proc. of the 4th International Conference on Principles of Knowledge Representation and Reasoning (KR'94), Bonn, Germany, USA*, pages 530–539, Bonn, FRG, May 1994. Morgan Kaufmann.

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

[76] Ph. Vincke. *Multicriteria Decision Aid.* Wiley, 1992.

[77] A. Wald. *Statistical Decision Funtions.* John Wiley, 1950.

[78] P. Weng. Qualitative decision making under possibilistic uncertainty: towards more discriminating criteria. In *Proc. of the 21st Conference in Uncertainty in Artificial Intelligence (UAI'05), Edinborough, UK*, pages 615–622, 2005.

[79] P. Weng. An axiomatic approach in qualitative decision theory with binary possibilistic utility. In *Proc. of the 17th European Conference on Artificial Intelligence (ECAI 2006), Riva del Garda, Italy*, pages 467–471. IOS Press, 2006.

[80] P. Weng. Axiomatic foundations for a class of generalized expected utility: Algebraic expected utility. In *Proc. of the 22nd Conference in Uncertainty in Artificial Intelligence, Cambridge, MA, USA*, pages 520–527, 2006.

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

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

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