Possibility Theory and its Applications: Where Do we Stand ?

Didier Dubois and Henri Prade

IRIT-CNRS, Université Paul Sabatier, 31062 Toulouse Cedex 09, France

December 19, 2011

Abstract

This paper provides an overview of possibility theory, emphasizing its historical roots and its recent developments. Possibility theory lies at the crossroads between fuzzy sets, probability and non-monotonic reasoning. Possibility theory can be cast either in an ordinal or in a numerical setting. Qualitative possibility theory is closely related to belief revision theory, and common-sense reasoning with exception-tainted knowledge in Artificial Intelligence. Possibilistic logic provides a rich representation setting, which enables the handling of lower bounds of possibility theory measures, while remaining close to classical logic. Qualitative possibility theory has been axiomatically justified in a decision-theoretic framework in the style of Savage, thus providing a foundation for qualitative decision theory. Quantitative possibility theory is the simplest framework for statistical reasoning with imprecise probabilities. As such it has close connections with random set theory and confidence intervals, and can provide a tool for uncertainty propagation with limited statistical or subjective information.

1   Introduction

Possibility theory is an uncertainty theory devoted to the handling of incomplete information. To a large extent, it is comparable to probability theory because it is based on set-functions. It differs from the latter by the use of a pair of dual set functions (possibility and necessity measures) instead of only one. Besides, it is not additive and makes sense on ordinal structures. The name "Theory of Possibility" was coined by Zadeh [142], who was inspired by a paper by Gaines and Kohout [91]. In Zadeh's view, possibility distributions were meant to provide a graded semantics to natural language statements. However, possibility and necessity measures can also be the basis of a full-fledged representation of partial belief that parallels probability. It can be seen either as a coarse, non-numerical version of probability theory, or a framework for reasoning with extreme probabilities, or yet a simple approach to reasoning with imprecise probabilities [74].

After reviewing pioneering contributions to possibility theory, we recall its basic concepts and present the two main directions along which it has developed: the qualitative and quantitative settings. Both approaches share the same basic "maxitivity" axiom. They differ when it comes to conditioning, and to independence notions. Then we discuss prospective lines of research in the area.

1

2 Historical Background

Zadeh was not the first scientist to speak about formalising notions of possibility. The modalities *possible* and *necessary* have been used in philosophy at least since the Middle-Ages in Europe, based on Aristotle's and Theophrastus' works [22]. More recently they became the building blocks of Modal Logics that emerged at the beginning of the XXth century from the works of C.I. Lewis (see Hughes and Cresswell [31]). In this approach, possibility and necessity are all-or-nothing notions, and handled at the syntactic level. More recently, and independently from Zadeh's view, the notion of possibility, as opposed to probability, was central in the works of one economist, and in those of two philosophers.

G. L. S. Shackle A graded notion of possibility was introduced as a full-fledged approach to uncertainty and decision in the 1940-1970's by the English economist G. L. S. Shackle [127], who called *degree of potential surprise* of an event its degree of impossibility, that is, the degree of necessity of the opposite event. Shackle's notion of possibility is basically epistemic, it is a "character of the chooser's particular state of knowledge in his present." Impossibility is understood as disbelief. Potential surprise is valued on a disbelief scale, namely a positive interval of the form $[0, y^*]$, where y^* denotes the absolute rejection of the event to which it is assigned. In case everything is possible, all mutually exclusive hypotheses have zero surprise. At least one elementary hypothesis must carry zero potential surprise. The degree of surprise of an event, a set of elementary hypotheses, is the degree of surprise of its least surprising realisation. Shackle also introduces a notion of conditional possibility, whereby the degree of surprise of a conjunction of two events A and B is equal to the maximum of the degree of surprise of A, and of the degree of surprise of B, should A prove true. The disbelief notion introduced later by Spohn [130] employs the same type of convention as potential surprise, but using the set of natural integers as a disbelief scale; his conditioning rule uses the subtraction of natural integers.

D. Lewis In his 1973 book [109] the philosopher David Lewis considers a graded notion of possibility in the form of a relation between possible worlds he calls *comparative possibility*. He equates this concept of possibility to a notion of similarity between possible worlds. This non-symmetric notion of similarity is also comparative, and is meant to express statements of the form: *a world j is at least as similar to world i as world k is.* Comparative similarity of j and k with respect to i is interpreted as the comparative possibility of j with respect to k viewed from world i. Such relations are assumed to be complete pre-orderings and are instrumental in defining the truth conditions of counterfactual statements. Comparative possibility relations $\geq_\Pi$ obey the key axiom: for all events A, B, C,

$$A \geq_\Pi B \text{ implies } C \cup A \geq_\Pi C \cup B.$$

This axiom was later independently proposed by the first author [42] in an attempt to derive a possibilistic counterpart to comparative probabilities. Independently, the connection

between numerical possibility and similarity was investigated by Sudkamp [131].

L. J. Cohen   A framework very similar to the one of Shackle was proposed by the philosopher L. J. Cohen [32] who considered the problem of legal reasoning. He introduced so-called *Baconian probabilities* understood as degrees of provability. The idea is that it is hard to prove someone guilty at the court of law by means of pure statistical arguments. The basic feature of degrees of provability is that a hypothesis and its negation cannot both be provable together to any extent (the contrary being a case for inconsistency). Such degrees of provability coincide with necessity measures.

L. A. Zadeh   In his seminal paper [142] Zadeh proposed an interpretation of membership functions of fuzzy sets as possibility distributions encoding flexible constraints induced by natural language statements. Zadeh articulated the relationship between possibility and probability, noticing that what is probable must preliminarily be possible. However, the view of possibility degrees developed in his paper refers to the idea of graded feasibility (degrees of ease, as in the example of "how many eggs can Hans eat for his breakfast") rather than to the epistemic notion of plausibility laid bare by Shackle. Nevertheless, the key axiom of "maxitivity" for possibility measures is highlighted. In two subsequent articles [143, 144], Zadeh acknowledged the connection between possibility theory, belief functions and upper/lower probabilities, and proposed their extensions to fuzzy events and fuzzy information granules.

3   Basic Notions of Possibility Theory.

The basic building blocks of possibility theory were first described in the authors' book [62], then more extensively in [67] and [105]. More recent accounts are in [74, 61][1]. Let S be a set of states of affairs (or descriptions thereof), or states for short. A possibility distribution is a mapping π from S to a totally ordered scale L, with top 1 and bottom 0, such as the unit interval. The function π represents the state of knowledge of an agent (about the actual state of affairs) distinguishing what is plausible from what is less plausible, what is the normal course of things from what is not, what is surprising from what is expected. It represents a flexible restriction on what is the actual state with the following conventions (similar to probability, but opposite to Shackle's potential surprise scale):

- $\pi(s) = 0$ means that state s is rejected as impossible;

- $\pi(s) = 1$ means that state s is totally possible (= plausible).

If S is exhaustive, at least one of the elements of S should be the actual world, so that $\exists s, \pi(s) = 1$ (normalisation). Distinct values may simultaneously have a degree of possibility equal to 1.

---

[1]See also $http://www.scholarpedia.org/article/Possibility\_theory$.

Possibility theory is driven by the principle of minimal specificity. It states that any hypothesis not known to be impossible cannot be ruled out. A possibility distribution π is said to be at least as specific as another π' if and only if for each state of affairs s: $\pi(s) \leq \pi'(s)$ (Yager [141]). Then, π is at least as restrictive and informative as π'.

In the possibilistic framework, extreme forms of partial knowledge can be captured, namely:

- Complete knowledge: for some $s_0, \pi(s_0) = 1$ and $\pi(s) = 0, \forall s \neq s_0$ (only s_0 is possible)

- Complete ignorance: $\pi(s) = 1, \forall s \in S$ (all states are possible).

Given a simple query of the form "does event A occur?" where A is a subset of states, the response to the query can be obtained by computing degrees of possibility and necessity, respectively (if the possibility scale $L = [0, 1]$):

$$\Pi(A) = \sup_{s \in A} \pi(s); \ N(A) = \inf_{s \notin A} 1 - \pi(s).$$

$\Pi(A)$ evaluates to what extent A is consistent with π, while $N(A)$ evaluates to what extent A is certainly implied by π. The possibility-necessity duality is expressed by $N(A) = 1 - \Pi(A^c)$, where A^c is the complement of A. Generally, $\Pi(S) = N(S) = 1$ and $\Pi(\emptyset) = N(\emptyset) = 0$. Possibility measures satisfy the basic "maxitivity" property $\Pi(A \cup B) = \max(\Pi(A), \Pi(B))$. Necessity measures satisfy an axiom dual to that of possibility measures, namely $N(A \cap B) = \min(N(A), N(B))$. On infinite spaces, these axioms must hold for infinite families of sets.

Human knowledge is often expressed in a declarative way using statements to which belief degrees are attached. It corresponds to expressing constraints the world is supposed to comply with. Certainty-qualified pieces of uncertain information of the form "A is certain to degree α" can then be modeled by the constraint $N(A) \geq \alpha$. The least specific possibility distribution reflecting this information is [67]:

$$\pi_{(A,\alpha)}(s) = \left\{ \begin{array}{ll} 1, & \text{if } s \in A \\ 1 - \alpha & \text{otherwise} \end{array} \right\} \tag{1}$$

This possibility distribution is a key-building lock to construct possibility distributions. Acquiring further pieces of knowledge leads to updating $\pi_{(A,\alpha)}$ into some $\pi < \pi_{(A,\alpha)}$.

Apart from Π and N, a measure of *guaranteed possibility* can be defined [71, 54] : $\Delta(A) = \inf_{s \in A} \pi(s)$. It estimates to what extent *all* states in A are actually possible according to evidence. $\Delta(A)$ can be used as a degree of evidential support for A. Uncertain statements of the form "A is possible to degree β" often mean that all realizations of A are possible to degree β. They can then be modeled by the constraint $\Delta(A) \geq \beta$. It corresponds to the idea of observed evidence. This type of information is better exploited by assuming an informational principle opposite to the one of minimal specificity, namely,

any situation not yet observed is tentatively considered as impossible. This is similar to closed-world assumption. The most specific distribution $\delta_{(A,\beta)}$ in agreement with $\Delta(A) \geq \beta$ is :

$$\delta_{(A,\beta)}(s) = \left\{ \begin{array}{ll} \beta, & \text{if } s \in A \\ 0 & \text{otherwise.} \end{array} \right\}$$

Acquiring further pieces of evidence leads to updating $\delta_{(A,\beta)}$ into some wider distribution $\delta > \delta_{(A,\beta)}$. Such evidential support functions do not behave with the same conventions as possibility distributions: $\delta(s) = 1$ means that S is guaranteed to be possible, because of a high evidential support, while $\delta(s) = 0$ only means that S has not been observed yet (hence is of unknown possibility). Distributions δ are generally not normalised to 1, and serve as lower bounds to possibility distributions π (because what is observed must be possible). Such a bipolar representation of information using pairs (δ, π) may provide a natural interpretation of interval-valued fuzzy sets [77]. Note that possibility distributions induced from certainty-qualified pieces of knowledge combine conjunctively, by discarding possible states, while evidential support distributions induced by possibility-qualified pieces of evidence combine disjunctively, by accumulating possible states.

Possibility theory has enabled a typology of fuzzy rules to be laid bare, distinguishing rules whose purpose is to propagate uncertainty through reasoning steps, from rules whose main purpose is similarity-based interpolation [72], depending on the choice of a many-valued implication connective that models a rule. The bipolar view of information based on (δ, π) pairs sheds new light on the debate between conjunctive and implicative representation of rules [88]. Representing a rule as a material implication focuses on counterexamples to rules, while using a conjunction between antecedent and consequent points out examples of the rule and highlights its positive content. Traditionally in fuzzy control and modelling, the latter representation is adopted, while the former is the logical tradition. Introducing fuzzy implicative rules in modelling accounts for constraints or landmark points the model should comply with (as opposed to observed data) [93]. The bipolar view of rules in terms of examples and counterexamples may turn out to be very useful when extracting fuzzy rules from data [57].

Notions of conditioning and independence were studied for possibility measures. Conditional possibility is defined similarly to probability theory using a Bayesian-like equation of the form [67]

$$\Pi(B \cap A) = \Pi(B \mid A) \star \Pi(A).$$

However, in the ordinal setting the operation $\star$ cannot be a product and is changed into the minimum. In the numerical setting, there are several ways to define conditioning, not all of which have this form, as seen later in this paper. There are also several variants of possibilistic independence [35, 34, 46]. Generally, independence in possibility theory is neither symmetric, nor insensitive to negation. For Boolean variables, independence between events is not equivalent to independence between variables.

An important example of a possibility distribution is the fuzzy interval, which is a fuzzy set of the real line whose cuts are intervals [62, 67]. The calculus of fuzzy intervals is an extension of interval arithmetics based on a possibilistic counterpart of a computation of random variable. To compute the addition of two fuzzy intervals A and B one has to compute the membership function of $A \oplus B$ as the degree of possibility $\mu_{A \oplus B}(z) = \Pi(\{(x,y) : x + y = z\})$, based on the possibility distribution $\min(\mu_A(x), \mu_B(y))$. There is a large literature on possibilistic interval analysis; see [58] for a survey of XXth century references.

4   Qualitative Possibility Theory

This section is restricted to the case of a finite state space S, supposed to be the set of interpretations of a formal propositional language. In other words, S is the universe induced by Boolean attributes. A plausibility ordering is a complete pre-order of states denoted by $\geq_\pi$, which induces a well-ordered partition $\{E_1, \cdots, E_n\}$ of S. It is the comparative counterpart of a possibility distribution π, i.e., $s \geq_\pi s'$ if and only if $\pi(s) \geq \pi(s')$. Indeed it is more natural to expect that an agent will supply ordinal rather than numerical information about his beliefs. By convention E_1 contains the most normal states of fact, E_n the least plausible, or most surprising ones. Denoting by $\max(A)$ any most plausible state $s_0 \in A$, ordinal counterparts of possibility and necessity measures [42] are then defined as follows: $\{s\} \geq_\Pi \emptyset$ for all $s \in S$ and

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

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

Possibility relations $\geq_\Pi$ are those of Lewis [109] and satisfy his characteristic property

$$A \geq_\Pi B \text{ implies } C \cup A \geq_\Pi C \cup B$$

while necessity relations can also be defined as $A \geq_N B$ if and only if $B^c \geq_\Pi A^c$, and satisfy a similar axiom:

$$A \geq_N B \text{ implies } C \cap A \geq_N C \cap B.$$

The latter coincide with epistemic entrenchment relations in the sense of belief revision theory [92, 69]. Conditioning a possibility relation $\geq_\Pi$ by an non-impossible event $C >_\Pi \emptyset$ means deriving a relation $\geq_\Pi^C$ such that

$$A \geq_\Pi^C B \text{ if and only if } A \cap C \geq_\Pi B \cap C.$$

The notion of independence for comparative possibility theory was studied in Dubois et al. [46], for independence between events, and Ben Amor et al. [11] between variables.

4.1 Nonmonotonic Inference

Suppose S is equipped with a plausibility ordering. The main idea behind qualitative possibility theory is that the state of the world is always believed to be as normal as possible, neglecting less normal states. $A \geq_\Pi B$ really means that there is a normal state where A holds that is at least as normal as any normal state where B holds. The dual case $A \geq_N B$ is intuitively understood as "A is at least as certain as B", in the sense that there are states where B fails to hold that are at least as normal as the most normal state where A does not hold. In particular, the events accepted as true are those which are true in all the most plausible states, namely the ones such that $A >_N \emptyset$. These assumptions lead us to interpret the plausible inference $A \mid\approx B$ of a proposition B from another A, under a state of knowledge $\geq_\Pi$ as follows: *B should be true in all the most normal states were A is true*, which means $B >_\Pi^A B^c$ in terms of ordinal conditioning, that is, $A \cap B$ is more plausible than $A \cap B^c$. $A \mid\approx B$ also means that the agent considers B as an accepted belief in the context A.

This kind of inference is nonmonotonic in the sense that $A \mid\approx B$ does not always imply $A \cap C \mid\approx B$ for any additional information C. This is similar to the fact that a conditional probability $P(B \mid A \cap C)$ may be low even if $P(B \mid A)$ is high. The properties of the consequence relation $\mid\approx$ are now well-understood, and are precisely the ones laid bare by Lehmann and Magidor [108] for their so-called "rational inference". Monotonicity is only partially restored: $A \mid\approx B$ implies $A \cap C \mid\approx B$ holds provided that $A \mid\approx C^c$ does not hold (i.e. that states were A is true do not typically violate C). This property is called *rational monotony*, and, along with some more standard ones (like closure under conjunction), characterizes default possibilistic inference $\mid\approx$. In fact, the set $\{B, A \mid\approx B\}$ of accepted beliefs in the context A is deductively closed, which corresponds to the idea that the agent reasons with accepted beliefs in each context as if they were true, until some event occurs that modifies this context. This closure property is enough to justify a possibilistic approach [52] and adding the rational monotonicity property ensures the existence of a single possibility relation generating the consequence relation $\mid\approx$[15]. Possibility theory has been studied from the point of view of cognitive psychology. Experimental results [124] suggest that there are situations where people reason about uncertainty using the rules or possibility theory, rather than with those of probability theory.

Plausibility orderings can be generated by a set of if-then rules tainted with unspecified exceptions. This set forms a knowledge base supplied by an agent. Each rule "if A then B" is understood as a constraint of the form $A \cap B >_\Pi A \cap B^c$ on possibility relations. There exists a single minimally specific element in the set of possibility relations satisfying all constraints induced by rules (unless the latter are inconsistent). It corresponds to the most compact plausibility ranking of states induced by the rules [15]. This ranking can be computed by an algorithm originally proposed by Pearl [118].

4.2 Possibilistic Logic

Qualitative possibility relations can be represented by (and only by) possibility measures ranging on any totally ordered set L (especially a finite one) [42]. This absolute representation on an ordinal scale is slightly more expressive than the purely relational one. When the finite set S is large and generated by a propositional language, qualitative possibility distributions can be efficiently encoded in possibilistic logic [90, 59, 75]. A possibilistic logic base K is a set of pairs (ϕ, α), where ϕ is a Boolean expression and α is an element of L. This pair encodes the constraint $N(\phi) \geq \alpha$ where $N(\phi)$ is the degree of necessity of the set of models of ϕ. Each prioritized formula (ϕ, α) has a fuzzy set of models (described in Section 3) and the fuzzy intersection of the fuzzy sets of models of all prioritized formulas in K yields the associated plausibility ordering on S.

Syntactic deduction from a set of prioritized clauses is achieved by refutation using an extension of the standard resolution rule, whereby $(\phi \vee \psi, \min(\alpha, \beta))$ can be derived from $(\phi \vee \xi, \alpha)$ and $(\psi \vee \neg \xi, \beta)$. This rule, which evaluates the validity of an inferred proposition by the validity of the weakest premiss, goes back to Theophrastus, a disciple of Aristotle. Possibilistic logic is an inconsistency-tolerant extension of propositional logic that provides a natural semantic setting for mechanizing non-monotonic reasoning [17], with a computational complexity close to that of propositional logic.

Another compact representation of qualitative possibility distributions is the possibilistic directed graph, which uses the same conventions as Bayesian nets, but relies on an ordinal notion of conditional possibility [67]

$$\Pi(B \mid A) = \left\{ \begin{array}{cc} 1, & \text{if } \Pi(B \cap A) = \Pi(A) \\ \Pi(B \cap A) & \text{otherwise.} \end{array} \right\}$$

Joint possibility distributions can be decomposed into a conjunction of conditional possibility distributions (using minimum) in a way similar to Bayes nets [14]. It is based on a symmetric notion of qualitative independence $\Pi(B \cap A) = \min(\Pi(A), \Pi(B))$ that is weaker than the causal-like condition $\Pi(B \mid A) = \Pi(B)$ [46]. Ben Amor and Benferhat [12] investigate the properties of qualitative independence that enable local inferences to be performed in possibilistic nets. Uncertainty propagation algorithms suitable for possibilistic graphical structures have been studied [13].

Other types of possibilistic logic can also handle constraints of the form $\Pi(\phi) \geq \alpha$, or $\Delta(\phi) \geq \alpha$ [75]. Possibilistic logic can be extended to logic programming [1, 10], similarity reasoning [2], and many-valued logic as extensively studied by Godo and colleagues [38].

4.3 Decision-theoretic foundations

Zadeh [142] hinted that "since our intuition concerning the behaviour of possibilities is not very reliable", our understanding of them "would be enhanced by the development of an axiomatic approach to the definition of subjective possibilities in the spirit of axiomatic

approaches to the definition of subjective probabilities". Decision-theoretic justifications of qualitative possibility were devised, in the style of Savage [125] more than 10 years ago. On top of the set of states, assume there is a set X of consequences of decisions. A decision, or act, is modeled as a mapping f from S to X assigning to each state S its consequence $f(s)$. The axiomatic approach consists in proposing properties of a preference relation $\succeq$ between acts so that a representation of this relation by means of a preference functional $W(f)$ is ensured, that is, act f is as good as act g (denoted $f \succeq g$) if and only if $W(f) \geq W(g)$. $W(f)$ depends on the agent's knowledge about the state of affairs, here supposed to be a possibility distribution π on S, and the agent's goal, modeled by a utility function u on X. Both the utility function and the possibility distribution map to the same finite chain L. A pessimistic criterion $W_\pi^-(f)$ is of the form:

$$W_\pi^-(f) = \min_{s \in S} \max(n(\pi(s)), u(f(s)))$$

where n is the order-reversing map of L. $n(\pi(s))$ is the degree of certainty that the state is not s (hence the degree of surprise of observing s), $u(f(s))$ the utility of choosing act f in state s. $W_\pi^-(f)$ is all the higher as all states are either very surprising or have high utility. This criterion is actually a prioritized extension of the Wald maximin criterion. The latter is recovered if $\pi(s) = 1$ (top of L) $\forall s \in S$. According to the pessimistic criterion, acts are chosen according to their worst consequences, restricted to the most plausible states $S^* = \{s, \pi(s) \geq n(W_\pi^-(f))\}$. The optimistic counterpart of this criterion is:

$$W_\pi^+(f) = \max_{s \in S} \min(\pi(s)), u(f(s))).$$

$W_\pi^+(f)$ is all the higher as there is a very plausible state with high utility. The optimistic criterion was first proposed by Yager [139] and the pessimistic criterion by Whalen [138]. These optimistic and pessimistic possibilistic criteria are particular cases of a more general criterion based on the Sugeno integral [97] specialized to possibility and necessity of fuzzy events [142, 62]:

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

where $F_\lambda = \{s \in S, u(f(s)) \geq \lambda\}$, γ is a monotonic set function that reflects the decision-maker attitude in front of uncertainty: $\gamma(A)$ is the degree of confidence in event A. If $\gamma = \Pi$, then $S_{\Pi,u}(f) = W_\pi^+(f)$. Similarly, if $\gamma = N$, then $S_{N,u}(f) = W_\pi^-(f)$.

For any acts f, g, and any event A, let fAg denote an act consisting of choosing f if A occurs and g if its complement occurs. Let $f \wedge g$ (resp. $f \vee g$) be the act whose results yield the worst (resp. best) consequence of the two acts in each state. Constant acts are those whose consequence is fixed regardless of the state. A result in [82, 83] provides an act-driven axiomatization of these criteria, and enforces possibility theory as a "rational" representation of uncertainty for a finite state space S:

Theorem 1. *Suppose the preference relation $\succeq$ on acts obeys the following properties:*

1. $(X^S, \succeq)$ is a complete preorder.

2. There are two acts such that $f \succ g$.

3. $\forall A, \forall g$ and h constant, $\forall f, g \succeq h$ implies $gAf \succeq hAf$.

4. If f is constant, $f \succ h$ and $g \succ h$ imply $f \wedge g \succ h$.

5. If f is constant, $h \succ f$ and $h \succ g$ imply $h \succ f \vee g$.

then there exists a finite chain L, an L-valued monotonic set-function γ on S and an L-valued utility function u, such that $\succeq$ is representable by a Sugeno integral of $u(f)$ with respect to γ. Moreover γ is a necessity (resp. possibility) measure as soon as property (4) (resp. (5)) holds for all acts. The preference functional is then $W_\pi^-(f)$ (resp. $W_\pi^+(f)$).

Axioms (4-5) contradict expected utility theory. They become reasonable if the value scale is finite, decisions are one-shot (no compensation) and provided that there is a big step between any level in the qualitative value scale and the adjacent ones. In other words, the preference pattern $f \succ h$ always means that f is significantly preferred to h, to the point of considering the value of h negligible in front of the value of f. The above result provides decision-theoretic foundations of possibility theory, whose axioms can thus be tested from observing the choice behavior of agents. See [49] for another approach to comparative possibility relations, more closely relying on Savage axioms, but giving up any comparability between utility and plausibility levels. The drawback of these and other qualitative decision criteria is their lack of discrimination power [47]. To overcome it, refinements of possibilistic criteria were recently proposed, based on lexicographic schemes [89]. These new criteria turn out to be representable by a classical (but big-stepped) expected utility criterion. Qualitative possibilistic counterparts of influence diagrams for decision trees have been recently investigated [98].

More recently, possibilistic qualitative bipolar decision criteria have been defined, axiomatized [48] and empirically tested [23]. They are qualitative counterparts of cumulative prospect theory criteria of Kahneman and Tverski [133].

5   Quantitative Possibility Theory

The phrase "quantitative possibility" refers to the case when possibility degrees range in the unit interval. In that case, a precise articulation between possibility and probability theories is useful to provide an interpretation to possibility and necessity degrees. Several such interpretations can be consistently devised: a degree of possibility can be viewed as an upper probability bound [70], and a possibility distribution can be viewed as a likelihood function [60]. A possibility measure is also a special case of a Shafer plausibility function [126]. Following a very different approach, possibility theory can account for probability distributions with extreme values, infinitesimal [130] or having big steps [16]. There are

finally close connections between possibility theory and idempotent analysis [113]. The theory of large deviations in probability theory [123] also handles set-functions that look like possibility measures [117]. Here we focus on the role of possibility theory in the theory of imprecise probability.

5.1 Possibility as upper probability

Let π be a possibility distribution where $\pi(s) \in [0, 1]$. Let $\mathbf{P}(\pi)$ be the set of probability measures P such that $P \leq \Pi$, i.e. $\forall A \subseteq S, P(A) \leq \Pi(A)$. Then the possibility measure Π coincides with the upper probability function P^* such that $P^*(A) = \sup\{P(A), P \in \mathbf{P}(\pi)\}$ while the necessity measure N is the lower probability function P_* such that $P_*(A) = \inf\{P(A), P \in \mathbf{P}(\pi)\}$; see [70, 36] for details. P and π are said to be consistent if $P \in \mathbf{P}(\pi)$. The connection between possibility measures and imprecise probabilistic reasoning is especially promising for the efficient representation of non-parametric families of probability functions, and it makes sense even in the scope of modeling linguistic information [136].

A possibility measure can be computed from nested confidence subsets $\{A_1, A_2, \ldots, A_m\}$ where $A_i \subset A_{i+1}, i = 1 \ldots m - 1$. Each confidence subset A_i is attached a positive confidence level λ_i interpreted as a lower bound of $P(A_i)$, hence a necessity degree. It is viewed as a certainty-qualified statement that generates a possibility distribution π_i according to Section 3. The corresponding possibility distribution is

$$\pi(s) = \min_{i=1,\ldots,m} \pi_i(s) = \left\{ \begin{array}{cc} 1 & \text{if } u \in A_1 \\ 1 - \lambda_{j-1} & \text{if } j = \max\{i : s \notin A_i\} > 1 \end{array} \right\}$$

The information modeled by π can also be viewed as a nested random set $\{(A_i, \nu_i), i = 1, \ldots, m\}$, where $\nu_i = \lambda_i - \lambda_{i-1}$. This framework allows for imprecision (reflected by the size of the A_i's) and uncertainty (the ν_i's). And ν_i is the probability that the agent only knows that A_i contains the actual state (it is not $P(A_i)$). The random set view of possibility theory is well adapted to the idea of imprecise statistical data, as developed in [94, 103]. Namely, given a bunch of imprecise (not necessarily nested) observations (called focal sets), π supplies an approximate representation of the data, as $\pi(s) = \sum_{i:s \in A_i} \nu_i$.

The set $\mathbf{P}(\pi)$ contains many probability distributions, arguably too many. Neumaier [116] has recently proposed a related framework, in a different terminology, for representing smaller subsets of probability measures using two possibility distributions instead of one. He basically uses a pair of distributions (δ, π) (in the sense of Section 3) of distributions, he calls "cloud", where δ is a guaranteed possibility distribution (in our terminology) such that $\pi \geq \delta$. A cloud models the (generally non-empty) set $\mathbf{P}(\pi) \cap \mathbf{P}(1 - \delta)$, viewing $1 - \delta$ as a standard possibility distribution. The precise connections between possibility distributions, clouds and other simple representations of numerical uncertainty is studied in [39].

11

5.2 Conditioning

There are two kinds of conditioning that can be envisaged upon the arrival of new information E. The first method presupposes that the new information alters the possibility distribution π by declaring all states outside E impossible. The conditional measure $\pi(. \mid E)$ is such that $\Pi(B \mid E) \cdot \Pi(E) = \Pi(B \cap E)$. This is formally Dempster rule of conditioning of belief functions, specialised to possibility measures. The conditional possibility distribution representing the weighted set of confidence intervals is,

$$\pi(s \mid E) = \left\{ \begin{array}{ll} \frac{\pi(s)}{\Pi(E)}, & \text{if } s \in E \\ 0 & \text{otherwise.} \end{array} \right\}$$

De Baets et al. [33] provide a mathematical justification of this notion in an infinite setting, as opposed to the min-based conditioning of qualitative possibility theory. Indeed, the maxitivity axiom extended to the infinite setting is not preserved by the min-based conditioning. The product-based conditioning leads to a notion of independence of the form $\Pi(B \cap E) = \Pi(B) \cdot \Pi(E)$ whose properties are very similar to the ones of probabilistic independence [34].

Another form of conditioning [73, 37], more in line with the Bayesian tradition, considers that the possibility distribution π encodes imprecise statistical information, and event E only reflects a feature of the current situation, not of the state in general. Then the value $\Pi(B \mid\mid E) = \sup\{P(B \mid E), P(E) > 0, P \leq \Pi\}$ is the result of performing a sensitivity analysis of the usual conditional probability over $\mathbf{P}(\pi)$ (Walley [135]). Interestingly, the resulting set-function is again a possibility measure, with distribution

$$\pi(s \mid\mid E) = \left\{ \begin{array}{ll} \max(\pi(s), \frac{\pi(s)}{\pi(s)+N(E)}), & \text{if } s \in E \\ 0 & \text{otherwise.} \end{array} \right\}$$

It is generally less specific than π on E, as clear from the above expression, and becomes non-informative when $N(E) = 0$ (i.e. if there is no information about E). This is because $\pi(\cdot \mid\mid E)$ is obtained from the focusing of the generic information π over the reference class E. On the contrary, $\pi(\cdot \mid E)$ operates a revision process on π due to additional knowledge asserting that states outside E are impossible. See De Cooman [37] for a detailed study of this form of conditioning.

5.3 Probability-possibility transformations

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). It is useful to cast all pieces of information in the same framework. The basic requirement is to respect the consistency principle $\Pi \geq P$. The problem is then either to pick a probability measure in $\mathbf{P}(\pi)$, or to construct a possibility measure dominating P.

There are two basic approaches to possibility/probability transformations, which both respect a form of probability-possibility consistency. One, due to Klir [106, 96] is based on a principle of information invariance, the other [84] is based on optimizing information content. Klir assumes that possibilistic and probabilistic information measures are commensurate. Namely, the choice between possibility and probability is then a mere matter of translation between languages "neither of which is weaker or stronger than the other" (quoting Klir and Parviz [107]). It suggests that entropy and imprecision capture the same facet of uncertainty, albeit in different guises. The other approach, recalled here, considers that going from possibility to probability leads to increase the precision of the considered representation (as we go from a family of nested sets to a random element), while going the other way around means a loss of specificity.

From possibility to probability The most basic example of transformation from possibility to probability is the Laplace principle of insufficient reason claiming that what is equally possible should be considered as equally probable. A generalised Laplacean indifference principle is then adopted in the general case of a possibility distribution π: the weights ν_i bearing the sets A_i from the nested family of levels cuts of π are uniformly distributed on the elements of these cuts A_i. Let P_i be the uniform probability measure on A_i. The resulting probability measure is $P = \sum_{i=1,\ldots,m} \nu_i \cdot P_i$. This transformation, already proposed in 1982 [63] comes down to selecting the center of gravity of the set $\mathbf{P}(\pi)$ of probability distributions dominated by π. This transformation also coincides with Smets' pignistic transformation [129] and with the Shapley value of the "unamimity game" (another name of the necessity measure) in game theory. The rationale behind this transformation is to minimize arbitrariness by preserving the symmetry properties of the representation. This transformation from possibility to probability is one-to-one. Note that the definition of this transformation does not use the nestedness property of cuts of the possibility distribution. It applies all the same to non-nested random sets (or belief functions) defined by pairs $\{(A_i, \nu_i), i = 1, \ldots, m\}$, where ν_i are non-negative reals such that $\sum_{i=1,\ldots,m} \nu_i = 1$.

From objective probability to possibility From probability to possibility, the rationale of the transformation is not the same according to whether the probability distribution we start with is subjective or objective [86]. In the case of a statistically induced probability distribution, the rationale is to preserve as much information as possible. This is in line with the handling of Δ-qualified pieces of information representing observed evidence, considered in section 3; hence we select as the result of the transformation of a probability measure P, the most specific possibility measure in the set of those dominating P [84]. This most specific element is generally unique if P induces a linear ordering on S. Suppose S is a finite set. The idea is to let $\Pi(A) = P(A)$, for these sets A having minimal probability among other sets having the same cardinality as A. If $p_1 > p_2 > \cdots > p_n$, then $\Pi(A) = P(A)$ for sets A of the form $\{s_i, \ldots, s_n\}$, and the possibility distribution is

defined as $\pi_P(s_i) = \sum_{j=i,...,m} p_j$, with $p_j = P(\{s_j\})$. Note that π_P is a kind of cumulative distribution of P, already known as a Lorentz curve in the mathematical literature [112]. If there are equiprobable elements, the unicity of the transformation is preserved if equipossibility of the corresponding elements is enforced. In this case it is a bijective transformation as well. Recently, this transformation was used to prove a rather surprising agreement between probabilistic indeterminateness as measured by Shannon entropy, and possibilistic non-specificity. Namely it is possible to compare probability measures on finite sets in terms of their relative *peakedness* (a concept adapted from Birnbaum [21]) by comparing the relative specificity of their possibilistic transforms. Namely let P and Q be two probability measures on S and π_P, π_Q the possibility distributions induced by our transformation. It can be proved that if $\pi_P \geq \pi_Q$ (i.e. P is less peaked than Q) then the Shannon entropy of P is higher than the one of Q [55]. This result give some grounds to the intuitions developed by Klir [106], without assuming any commensurability between entropy and specificity indices.

Possibility distributions induced by prediction intervals In the continuous case, moving from objective probability to possibility means adopting a representation of uncertainty in terms of prediction intervals around the mode viewed as the "most frequent value". Extracting a prediction interval from a probability distribution or devising a probabilistic inequality can be viewed as moving from a probabilistic to a possibilistic representation. Namely suppose a non-atomic probability measure P on the real line, with unimodal density p, and suppose one wishes to represent it by an interval I with a prescribed level of confidence $P(I) = \gamma$ of hitting it. The most natural choice is the most precise interval ensuring this level of confidence. It can be proved that this interval is of the form of a cut of the density, i.e. $I_\gamma = \{s, p(s) \geq \theta\}$ for some threshold θ. Moving the degree of confidence from 0 to 1 yields a nested family of prediction intervals that form a possibility distribution π consistent with P, the most specific one actually, having the same support and the same mode as P and defined by ([84]):

$$\pi(\inf I_\gamma) = \pi(\sup I_\gamma) = 1 - \gamma = 1 - P(I_\gamma)$$

This kind of transformation again yields a kind of cumulative distribution according to the ordering induced by the density p. Similar constructs can be found in the statistical literature (Birnbaum [21]). More recently Mauris et al. [81] noticed that starting from any family of nested sets around some characteristic point (the mean, the median,...), the above equation yields a possibility measure dominating P. Well-known inequalities of probability theory, such as those of Chebyshev and Camp-Meidel, can also be viewed as possibilistic approximations of probability functions. It turns out that for symmetric unimodal densities, each side of the optimal possibilistic transform is a convex function. Given such a probability density on a bounded interval $[a, b]$, the triangular fuzzy number whose core is the mode of p and the support is $[a, b]$ is thus a possibility distribution

14

dominating P regardless of its shape (and the tightest such distribution). These results justify the use of symmetric triangular fuzzy numbers as fuzzy counterparts to uniform probability distributions. They provide much tighter probability bounds than Chebyshev and Camp-Meidel inequalities for symmetric densities with bounded support. This setting is adapted to the modelling of sensor measurements [115]. These results are extended to more general distributions by Baudrit et al., [7], and provide a tool for representing poor probabilistic information. More recently, Mauris [114] unifies, by means of possibility theory, many old techniques independently developed in statistics for one-point estimation, relying on the idea of dispersion of an empirical distribution. The efficiency of different estimators can be compared by means of fuzzy set inclusion applied to optimal possibility transforms of probability distributions. This unified approach does not presuppose a finite variance.

Subjective possibility distributions The case of a subjective probability distribution is different. Indeed, the probability function is then supplied by an agent who is in some sense forced to express beliefs in this form due to rationality constraints, and the setting of exchangeable bets. However his actual knowledge may be far from justifying the use of a single well-defined probability distribution. For instance in case of total ignorance about some value, apart from its belonging to an interval, the framework of exchangeable bets enforces a uniform probability distribution, on behalf of the principle of insufficient reason. Based on the setting of exchangeable bets, it is possible to define a subjectivist view of numerical possibility theory, that differs from the proposal of Walley [135]. The approach developed by Dubois, Prade and Smets [87] relies on the assumption that when an agent constructs a probability measure by assigning prices to lotteries, this probability measure is actually induced by a belief function representing the agents actual state of knowledge. We assume that going from an underlying belief function to an elicited probability measure is achieved by means of the above mentioned pignistic transformation, changing focal sets into uniform probability distributions. The task is to reconstruct this underlying belief function under a minimal commitment assumption. In the paper [87], we pose and solve the problem of finding the least informative belief function having a given pignistic probability. We prove that it is unique and consonant, thus induced by a possibility distribution. The obtained possibility distribution can be defined as the converse of the pignistic transformation (which is one-to-one for possibility distributions). It is subjective in the same sense as in the subjectivist school in probability theory. However, it is the least biased representation of the agents state of knowledge compatible with the observed betting behaviour. In particular it is less specific than the one constructed from the prediction intervals of an objective probability. This transformation was first proposed in [64] for objective probability, interpreting the empirical necessity of an event as summing the excess of probabilities of realizations of this event with respect to the probability of the most likely realization of the opposite event.

Possibility theory and defuzzification Possibilistic mean values can be defined using Choquet integrals with respect to possibility and necessity measures [65, 37], and come close to defuzzification methods [134]. A fuzzy interval is a fuzzy set of reals whose membership function is unimodal and upper-semi continuous. Its α-cuts are closed intervals. Interpreting a fuzzy interval M, associated to a possibility distribution μ_M, as a family of probabilities, upper and lower mean values $E^*(M)$ and $E_*(M)$, can be defined as [66]:

$$E_*(M) = \int_0^1 \inf M_\alpha d\alpha; \quad E^*(M) = \int_0^1 \sup M_\alpha d\alpha$$

where M_α is the α-cut of M.

Then the mean interval $E(M) = [E_*(M), E^*(M)]$ of M is the interval containing the mean values of all random variables consistent with M, that is $E(M) = \{E(P) \mid P \in \mathbf{P}(\mu_M)\}$, where $E(P)$ represents the expected value associated to the probability measure P. That the "mean value" of a fuzzy interval is an interval seems to be intuitively satisfactory. Particularly the mean interval of a (regular) interval $[a, b]$ is this interval itself. The upper and lower mean values are linear with respect to the addition of fuzzy numbers. Define the addition $M + N$ as the fuzzy interval whose cuts are $M_\alpha + N_\alpha = \{s+t, s \in M_\alpha, t \in N_\alpha\}$ defined according to the rules of interval analysis. Then $E(M + N) = E(M) + E(N)$, and similarly for the scalar multiplication $E(aM) = aE(M)$, where aM has membership grades of the form $\mu_M(s/a)$ for $a \neq 0$. In view of this property, it seems that the most natural defuzzication method is the middle point $\hat{E}(M)$ of the mean interval (originally proposed by Yager [140]). Other defuzzification techniques do not generally possess this kind of linearity property. $\hat{E}(M)$ has a natural interpretation in terms of simulation of a fuzzy variable [28], and is the mean value of the pignistic transformation of M. Indeed it is the mean value of the empirical probability distribution obtained by the random process defined by picking an element α in the unit interval at random, and then an element s in the cut M_α at random.

6   Some Applications

Possibility theory has not been the main framework for engineering applications of fuzzy sets in the past. However, on the basis of its connections to symbolic artificial intelligence, to decision theory and to imprecise statistics, we consider that it has significant potential for further applied developments in a number of areas, including some where fuzzy sets are not yet always accepted. Only some directions are pointed out here.

1. Possibility theory also offers a framework for preference modeling in constraint-directed reasoning. Both prioritized and soft constraints can be captured by possibility distributions expressing degrees of feasibility rather than plausibility [51]. Possibility offers a natural setting for fuzzy optimization whose aim is to balance the levels of satisfaction of multiple fuzzy constraints (instead of minimizing an overall

cost) [53]. Qualitative decision criteria are particularly adapted to the handling of uncertainty in this setting. Applications of possibility theory-based decision-making can be found in scheduling [50, 128, 29, 30]. Possibility distributions can also model ill-known constraint coefficients in linear and non-linear programming, thus leading to variants of chance-constrained programming [102]. Besides, the possibilistic logic setting provides a compact representation framework for preferences, which is more expressive than the CP-net approach [104].

2. Quantitative possibility theory is the natural setting for a reconciliation between probability and fuzzy sets. An important research direction is the comparison between fuzzy interval analysis [58] and random variable calculations with a view to unifying them [68]. Indeed, a current major concern, in for instance risk analysis studies, is to perform uncertainty propagation under poor data and without independence assumptions (see the papers in the special issue [100]). Finding the potential of possibilistic representations in computing conservative bounds for such probabilistic calculations is certainly a major challenge [99]. Methods for joint propagation of possibilistic and probabilistic information have been devised [9], based on casting both in a random set setting [6]; the case of probabilistic models with fuzzy interval parameters has also been dealt with [8]. The active area of fuzzy random variables is also connected to this question [95].

Other applications of possibility theory can be found in fields such as data analysis [137, 132, 24], database querying [25], diagnosis [27, 26], belief revision [18], argumentation [4, 3], case-based reasoning [56, 101], learning [120, 121], and information merging [19] (taking advantage of the bipolar representation setting which distinguishes between positive information of the form $\Delta(\phi) \geq \alpha$ and negative information expressing impossibility under the form $N(\phi) \geq \alpha \Leftrightarrow 1 - \Pi(\neg\phi) \geq \alpha$ [20]).

7 Some current research lines

A number of on-going works deal with new research lines where possibility theory is central. In the following we outline a few of those:

- *Formal concept analysis*: Formal concept analysis (FCA) studies Boolean data tables relating objects and attributes. The key issue of FCA is to extract so-called concepts from such tables. A concept is a maximal set of objects sharing a maximal number of attributes. The enumeration of such concepts can be carried out via a Galois connection between objects and attributes, and this Galois connection uses operators similar to the Δ function of possibility theory. Based on this analogy, other correspondences can be laid bare using the three other set-functions of possibility theory [45, 41]. In particular, one of these correspondences detects independent subtables

[79]. This approach can be systematized to fuzzy or uncertain versions of formal concept analysis.

- *Generalised possibilistic logic* Possibilistic logic, in its basic version, attaches degrees of necessity to formulas, which turn them into graded modal formulas of the necessity kind. However only conjunction of weighted formulas are allowed. Yet very early we noticed that it makes sense to extend the language towards handing constraints on the degree of possibility of a formula. This requires allowing for negation and disjunctions of necessity-qualified proposition. This extension, still under study [78], puts together the KD modal logic and basic possibilistic logic. Recently it has been shown that non-monotonic logic programing languages can be translated into generalized possibilistic logic, making the meaning of negation by default in rule much more transparent [85]. This move from basic to generalized possibilistic logic also enables further extensions to the multi-agent and the multi-source case [76] to be considered. Besides, it has been recently shown that a Sugeno integral can be also represented in terms of possibilistic logic, which enables us to lay bare the logical description of an aggregation process [80].

- *Qualitative capacities and possibility measures.* While a numerical possibility measure is equivalent to a convex set of probability measures, it turns out that in the qualitative setting, a monotone set-function can be represented by means of a family of possibility measures [5, 43]. This line of research enables qualitative counterparts of results in the study of Choquet capacities in the numerical settings to be established. Especially, a monotone set-function can be seen as the counterpart of a belief function, and various concepts of evidence theory can be adapted to this setting [119]. Sugeno integral can be viewed as a lower possibilistic expectation in the sense of section 4.3 [43]. These results enable the structure of qualitative monotonic set-functions to be laid bare, with possible connection with neighborhood semantics of non-regular modal logics.

- *Regression and kriging* Fuzzy regression analysis is seldom envisaged from the point of view of possibility theory. One exception is the possibilistic regression initiated by Tanaka and Guo [132], where the idea is to approximate precise or set-valued data in the sense of inclusion by means of a set-valued or fuzzy set-valued linear function obtained by making the linear coefficients of a linear function fuzzy. The alternative approach is the fuzzy least squares of Diamond [40] where fuzzy data are interpreted as functions and a crisp distance between fuzzy sets is often used. However, fuzzy data are questionably seen as objective entities [110]. The introduction of possibility theory in regression analysis of fuzzy data comes down to an epistemic view of fuzzy data whereby one tries to construct the envelope of all linear regression results that could have been obtained, had the data been precise[44]. This view has been applied to the kriging problem in geostatistics [111]. Another use of possibility theory consists in

exploiting possibility-probability transforms to develop a form of quantile regression on crisp data [122], yielding a fuzzy function that is much more faithful to the data set than what a fuzzified linear function can offer.

References

[1] T. Alsinet, L. Godo: Towards an automated deduction system for first-order possibilistic logic programming with fuzzy constants. *Int. J. Intell. Syst.*,17, 887-924, 2002.

[2] T. Alsinet, L. Godo: Adding similarity-based reasoning capabilities to a Horn fragment of possibilistic logic with fuzzy constants. *Fuzzy Sets and Systems*, 144, 43-65, 2004.

[3] T. Alsinet, C. Chesñevar, L. Godo, S. Sandri, G. Simari: Formalizing argumentative reasoning in a possibilistic logic programming setting with fuzzy unification. *Int. J. Approx. Reasoning*, 48, 711-729, 2008.

[4] L. Amgoud, H. Prade. Reaching agreement through argumentation: A possibilistic approach. *Proc. of the 9th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'04)*(Whistler, BC, Canada), AAAI Press, 175-182, 2004.

[5] G. Banon, Constructive decomposition of fuzzy measures in terms of possibility and necessity measures. *Proc. VIth IFSA World Congress*, So Paulo, Brazil, vol. I, 217-220, 1995.

[6] C. Baudrit, I. Couso, D. Dubois. Joint propagation of probability and possibility in risk analysis: Towards a formal framework. *Inter. J. of Approximate Reasoning*, 45, 82-105, 2007.

[7] C. Baudrit, D. Dubois, Practical representations of incomplete probabilistic knowledge *Computational Statistics & Data Analysis*, 51 86-108, 2006.

[8] C. Baudrit, D. Dubois, N. Perrot, Representing parametric probabilistic models tainted with imprecision. *Fuzzy Sets and Systems*, 159, 2008, 1913-1928.

[9] C. Baudrit, D. Guyonnet, D. Dubois, Joint propagation and exploitation of probabilistic and possibilistic information in risk assessment, *IEEE Trans. Fuzzy Systems*, 14, 593-608, 2006.

[10] K. Bauters, S. Schockaert, M. De Cock, D. Vermeir, D., Possibilistic answer set programming revisited. *Proc. 26th Conf. on Uncertainty in Artificial Intelligence (UAI'10)*, Catalina Island, 48-55, 2010.

[11] N. Ben Amor, K. Mellouli, S. Benferhat, D. Dubois, H. Prade, A theoretical framework for possibilistic independence in a weakly ordered setting. *Int. J. Uncert. Fuzz. & Knowl.-B. Syst.* 10, 117-155, 2002.

[12] N. Ben Amor, S. Benferhat, Graphoid properties of qualitative possibilistic independence relations. *Int. J. Uncert. Fuzz. & Knowl.-B. Syst.* 13, 59-97, 2005.

[13] N. Ben Amor, S. Benferhat, K. Mellouli, Anytime propagation algorithm for min-based possibilistic graphs. *Soft Comput.*, 8, 50-161, 2003.

[14] S. Benferhat, D. Dubois, L. Garcia and H. Prade, On the transformation between possibilistic logic bases and possibilistic causal networks. *Int. J. Approximate Reasoning*, 29:135-173, 2002.

[15] S. Benferhat, D. Dubois and H. Prade, Nonmonotonic reasoning, conditional objects and possibility theory, *Artificial Intelligence*, 92: 259-276, 1997.

[16] S. Benferhat, D. Dubois and H. Prade Possibilistic and standard probabilistic semantics of conditional knowledge bases, *J. Logic Comput.*, 9, 873-895, 1999.

[17] S. Benferhat, D. Dubois, and H. Prade, Practical Handling of exception-tainted rules and independence information in possibilistic logic *Applied Intelligence*, 9, 101-127,1998.

[18] S. Benferhat, D. Dubois, H. Prade and M.-A. Williams. A practical approach to revising prioritized knowledge bases,*Studia Logica*, 70, 105-130, 2002.

[19] S. Benferhat, D. Dubois, S. Kaci, H. Prade, Bipolar possibility theory in preference modeling: Representation, fusion and optimal solutions. *Information Fusion*, 7, 135-150, 2006.

[20] S. Benferhat, D. Dubois, S. Kaci, H. Prade, Modeling positive and negative information in possibility theory. *Int. J. Intell. Syst.*, 23, 1094-1118, 2008.

[21] Z.W. Birnbaum On random variables with comparable peakedness. *Ann. Math. Stat.* 19, 76-81, 1948.

[22] I. M. Bocheński, *La Logique de Théophraste.* Librairie de l'Université de Fribourg en Suisse. 1947.

[23] J.-F. Bonnefon, D. Dubois, H. Fargier, S. Leblois, Qualitative heuristics for balancing the pros and the cons, *Theory and Decision*, 65, 71-95, 2008.

[24] C. Borgelt, J. Gebhardt and R. Kruse, Possibilistic graphical models. In G. Della Riccia et al. editors, *Computational Intelligence in Data Mining*, Springer, Wien, pages 51-68, 2000.

[25] P. Bosc and H. Prade, An introduction to the fuzzy set and possibility theory-based treatment of soft queries and uncertain of imprecise databases. In: P. Smets, A. Motro, eds. *Uncertainty Management in Information Systems*, Dordrecht: Kluwer, 285-324, 1997.

[26] S. Boverie et al. Online diagnosis of engine dyno test benches: a possibilistic approach *Proc. 15th. Eur. Conf. on Artificial Intelligence*, Lyon. Amsterdam: IOS Press, 658-662, 2002.

[27] D. Cayrac, D. Dubois and H. Prade, Handling uncertainty with possibility theory and fuzzy sets in a satellite fault diagnosis application. *IEEE Trans. on Fuzzy Systems*, 4, 251-269, 1996.

[28] S. Chanas and M. Nowakowski, Single value simulation of fuzzy variable, *Fuzzy Sets and Systems*, 25, 43-57, 1988.

[29] S. Chanas and P. Zielinski, Critical path analysis in the network with fuzzy activity times, *Fuzzy Sets and Systems*, 122, 195-204, 2001.

[30] S. Chanas, D. Dubois, and P. Zielinski, Necessary criticality in the network with imprecise activity times. *IEEE transactions on Man, Machine and Cybernetics*, 32:393-407, 2002.

[31] B.F. Chellas *Modal Logic, an Introduction*, Cambridge University Press, Cambridge, 1980.

[32] L. J. Cohen, *The Probable and the Provable.* Oxford: Clarendon, 1977.

[33] B. De Baets, E. Tsiporkova and R. Mesiar Conditioning in possibility with strict order norms, *Fuzzy Sets and Systems*, 106, 221-229, 1999.

[34] L.M. De Campos and J.F. Huete. Independence concepts in possibility theory, *Fuzzy Sets and Systems*, 103, 127-152 & 487-506, 1999.

[35] G. De Cooman, Possibility theory. Part I: Measure- and integral-theoretic groundwork; Part II: Conditional possibility; Part III: Possibilistic independence. *Int. J. of General Syst.*, 25: 291-371, 1997.

[36] G. De Cooman and D. Aeyels, Supremum-preserving upper probabilities. *Information Sciences*, 118, 173 -212, 1999.

[37] G. De Cooman, Integration and conditioning in numerical possibility theory. *Annals of Math. and AI*, 32, 87-123, 2001.

[38] P. Dellunde, L. Godo, E. Marchioni: Extending possibilistic logic over Gdel logic. *Int. J. Approx. Reasoning*, 52, 63-75, 2011.

[39] S. Destercke D.Dubois, E. Chojnacki, Unifying practical uncertainty representations Part I: Generalized p-boxes. *Inter. J. of Approximate Reasoning*, 49, 649-663, 2008; Part II: Clouds. *Inter. J. of Approximate Reasoning*, 49, 664-677, 2008.

[40] P. Diamond (1988) Fuzzy least squares. *Information Sciences*, 46, 141-157

[41] Y. Djouadi, H. Prade: Possibility-theoretic extension of derivation operators in formal concept analysis over fuzzy lattices. *Fuzzy Optimization & Decision Making* 10, 287-309, 2011.

[42] D. Dubois, Belief structures, possibility theory and decomposable measures on finite sets. *Computers and AI*, 5, 403-416, 1986.

[43] D. Dubois Fuzzy measures on finite scales as families of possibility measures, *Proc. 7th Conf. of the Europ. Soc. for Fuzzy Logic and Technology (EUSFLAT'11)*, Annecy, Atlantis Press, 822-829, 2011.

[44] D. Dubois Ontic vs. epistemic fuzzy sets in modeling and data processing tasks, *Proc. Inter. Joint Conf. on Computational Intelligence*, Paris, pp. IS13-IS19.

[45] D. Dubois, F. Dupin de Saint-Cyr, H. Prade: A possibility-theoretic view of formal concept analysis. *Fundam. Inform.*, 75, 195-213, 2007.

[46] D. Dubois, L. Farinas del Cerro, A. Herzig and H. Prade, Qualitative relevance and independence: A roadmap, *Proc. of the 15h Inter. Joint Conf. on Artif. Intell.*, Nagoya, 62-67, 1997.

[47] D. Dubois, H. Fargier. Qualitative decision rules under uncertainty. In G. Della Riccia, et al. editors, *Planning Based on Decision Theory*, CISM courses and Lectures 472, Springer Wien, 3-26, 2003.

[48] D. Dubois, H. Fargier, J.-F. Bonnefon. On the qualitative comparison of decisions having positive and negative features. *J. of Artificial Intelligence Research*, AAAI Press, 32, 385-417, 2008.

[49] D. Dubois, H. Fargier, and P. Perny H. Prade, Qualitative decision theory with preference relations and comparative uncertainty: An axiomatic approach. *Artificial Intelligence*, 148, 219-260, 2003.

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

[51] D. Dubois, H. Fargier, and H. Prade, Possibility theory in constraint satisfaction problems: Handling priority, preference and uncertainty. *Applied Intelligence*, 6: 287-309, 1996.

[52] D. Dubois, H. Fargier, and H. Prade, Ordinal and probabilistic representations of acceptance. *J. Artificial Intelligence Research*, 22, 23-56, 2004

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

[54] D. Dubois, P. Hajek and H. Prade, Knowledge-driven versus data-driven logics. *J. Logic, Lang. and Inform.*, 9: 65–89, 2000.

[55] D. Dubois, E. Huellermeier Comparing probability measures using possibility theory: A notion of relative peakedness. Inter. J. of Approximate Reasoning, 45, 364-385, 2007

[56] D. Dubois, E. Huellermeier and H. Prade. Fuzzy set-based methods in instance-based reasoning, *IEEE Trans. on Fuzzy Systems*, 10, 322-332, 2002.

[57] D. Dubois, E. Huellermeier and H. Prade. A systematic approach to the assessment of fuzzy association rules, *Data Mining and Knowledge Discovery*, 13, 167-192, 2006

[58] D. Dubois, E. Kerre, R. Mesiar, H. Prade, Fuzzy interval analysis. In: D. Dubois and H. Prade, editors, *Fundamentals of Fuzzy Sets*. Boston, Mass: Kluwer, pages 483-581, 2000.

[59] D. Dubois, J. Lang and H. Prade, Possibilistic logic. In D.M. Gabbay, et al, editors, *Handbook of Logic in AI and Logic Programming*, Vol. 3, Oxford University Press, pages 439-513, 1994.

[60] D. Dubois, S. Moral and H. Prade, A semantics for possibility theory based on likelihoods, *J. Math. Anal. Appl.*, 205, 359-380, 1997.

[61] D. Dubois, H.T. Nguyen, H. Prade, Fuzzy sets and probability: misunderstandings, bridges and gaps. In D. Dubois and H. Prade, editors, *Fundamentals of Fuzzy Sets*. Boston, Mass: Kluwer, pages 343-438, 2000

[62] D. Dubois, H. Prade, *Fuzzy Sets and Systems: Theory and Applications*. New York: Academic Press, 1980.

[63] D. Dubois, H. Prade, On several representations of an uncertain body of evidence. In: M. Gupta, E. Sanchez, editors, *Fuzzy Information and Decision Processes*, North-Holland: Amsterdam, pages 167-181,1982.

[64] D. Dubois, H. Prade Unfair coins and necessity measures: a possibilistic interpretation of histograms, *Fuzzy Sets and Systems*, 10(1), 15-20, 1983.

[65] D. Dubois and H. Prade, Evidence measures based on fuzzy information, *Automatica*, 21: 547-562, 1985.

[66] D. Dubois and H. Prade, The mean value of a fuzzy number, *Fuzzy Sets and Systems*, 24: 279-300, 1987.

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

[68] D. Dubois and H. Prade, Random sets and fuzzy interval analysis. *Fuzzy Sets and Systems*, 42: 87-101, 1991.

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

[70] D. Dubois and H. Prade, When upper probabilities are possibility measures, *Fuzzy Sets and Systems*, 49:s 65-74, 1992.

[71] D. Dubois, H. Prade: Possibility theory as a basis for preference propagation in automated reasoning. *Proc. of the 1st IEEE Inter. Conf. on Fuzzy Systems (FUZZ-IEEE'92)*, San Diego, Ca., March 8-12, 1992, 821-832.

[72] D. Dubois and H. Prade, What are fuzzy rules and how to use them. *Fuzzy Sets and Systems*, 84: 169-185, 1996.

[73] D. Dubois and H. Prade, Bayesian conditioning in possibility theory, *Fuzzy Sets and Systems*, 92: 223-240, 1997.

[74] D. Dubois and H. Prade, Possibility theory: Qualitative and quantitative aspects. In: D. M. Gabbay and P. Smets P., editors *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, Vol. 1., Dordrecht: Kluwer Academic, 169-226, 1998.

[75] D. Dubois, H. Prade. Possibilistic logic: A retrospective and prospective view. *Fuzzy Sets and Systems*, 144, 3-23, 2004.

[76] D. Dubois, H. Prade. Toward multiple-agent extensions of possibilistic logic. *Proc. IEEE Inter. Conf. on Fuzzy Systems (FUZZ-IEEE'07)*, London, July 23-26,187-192, 2007.

[77] D. Dubois, H. Prade. An overview of the asymmetric bipolar representation of positive and negative information in possibility theory. *Fuzzy Sets and Systems*, 160, 1355-1366, 2009.

[78] D. Dubois, H. Prade: Generalized possibilistic logic. *Proc. 5th Inter. Conf. on Scalable Uncertainty Management (SUM'11)*, Dayton, Springer, LNCS 6929, 428-432, 2011.

[79] D. Dubois, H. Prade. Possibility theory and formal concept analysis: Characterizing independent sub-contexts to appear in *Fuzzy Sets and Systems*.

[80] D. Dubois, H. Prade, A. Rico. A possibilistic logic view of Sugeno integrals. *Proc. Eurofuse Workshop on Fuzzy Methods for Knowledge-Based Systems (EUROFUSE*

2011), Sept. 21-23, Régua, Portugal, (P. Melo-Pinto, P. Couto, C. Serôdio, J. Fodor, B. De Baets, eds.), Advances in Intelligent and Soft Computing, vol. 107, Springer-Verlag, Berlin, 2011, 19-30.

[81] D. Dubois, L. Foulloy, G. Mauris, H. Prade, Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities. *Reliable Computing*, 10, 273-297, 2004.

[82] D. Dubois, H. Prade and R. Sabbadin, Qualitative decision theory with Sugeno integrals In [97], 314-322, 2000.

[83] D. Dubois, H. Prade and R. Sabbadin, Decision-theoretic foundations of possibility theory. *Eur. J. Operational Research*, 128, 459-478, 2001

[84] D. Dubois, H. Prade and S. Sandri On possibility/probability transformations. In R. Lowen, M. Roubens, editors. *Fuzzy Logic: State of the Art*, Dordrecht: Kluwer Academic Publ.,103-112, 1993.

[85] D. Dubois, H. Prade, S. Schockaert. Rules and meta-rules in the framework of possibility theory and possibilistic logic. *Scientia Iranica*, 18 (3), 566-573, 2011.

[86] D. Dubois, H. Prade and Smets, New semantics for quantitative possibility theory. *Proc. 6th Europ. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'01)*, Toulouse, Sept. 19-21, LNAI 2143, Springer-Verlag, 410-421, 2001.

[87] D. Dubois, H. Prade and P. Smets, A definition of subjective possibility, *Badania Operacyjne i Decyzije* (Wroclaw) 4, 7-22, 2003.

[88] D. Dubois, H. Prade, and L. Ughetto. A new perspective on reasoning with fuzzy rules. *Int. J. of Intelligent Systems*, 18 : 541-567, 2003.

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

[90] H. Farreny, H. Prade, Default and inexact reasoning with possibility degrees. *IEEE Trans. Systems, Man and Cybernetics*, 16(2), 270-276, 1986.

[91] B. R Gaines and L. Kohout, Possible automata. *Proc. Int. Symp. Multiple-Valued logics*, Bloomington, IN, pages 183-196, 1975.

[92] P. Gärdenfors, *Knowledge in Flux*, Cambridge, MA.: MIT Press, 1988.

[93] S. Galichet, D. Dubois, H. Prade. Imprecise specification of ill-known functions using gradual rules. *Int. J. of Approximate reasoning*, 35, 205-222, 2004.

[94] J. Gebhardt and R. Kruse, The context model. I*Int. J. Approximate Reasoning*, 9, 283-314, 1993.

[95] M. Gil, ed., Fuzzy Random Variables, Special Issue, *Information Sciences*, 133, 2001.

[96] J. F. Geer and G.J. Klir, A mathematical analysis of information-preserving transformations between probabilistic and possibilistic formulations of uncertainty, *Int. J. of General Systems*, 20, 143-176, 1992.

[97] M. Grabisch, T. Murofushi and M. Sugeno, editors, *Fuzzy measures and Integrals Theory and Applications*, Heidelberg: Physica-Verlag, 2000.

[98] W. Guezguez, N. Ben Amor, K. Mellouli, Qualitative possibilistic influence diagrams based on qualitative possibilistic utilities. *Europ. J. of Operational Research*, 195, 223-238, 2009.

[99] D. Guyonnet, B. Bourgine, D. Dubois, H. Fargier, B. Côme, J.-P. Chilès, Hybrid approach for addressing uncertainty in risk assessments. *J. of Environ. Eng.*, 129, 68-78, 2003.

[100] J.C. Helton, W.L. Oberkampf, editors (2004) Alternative Representations of Uncertainty. *Reliability Engineering and Systems Safety*, 85, 369 p.

[101] E. Huellermeier, D. Dubois and H. Prade. Model adaptation in possibilistic instance-based reasoning. *IEEE Trans. on Fuzzy Systems*, 10, 333-339, 2002.

[102] M. Inuiguchi, H. Ichihashi, Y. Kume Modality constrained programming problems: A unified approach to fuzzy mathematical programming problems in the setting of possibility theory, *Information Sciences*, 67, 93-126, 1993.

[103] C. Joslyn, Measurement of possibilistic histograms from interval data, *Int. J. of General Systems*, 26: 9-33, 1997.

[104] S. Kaci, H. Prade, Mastering the processing of preferences by using symbolic priorities in possibilistic logic. *Proc. 18th Europ. Conf. on Artificial Intelligence (ECAI'08)*, Patras, July 21-25, 376-380, 2008.

[105] G. J Klir and T. Folger *Fuzzy Sets, Uncertainty and Information*. Englewood Cliffs, NJ: Prentice Hall, 1988.

[106] G. J. Klir, A principle of uncertainty and information invariance, *Int. J. of General Systems*, 17: 249-275, 1990.

[107] G. J. Klir and B. Parviz B. Probability-possibility transformations: A comparison, *Int. J. of General Systems*, 21: 291-310, 1992.

[108] D. Lehmann, M. Magidor, What does a conditional knowledge base entail? *Artificial Intelligence*, 55, 1-60, 1992.

[109] D. L. Lewis, *Counterfactuals*. Oxford: Basil Blackwell, 1973.

[110] K. Loquin, D. Dubois, Kriging and epistemic uncertainty: A critical discussion. In: *Methods for Handling Imperfect Spatial Information*, (R. Jeansoulin, O. Papini, H. Prade, S. Schockaert, eds.), Springer-Verlag, STUDFUZZ 256, 269305, 2010.

[111] K. Loquin, D. Dubois, A fuzzy interval analysis approach to kriging with ill-known variogram and data, to appear in *Soft Computing*, 2012.

[112] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, 1979.

[113] V. Maslov, *Méthodes Opératorielles*, Mir Publications, Moscow, 1987.

[114] G. Mauris Possibility distributions: A unified representation of usual direct-probability-based parameter estimation methods Int. J. of Approximate Reasoning, 52:1232-1242, 2011.

[115] G. Mauris, V. Lasserre, L. Foulloy, Fuzzy modeling of measurement data acquired from physical sensors. *IEEE Trans. on Measurement and Instrumentation*, 49: 1201-1205, 2000.

[116] A. Neumaier, Clouds, fuzzy sets and probability intervals. *Reliable Computing*, 10, 249-272, 2004.

[117] H. T. Nguyen, B. Bouchon-Meunier, Random sets and large deviations principle as a foundation for possibility measures, *Soft Computing*, 8:61-70, 2003.

[118] J. Pearl, System Z: A natural ordering of defaults with tractable applications to default reasoning, *Proc. 3rd Conf. Theoretical Aspects of Reasoning About Knowledge*. San Francisco: Morgan Kaufmann, pages 121-135, 1990.

[119] H. Prade, A. Rico, Possibilistic evidence, *Proc. 11th Europ. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'11*, Belfast, UK, June 29-July 1, Springer, LNAI 6717, 713-724, 2011.

[120] H. Prade, M. Serrurier, Introducing possibilistic logic in ILP for dealing with exceptions. *Artif. Intell.* , 171, 939-950, 2007.

[121] H. Prade, M. Serrurier, Bipolar version space learning. *Int. J. Intell. Syst.*, 23,1135-1152, 2008.

[122] H. Prade, M. Serrurier. Maximum-likelihood principle for possibility distributions viewed as families of probabilities. *Proc. IEEE Inter. Conf. on Fuzzy Systems (FUZZ-IEEE'11)*, Taipei, 2987-2993, 2011.

[123] A. Puhalskii, *Large Deviations and Idempotent Probability*, Chapman and Hall, 2001

[124] E. Raufaste, R. Da Silva Neves, C. Mariné Testing the descriptive validity of possibility theory in human judgements of uncertainty. *Artificial Intelligence*, 148: 197-218, 2003.

[125] L.J. Savage *The Foundations of Statistics*, New York: Dover, 1972.

[126] G. Shafer, Belief functions and possibility measures. In J.C. Bezdek, editor, *Analysis of Fuzzy Information Vol. I: Mathematics and Logic*, Boca Raton, FL : CRC Press, pages 51-84, 1987.

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

[128] R. Slowinski and M. Hapke, editors (2000) *Scheduling under Fuzziness*, Heidelberg: Physica-Verlag, 2000.

[129] P. Smets (1990). Constructing the pignistic probability function in a context of uncertainty. In: M. Henrion et al., editors, *Uncertainty in Artificial Intelligence*, vol. 5, Amsterdam: North-Holland, pages 29-39, 1990.

[130] W. Spohn, A general, nonprobabilistic theory of inductive reasoning. In R.D. Shachter, et al., editors, *Uncertainty in Artificial Intelligence*, Vol. 4. Amsterdam: North Holland. pages 149-158, 1990.

[131] T. Sudkamp, Similarity and the measurement of possibility. *Actes Rencontres Francophones sur la Logique Floue et ses Applications* (Montpellier, France), Toulouse: Cepadues Editions, pages 13-26, 2002.

[132] H. Tanaka, P.J. Guo, *Possibilistic Data Analysis for Operations Research*, Heidelberg: Physica-Verlag, 1999.

[133] A. Tversky and D. Kahneman. Advances in prospect theory: cumulative representation of uncertainty. *J. of Risk and Uncertainty*, 5, 297-323, 1992.

[134] W. Van Leekwijck and E. E. Kerre, Defuzzification: criteria and classification, *Fuzzy Sets and Systems*, 108: 303-314, 2001.

[135] P. Walley, *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall, 1991.

[136] P. Walley and G. De Cooman, A behavioural model for linguistic uncertainty, *Information Sciences*, 134; 1-37, 1999.

[137] O. Wolkenhauer, *Possibility Theory with Applications to Data Analysis*. Chichester: Research Studies Press, 1998.

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

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

[140] R.R. Yager, A procedure for ordering fuzzy subsets of the unit interval, *Information Sciences*, 24:143-161, 1981.

[141] R.R. Yager, An introduction to applications of possibility theory. *Human Systems Management*, 3: 246-269, 1983.

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

[143] L. A. Zadeh. Fuzzy sets and information granularity. In M.M. Gupta, R. Ragade, R.R. Yager, editors, *Advances in Fuzzy Set Theory and Applications*, Amsterdam: North-Holland, 1979, pages 3-18.

[144] L. A. Zadeh. Possibility theory and soft data analysis. In L. Cobb, R. Thrall, editors, *Mathematical Frontiers of Social and Policy Sciences*, Boulder, Co. : Westview Press, pages 69-129, 1982.