POSSIBILITY THEORY

In common use, the word *possibility* conveys two meanings. One is physical and refers to the idea of feasibility. Then, *possible* means *achievable,* as in the sentence "it is possible *for* Hans to eat six eggs for breakfast." The other is epistemic and refers to the idea of plausibility. There, *possible* means logically *consistent* with (i.e., not contradicting) the available information, as in the sentence "it is possible *that* it will rain tomorrow." These two meanings correspond to the difference between realizability and plausibility and are mostly unrelated.

Moreover, the idea of attainment often occurs together with the idea of preference: considering mutually exclusive alternatives, the most feasible one (in some sense) is usually preferred. Then we may view preference as subjective feasibility (physical realizability corresponds clearly to objective feasibility). The epistemic understanding of *possible* as plausible that was just mentioned is subjective and goes along with the idea that something is possible insofar as it is not surprising (because consistent) with respect to what is known. An interpretation in terms of objective plausibility can be also encountered when possibility refers to (upper bounds of) frequencies, as we shall see.

Although many people consider, mainly by habit, that possibility is always a binary notion (things are possible or are not possible), it makes sense, both with the feasibility and the plausibility interpretations, to consider that possibility may be a matter of degree and that one thing may be estimated or perceived as being more possible than another.

Possibility theory was coined by L. A. Zadeh (1) in the late seventies as an approach to modeling flexible restrictions on the value of variables of interest constructed from linguistic pieces of information, described by fuzzy sets and representing the available knowledge. This approach offers a graded modeling of the idea of possibility. Physical possibility has been advocated by Zadeh (1) to justify the axiomatic rule of possibility measures, expressing that the possibility of A OR B should be equated to the maximum of the possibility of A and of the possibility of B (because the degree of ease of some action that produces A OR B is given by the easiest of two actions which produce A and B, respectively). However, the intended use of possibility theory as a nonclassical theory of uncertainty, different from probability theory, and its application to approximate reasoning advocated by Zadeh rather agrees with the epistemic interpretation. In fact, as we advocate, possibility theory can be especially useful for modeling plausibility and preference.

This brief introduction is organized in three main parts. The basic elements of the theory are given first. Then the relationships and differences with the other uncertainty frameworks are discussed, and lastly applications to approximate reasoning are presented.

FUNDAMENTALS OF POSSIBILITY THEORY

Possibility Distribution

When a fuzzy set is used to express an incomplete piece of information E about the value of a single-valued variable, the degree attached to a value expresses the level of possibility that this value is indeed *the value* of the variable. This is what happens if the available information is couched in words, for example, "Tom is young." Here the fuzzy set *young* represents the set of possible values of the variable x = "age of Tom." The fuzzy set E is then interpreted as a *possibility distribution* (1), which expresses the levels of plausibility of the possible values of the ill-known variable x. If, rather, we are stating a requirement under the form of a flexible constraint, for example, we are looking for a young person, then the possibility distribution would represent our preference profile on the age of the person to be recruited. If the only available knowledge about x is that "x lies in E" where $E \subseteq U$, then the possibility distribution of x is defined by Eq. (1)

$$\pi_x(u) = \mu_E(u), \forall\, u \in U \tag{1}$$

where E (with membership function μ_E) is interpreted as the fuzzy set of (more or less) possible values of x and where π_x ranges on [0,1]. More generally, the range of a possibility distribution can be any bounded linearly ordered scale (which may be discrete with a finite number of levels). Fuzzy sets, viewed as possibility distributions, act as flexible constraints on the values of variables referred to in natural language sentences. Equation (1) represents a statement of the form "x lies in E" or more informally "x is E." It does not mean, however, that possibility distributions are the same as membership functions. Equation (1) is an assignment statement because it means that given that the only available knowledge is "x lies in E," the degree of possibility that $x = u$ is evaluated by the degree of membership $\mu_E(u)$. Note that distinct values may simultaneously have a degree of possibility equal to 1. $\pi_x(u) = 0$ means that u is completely impossible as a value for x.

If two possibility distributions π_x and π'_x pertaining to the same variable x, are such that $\pi_x < \pi'_x$, π_x is said to be more *specific* than π'_x in the sense that no value u is considered less possible for x according to π'_x than to π_x. This concept of specificity whose importance has been first stressed by Yager (2) underlies the idea that any possibility distribution π_x is provisional in nature and likely to be improved by further information, when the available information is not complete. When $\pi_x < \pi'_x$, the information π'_x is redundant and can be dropped. Numerical measures of (non) specificity have been introduced (see Refs. 3 and 4). In possibility theory, specificity plays a role similar to entropy in probability theory.

When the available information stems from several sources that are considered reliable, the possibility distribution that accounts for it is the least specific possibility distribution to satisfy the set of constraints induced by the pieces of information given by the different sources. This is the principle of minimum specificity. Particularly, it means that given a statement "x is E," then any possibility distribution π such that $\forall u$, $\pi(u) \le \mu_E(u)$ is in accordance with "x is E." However, choosing a particular π, such that $\forall u$, $\pi(u) \le \mu_E(u)$, to represent our knowledge about x, would be arbitrarily too precise. Hence Eq. (1) is naturally adopted if "x is E" is the only available knowledge, and already embodies the principle of minimum specificity.

Possibility and Necessity Measures

The extent to which the information "x is E" is consistent with a statement like "the value of X is in subset A" is estimated by means of the possibility measure Π, defined from $\forall u$, $\pi_x(u) = \mu_E(u)$, by (1)

$$\Pi(A) = \sup_{u \in A} \pi_x(u) \tag{2}$$

where A is a classical subset of U. The value of $\Pi(A)$ corresponds to the element(s) of A having the greatest possibility degree according to π_x. In the finite case, 'sup' can be changed into 'max' in Eq. (2). $\Pi(A) = 0$ means that $x \in A$ is impossible knowing that "x is E." $\Pi(A)$ estimates the *consistency* of the statement '$x \in A$' with what we know about the possible values of x, as emphasized in Refs. 5 and 7. It corresponds to the epistemic view of possibility. Indeed, if π_x models a nonfuzzy piece of incomplete information represented by an ordinary subset E, Eq. (2) reduces to

$$\begin{aligned} \Pi_E(A) &= 1 \text{ if } A \cap E \neq \varnothing \quad (x \in A \text{ and } x \in E \text{ are consistent}) \\ &= 0 \text{ otherwise} \quad (A \text{ and } E \text{ are mutually exclusive}) \end{aligned} \tag{3}$$

Any possibility measure Π satisfies the following max-decomposability characteristic property

$$\Pi(A \cup B) = \max[\Pi(A), \Pi(B)] \tag{4}$$

When U is not finite, the axiom in Eq. (4) is replaced by $\Pi(\cup_{i \in I} A_i) = \sup_{i \in I} \Pi(A_i)$ for any index set I.

Among the features of possibility measures that contrast with probability measures, let us point out the weak relationship between the possibility of an event A and that of its complement $\overline{A}$ ('not A'). Either A or $\overline{A}$ must be possible, that is, $\max[\Pi(A), \Pi(\overline{A})] = 1$ due to $A \cup \overline{A} = U$ and $\Pi(U) = 1$ (normalization of Π). In the case of total ignorance, both A and $\overline{A}$ are fully possible: $\Pi(A) = 1 - \Pi(\overline{A})$. Note that this leads to a representation of ignorance ($E = U$ and $\forall\, A \neq \emptyset$, $\Pi_E(A) = 1$) which presupposes nothing about the number of elements in the reference set U (elementary events), whereas the latter aspect plays a crucial role in probabilistic modeling. The case when $\Pi(A) = 1$, $\Pi(\overline{A}) > 0$ corresponds to partial ignorance about A. Besides, $\Pi(\varnothing) = 0$ is a natural convention since $\Pi(A) = \Pi(A \cup \varnothing) = \max[\Pi(A), \Pi(\varnothing)]$ entails $\forall A$, $\Pi(A) \ge \Pi(\varnothing)$. Note that we have *only* $\Pi(A \cap B) \le \min[\Pi(A), \Pi(B)]$. It agrees with the fact that in the case of total ignorance about A, for $B = \overline{A}$, $\Pi(A \cap B) = 0$ whereas $\Pi(A) = \Pi(\overline{A}) = 1$.

The weak relationship between $\Pi(A)$ and $\Pi(\overline{A})$ forces us to consider both quantities to describe uncertainty about the occurrence of A. $\Pi(\overline{A})$ tells us about the possibility of 'not A', hence about the *certainty* (or necessity) of the occurrence of A since, when 'not A' is impossible, then A is certain. Thus it is natural to use this duality and define the degree of necessity of A [Dubois and Prade (8), Zadeh (9)] as

$$\mathrm{N}(A) = 1 - \Pi(\overline{A}) = \inf_{u \notin A} 1 - \pi_x(u) \tag{5}$$

The duality relationship Eq. (5) between Π and N expresses that A is all the more certain as $\overline{A}$ is less consistent with the available knowledge. In other words, A is all the more necessarily true as 'not A' is more impossible. This is a gradual version of the duality between possibility and necessity in modal logic.

When π_x models a nonfuzzy piece of incomplete information represented by an ordinary subset E, Eq. (5) reduces to

$$\begin{aligned} \mathrm{N}_E(A) &= 1 \text{ if } E \subseteq A \quad (\text{information } x \in E \text{ logically entails } x \in A) \\ &= 0 \text{ otherwise} \end{aligned} \tag{6}$$

In the case of complete knowledge, that is $E = \{u_0\}$ for some u_0, $\Pi_E(A) = N_E(A) = 1$ if and only if $u_0 \in A$. Complete ignorance corresponds to $E = U$, and then $\forall\ A \neq \emptyset$, $\Pi_E(A) = 1$ and $\forall\ A \neq U$, $N_E(A) = 0$ (everything is possible and nothing is certain).

In the general case, N(A) estimates to what extent each value outside A, that is, in the complement $\overline{A}$ of A, has a low degree of possibility. Thus the values with the higher degrees of possibility should be included among the elements of A, which makes us somewhat *certain* that, indeed, x belongs to A. It is easy to verify that $\mathrm{N}(A) > 0$ implies $\Pi(A) = 1$, that is an event is completely possible (completely consistent with what is known) before being somewhat certain. This property ensures the natural inequality $\Pi(A) \geq \mathrm{N}(A)$. $\mathrm{N}(A) > 0$ corresponds to the idea of (provisionally) accepting A as a belief. The above definition of N from Π makes sense only if Π, and thus π_x, are normalized, that is, $\sup_{u \in U} \pi_x(u) = 1$. It means that U is fully consistent with the available knowledge, meaning that this knowledge itself is consistent if U is exhaustive as a referential.

Necessity measures satisfy an axiom dual of Eq. (4), namely,

$$\mathrm{N}(A \cap B) = \min[\mathrm{N}(A), \mathrm{N}(B)] \tag{7}$$

It expresses that the conjoint event 'A and B' is all the more certain as A is certain and B is certain. This is the characteristic axiom of necessity measures. Mind that we have only $\mathrm{N}(A \cup B) \geq \max[\mathrm{N}(A), \mathrm{N}(B)]$. Indeed, we may be somewhat certain that x lies in $A \cup B$ without knowing at all if x is in A or is rather in B (at least for $A \cup B = U$, $\mathrm{N}(U) = 1$).

Fuzzy Events

Possibility and necessity measures naturally extend to fuzzy events. Using the ideas of consistency and entailment as the basis for possibility and necessity, the following extensions are obtained. If F is fuzzy set with membership function μ_F,

$$\Pi(F) = \sup_u \mu_F(u) * \pi_x(u) \tag{8}$$

where $*$ is a monotonic conjunctive-like operation (such that $1 * t = t$ and $0 * t = 0$ to recover Eq. (2) when $F = A$ is nonfuzzy), and

$$\mathrm{N}(F) = \inf_u I[\pi_x(u), \mu_F(u)] \tag{9}$$

where I is an implication operation of the form $I(a, b) = 1 - [a * (1 - b)]$, which reduces to Eq. (5) when $F = A$ is nonfuzzy. This choice preserves the identity $\mathrm{N}(F) = 1 - \Pi(\overline{F})$ for the usual fuzzy set complementation [$\forall u$, $\mu_{\overline{F}}(u) = 1 - \mu_F(u)$]. Equation (8) has been proposed by (1) with $* = \min$. Taking $* = \min$, the characteristic properties Eqs. (4) and (7) of possibility and necessity measures with respect to disjunction and conjunction still hold. $\Pi(F) \geq \mathrm{N}(F)$, $\forall F$ still holds (if π_x is normalized).

The possibility of a fuzzy set (with $* = \min$) is a remarkable example of a Sugeno integral (10) since

$$\Pi(F) = \sup_{\alpha \in (0,1]} \min[\alpha, \Pi(F_\alpha)] \tag{10}$$

where $F_\alpha = \{u, \mu_F(u) \geq \alpha\}$. The same formula holds (6,11) for the necessity measure N in place of Π. It points out that these definitions are compatible with the α-cut view of a fuzzy set.

Other Set Functions—Certainty and Possibility Qualification

Apart from Π and N, two other set functions can be defined using sup or inf, namely, a measure of "guaranteed possibility" (12):

$$\Delta(A) = \inf_{u \in A} \pi_x(u) \tag{11}$$

which estimates to what extent *all* the values in A are actually possible for x according to what is known, that is, each value in A is at least possible for x at the degree $\Delta(A)$. Clearly Δ is a stronger measure than Π, that is, $\Delta \leq \Pi$, since Π estimates only the existence of at least one value in A compatible with the available knowledge, whereas the evaluation provided by Δ concerns *all* the values in A. Note also that Δ and N are unrelated. A dual measure of potential certainty, $\nabla(A) = 1 - \Delta(\overline{A})$, estimates to what extent there exists at least one value in the complement of A which has a low degree of possibility. Δ and ∇ are monotonically *decreasing* set functions (in the wide sense) with respect to set inclusion, for example $\Delta(A \cup B) = \min[\Delta(A), \Delta(B)]$. It contrasts with Π and N which are monotonically increasing. Δ agrees with the idea of explicit permission: if A or B is permitted, both A *and* B are permitted.

The set function Δ plays an important role in the representation of possibility-qualified statements, as we are going to see. A possibility distribution can be implicitly specified through the qualification of ordinary, or fuzzy, subsets of the referential U, in terms of either certainty or possibility. Let A be an *ordinary* subset of U (with characteristic function μ_A), namely,

1. The statement "it is certain at least to the degree α that the value of x is in A" will be interpreted as "any value outside A is at most possible at the complementary degree, namely $1 - \alpha$," that is, $\forall u \notin A$, $\pi_x(u) \leq 1 - \alpha$, which leads to the following (7):

   "A is α-certain for x" is translated by
   $$\forall u \in U, \pi_x(u) \leq \max[\mu_A(u), 1 - \alpha] \tag{12}$$

   It can be verified that this is equivalent to $\mathrm{N}(A) \geq \alpha$. Note that for $\alpha = 1$, the inequality $\forall u$, $\pi_x(u) \leq \mu_A(u)$ is recovered. When α decreases from 1 to 0, our knowledge evolves from complete certainty in A to acknowledged ignorance about x.

2. The statement "A is a possible range for x at least at the degree α" will be understood as $\forall u \in A$, $\pi_x(u) \geq \alpha$, which leads to the following:

"A is α-possible for x" is translated by

$$\forall u \in U, \min[\mu_A(u), \alpha] \leq \pi_x(u) \quad (13)$$

It can be verified that this is equivalent to $\Delta(A) \geq \alpha$. Note that for $\alpha = 1$, Eq. (13) reduces to $\forall u$, $\pi_x(u) \geq \mu_A(u)$. When α decreases from 1 to 0, our knowledge evolves from the certainty that A is the minimal range of our ignorance to a total lack of information whatsoever. The representation of possibility-qualified *fuzzy* statements by Eqs. (12) and (13), respectively, can be still justified when A becomes a fuzzy set; see (12).

Qualitative Versus Quantitative Possibility Theory—Conditioning

Besides, an ordinal uncertainty scale is sufficient for defining possibility (and necessity) measures, because possibility theory uses only max, min, and the order-reversing operation of the scale ($1 - (\cdot)$ on [0, 1]). This agrees with a rather qualitative view of uncertainty. Indeed it has been shown that possibility measures are the unique numerical counterparts of qualitative possibility relationships, defined as nontrivial, reflexive, complete, transitive binary relationships "$B \succeq A$," expressing that an event B is at least as possible as another event A and satisfying the characteristic requirement $\forall A, B, C$, $B \succeq A \Rightarrow B \cup C \succeq A \cup C$ (13). Such relationships were introduced by Lewis in Ref. 14.

Possibility theory can be interpreted either as a model of ordinal uncertainty based on a linear ordering (where statements can be ranked only according to their levels of possibility and their level of necessity in a given scale), or as a numerical model of uncertainty which can then be related to probability theory and other uncertainty frameworks. This distinction affects how conditioning is defined.

Conditioning in possibility theory can be defined similarly to that in probability theory, namely, through an equation of the form $\Pi(A \cap B) = \Pi(B|A) * \Pi(A)$. Clearly, the choice for $*$ should be compatible with the nature of the scale which is used. A possible choice for $*$, in agreement with an ordinal scale, is min (15), namely,

$$\forall B, B \cap A \neq \varnothing, \Pi(A \cap B) = \min\left[\Pi(B|A), \Pi(A))\right] \quad (14)$$

This equation has more than one solution. Dubois and Prade (7) proposed selecting the least specific one, that is (for $\Pi(A) > 0$),

$$\begin{aligned}\Pi(B|A) &= 1 \text{ if } \Pi(A \cap B) = \Pi(A) \\ &= \Pi(A \cap B) \text{ otherwise}\end{aligned} \quad (15)$$

The conditional necessity function is defined by $\text{N}(B|A) = 1 - \Pi(\overline{B}|A)$, by duality. Note that $\text{N}(B|A) > 0 \Leftrightarrow \Pi(A \cap B) > \Pi(A \cap \overline{B})$, which expresses that B is an accepted belief in the context A if and only if B is more plausible than $\overline{B}$ when A is true. A notion of qualitative independence has been recently introduced (16), namely, $\text{N}(B|A) > 0$ and $\text{N}(B|A \cap C) > 0$ hold simultaneously. Thus it expresses that B is believed in context A independently of C. Letting $A = U$, it can be verified that this notion is stronger than the condition $\Pi(B \cap C) = \min[\Pi(B), \Pi(C)]$, which expresses a form of unrelatedness; see Ref. 17.

The previous definitions require only a purely ordinal view of possibility theory. In the case of a numerical scale, we can also use the product instead of min in the conditioning Eq. (14). It leads to

$$\forall B, B \cap A \neq \varnothing, \Pi(B|A) = \frac{\Pi(A \cap B)}{\Pi(A)} \quad (16)$$

provided that $\Pi(A) \neq 0$. This is the Dempster rule of conditioning specialized to possibility measures, that is, consonant plausibility measures of Shafer (18).

POSSIBILITY THEORY VERSUS OTHER UNCERTAINTY FRAMEWORKS

Relationships with Probabilities and Belief Functions

Formally speaking, possibility measures clearly depart from probability measures in various respects. The former are max-decomposable for the disjunction of events, whereas the latter are additive (for mutually exclusive events). Dually, necessity measures are min-decomposable for conjunction. However, possibility (respectively, necessity) measures are not compositional for conjunction (respectively, disjunction) or for negation (whereas probabilities are compositional only for negation). In possibility theory, the representation of the uncertainty of A requires two weakly related numbers, namely, $\Pi(A)$ and $\text{N}(A) = 1 - \Pi(\overline{A})$, which contrasts with probabilities. Thus a distinction can be made between the impossibility of $A(\Pi(A) = 0 \Leftrightarrow \text{N}(\overline{A}) = 1)$ and the total lack of certainty about $A(\text{N}(A) = 0)$, which is entailed by (but not equivalent to) the impossibility of A (whereas in probability theory, $\text{Prob}(\overline{A}) = 1 \Leftrightarrow \text{Prob}(A) = 0$).

Despite the difference between probability and possibility, it is noteworthy that possibility measures can be given a purely frequentist interpretation (19,20). Given any statistical experiment with outcomes in U, assume that a precise observation of outcomes is out of reach for some reason. For instance, U is a set of candidates to an election, and the experiment is an opinion poll where individuals have not yet made up their minds completely and are allowed to express them by proposing a subset of candidates, containing their future choice. Let $\mathscr{F} \subseteq 2^U$ be the set of observed responses, and $m(E)$, $E \in \mathscr{F}$, be the proportion of responses of the form of a crisp set $E(\sum_{E \in \mathscr{F}} m(E) = 1, m(\emptyset) = 0)$. $(\mathscr{F}, m)$ defines a basic probability assignment in the sense of Shafer (18). It can be also viewed as a *random set*. In that case

$$P^*(A) = \textstyle\sum_{E \in \mathscr{F}} \Pi_E(A) \cdot m(E) = \sum_{A \cap E \neq \varnothing} m(E) \quad (17)$$

is the expected possibility of A in the sense of logical possibility. Formally, $P^*(A)$ is mathematically identical to an upper probability in the sense of Dempster (21) or to a plausibility function in the sense of Shafer (18), and $P_*(A) = 1 - P^*(\overline{A}) = \sum_{E \in \mathscr{F}} \text{N}_E(A) \cdot m(E) = \sum_{E \subseteq A} m(E)$ has the same property as a Shafer belief function. To recover the max-decomposability (1) for P^*, it is necessary and sufficient to assume that $\mathscr{F}$ defines a *nested* sequence of sets (18). Then P^* is a possibility measure. Moreover, the guaranteed possibility function is a special case of the Shafer commonality function. Hence, possibility measures correspond to imprecise but coherent (due to the nestedness property), statistical evidence, that is, an ideal situation opposite to the case of probability measures (ideal, too) where outcomes form a partition of Ω. More generally, necessity (and possibility) measures provide inner and outer approximation of belief (and plausibility) functions; see (22).

Possibility measures can be also viewed as a particular type of upper probability system. An upper probability P^* induced by a set of lower bounds $\{P(A_i) \geq \alpha_i, i = 1, \ldots, n\}$ (i.e., $P^*(B) = \sup\{P(B)|P \in \mathscr{P}\}$ where $\mathscr{P} = \{P|P(A_i) \geq \alpha_i, i = 1, \ldots, n\}$), is a possibility measure if the set $\{A_1, \ldots, A_n\}$ is nested, that is, after a suitable renumbering $A_1 \subseteq A_2 \subseteq \cdots \subseteq A_n$. Conversely, any possibility measure on a finite set can be induced by such a set of lower bounds with nested A_i's; see (23). This view leads to a definition of conditioning different from Eqs. (15) or (16); see (24).

The problem of transforming a possibility distribution into a probability distribution and conversely is meaningful in the scope of uncertainty combination with heterogeneous sources (some supplying statistical data, other linguistic data, for instance). However, raising the issue means that some consistency exists between possibilistic and probabilistic representations of uncertainty. The basic question is whether this is a mere matter of translation between languages "neither of which is weaker or stronger than the other" (25). Adopting this assumption leads to transformations that respect a principle of uncertainty and information invariance. However, if we accept the fact that possibility distributions are weaker representations of uncertainty than probability distributions, the transformation problem must be stated otherwise. Going from possibility to probability increases the informational content of the considered representation, whereas going the other way around means an informational loss. Hence the principles behind the two transformations are different, and asymmetric transformations are obtained (26): *From possibility to probability,* a generalized Laplacean indifference principle is adopted. *From probability to possibility,* the rationale is to preserve as much information as possible which leads to selecting the most specific upper approximation of the probability.

Taking advantage of the inequalities $\inf\{P(A|b), b \in B\} \leq P(A|B) \leq \sup\{P(A|b), b \in B\}$, it is possible to see a possibility measure (respectively, a Δ function) as the upper (respectively, lower) envelope of a family of likelihood functions (27). Several authors have previously suggested viewing likelihood functions as possibility distributions [e.g., (28,29)].

Lastly, Spohn (30) has proposed a theory of epistemic states with strong similarities to possibility theory, with which it shares the idea of ordering between possible worlds. What he calls an ordinal conditional function is a mapping from a finite set of events to the set of positive integers, denoted κ such that $\kappa(A \cup B) = \min[\kappa(A), \kappa(B)]$. $\kappa(A)$ expresses a degree of disbelief in A and grows as A becomes less plausible. Moreover, there is an elementary event $\{u\}$ such that $\kappa(\{u\}) = 0$. It is easy to check that for any real number $c > 1$, $1 - c^{-\kappa(\overline{A})}$ is a degree of necessity [see (31)]. A probabilistic interpretation of $\kappa(A)$ has been suggested by Spohn (30). $\kappa(A) = n$ is interpreted as a small probability of the form ϵ^n, that is, the probability of a rare event. Indeed, if A has a small probability with order of magnitude ϵ^n and B has also a small probability of the form ϵ^m, then $P(A \cup B)$ is of the order of magnitude of $\epsilon^{\min(m,n)}$. These remarks may lead to an interpretation of possibility and necessity measures in terms of probabilities of rare events.

Rough Sets

Rough set theory (32) captures the idea of indiscernibility. Indiscernibility means the lack of discriminatory power between elements in a set. At a very primitive level, this aspect can be captured by an equivalence relationship R on the set U, such that $u\ R\ u'$ means that u and u' cannot be told apart. Then R induces a partition on U, made of the elements of the quotient space U/R. As a consequence, any subset A of U can be described only by means of clusters in U given here by the equivalence classes $[u]$ of R, namely,

$$\begin{cases} \text{the lower image of } A: A_* = \{[u] | [u] \subseteq A\} \\ \text{the upper image of } A: A^* = \{[u] | A \cap [u] \neq \varnothing\} \end{cases} \tag{18}$$

This aspect has been studied by Pawlak (32) under the name "rough set" and is also studied by Shafer (18) when he considers coarsening, refinements, and compatibility relationships between frames. Although possibility and fuzzy set theory are not directly concerned with indistinguishability, A_* and A^* can easily be interpreted in terms of necessity and possibility: $[u] \in A_* \Leftrightarrow N_{[u]}(A) = 1$, and $[u] \in A^* \Leftrightarrow \Pi_{[u]}(A) = 1$, where $E = [u]$ (using the notations of Eqs. (3) and (6)). $[u]$ belongs to A_* (respectively, A^*) if and only if it is certain (respectively, possible) that any (respectively, some) element close to u (in the sense of R) belongs to A.

It is possible to extend the rough set framework with (fuzzy) similarity relationships or fuzzy partitions (33). Indiscernibility, which is also linked to Poincaré's paradox of mathematical continuum, is clearly an important issue in knowledge representation, where information appears in a granular form whereas partial belief is measured on a continuous scale. This question is clearly orthogonal to that of modeling partial belief, because it affects the definition and the structure of frames of discernment. The coarsening of the referential U into the quotient space U/R (where R is a classical equivalence relationship) thus induces lower and upper approximations π_* and π^* for a possibility distribution π, defined by

$$\begin{aligned} \forall \omega \in U/R, \pi_*(\omega) &= \inf_{u \in \omega} \pi(u) \\ \pi^*(\omega) &= \sup_{u \in \omega} \pi(u) \end{aligned} \tag{19}$$

Here what is fuzzified is the subset A to be approximated, agreeing with Eq. (18). Indeed letting $\mu_F(u) = \pi(u)$, Eq. (19) corresponds to the necessity and the possibility of the fuzzy event F based on the nonfuzzy possibility distribution μ_ω. These lower and upper approximations can be generalized by using fuzzy equivalence classes based on similarity relationships instead of the crisp equivalence classes μ_ω. See Ref. 33.

APPROXIMATE REASONING

The theory of approximate reasoning, whose basic principles have been formulated by Zadeh (34) can be viewed as a direct application of possibility theory. Indeed, it is essentially a methodology for representing fuzzy and incomplete information in terms of unary or joint possibility distributions and inferring the values of variables of interest by applying the rules of possibility theory. What Zadeh's representation and approximate reasoning theory mainly provides is a powerful tool for interfacing symbolic knowledge and numerical variables that has proved very useful in applications where qualitative knowledge pertains to numerical quantities (e.g., fuzzy logic controllers where precise input values are matched against the fuzzy conditions of rules). Neither classical rule-based systems nor classical logic are fully adapted to properly handling the interface between numbers and symbols without resorting to arbitrary thresholds for describing predicate extensions.

Joint Possibility Distributions

Let x and y be two variables (on U and V, respectively) linked together through a fuzzy restriction on the Cartesian product $U \times V$, encoded by a possibility distribution $\pi_{x,y}$, called a joint possibility distribution (the following can be easily extended to more than two variables). The unary possibility distribution π_x, representing the induced restriction on the possible values of x, can be calculated as the *projection* of $\pi_{x,y}$ on U defined in Refs. 34 and 35 by

$$\begin{aligned}\pi_x(u) &= \Pi(\{u\} \times V)\\ &= \sup_v \pi_{x,y}(u,v)\end{aligned} \tag{20}$$

Generally, $\pi_{x,y} \le \min(\pi_x, \pi_y)$. When equality holds, $\pi_{x,y}$ is then said to be min-separable, and the variables are said to be *noninteractive* (35). It is in accordance with the principle of minimal specificity, since $\pi_x(u)$ is calculated from the highest possibility value of pairs (x, y) where $x = u$. When modeling incomplete information, noninteractivity expresses a lack of knowledge about the links between x and y. If we start with the two pieces of knowledge represented by π_x and π_y and if we do not know whether or not x and y are interactive that is, π_{xy} is not known, we can use the upper bound $\min(\pi_x, \pi_y)$ instead, which is less informative (but which agrees with the available knowledge). This is minimal specificity again.

Note that whereas in probability theory, independence plays the same role as noninteractivity in possibility theory (probabilistic variables can be assumed to be stochastically independent by virtue of the principle of maximum entropy), stochastic independence does not lead to bounding properties as noninteractivity does because stochastic independence assumes an actual absence of correlation whereas noninteractivity expresses a lack of knowledge.

The noninteractivity of variables implies that the possibility and necessity measures become decomposable with respect to Cartesian products $(A \times B)$ and the associated coproducts $(A + B = \overline{\overline{A} \times \overline{B}})$, respectively:

$$\Pi(A \times B) = \min\left[\Pi_x(A), \Pi_y(B)\right] \tag{21}$$

$$\mathrm{N}(A + B) = \max[\mathrm{N}_x(A), \mathrm{N}_y(B)] \tag{22}$$

where A and B are subsets of the universes U and V (the ranges of x and y) and Π_x and N_x are possibility and necessity measures based on the normalized possibility distribution π_x. This is useful in fuzzy pattern matching evaluation where compound requirements of the form 'A and B' or 'A or B' are matched again fuzzy data pertaining only to different attributes x.

Representing Fuzzy Rules

A general approach to modeling fuzzy statements has been outlined by Zadeh (36) via the representation language PRUF. Of particular interest are fuzzy rules, which are rules whose conditions and/or conclusions have the form x is F. Simple fuzzy rules have the generic form "if x is F, then y is G" and express a fuzzy link between x and y. Several interpretation of fuzzy rules exist (37). An immediate application of possibility and certainty qualification is the representation of two kinds of fuzzy rules, called certainty and possibility rules (12). The fuzzy rule "the more x is F, the more certain y is G" can be represented by

$$\pi_{x,y}(u,v) \le \max[\mu_G(v), 1 - \mu_F(u)] \tag{23}$$

and the fuzzy rule "the more x is F, the more possible y is G" by

$$\pi_{x,y}(u,v) \ge \min[\mu_G(v), \mu_F(u)] \tag{24}$$

Indeed, letting $\alpha = \mu_F(u)$ and changing $\mu_A(u)$ into $\mu_G(v)$ in the expressions of certainty qualification [Eq. (12)] and possibility qualification [Eq. (13)], these rules can be understood as "if $x = u$, then y is G is $\mu_F(u)$-certain" (respectively, $\mu_F(u)$-possible). Note that Eq. (24) is the representation of a fuzzy rule originally proposed by Mamdani (38) in fuzzy logic controllers. The principle of minimal specificity leads to representing the certainty rules by $\pi_{x,y}(u, v) = \mu_F(u) \rightarrow \mu_G(v)$, where $a \rightarrow b = \max(1 - a, b)$ is known as the Dienes implication. The principle of minimal specificity cannot be applied to Eq. (24). On the contrary, the rule expresses that the values v, for instance, in the core of G (such that $\mu_G(v) = 1$) are possible at least at the degree $\alpha = \mu_F(u)$. This does not forbid other values outside G from being possible also. The inequality in Eq. (24) explains why conclusions are combined *disjunctively* in Mamdani's treatment of fuzzy rules (and not conjunctively as with implication-based representations of fuzzy rules). Here a maximal informativeness principle consists in considering that only the values in G are possible and only at the degree $\mu_F(u)$ (for the values in the core of G). This leads to $\pi_{x,y}(u, v) = \min[\mu_G(v), \mu_F(u)]$, that is, only what is explicitly stated as possible is assumed possible.

Another type of fuzzy rule, which is of interest in interpolative reasoning, is gradual rules. They are of the form "the more x is F, the more y is G." This translates into a constraint on $\pi_{x,y}$ acknowledging that the image of 'x is F' by $\pi_{x,y}$ (defined by combination and projection) is included in G, that is,

$$\forall u, \sup_u \min[\mu_F(u), \pi_{x,y}(u,v)] \le \mu_G(v) \tag{25}$$

Then the principle of minimal specificity leads to a representation by $\pi_{x,y}(u, v) = 1$ if $\mu_F(u) \le \mu_G(v)$. When $\mu_F(u) > \mu_G(v)$, $\pi_{x,y}(u, v)$ can be taken equal to 0 or $\mu_G(v)$ depending whether or not the relationship between x and y is fuzzy.

The Approximate Reasoning Methodology

Inference in the framework of possibility theory is based on a combination/projection principle stated by Zadeh (35) for fuzzy constraints, namely, given a set of n statements $S_1 \ldots S_n$ that form a knowledge base, inference proceeds in three steps:

1. Translate $S_1, \ldots, S_n$ into possibility distributions restricting the values of involved variables. Facts of the form 'x is F' translate into $\pi_x = \mu_F$, and rules of the form "if x is F, then y is G" translate into possibility distributions $\pi_{x,y} = \mu_R$ where μ_R derives from the semantics of the rule, as discussed previously.
2. Combine the various possibility distribution conjunctively to build a joint possibility distribution expressing the meaning of the knowledge base, that is,

$$\pi = \min(\pi_1, \ldots, \pi_n)$$

3. Project π on the universe corresponding to some variable of interest.

The combination/projection principle is an extension of classical deduction. For instance, if S_1 = "x is F' " and S_2 = "if x is F, then y is G",

$$\pi_{x,y}(u, v) = \min[\mu_F(u), \mu_R(u, v)]$$

where μ_R represents the rule S_2. Then, the fact "y is G'" is inferred such that $\mu_{G'}(v) = \sup_u \min[\mu_{F'}(u), \mu_R(u, v)]$. This is called the *generalized modus ponens* and was proposed by Zadeh (39). This approach has found applications in many systems implementing fuzzy logic and in possibilistic belief networks, such as POSSINFER (40).

Possibilistic Logic and Nonmonotonic Reasoning

An important particular case of approximate reasoning considers only necessity-qualified classical propositions. It gives birth to possibilistic logic (41). A possibilistic knowledge base K is a set of pairs (p, s) where p is a classical logic formula and s is a lower bound of a degree of necessity ($N(p) \geq s$). It can be viewed as a stratified deductive data base where the higher s is, the safer is the piece of knowledge p. Reasoning from K means using the safest part of K to make inference, whenever possible. Denoting $K_\alpha = \{p, (p, s) \in K, s \geq \alpha\}$, the entailment $K \vdash (p, \alpha)$ means that $K_\alpha \vdash p$. K can be inconsistent and its inconsistency degree is $\text{inc}(K) = \sup\{\alpha, K \vdash (\bot, \alpha)\}$ where $\bot$ denotes the contradiction. In contrast with classical logic, inference in the presence of inconsistency becomes nontrivial. This is the case when $K \vdash (p, \alpha)$ where $\alpha > \text{inc}(K)$. Then it means that p follows from a consistent and safe part of K (at least at level α). Moreover, adding p to K and nontrivially entailing q from $K \cup \{p\}$ corresponds to revising K upon learning p and having q as a consequence of the revised knowledge base. This notion of revision is exactly that studied by Gärdenfors (42) at the axiomatic level (31). This kind of syntactic nontrivial inference is sound and complete with respect to a so-called preferential entailment at the semantic level. p is said to preferentially entail q if all the preferred interpretations, which make p true, make q true also. A preferred interpretation is one which maximizes the possibility distribution π_K, the least specific possibility distribution satisfying the set of constraints $N(p) \geq s$ where $(p, s) \in K$; see (43).

Possibilistic logic does not allow for directly encoding pieces of generic knowledge, such as "birds fly." However, it provides a target language in which plausible inference from generic knowledge can be achieved in the face of incomplete evidence. In possibility theory, "p generally entails q" is understood as "$p \wedge q$ is a more plausible situation than $p \wedge \neg q$." It defines a constraint of the form $\Pi(p \wedge q) > \Pi(p \wedge \neg q)$ that restricts a set of possibility distributions. Given a set S of generic knowledge statements of the form "p_i generally entails q_i," a possibilistic base can be computed as follows. For each interpretation ω of the language, the maximal possibility degree $\pi(\omega)$ is computed that obeys the set of constraints in S. This is done by virtue of the principle of minimal specificity (or commitment) that assumes each situation as possible insofar as it has not been ruled out. Then each generic statement is turned into a material implication $\neg p_i \vee q_i$, to which $N(\neg p_i \vee q_i)$ is attached. It comes down, as shown by Benferhat et al. (44), to rank ordering the generic rules giving priority to the most specific ones, as done in Pearl's system Z (45). A very important property of this approach is that it is exception-tolerant. It offers a convenient framework for implementing a basic form of nonmonotonic system called *rational closure* (46) and addresses a basic problem in the expert system literature, that is, handling exceptions in uncertain rules.

Computing with Fuzzy Numbers

Another important application of the combination/projection principle is computation with ill-known quantities. Given two ill-known quantities represented by the possibility distributions π_x and π_y on the real line, the possibility distribution restricting the possible values of $f(x, y)$ where f is some function (e.g., $f(x, y) = x + y$), is given by

$$\begin{aligned}\pi_{f(x,y)}(w) &= \Pi[f^{-1}(w)] \\ &= \sup_{(u,v)\in f^{-1}(w)} \min[\pi_x(u), \pi_y(v)]\end{aligned} \tag{26}$$

where $f^{-1}(w) = \{(u, v)|f(u, v) = w\}$ and x and y are assumed to be noninteractive. This is the basis for extending arithmetic operations to fuzzy numbers and their application to mathematical programming. Simple formulas exist for practically computing the sum of parameterized fuzzy intervals and other arithmetic operations; see (47–49). It also applies in flexible constraint satisfaction problems. Constraint satisfaction problems, which are a basic paradigm in artificial intelligence, can be extended to flexible and prioritized constraints in the setting of possibility theory. In this case, possibility distributions model preferences about the way constraints have to be satisfied. Ill-known parameters in the description of the problems can be also dealt with. Thus preferences and uncertainty are handled in the same framework (50).

Defuzzification

Fuzzy-set-based approximate reasoning usually yields conclusions in the form of possibility distributions. Then to make them more easy to interpret for the end user, we may approximate them linguistically in terms of fuzzy sets which represent terms in a prescribed vocabulary. We may also defuzzify these conclusions. A well-known method, often used in fuzzy control applications, is the center of gravity method which computes the scalar value

$$\text{defuzz}(\pi) = \int u \cdot \pi(u)\, du \Big/ \int \pi(u)\, du \tag{27}$$

if π is the result to be defuzzified. However, this method has no justification in the framework of possibility theory. Another method, which agrees with this framework is the selection of the value(s) which maximize(s) π, but it generally yields a set larger than a singleton.

Yet another method, which is especially of interest for fuzzy intervals of the real line, consists of computing lower and upper expectations of the variable restricted by the fuzzy interval. It is thus summarized by a pair of numbers interpretable as an ordinary interval. For example, let $\pi = \mu_F$ be a fuzzy interval. Its α-cut $F_\alpha = \{u, \mu_F(u) \geq \alpha\}$ is by definition an interval [inf F_α, sup F_α]. Then the lower and upper expectations can be defined by

$$E_*(F) = \int_0^1 \inf F_\alpha \cdot d\alpha$$

and

$$E^*(F) = \int_0^1 \sup F_\alpha \cdot d\alpha \tag{28}$$

See (51) for a justification. Thus, the interval $[E_*(F), E^*(F)]$ summarizes the imprecision of the fuzzy interval F.

It can be also verified that

$$\frac{E_*(F) + E^*(F)}{2} = \int_{-\infty}^{+\infty} x \cdot p_F(x)\,dx$$

with

$$p_F(x) = \int_0^1 \mu_{F_\alpha}(x) \frac{d\alpha}{|F_\alpha|} \tag{29}$$

which provides a justification of the middle of the interval as the result of the defuzzification of F, since the density is based on a uniform probability on each α-cut $F_\alpha(|F_\alpha|$ is the length of $F_\alpha)$, just applying a generalized Laplacean indifference principle. With this definition, the defuzzification of the sum of two fuzzy intervals [in the sense of Eq. (26)] yields the sum of the values obtained by defuzzifying each fuzzy interval.

CONCLUDING REMARKS

This article provides a brief overview of the basic features of possibility theory, of its relationships with other uncertainty calculi, and of its applications to various types of approximate reasoning. However, some applications, such as data fusion, in which possibility theory offers a variety of combination modes (including weighted, prioritized, and adaptive aggregation rules) in poorly informed environments (52,53), and decision under uncertainty, where a (necessity-based) pessimistic and a (possibility-based) optimistic qualitative counterpart of classical expectation (in terms of Sugeno integrals), have been proposed in (54) but are not reviewed here. Recent theoretical developments of possibility theory can be found in (55), and a thorough presentation and discussion of possibility theory can be found in Ref. 56.

BIBLIOGRAPHY

1. L. A. Zadeh, Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets Syst.,* **1**: 3–28, 1978.
2. R. R. Yager, An introduction to applications of possibility theory, *Human Syst. Manage.,* **3**: 246–269, 1983.
3. R. R. Yager, On the specificity of a possibility distribution, *Fuzzy Sets Syst.,* **50** (3): 279–292, 1992.
4. G. J. Klir and Y. Bo, *Fuzzy Sets and Fuzzy Logic—Theory and Applications,* Upper Saddle River, NJ: Prentice-Hall, 1995.
5. R. R. Yager, A foundation for a theory of possibility, *J. Cybernetics,* **10**: 177–204, 1980.
6. H. Prade, Modal semantics and fuzzy set theory. In R. R. Yager (ed.), *Fuzzy Set and Possibility Theory—Recent Developments,* New York: Pergamon Press, 1982, p. 232–246.
7. D. Dubois and H. Prade, (with the collaboration of H. Farreny, R. Martin–Clouaire, and C. Testemale), *Possibility Theory—An Approach to Computerized Processing of Uncertainty,* New York: Plenum Press, 1988.
8. D. Dubois and H. Prade, *Fuzzy Sets and Systems—Theory and Applications,* New York: Academic Press, 1980.
9. L. A. Zadeh, Fuzzy sets and information granularity. In M. M. Gupta, R. K. Ragade, and R. R. Yager (eds.), *Advances in Fuzzy Set Theory and Applications,* Amsterdam: North-Holland, 1979, pp. 3–18.
10. M. Sugeno, Fuzzy measures and fuzzy integrals—A survey, In M. M. Gupta, G. N. Saridis, and B. R. Gaines (eds.), *Fuzzy Automata and Decision Processes,* Amsterdam: North-Holland, 1977, pp. 89–102.
11. M. Inuiguchi, H. Ichihashi, and H. Tanaka, Possibilistic linear programming with measurable multiattribute value functions, *ORSA J. Comput.,* **1** (3): 146–158, 1989.
12. D. Dubois and H. Prade, Fuzzy rules in knowledge-based systems—Modelling gradedness, uncertainty and preference. In R. R. Yager and L. A. Zadeh (eds.), *An Introduction to Fuzzy Logic Applications in Intelligent Systems,* Dordrecht. The Netherlands: Kluwer Academic Publ., 1992, pp. 45–68.
13. D. Dubois, Belief structures, possibility theory and decomposable confidence measures on finite sets, *Comput. Artificial Intell.,* **5** (5): 403–416, 1986.
14. D. K. Lewis, *Counterfactuals,* Oxford: Basil Blackwell, 1973. 2nd edition, Worcester, UK: Billing and Sons, 1986.
15. E. Hisdal, Conditional possibilities—Independence and non-interactivity, *Fuzzy Sets Syst.,* **1**: 283–297, 1978.
16. D. Dubois et al., An ordinal view of independence with application to plausible reasoning. In R. Lopez de Mantaras and D. Poole (eds.), *Proc. 10th Conf. Uncertainty Artificial Intell.,* Seattle, WA, July 29–31, 1994, pp. 195–203.
17. S. Nahmias, Fuzzy variables, *Fuzzy Sets Syst.,* **1**: 97–110, 1978.
18. G. Shafer, *A Mathematical Theory of Evidence,* Princeton, NJ: Princeton University Press, 1976.
19. P. Z. Wang and E. Sanchez, Treating a fuzzy subset as a projectable random subset. In M. M. Gupta and E. Sanchez (eds.), *Fuzzy Information and Decision Processes,* Amsterdam: North-Holland, 1982, pp. 213–219.
20. D. Dubois and H. Prade, Fuzzy sets, probability and measurement, *Eur. J. Operations Res.,* **40**: 135–154, 1989.
21. A. P. Dempster, Upper and lower probabilities induced by a multiple-valued mapping. *Ann. Math. Statistics,* **38**: 325–339, 1967.
22. D. Dubois and H. Prade, Consonant approximations of belief functions, *Int. J. Approximate Reasoning,* **4** (5/6): 419–449, 1990.
23. D. Dubois and H. Prade, When upper probabilities are possibility measures, *Fuzzy Sets Syst.,* **49**: 65–74, 1992.
24. D. Dubois and H. Prade, Focusing vs. revision in possibility theory, in *Proc. 5th Int. Conf. Fuzzy Syst. (FUZZ-IEEE),* New Orleans, Sept. 8–11, 1996, pp. 1700–1705.
25. G. J. Klir and B. Parviz, Probability-possibility transformation: A comparison, *Int. J. General Syst.,* **21**: 291–310, 1992.

26. D. Dubois, H. Prade, and S. Sandri, On possibility/probability transformations. In R. Lowen and M. Lowen (eds.), *Fuzzy Logic: State of the Art,* Kluwer Academic Publ., 1993, pp. 103–112.
27. D. Dubois, S. Moral, and H. Prade, A semantics for possibility theory based on likelihoods, *J. Math. Anal. Appl.,* **205**: 359–380, 1997.
28. B. Natvig, Possibility versus probability, *Fuzzy Sets Syst.,* **10**: 31–36, 1983.
29. S. F. Thomas, *Fuzziness and Probability,* Wichita, KS: ACG Press, 1995.
30. W. Spohn, Ordinal conditional functions: A dynamic theory of epistemic states. In W. Harper and B. Skyrms (eds.), *Causation in Decision, Belief Change and Statistics,* Kluwer Acad. Publishers, Dordecht, Netherlands, 1988, pp. 105–134.
31. D. Dubois and H. Prade, Epistemic entrenchment and possibilistic logic, *Artificial Intell.,* **50**: 223–239, 1991.
32. Z. Pawlak, *Rough Sets—Theoretical Aspects of Reasoning about Data,* Dordrecht: Kluwer Academic Publ., 1991.
33. D. Dubois and H. Prade, Rough fuzzy sets and fuzzy rough sets, *Int. J. General Syst.,* **17** (2–3): 191–209, 1990.
34. L. A. Zadeh, A theory of approximate reasoning. In J. E. Hayes, D. Michie, and L. I. Mikulich (eds.), *Machine Intelligence,* New York: Elsevier, 1979, Vol. 9, pp. 149–194.
35. L. A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning, *Information Sciences,* Part 1: **8**: 199–249; Part 2: **8**: 301–357; Part 3: **9**: 43–80, 1975.
36. L. A. Zadeh, PRUF—A meaning representation language for natural languages, *Int. J. Man-Machine Studies,* **10**: 395–460, 1978.
37. D. Dubois and H. Prade, What are fuzzy rules and how to use them, *Fuzzy Sets Syst.,* **84**: 169–185, 1996.
38. E. H. Mamdani, Application of fuzzy logic to approximate reasoning using linguistic systems, *IEEE Trans. Comput.,* **26**: 1182–1191, 1977.
39. L. A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Trans. Systems, Man Cybernetics,* **3**: 28–44, 1973.
40. J. Gebhardt and R. Kruse, POSSINFER: A software tool for possibilistic inference. In D. Dubois, H. Prade, and R. R. Yager (eds.), *Fuzzy Information Engineering—A Guided Tour of Applications,* New York: Wiley, 1997, pp. 405–415.
41. D. Dubois, J. Lang, and H. Prade, Automated reasoning using possibilistic logic: Semantics, belief revision and variable certainty weights, *IEEE Trans. Knowl. Data Eng.,* **6** (1): 64–71, 1994.
42. P. Gärdenfors, *Knowledge in Flux—Modeling the Dynamics of Epistemic States,* Cambridge, MA: The MIT Press, 1988.
43. D. Dubois and H. Prade, Possibilistic logic, preferential models, non-monotonicity and related issues, *Proc. Int. Joint Conf. Artificial Intell. (IJCAI'91),* Sydney, Australia, Aug. 24–30, 1991, pp. 419–424.
44. S. Benferhat, D. Dubois, and H. Prade, Representing default rules in possibilistic logic, *Proc. 3rd Int. Conf. Principles Knowl. Representation Reasoning (KR'92),* Cambridge, MA, Oct. 26–29, 1992, pp. 673–684.
45. J. Pearl, System Z: A natural ordering of defaults with tractable applications to default reasoning, *Proc. 3rd Conf. Theoretical Aspects Reasoning About Knowl. (TARK'90)* (R. Parikh, ed.), Morgan and Kaufmann, 1990, pp. 121–135.
46. D. Lehmann and M. Magidor, What does a conditional knowledge base entail?, *Artificial Intell.,* **55** (1): 1–60, 1992.
47. A. Kaufmann and M. M. Gupta, *Introduction to Fuzzy Arithmetic—Theory and Applications,* New York: Van Nostrand Reinhold, 1985.
48. D. Dubois and H. Prade, Fuzzy numbers: An overview. In J. C. Bezdek (ed.), *The Analysis of Fuzzy Information—Vol. 1: Mathematics and Logic,* Boca Raton, FL: CRC Press, 1987, pp. 3–39.
49. R. Slowinski and J. Teghem (eds.), *Stochastic Versus Fuzzy Approaches to Multiobjective Mathematical Programming Under Uncertainty.* Dordrecht: Kluwer Academic Publ., 1990.
50. D. Dubois, H. Fargier, and H. Prade, Possibility theory in constraint satisfaction problems: Handling priority, preference and uncertainty, *Appl. Intell.,* **6**: 287–309, 1996.
51. D. Dubois and H. Prade, The mean value of a fuzzy number, *Fuzzy Sets Syst.,* **24**: 279–300, 1987.
52. D. Dubois and H. Prade, Possibility theory and data fusion in poorly informed environments, *Control Eng. Practice,* **2** (5): 811–823, 1994.
53. R. R. Yager, A general approach to the fusion of imprecise information, *Int. J. Intell. Syst.,* **12**: 1–29, 1997.
54. D. Dubois and H. Prade, Possibility theory as a basis for qualitative decision theory, *Proc. 14th Int. Joint Conf. Artificial Intell. (IJCAI'95),* Montréal, Canada, Aug. 20–25, 1995, pp. 1924–1930.
55. G. De Cooman, Possibility theory—Part I: Measure- and integral-theoretics groundwork; Part II: Conditional possibility; Part III: Possibilistic independence, *Int. J. General Syst.,* **25**: 291–371, 1997.
56. D. Dubois and H. Prade, Possibility theory: Qualitative and quantitative aspects, in *Handbook of Defeasible Reasoning and Uncertainty Management. Volume I: Quantified Representation of Uncertainty and Imprecision,* Dordrecht: Kluwer Academic Publ., 1998.

DIDIER DUBOIS
HENRI PRADE
Paul Sabatier University

POSSIBILITY THEORY DATABASES.

See FUZZY INFORMATION RETRIEVAL AND DATABASES.