Possibility Theory as a Basis for Qualitative Decision Theory

Didier Dubois and Henri Prade
Institut de Recherche en Informatique de Toulouse (I.R.I.T.) - C.N.R.S.
Universite Paul Sabatier - Bat. 1R3, 118 route de Narbonne
31062 Toulouse Cedex, France

Abstract

A counterpart to von Neumann and Morgenstern' expected utility theory is proposed in the framework of possibility theory. The existence of a utility function, representing a preference ordering among possibility distributions (on the consequences of decision-maker's actions) that satisfies a series of axioms pertaining to decision-maker's behavior, is established. The obtained utility is a generalization of Wald's criterion, which is recovered in case of total ignorance; when ignorance is only partial, the utility takes into account the fact that some situations are more plausible than others. Mathematically, the qualitative utility is nothing but the necessity measure of a fuzzy event in the sense of possibility theory (a so-called Sugeno integral). The possibilistic representation of uncertainty, which only requires a linearly ordered scale, is qualitative in nature. Only max, min and order-reversing operations are used on the scale. The axioms express a risk-averse behavior of the decision maker and correspond to a pessimistic view of what may happen. The proposed qualitative utility function is currently used in flexible constraint satisfaction problems under incomplete information. It can also be used in association with possibilistic logic, which is tailored to reasoning under incomplete states of knowledge.

1 Introduction

Standard approaches to decision under uncertainty are based on maximum expected utility theory. The expected utility criterion is particularly appealing since it can be justified on the basis of an axiomatic approach [von Neumann & Morgenstern, 1944]. However, its application requires that both numerical probabilities and utilities about the consequences of actions are available. The representation of incomplete states of knowledge has led Artificial Intelligence to introduce non-probabilistic models ot" uncertainty such as Shafer's theory of evidence, possibility theory, and nonmonotonic logics. These approaches seem particularly suitable for the representation of states of partial ignorance in an unbiased way; see [Dubois et al., 1994c] for a discussion. On the basis of a careful distinction between reasoning tasks (where consequences of the actual state of information are only propagated) and decision tasks (where choices are elaborated taking into account both uncertainty and preferences), it can be advocated that the probabilistic approach can still be used at the decision level even if another framework is used for knowledge representation and reasoning purposes; thus Smets [1990] has proposed a so-called "pignistic" transformation for computing meaningful probabilities (in a decision-theoretic perspective) from belief functions representing the available information.

Possibility theory provides a faithful representation of partial ignorance, but its core is also qualitative in nature since it only requires a scale where max, min and order-reversing operations can be defined. This qualitative nature agrees with the fact that only poor and incomplete information is available in many practical situations. In order to cope with such situations, several proposals for a qualitative decision theory have been recently presented by Pearl [1993], Tan & Pearl [1994a,b] on the one hand and by Boutilier [1994] on the other hand. Both approaches are connected with default reasoning for the handling of uncertain pieces of knowledge; the former, which relates to Spohn [1988] ordinal conditional "kappa" functions, requires the use of scales where the sum or the product are meaningful, while the latter somewhat gets rid of uncertainty by considering the most plausible states of the world only, when making a decision. The possibilistic approach which is presented in this paper is different in various respects. First, possibilistic utility obeys a series of axioms which may be regarded as a qualitative counterpart to von Neumann & Morgenstern [1944] axioms. Second, uncertainty and preferences are both estimated on ordinal scales where the only meaningful operations are max, min and the reversing of the ordering. Third, the proposed decision theory is closely associated with an approach to the modelling of uncertainty, here the theory of possibility introduced by Zadeh [1978]. Fourth, a kind of commensurateness assumption between possibility levels and preference levels is made, such that a decision rates all the better as it makes undesirable states of affairs less possible. This assumption leads to a framework where decisions under incomplete information can be completely ordered, although without resorting to numerical representation.

The next section presents the axioms proposed as a basis for qualitative utility when uncertainty is modelled by means of possibility distributions. Section 3 establishes the existence of a qualitative utility function and explains how to compute it. Section 4 provides a brief discussion of the qualitative expected utility which has been obtained. Section 5 illustrates the usefulness of the qualitative utility function in order to estimate the degree of satisfaction of a flexible constraint under uncertainty. Section 6 discusses how uncertainty and preference can be jointly handled using possibilistic logic and possibilistic utility theory.

2 Axioms for a Qualitative Utility Theory

Let X be a set of situations (states of the world), supposedly finite. After Savage [1974], an act is a function f from X to C, the set of possible consequences of the act. This function is attached to a particular decision and specifies what is the consequence of being in situation x when the decision is made; f(x) is sometimes interpreted as the expected pay-off of the act, when $x \in X$ is the situation. It is supposed that there exists a linear (i.e., complete partial) ordering $\geq$ over C expressing preference. Thus, a given act induces a preference over X, whereby $x \geq y$ iff $f(x) \geq f(y)$. This preference relation between precisely-known situations will be extended to incomplete belief states pervaded with uncertainty by means of a relation obeying a series of axioms given in the following.

Indeed, we may have incomplete information about the actual situation in X. The *belief state* about which situation in X is the actual one is supposed to be represented by a possibility distribution π. A possibility distribution π defined on X takes its values on a valuation scale V, where V is supposed to be linearly ordered. V is assumed to be bounded and we take $\sup V=1$ and $\inf V=0$. The inequality $\pi(x) \leq \pi(x')$ means that x' is at least as plausible (normal) as x as being the actual situation. $\pi(x)=0$ means that x is impossible, i.e., definitely excluded by π as being the true situation. $\pi(x)=1$ means that x is normal, unsurprizing. We restrict ourselves to *consistent* belief states π which are such that $\exists x \in X$, $\pi(x)=1$ (i.e., there exists at least one situation which is completely possible); how this closely relates to the notion of logical consistency can be fully understood in the framework of possibilistic logic (e.g., [Dubois et al., 1994b, c]). The state of total ignorance $\pi_?$ is represented by $\forall x \in X$, $\pi_?(x)=1$ where any situation is found totally possible. We shall denote by $Pi(X) = \{\pi, X \rightarrow V\}$ the set of consistent possibility distributions over X. The cardinality of V is supposed to be at least as large as the one of X, so as to account for total plausibility orderings on X. Clearly, $X \subseteq Pi(X)$ using the identification function $x_0 \rightarrow \mu_{\{x_0\}}$ where $\pi(x)=\mu_{\{x_0\}}(x)=1$ if $x=x_0$ and $\mu_{\{x_0\}}(x)=0$ otherwise. $\mu_{\{x_0\}}$ is called a precise belief state and we shall write, $\pi=\{x_0\}$, or even $\pi=x_0$ for the sake of notational simplicity. More generally, if π is the characteristic function of a subset A of X, we write π=A. For instance $\pi_? = X$.

A possibility distribution representing a belief state involves a set of mutually exclusive alternatives, where each element can be ranked according to its level of plausibility to be the true situation. Let x and y be two elements of X, the possibility distribution π defined by $\pi(x)=\lambda$, $\pi(y)=\mu$, $\pi(z)=0$ for $z \neq x$, $z \neq y$ with $\max(\lambda,\mu)=1$ (in order to have π normalized), will be called a *qualitative lottery* and will be denoted by $(\lambda/x, \mu/y)$, which means that we are either in situation x or in situation y with the respective levels of possibility λ and μ. More generally, any possibility distribution π can be viewed as a multiple-consequence lottery $(\lambda_1/x_1,\ldots,\lambda_n/x_n)$ where $X=\{x_1,\ldots,x_n\}$ and $\lambda_i=\pi(x_i)$. We will also use the notation $(\lambda/\pi_1, \mu/\pi_2)$ (with $\max(\lambda,\mu)=1$) for denoting the compound possibility distribution $\pi=\max(\min(\pi_1,\lambda),\min(\pi_2,\mu))$. This can be viewed as a lottery over multiple-consequence lotteries corresponding to π_1 and π_2. The lottery $(\lambda/x, \mu/y)$ can be viewed as a particular case of it when π_1 and π_2 are possibility distributions focusing on singletons. The resulting possibility distribution $\pi=\max(\min(\pi_1,\lambda),\min(\pi_2,\mu))$, with $\max(\lambda,\mu)=1$, is here the qualitative counterpart of probabilistic mixtures $\lambda p_1+(1-\lambda)p_2$; see [Dubois & Prade, 1990], [Dubois et al., 1993].

Lastly, the decision maker (DM) is supposed to express a preference between (consistent) belief states. This preference relation, denoted by $\succeq$, should be understood in the following way: $\pi \succeq \pi'$ means that DM expects at least as much in terms of pay-off when he believes π than when be believes π'. We suppose that

Axiom 1: $\succeq$ is a complete partial ordering.

$\succ$ shall denote the strict relation associated with $\succeq$, and $\pi \sim \pi'$ will mean that $\pi \succeq \pi'$ and $\pi' \succeq \pi$. A *qualitative utility function* is a mapping u from Pi(X) to a linearly ordered scale U. Thus, $u(\pi)$ will denote the ordinal evaluation of an act in the state of belief π. We shall simply write u(x) (instead of $u(\{x\})$, when $\pi=\{x\}$, for denoting the utility of the considered act when the situation is precisely known as being x. We also write u(A) instead of $u(\pi)$ with $\pi=\mu_A$ (characteristic function of subset A). As usual, a (qualitative) utility function u will be said to represent a preference ordering $\succeq$ if and only if $u(\pi) \geq u(\pi') \Leftrightarrow \pi \succeq \pi'$.

Together with Axiom 1, we propose the following axioms as a basis of our decision theory.

Axiom 2 (certainty equivalence):

If the belief state is a crisp set $A \subseteq X$,
then there is $x \in A$ such that $\{x\} \sim A$.

The intuition behind Axiom 2 is the following one. The decision-maker DM only knows that one of the situations in A is the true situation. Hence, the utility of the act should lie in the set $\{u(x), x \in A\}$, where u(x) is the utility of the act in situation x. The choice of the state x such that $u(x)=u(A)$ reflects DM's attitude towards risk.

Axiom 3 (risk aversion, or "precision is safer"):

$$\pi \leq \pi' \Rightarrow \pi \succeq \pi'.$$

The motivation underlying this axiom is that DM prefers the belief state which are more precise. Indeed if $\pi \leq \pi'$ (i.e., $\forall x \in X$, $\pi(x) \leq \pi'(x)$), it means that belief state π' is less informed than belief state π (since according to π' any situation x is found to be at least as much possible as according to π). Thus, π' is viewed as a more risky belief state than π, since the worst situation pertaining to π' is not less possible than the worst situation pertaining to π. Behaving in this way DM is cautious. This is a risk-averse attitude since, for all x in A $\{x\} \succeq A$, i.e., in terms of a utility function $u(x) \geq u(A)$. This means that DM prefers to be in situation x for sure, than only knowing that he is in A and having the possibility of receiving less, in case $\exists y \in A$ and $u(y)<u(x)$. Total ignorance is always the worst situation. At this point, it should be clear that the utility function we look for is in accordance with Wald maximin criterion.

Axiom 4 ("independence"):

If $\pi_1 \sim \pi_2$ then $(\lambda/\pi_1, \mu/\pi') \sim (\lambda/\pi_2, \mu/\pi')$.

It is assumed that $\max(\lambda,\mu)=1$ in order to ensure that the new possibility distributions built from π_1, π_2, π' are still normalized. Axiom 4 means that if two belief states are judged to be equivalent from a pay-off point of view, this equivalence still holds when we substitute one of these belief states by the other in compound lotteries, thus expressing a form of independence. Indeed this axiom is a possibilistic counterpart of the so-called linearity or independence axiom in von Neumann & Morgenstern' utility theory.

Axiom 5 (reduction of lotteries):

$(\lambda/x, \mu/(\alpha/x, \beta/y)) \sim (\max(\lambda,\min(\mu,\alpha))/x, \min(\mu,\beta)/y)$.

This axiom reduces meta lotteries to ordinary ones. Again it is assumed that $\max(\lambda,\mu)=1=\max(\alpha,\beta)$ in order to guarantee normalization preservation. This axiom, which is the counterpart of von Neumann & Morgenstern' substitutability axiom, can be better understood by visualizing what it expresses in terms of qualitative lotteries, as shown on Figure 1. Namely, if two steps are needed to reach situation y, one which succeeds with possibility μ, and the next one that succeeds with possibility β, then the possibility of reaching situation y is not higher than $\min(\mu,\beta)$. However, if alternative paths exist, leading to x, with respective possibility levels λ and μ, then the possibility of reaching x is at least $\max(\lambda,\mu)$.

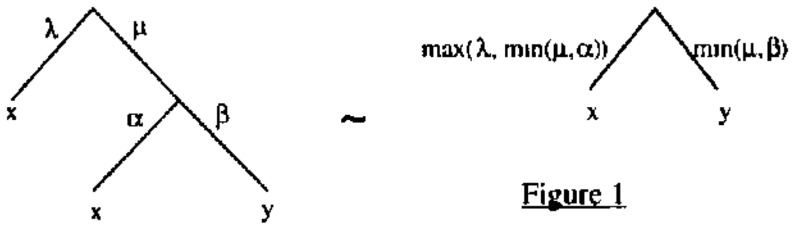


Figure 1

Axiom 6 (continuity):

If $\pi \succeq \pi'$ then $\exists \lambda \in V$, $\pi' \sim (1/\pi, \lambda/X)$.

This axiom is reminding of continuity axioms in classical utility theory. $(1/\pi,\lambda/X)$ is a meta-lottery whereby it is slightly possible to be in a state of total ignorance instead of in belief state π. Clearly, the lottery $(1/\pi,\lambda/X)$ corresponds to the possibility distribution $\pi''=\max(\pi,\lambda)$ and is a state of belief less informed than π. When $\lambda=1$ $\pi''=\pi_?$ and we know by Axiom 3 that $u(\pi_?)\leq u(\pi)$, $\forall \pi \in Pi(X)$. Moving λ in V from 0 to 1, when $\pi \succeq \pi'$, leads to decrease the preference down to a possibility distribution which is not preferred to π'. The idea behind Axiom 6 is that π' is hit for some value of λ.

3 Form of the Qualitative Utility Function

Theorem: Given a preference relation $\succeq$ on Pi(X) verifying Axioms 1 to 6, there exists a fuzzy set F on X and a utility function u from Pi(X) to a totally ordered set U representing $\succeq$ such that for each $\pi \in Pi(X)$, we have

$$u(\pi)=\min_{x\in X} \max(n(\pi(x)), \mu_F(x)) \qquad (1)$$

where n is an order reversing function from the possibility scale V to the preference scale U such that $n(0)=1$ and $n(1)=0$ where 1 denotes the top elements of U and V and 0 their bottom elements.

Note that (1) yields $u(x)=\mu_F(x)$.

Proof:

I) u defined by (1) satisfies Axioms 1 to 6.

This is obvious for Axioms 1, 3 and 4

Axiom 2: If $\pi=\mu_A$, $\forall x\in A$, $n(\pi(x))=0$ and $\forall x \notin A$, $n(\pi(x))=1$. Thus $u(\pi)=\min_{x\in A} u(x)$ and then $\exists x\in A$, $u(\pi)=u(x)$. Note that $\pi_?$ is such that $u(\pi_?)=\min_{\pi\in Pi(X)} u(\pi)=\inf_{x\in X} u(x)$.

Axiom 5: Note that

$u((\lambda/x, \mu/y)) = \min(\max(u(x),n(\lambda)),\max(u(y),n(\mu)))$

and that

$$\begin{aligned}
u((\lambda/\pi_1, \mu/\pi_2)) &= u(\max(\min(\pi_1,\lambda),\min(\pi_2,\mu)))\\
&= \min_x \max(n(\max(\min(\pi_1(x),\lambda),\min(\pi_2(x),\mu))),u(x))\\
&= \min_x \max[\min(\max(n(\pi_1(x)),n(\lambda)),\\
&\qquad \max(n(\pi_2(x)),n(\mu))),u(x)]\\
&= \min_x \min(\max(n(\pi_1(x)),u(x),n(\lambda)),\\
&\qquad \max(n(\pi_2(x)),u(x),n(\mu)))\\
&= \min(\max(u(\pi_1),n(\lambda)), \max(u(\pi_2),n(\mu))). \qquad (2)
\end{aligned}$$

Then, it can be easily checked that

$$\begin{aligned}
u((\lambda/x, \mu/(\alpha/x, \beta/y))&\\
&= \min(\max(u(x),n(\lambda)),\max(u((\alpha/x, \beta/y)),n(\mu)))\\
&= \min(\max(u(x),n(\lambda)),\max(\min(\max(u(x),n(\alpha)),\\
&\qquad \max(u(y),n(\beta)),n(\mu)))\\
&= \min(\max(u(x),\min(n(\lambda),\max(n(\mu),n(\alpha)))),\\
&\qquad \max(u(y),n(\mu),n(\beta)))\\
&= \min(\max(u(x), n(\max(\lambda,\min(\mu,\alpha))),\\
&\qquad \max(u(y),n(\min(\mu,\beta))))\\
&= u((\max(\lambda,\min(\mu,\alpha))/x,\min(\mu,\beta)/y).
\end{aligned}$$

Axiom 6: $u(\pi)\geq u(\pi')$. Let $\lambda=n^{-1}(u(\pi'))$. Then

$u(1/\pi,\lambda/X) = \min(u(\pi),\max(u(\pi'),u(\pi_?)) = u(\pi')$.

Remark: Note that we have the following consequence: $\pi\sim\pi' \Rightarrow \pi\sim\max(\pi,\pi')$.

It means that if two belief states are considered as equally risky, it is indifferent for the DM to be in any of these belief states. For instance, it $\{x\}\sim\{x'\}$, then the DM does not wish to know whether the situation is x or x' since his pay-off will be the same in both cases; hence we shall have $u(\{x,x'\})=u(x)=u(x')$.

Indeed the form (2) of the utility function leads to: $u(\max(\pi,\pi'))=\min(u(\pi),u(\pi'))=u(\pi)=u(\pi')$.

II) Existence of a utility function u of the form (1), given Pi(X) equipped with $\succeq$.

First, consider the restriction of $(Pi(X), \succeq)$ to $(X,\succeq)$. Let us map $(X,\succeq)$ to an ordinal scale U such that inf(X) is mapped to 0 and sup(X) is mapped to 1.

Let u be a function from X to U such that $u(x)\geq u(y) \Leftrightarrow x\succeq y$. The proof then goes as follows:

a) extension of the utility function from X to its power set 2^X. This enables us to compare states of incomplete information;
b) definition of the utility of the elementary lottery (1/supX, λ/infX) (or equivalently (1/1, λ/0) since by definition u(supX)=1 and u(infX)=0), for any $\lambda\in V$;
c) computation of the utility of the qualitative lottery (1/x, λ/y) for any $x, y\in X$;
d) prove that $u((1/\pi, 1/\pi'))=\min(u(\pi),u(\pi'))$;
e) computation of $u(\pi)$ for any state π of incomplete information.

a) Function u can be extended to subsets A of X in the following way $u(A)=\min_{x\in A} u(x)$. Indeed, if $x_A\sim A$, where $x_A\in A$, is the certainty equivalent of A (according to Axiom 2), then $\forall x\in A$, $x\succeq x_A$. This can be easily shown in the following way. Assume $\exists x\in A$, $x_A\succ x$; then by Axiom 3, $x_A\succ x\succeq A$ (since $\mu_{\{x\}}\leq\mu_A$), hence

$x_A \succ A$ which contradicts the hypothesis. Hence $u(A)=u(x_A)=\min_{x\in A} u(x)$.

Note that A denotes a state of information where the consequence of the act under evaluation is only known to lie in A. u(A) is thus the pessimistic utility criterion proposed by Wald when the probability of outcomes in A are not available.

b) Let $\underline{x}=\inf X$ such that $u(\underline{x})=0$ and $\bar{x}=\sup X$, such that $u(\bar{x})=1$. Consider $\lambda \in V$ a level of possibility, and the lottery $(1/\bar{x}, \lambda/\underline{x})$. Let n be a bijection $V \to U$ such that $\forall u_1,u_2 \in V,\ u_1<u_2 \Rightarrow n(u_1)>n(u_2),\ n(0)=1,\ n(1)=0$. Then let us define $u(1/\bar{x}, \lambda/\underline{x})=n(\lambda)$. This definition agrees with any preference ordering $\succeq$ on Pi(X). Indeed

if $\lambda=0$ then $(1/\bar{x}, 0/\underline{x}) \sim \bar{x}$, and $n(0)=1=u(\bar{x})$

if $\lambda=1$ then $(1/\bar{x}, 1/\underline{x}) \sim \{\bar{x},\underline{x}\} \sim \underline{x}$ (as shown in a)) and $n(1)=0=u(\underline{x})$.

Moreover Axiom 3 enforces $(1/\bar{x}, \lambda/\underline{x}) \succeq (1/\bar{x}, \lambda'/\underline{x})$ whenever $\lambda \le \lambda'$ (since the lottery $(1/\bar{x},\lambda/\underline{x})$ is more informed than the other), and we do have that $u(1/\bar{x}, \lambda/\underline{x}) \ge u(1/\bar{x}, \lambda'/\underline{x})$. Hence the definition of the qualitative utility is in agreement with the axioms. It is clearly a special case of (1) for $\pi(x)=1$ if $x=\bar{x}$, λ if $x=\underline{x}$, and 0 otherwise, and $\mu_F(x)=u(x)$.

c) Now consider the lottery $(1/x, \lambda/y)$ for $x, y \in X$.

• Assume first that $u(x) \le u(y)$. Let us prove that $u(1/x, \lambda/y)=u(x)$. Indeed

$$\begin{aligned} (1/x, \lambda/y) &\succeq \{x,y\} && \text{(Axiom 3)}\\ \{x,y\} &\sim x && \text{(Axiom 2)}\\ x &\succeq (1/x, \lambda/y) && \text{(Axiom 3)} \end{aligned}$$

hence $(1/x, \lambda/y) \sim x$ and they have the same utility values.

• Assume now that $u(x) \ge u(y)$. First note that $y \sim (1/\bar{x}, n^{-1}(u(y))/\underline{x})$. Indeed $u(y)=n \circ n^{-1}(u(y))$ by definition. If $x=\bar{x}$, then

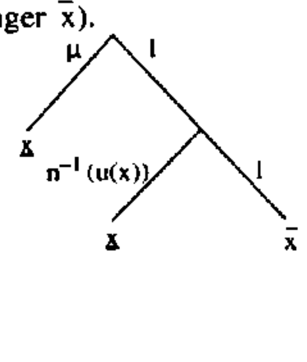
$$\begin{aligned} (1/\bar{x}, \lambda/y) &\sim (1/\bar{x}, \lambda/(1/\bar{x}, n^{-1}(u(y))/\underline{x})) && \text{(independence Axiom 4)}\\ &\sim (1/\bar{x}, \max(n^{-1}(u(y)),\lambda)/\underline{x}) && \text{(Axiom 5).} \end{aligned}$$

Hence
$$\begin{aligned} u(1/\bar{x}, \lambda/y) &= n(\max(n^{-1}(u(y)),\lambda))\\ &= \min(u(y),n(\lambda)). \end{aligned}$$

• In the general case, if $x \ge y$ then $y \sim (1/x, \mu/X)$ (Axiom 6) for some $\mu \in V$ and since X contains $\underline{x}$ (at worst $\underline{x}=y$), $X \sim \underline{x}$. Hence, using the independence Axiom 4, $y \sim (1/x, \mu/\underline{x})$. Now we note that $x \sim (1/\bar{x}, n^{-1}(u(x))/\underline{x})$, and we can substitute:

$$\begin{aligned} y &\sim (1/(1/\bar{x}), n^{-1}(u(x))/\underline{x}), \mu/\underline{x})\\ &\sim (1/\bar{x}, \max(n^{-1}(u(x)), \mu), \underline{x}) \text{ using Axiom 5 on the lottery below} \end{aligned}$$

Hence $u(y)=\min(u(x),n(\mu))=n(\mu)$ since $x \ge y$. Hence $\mu=n^{-1}(u(y))$. So $y \sim (1/x, n^{-1}(u(y))/\underline{x})$ (note the presence of x, and no longer $\bar{x}$).



Now
$$\begin{aligned} (1/x, \lambda/y) &\sim (1/x, \lambda/(1/x, n^{-1}(u(y))/\underline{x})) && \text{(Axiom 4)}\\ &\sim (1/x, \min(\lambda, n^{-1}(u(y))/\underline{x}) && \text{(Axiom 5)}\\ &\sim ((1/\bar{x}, n^{-1}(u(x))/\underline{x}),\ \min(\lambda, n^{-1}(u(y))/\underline{x}) && \text{(Axiom 4)}\\ &\sim (\bar{x}, \max(n^{-1}(u(x)),\ \min(\lambda, n^{-1}(u(y))/\underline{x}) && \text{(Axiom 5)} \end{aligned}$$

Hence $u(1/x, \lambda/y)=\min(u(x),\max(n(\lambda),u(y)))$ when $x \ge y$. Note that if $x \le y$, the above expression yields $u(x)$; it is thus valid regardless of the preference on $\{x,y\}$.

d) Consider two possibility distributions π, π' with $\pi \ge \pi'$. Then from Axiom 6, $\exists \lambda \in V$, $\pi' \sim (1/\pi,\lambda/X)$, hence

$$\begin{aligned} (1/\pi, 1/\pi') &\sim (1/\pi, 1/(1/\pi, \lambda/X)) && \text{(Axiom 4)}\\ &\sim (1/\pi, \lambda/X) \sim \pi'. && \text{(Axiom 5)} \end{aligned}$$

Hence $u(\max(\pi,\pi'))=u(\pi')$ when $u(\pi') \le u(\pi)$, noticing that the lottery $(1/\pi, 1/\pi')$ encodes the possibility distribution $\max(\pi,\pi')$.

e) Lastly we can compute $u(\pi)$ for any $\pi \in Pi(X)$. Let $\pi \in Pi(X)$, and assume that $\pi(x_1)=1 \ge \pi(x_2) \ge \ldots \ge \pi(x_n)$. Let $\pi_i \in Pi(X)$ be such that $\pi_i(x_1)=1$, $\pi_i(x_i)=\pi(x_i)$, and $\pi_i(x)=0$ *if* $x \notin \{x_1,x_i\}$. *Then* $\pi \sim (1/\pi_1, 1/\pi_2, \ldots, 1/\pi_n)$, since $\pi=\max_{i=1,n} \pi_i$. Hence $u(\pi)=\min_{i=1,n} u(\pi_i)$ from d). Now $u(\pi_i)=\min(u(x_1),\max(n(\pi(x_i)),u(x_i)))$ from c). Hence

$$\begin{aligned} u(\pi) &= \min_{i=1,n} \min(u(x_1),\max(n(\pi(x_1)),u(x_1)))\\ &= \min_{i=1,n} \max(n(\pi(x_i)),u(x_i)) \text{ noticing that } \pi(x_1)=1. \end{aligned}$$

The fuzzy set F is such that $\mu_F(x)=u(x)$, $\forall x \in X$.

4 Properties of the Qualitative Expected Utility

Interestingly enough the qualitative utility introduced in the previous section, $u(\pi)=\min_{x\in X} \max(n(\pi(x)),u(x))$ is the necessity of a fuzzy event [Dubois & Prade, 1980] in the sense of possibility theory, namely $u(\pi)=N_\pi(F)$ where F is the fuzzy set of preferred situations ($\mu_F(x)=u(x)$, $\forall x \in X$) and N_π is the necessity measure based on the possibility distribution π. Usually, when $V=U=[0,1]$, $n(t)=1-t$ in the above expression. $N_\pi(F)$ can be viewed as a degree of inclusion of the fuzzy set of more or less possible situations in the fuzzy set F of preferred outcomes, i.e., it estimates the certainty that the belief state π corresponds to the preferred situations described by F. As already said, there is a commensurability assumption made between the uncertainty scale and the preference scale, since possibility degrees and utility degrees are aggregated in the expression of $u(\pi)$. Note that

• $N_\pi(F)=u(\pi)=1$ iff $\{x\in X, \pi(x)>0\} \subseteq \{x \in X, u(x)=1\}$

i.e., the utility of π is maximum if all the more or less possible situations encompassed by π are among the most preferred ones.

• $N_\pi(F)=u(\pi)=0$ iff $\{x\in X, \pi(x)=1\} \cap \{x \in X, u(x)=0\} \neq \emptyset$

i.e., the utility of π is minimum if there is one of the most plausible situations whose pay-off is minimum (we recognize the risk-aversion of the approach). $N_\pi(F)=u(\pi)$ is all the greater as there is no situations with a high plausibility and low utility value.

When π is the characteristic function of an ordinary subset A of X, i.e., when all the situations encompassed by the belief state are equally plausible, as already said the utility $u(\pi)$ simplifies into $u(\pi)=\min_{x\in A} u(x)$ where we recognize Wald [1950]'s pessimistic criterion which leads to decisions maximizing the minimal pay-off. In the general case, $u(\pi)$ takes into account the fact that all the situations arc not equally plausible in the set $\{x\in X, \pi(x)>0\}$

Several authors have proposed definitions of utility functions in the presence of possibilistic uncertainty, including the form described in the theorem. Yager [1979] has introduced the possibilistic extension of the optimistic maximax criterion of the form dual to (1), i.e., $u(\pi)=\max_{x\in X} \min(\pi(x),\mu_F(x))$ which is the degree of possibility of a fuzzy set [Zadeh, 1978]. The possibilistic counterpart of Wald maximin criterion of the form proposed here, has been introduced by Whalen [1984], in terms of "disutility" function $D(\pi)=n^{-1}(u(\pi))$ where $u(\pi)$ is given by (1). $D(\pi)$ takes the form of the degree of possibility of the fuzzy set F (the fuzzy complement of F) of less preferred situations.

As already pointed out (e.g., [Inuiguchi et al., 1989]), the expression of the necessity of a fuzzy event is a particular case of a fuzzy integral in the sense of Sugeno [1974]. Namely $N_\pi(F)$ can be shown to be equal to (for V= U=[0,1])

$$N_\pi(F)=\sup_{\alpha\in(0,1]} \min(\alpha,N_\pi(F_\alpha))$$

with $F_\alpha=\{x\in X, \mu_F(x)\geq\alpha\}$, which is a particular case of Sugeno integral

$$\int_X h(x)\circ g(\cdot)=\sup_{\alpha\in(0,1]} \min(\alpha,g(H_\alpha))$$

with $H_\alpha=\{x\in X, h(x)\geq\alpha\}$ and g is a set function monotonic with respect to set inclusion, such that $g(\emptyset)=0$ and $g(X)=1$ Sugeno integrals can be regarded as qualitative counterparts to Choquet integrals of the form $\int_0^1 g(H_\alpha)d\alpha$. Sugeno integrals in general have been recently considered by Hougaard & Keiding [1994] for the utility representation of preferences on the set of non-additive set functions.

The utility function advocated in this paper relies on the notion of possibilistic mixture (as it can be seen in particular in Axiom 5), the result of the possibilistic mixture of π_1 and π_2, with $\max(\alpha,\beta)=1$, being equal to $\pi=\max(\min(\pi_1,\alpha),\min(\pi_2,\beta))$. Namely if Π_i is a possibility measure such that for all events A, B, $\Pi_i(A\cup B)=\max(\Pi_i(A),\Pi_i(B))$, then $\max(\min(\Pi_1,\alpha),\min(\Pi_2,\beta))$ is again a possibility measure; see Dubois & Prade [1990]. This is a particular case of extended mixtures of decomposable measures (which are a family of set functions encompassing probability measures and necessity and possibility measures as particular cases), as studied in [Dubois et al., 1993] where application to utility theory is pointed out.

Lastly, it should be emphasized that the proposed approach to qualitative utility closely parallels von Neumann & Morgenstern' theory and that there is some similarity between the two sets of axioms underlying the two approaches. Clearly, there are other ways of "distorting" classical utility theory. For instance, one of the authors [Dubois, 1986] has relaxed the reduction of lotteries axiom in a purely probabilistic framework by using a special family of operators in place of multiplication for expressing a relaxation of probabilistic independence. We might also think of using the multiplication instead of the min operation when simplifying numerical possibilistic lotteries, i.e., using mixtures of the form $\max(\alpha\cdot\pi_1, \beta\cdot\pi_2)$ with $\max(\alpha,\beta)=1$. This would lead to a utility function of the form $u(\pi)=\min_{x\in X} [1-(1-u(x))\cdot\pi(x)]$.

5 Handling Uncertainty in Flexible Constraint Satisfaction Problems

A constraint satisfaction problem is defined as a set $V=\{v_1,\ldots,v_n\}$ of decision variables, where D_i is the range of v_i, for all i, and a set of constraints $\{C_1,\ldots,C_m\}$. Each constraint C_j refers to a relation R_j that links a subset $V_j\subseteq V$ of variables, i.e., Rj is a subset of the Cartesian product $\times_{v_i\in V_j} D_i$, such that the restriction $(v_1,\ldots,v_n)\downarrow V_j$ of $(v_1,\ldots,v_n)$ to the variables in V_j should belong to R_j for all constraints $C_1,\ldots,C_m$. The constraint satisfaction problem consists in finding a feasible solution $(d_1,\ldots,d_n)$ to the set of constraints $\{C_1,\ldots,C_m\}$.

A flexible constraint is a constraint that can be violated to some extent. A flexible constraint can be modelled as a set of ordinary relations $R^1\ldots R^k$, such that $R^1\subseteq R^2\subseteq\ldots\subseteq R^k$. Namely R^2 is a relaxation of R^1, R^3 is a relaxation of R^2, etc. [Freuder, 1989]. An equivalent modelling consists in viewing a flexible constraint as a fuzzy relation $\mathcal{R}$, where $\mu_R(d_1,\ldots,d_n)$ is the degree of satisfaction of solution $(d_1,\ldots,d_n)$ [Zadeh, 1975]. The two views are equivalent if we attach satisfaction weights u_j, j=1,k to R^j such that $u_1=1\geq u_2\geq\ldots\geq u_k>0$, all weights belonging to a totally ordered scale U. Then define $\mu_R(d_1,\ldots,d_n)$ as follows $\mu_R(d_1,\ldots,d_n)=\max\{u_i, (d_1,\ldots,d_n)\in R^i\}$. Conversely $R^i=\{(d_1,\ldots,d_n),\mu_R(d_1,\ldots,d_n)\geq u_i\}$. It is possible to view the set $\{R^1,\ldots,R^k\}$ as k prioritized constraints, where R^j has a higher priority than R^i when j<i [Dubois et al., 1994a].

It often happens that uncertain parameters are involved in constraint satisfaction problems. For instance, in scheduling problems, the starting time of an activity is controlable and is a decision variable, but the ending time of the activity is partially unknown even if the starting time has been decided and is known, because the duration of the activity is partially uncontrollable. If there is a flexible constraint on the ending time of the activity, the degree of satisfaction of the decision consisting in the choice of a starting time is not precisely known. Let $z_1,\ldots,z_p$ be a set of uncertain parameters involved in a constraint satisfaction problem. A constraint C (a flexible one, generally) relates decision variables $v_1,\ldots,v_n$ *and* uncertain parameters $z_1,\ldots,z_p$. Let E_k the domain of z_k. Then the lack of knowledge on the parameter z_k is modelled by a possibility distribution π_k, and the possibility distribution π attached to $(z_1,\ldots,z_p)$ is $\pi=\min_{k=1,p} \pi_k$ [Dubois & Prade, 1988]. Let $z=(z_1,\ldots,z_p)$ and $d=(d_1,\ldots,d_n)$. The degree of satisfaction of

C by (d,z) is $\mu_R(d,z)$ and is ill-known, because so is z. A *robust* solution to the constraint C is d such that (d,z) satisfies C for all possible values of z.

In the non-flexible case, and assuming that the possible values of z form a subset A of $E_1 \times \ldots \times E_p$, v=d is a robust solution to constraint C if and only if $\forall z \in A$, $(d,z) \in R$ that is, $d \in R(v)=\{d,\{d\} \times A \subseteq R\}$.

In the general case let $A^1=\{z, \pi(z)=1\}$ be the set of normal values of z and $\underline{A}$ be the set of possible values of z (i.e., no value is possible outside $\underline{A}$). Similarly let R^1 be the set of preferred values of (v,z) and $\underline{R}$ be the set of feasible values of (v,z). Then a *completely robust solution* d is one that totally satisfies C for all possible values of Z, i.e.,

$$\forall z \in \underline{A},\ (d,z) \in R^1.$$

A *partially robust solution* d is one that partially satisfies C for all normal values of z, i.e.,

$$\forall z \in A^1,\ (d,z) \in \underline{R}.$$

The qualitative utility function enables a degree to which d is a robust solution to C to be computed as follows

$$\mu_C(d)=N_z(R(v))=\inf_z \max(n(\pi(z)),\mu_R(d,z)).$$

Accepting a solution such that $\mu_C(d) \geq \alpha$, where $\alpha \in U$, means that

- one accepts to assume that the actual value of z will lie in $A^{n(\alpha)}=\{z \mid \pi(z) \geq n(\alpha)\}$ where n is the order-reversing function that maps degrees of preference $\alpha \in U$ to degrees of possibility in V ;
- taking this assumption for granted, any solution such that $\mu_C(d) \geq \alpha$ satisfies for sure the constraint C to level α for any eventual value of z in $A^{n(\alpha)}$.

Again, the qualitative utility function tries to compromise between uncertainty and preference via a commensurateness assumption. A solution is all the better if it can cope with more implausible values of uncertain parameters.

Example: Tom wants to attend a meeting that starts at eight o'clock, and wants to decide when to get up so as to arrive on time. He has to take a bus, and the ride takes about 1 hour. He does not want to leave his home too early, say before 7 but not before 6.30. This is modelled by a fuzzy interval M such that $\underline{M}=[6.5, 7]$ and $M^1=\{7\}$. The constraint on the arrival time is that not more than 1/4 hour delay is acceptable, so that it is a fuzzy interval N with $N^1=[7, 8]$ and $\underline{N}=[7, 8.25]$. The uncertainty about the trip duration is modelled by a possibility distribution $\pi=\mu_A$ with $A^1=\{1\}$ and $\underline{A}=[0.75, 1.25]$ because if Tom is lucky (no wait, no traffic) the trip takes 3/4 hour, while if he is unlucky, it may take 1/2 hour more. The problem is to find a starting time s for the trip, such that s lies in M (constraint about getting up), and so that whatever the duration z of the trip the arrival time s+z is acceptable, i.e., lies in N. Applying the above qualitative utility theory with U=V=[0,1], and $n(\alpha)=1-\alpha$, it comes down to finding s which maximizes $\min(\mu_M(s), \inf_z \max(1-\mu_A(z),\mu_N(s+z))$ where μ_M, μ_A, μ_N are membership functions. To understand this expression, it is enough to see it as a multiple-valued evaluation of the sentence $\exists s$, $s \in M$ and $\forall z$ if $z \in A$ then $s+t \in N$. The maximization and the minimization are multiple-valued counterparts of the universal and the existential quantifiers respectively, and max(1−a,b) is a multiple-valued implication. The term $\inf_z (\max(1-\mu_A(z),\mu_N(s+z))$ is the qualitative utility of choosing s as a starting time decision, in a situation where the incomplete knowledge about z is defined by $\pi=\mu_A$, and where the preference level of decision s in situation z is $\mu_F(z)=\mu_N(s+z)$ (evaluating to what extent the arrival time constraint is violated). With linear membership functions it is found that Tom should leave his home at 6:52'30" am; then provided that the trip does not exceed 1:11'20", it ensures that Tom will not be more than 3'50" late. Note that the trip length estimate is rather safe. See Dubois et al. [1995] for a full treatment of this example, and the application of this approach to job-shop scheduling.

6 Decision with Generic Knowledge and Generic Preference Information

Belief states are often incompletely specified through the use of pieces of knowledge and information, some of which being uncertain. In other words, a possibility distribution over the possible situations is not always explicitly specified. Besides, preferences might also be incompletely specified. An instance of such problems is given by the following motivating example used in [Tan & Pearl, 1994b] and Boutilier [1994]. An agent is supposed i) to know that "if I have the umbrella, then I will be dry", $u \rightarrow d$; "if it rains and I do not have the umbrella, then I will be wet", $r \wedge \neg u \rightarrow \neg d$; and "typically if it is cloudy, it will rain", $c \rightarrow r$, (the latter rule is uncertain) ii) to observe that the sky is cloudy, iii) to prefer being dry rather being not dry $d \succ \neg d$ and carrying no umbrella $\neg u \succ u$. It is further assumed that "being dry is more important than not carrying an umbrella" $(d \succ \neg u)$. The problem is then for the agent to decide whether or not to take an umbrella.

Interpreting rules via material implication, knowledge expressed by (i) can be encoded in possibilistic logic [Dubois et al., 1994b] by assuming $N(\neg u \vee d)=1$, $N(\neg r \vee u \vee \neg d)=1$, $N(\neg c \vee r)=1-\lambda<1$ and $N(c)=1$, where N is a necessity measure. The least informative possibility distribution (on interpretations) obeying these constraints is such that the situations $r \wedge u \wedge d$ and $r \wedge \neg u \wedge \neg d$ have possibility 1, $\neg r \wedge u \wedge d$ and $\neg r \wedge \neg u \wedge d$ and $\neg r \wedge \neg u \wedge \neg d$ have a smaller possibility, say λ, with $0<\lambda<1$, while the other situations are virtually impossible. Note that the precise value of level λ remains unknown in our example. We have also to define the fuzzy sets of preferences F_u and $F_{\neg u}$, associated with the actions 'taking an umbrella' and 'not taking an umbrella', over the set of relevant situations. F_u will be a non-normalized fuzzy set $(\mu/r \wedge u \wedge d,\ \mu/\neg r \wedge u \wedge d)$ with $\mu<1$ caused by agent's reluctance to carry an umbrella. $F_{\neg u}$ will be the fuzzy set $(\alpha/r \wedge \neg u \wedge \neg d,\ \beta/\neg r \wedge \neg u \wedge d,\ \alpha/\neg r \wedge \neg u \wedge \neg d)$ with $\beta>\alpha$. Moreover, $\alpha<\mu$ since it is more important to be dry than not to carry an umbrella. It can be checked that

$$N_\pi(F_u)=\min(\mu,\max(\mu,n(\lambda))=\mu$$

and $N_\pi(F_{\neg u})=\min(\alpha,\max(n(\lambda),\beta),\max(n(\lambda),\alpha)=\alpha<N_\pi(F_u)$.

It leads to the decision: "take an umbrella". The value of λ does not influence the result since the important point in this very elementary example is that it is strictly more plausible that it rains rather than it does not rain, and our approach is risk-averse (that is, here, rain-averse). Note that in the above example, the rules are not conflicting. In the case of conflicting rules one may use the method developed by Benferhat et al. [1992] for encoding exception-prone rules in possibilistic logic.

7 Concluding Remarks

We have proposed a utility-based, axiomatically-grounded, decision theory which only requires ordinal scales for the assessment of uncertainty and preferences. The proposed approach opens the road to a genuine decision theory in the framework of possibility theory, a long term goal which was already at the basis of the work of the English economist Shackle [1961]. It can be applied in problems where the information is very rough and qualitative, including decision-theoretic planning. If we can afford scales with a somewhat richer structure, we may, for instance, as suggested at the end of Section 4, use a product-based rather than a min-based approach. Due to the equivalence between "max-product" possibility theory and Spohn [1988]'s ordinal conditional "kappa" functions up to a rescaling (see [Dubois & Prade, 1991]), it would lead to a decision-theoretic framework for kappa functions, whose axiomatization could be investigated both in the possibilistic setting and in the von Neumann-Morgenstern framework (with infinitesimal lotteries). This is a topic for further research.

References

[Benferhat et al., 19921 S. Benferhat. D. Dubois, and H. Prade. Representing default rules in possibilistic logic. In *Proc. 3rd Inter. Conf. on Principles of Knowledge Representation and Reasoning (KR'92),* pages 673-684, Cambridge, MA, 1992.

[Boutilier, 1994] C. Boutilier, Toward a logic for qualitative decision theory. In *Proc. 4th Inter. Conf. on Principles of Knowledge Representation and Reasoning (KR'94)* (J. Doyle et al., eds.), pages 75-86, Bonn, Germany, 1994.

[Dubois, 1986] D. Dubois. Generalized probabilistic independence and its implications for utility. *Operations Research Letters,* 5(5):255-260. 1986.

[Dubois & Prade 1980] D. Dubois, and H. Pradc. *Fuzzy Sets and Systems.* Academic Press, New York, 1980.

[Dubois & Prade, 1988] D. Dubois, and H. Prade. *Possibility Theory.* Plenum Press, New York, 1988.

[Dubois & Prade, 1990a] D. Dubois, and H. Prade. Aggregation of possibility measures. In *Multiperson Decision Making Using Fuzzy Sets and Possibility Theory* (J. Kacprzyk, and M. Fedrizzi, eds.), pages 55-63, Kluwer Academic Publ., Dordrecht, 1990.

[Dubois & Prade, 1991] D. Dubois, and H. Prade. Epistemic entrenchment and possibilistic logic. *Artificial Intelligence,* 50:223-239, 1991.

[Dubois et al., 1993] D. Dubois, J.C. Fodor, H. Prade, and M. Roubens. Aggregation of decomposable measures with application to utility theory. Tech. Report IRIT/93-55-R, IRIT, Univ. P. Sabatier, Toulouse, 1993. To appear in Theory and Decision.

[Dubois et al., 1994a] D. Dubois, H. Fargier, and H. Prade. Propagation and satisfaction of flexible constraints. In *Fuzzy Sets, Neural Networks, and Soft Computing* (R.R. Yager, and L.A. Zadeh, eds.), pages 166-187. Van Nostrand Reinhold, New York, 1994.

[Dubois et al., 1994b] D. Dubois, J. Lang, and H. Prade. Automated reasoning using possibilistic logic: Semantics, belief revision, and variable certainty weights. *IEEE Trans. on Knowledge and Data Engineering,* 6(l):64-69, 1994.

[Dubois et al., 1994c] D. Dubois, H. Prade, and Ph. Smets. Representing partial ignorance. In: Tech. Report IRIT/94-40-R, IRIT, Univ. P. Sabatier, Toulouse, France, 1994. Also in *IEEE Trans, on Systems, Man and Cybernetics,* to appear.

[Dubois et al., 1995] D. Dubois, H. Fargier. and H. Prade. Fuzzy constraints in job-shop scheduling. To appear in *J. Intelligent Manufacturing,* 1995.

[Freuder 1989] *B.C.* Freuder. Partial constraint satisfaction. In *Proc. llth Inter. Joint Conf. on Artificial Intelligence (IJCAI'89),* pages 278-283, Detroit, Michigan, 1989.

[Hougaard and Keiding, 1994] J.L. Hougaard, and H. Keiding. Representation of preferences on fuzzy measures by a fuzzy integral. Institute of Economics, University of Copenhagen, 1994.

[Inuiguchi et al., 1989] M. Inuiguchi, H. Ichihashi, and H. Tanaka. Possibilistic linear programming with measurable mulliattribute value functions. *ORSA J. on Computing,* 1(3):146-158, 1989.

[Pearl. 1993] J. Pearl. From qualitative utility to conditional "ought to". In: *Proc. 9th Inter. Conf on Uncertainty in Artificial Intelligence* (D. Heckerman, and H. Mamdani, eds.), pages 12-20, Morgan & Kaufmann, San Mateo, CA, 1993.

[Savage, 1972] LJ. Savage. *The Foundations of Statistics.* Dover. New York. 1972.

[Shackle, 1961] G.L.S. Shackle. *Decision, Order and Time in Human Affairs.* (2nd edition) Cambridge University Press, Cambridge, UK. 1961.

|Smets, 1990] Ph. Smets. Constructing the pignistic probability function in a context of uncertainty. In *Uncertainty in Artificial Intelligence, Vol. 5* (M. Henrion et al., eds.). pages 29-39, North-Holland, Amsterdam, 1990.

[Spohn, 1988] W. Spohn. Ordinal conditional functions: A dynamic theory of epistemic states. In *Causation in Decision, Belief Change, and Statistics* (W.L. Harper, and B. Skyrms, eds.), pages 105-134, Kluwer Academic Publ., 1988.

[Sugeno, 1974] M. Sugeno. Theory of fuzzy integrals and its applications. Doctoral thesis. Tokyo Institute of Technology, 1974.

[Tan & Pearl, 1994a] S.W. Tan, and J. Pearl. Specification and evaluation of preferences under uncertainty. In *Proc. 4th Inter. Conf on Principles of Knowledge Representation and Reasoning (KR'94)* (J. Doyle et al., eds.). pages 530-539, Bonn, Germany. 1994.

[Tan & Pearl, 1994b] S.W. Tan, and J. Pearl. Qualitative decision theory. In *Proc. llth National Conf on Artificial Intelligence (AAAI'94),* pages 928-933, Seattle, WA, 1994.

[von Neumann & Morgenstern, 1944] J. von Neumann, and O. Morgenstern. *Theory of Games and Economic Behavior.* Princeton University Press, Princeton, NJ, 1944.

[Wald, 1950] A. Wald. *Statistical Decision Functions.* Wiley & Sons, New York, 1950.

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

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

[Zadeh, 1975] L.A. Zadeh. Calculus of fuzzy restrictions. In *Fuzzy Sets and Their Applications to Cognitive and Decision Processes* (L.A. Zadeh et al., eds.), pages 1-39, Academic Press, New York. 1975.

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