Annals of Mathematics and Artificial Intelligence **32:** 35–66, 2001.
© 2001 *Kluwer Academic Publishers. Printed in the Netherlands.*

Possibility theory, probability theory and multiple-valued logics: A clarification *

Didier Dubois and Henri Prade
Institut de Recherche en Informatique de Toulouse (I.R.I.T.) – C.N.R.S., Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 4, France
E-mail: {dubois, prade}@irit.fr

There has been a long-lasting misunderstanding in the literature of artificial intelligence and uncertainty modeling, regarding the role of fuzzy set theory and many-valued logics. The recurring question is that of the mathematical and pragmatic meaningfulness of a compositional calculus and the validity of the excluded middle law. This confusion pervades the early developments of probabilistic logic, despite early warnings of some philosophers of probability. This paper tries to clarify this situation. It emphasizes three main points. First, it suggests that the root of the controversies lies in the unfortunate confusion between degrees of belief and what logicians call "degrees of truth". The latter are usually compositional, while the former cannot be so. This claim is first illustrated by laying bare the non-compositional belief representation embedded in the standard propositional calculus. It turns out to be an all-or-nothing version of possibility theory. This framework is then extended to discuss the case of fuzzy logic versus graded possibility theory. Next, it is demonstrated that any belief representation where compositionality is taken for granted is bound to at worst collapse to a Boolean truth assignment and at best to a poorly expressive tool. Lastly, some claims pertaining to an alleged compositionality of possibility theory are refuted, thus clarifying a pervasive confusion between possibility theory axioms and fuzzy set basic connectives.

1. Introduction

In the past twenty years, the necessity of handling uncertainty in problem-solving has generated an impressive amount of literature. Interestingly, traditional tools issued from probability theory were felt to be insufficient to handle all facets of uncertainty. New uncertainty calculi have been proposed such as belief functions [76,77], probability intervals [79], and possibility theory [37,85]. Their main motivation is a more faithful treatment of incomplete knowledge. Moreover, the emergence of fuzzy set theory, and the (now declining) fashion of expert systems [13] focused interest on weighted logics, that is, logics where formulas are assigned weights (usually taken in a totally ordered

---

* This work is partially based on a paper entitled "Can we enforce full compositionality in uncertainty calculi?", presented at the 12th National Conf. on Artificial Intelligence (AAAI'94) (Seattle, WA, July 31–August 4, 1994), see *Proceedings*, pp. 149–154. Some results were also presented at the Seminar on Enriched Lattice Structures for Many-Valued and Fuzzy Logics (LINZ'97) (Linz, Austria, February 25–March 1st, 1997).

scale). The connection with older many-valued logics dating back to the first half of the twentieth century was naturally made.

However, this abundant literature is pervaded with a lot of misunderstandings as to the role, the meaning and the properties of weights attached to propositions. Such a confusing state of the art seems to have its origin in the very way some pioneers of many-valued logics viewed their own calculi. For instance, in the 3-valued logics of Lukasiewicz, the third truth-value is often wrongly understood as expressing a lack of knowledge about truth. This type of assumption has led to many confusions and controversies until today, regarding the appropriateness of various calculi for uncertainty modeling, and even regarding the mathematical consistency of some theories, like fuzzy set theory [45] or possibility theory [26]. For instance, probabilistic logic was sometimes misleadingly viewed as a special kind of many-valued logics (especially [73]); fuzzy set theory has been criticized as an adhoc surrogate of probability theory built by applied researchers having a limited knowledge of the latter [60]; possibility theory has sometimes been viewed as just another way of presenting fuzzy set theory, thus adopting for the former all connectives of the latter, which is mathematically inconsistent (for instance, [17]). All these confusions have their origin in lack of distinction between partial truth and partial belief. This distinction has been neglected, as a mere philosophical debate with no technical consequence. However, it greatly affects the issue of compositionality with respect to logical connectives.

This paper tries to clarify the role of many-valued logics, fuzzy sets, and possibility theory for uncertainty modeling, in contrast to probability theory. First, a brief historical account of some of the above mentioned controversies and confusions is given in section 2. It involves major philosophers of logic and probability in the twentieth century, and expands to contemporary research in artificial intelligence, especially expert systems. The ideas summarized in this section are discussed in greater details in [34]. The next section explains how standard propositional calculus handles uncertainty, shedding light on the difference between "true " and "certainly true", and pointing out the fact that while truth in propositional calculus is binary and compositional, uncertainty is ternary and not compositional. It is also pointed out that uncertainty in propositional logic is captured by a 0–1-valued special case of possibility theory (and not probability). Section 4 envisages the many-valued case. At the semantic level, weights attached to propositions may have two distinct meanings: either they are truth-values or they are degrees of confidence. In the first case, logical propositions are fuzzy since truth is a matter of degree. The algebraic setting of fuzzy propositions is not Boolean. In the second case, most of the time, truth remains binary and weights may express the more or less strong inability, for an agent, to know whether a proposition is true or false. In other words, the algebraic setting remains Boolean and the weights are located at the meta-level. More generally, uncertainty is due to incomplete knowledge, while graded truth just refers to many-valued propositional variables.

Many-valued logics are basically construed as truth-functional calculi: the degree of truth of a formula can be calculated from the degrees of truth of its constituents. It has been tempting for many authors to exploit many-valued logic calculi for belief propagation. However, even if this approach may properly work in some expert-systems under severe restrictions of the structure of the knowledge base, the notion of compositional logic of uncertainty leads to mathematical difficulties in general. Namely, in a Boolean setting, the uncertainty of a formula cannot be a function of the uncertainties of its constituents. Section 5 describes existing trivialization results pertaining to this question, points out some open problems, and draws some conclusions pertaining to the debatable use of truth-functional logics for uncertainty handling. We investigate the possibility of an *almost* fully compositional uncertainty calculus, but its expressiveness turns to be very limited. In contrast, probability and possibility theories adopt an assumption of compositionality pertaining to *one connective only* (negation for probability functions, and disjunction for possibility functions). These theories do not model the same facet of uncertainty, as explained in section 6. They appear to be complementary theories rather than full-fledged rivals.

2. From many-valued logics to uncertainty modeling: a confusing tradition

Classically, the truth of a statement or proposition is defined as the perfect agreement of this statement with the actual state of facts. The basic principle of classical logic, known as the Principle of Bivalence, asserts the following:

Every proposition is either true or false.

The principle of bivalence, formulated and strongly defended by Chrisippus and his school in ancient Greece, was originally questioned by Epicureans, and even misleadingly rejected by them in the case of propositions referring to future contingencies. At the beginning of the twentieth century, Lukasiewicz had proposed to replace this principle by the more general Principle of Valence, which says the following:

Every proposition has a truth-value.

For Lukasiewicz, propositions are not only either true or false but they can have an intermediary truth-value. Truth-values are then usually modeled by numbers in the unit interval. In that author's mind, the third truth-value seems to express the idea of "unknown", or "possible" (see the English translation of [64]). This view, supported by an early attempt to encode three-valued logic into modal logic (see again the same volume of Lukasiewicz selected works), seems to have been shared by subsequent scholars in many-valued logic. For instance, according to the truth tables of the 3-valued logic of Kleene [58], the disjunction of two unknown propositions is unknown. This understanding of the third truth-value induces a confusion between truth-values and modalities such as probability and possibility. It leads to debatable interpretations of the law of excluded middle and the law of non-contradiction, that bear on the connections between many-valued logics and modal logics. Let us take an example, considered already by Aristotle, namely the proposition:

> "There will be a sea battle tomorrow (p) and there will not be a sea battle tomorrow ($\neg p$)".

This proposition, of the form "p and $\neg p$" is always false, because of the non-contradiction law. Similarly, the proposition "p or $\neg p$" is always true, because *tertium non datur*. But we may *fail to know* if the proposition "there will be a sea battle tomorrow" is true, and also if the proposition "there will not be a sea battle tomorrow" is true. In this case, at least intuitively, it seems reasonable to say that it is *possible* that there will be a sea battle tomorrow but at the same time, it is *possible* that there will not be a sea battle tomorrow. There was a recurrent tendency, until the twentieth century many-valued logic tradition, to claim the failure of the bivalence principle on such grounds, and to consider the *possible* as a third truth-value. However, the proposition *possible* p is not the same as p, and *possible* $\neg p$ is not the negation of *possible* p. Hence the fact that the proposition

"possible p and possible $\neg p$"

may be true does not question the law of non-contradiction. On the contrary, many-valued (= fuzzy) propositions are ones such that, due to the gradual boundary of their sets of models, proposition "p and $\neg p$" is not completely false in some interpretations. This is why Moisil [65] speaks of fuzzy logics as non-Chrisippean logics.

Besides, the probabilistic tradition considers equipossibility as identical to equiprobability (after Laplace). This assumption may have misled researchers into viewing many-valued logics as a setting for probabilistic reasoning. Many attempts were undertaken to import probability theory into the realm of logic or to reconstruct logic in accordance with the probabilistic laws by means of many-valued logics. For instance, while trying to develop a quantitative concept of truth, Reichenbach proposed his "probability logic" in which the alternative true-false is allegedly replaced by a continuous scale of truth-values. Reichenbach [73] begins with a simple illustrative example of the statement "I shall hit the center". As a measure of the degree of truth of this statement, Reichenbach proposes to measure the distance r of the hit to the center and to take the truth-value as equal to $1/(1+r)$. This can be done of course *after* the shot. Unfortunately, we cannot find advice in his work how to evaluate the truth of the above sentence *before* the shot. However, retrospectively, it is easy to figure out that this method is actually evaluating the fuzzy proposition "I shall hit *close to* the center". It can only be done after the hit. Quantifying the proposition before the hit is a matter of belief assessment, not of measuring graded truth.

More recently, in the last thirty years, there has been a very active research trend in Artificial Intelligence concerning the management of uncertainty in knowledge-based systems. This trend was originally influenced by the MYCIN experiments [13], where a basic idea was to attach weights expressing uncertainty to facts and rules in a knowledge base. A weighted logic or inference system is then said to be *compositional* if and only if the weight of a complex formula can be calculated by combining the weights of its atomic constituents. The compositionality property is used to propagate weights in reasoning procedures. Inference engines of expert systems usually work on a rule-by-rule basis, and the weight propagation schemes consist of three steps:

(i) computing the weight bearing on a composite fact from the weights bearing on the elementary parts of this fact (due to compositionality),

(ii) propagating the weight bearing on the conditions of the rule to its conclusion, by integrating the weight bearing on the rule, and

(iii) combining the weights bearing on partial conclusions pertaining to the same matter.

Many-valued logics are basically construed as truth-functional calculi, just like classical logic: the degree of truth of a formula can be calculated from the degrees of truth of its constituents. They are examples of compositional systems. It was thus tempting for many authors to develop an uncertainty calculi for expert systems on the basis of the truth-functional calculus of a many-valued logic, viewing degrees of truth as certainty factors; see [31] for a survey of such compositional uncertainty calculi.

However, investigating the validity of such a methodology requires a proper interpretation of the weights. Reading the literature in the expert system area, it appears that these weights may have two interpretations: degrees of belief and degrees of partial truth and that scholars again tend to make a confusion between these two notions. One of the reasons why this confusion was made is the need for simple calculations that made compositionality assumptions attractive. Even degrees of probability are sometimes called degrees of truth (e.g., [67]). Some papers presenting the expert system PROSPECTOR (see [44]) do refer to fuzzy logic connectives as a tool for combining probabilities. But probability theory is not compositional. It was only with the advent of the Bayes network methodology [70] that the correct approach to probability handling in Artificial Intelligence became clear. Pearl clearly explains the deficiencies of compositional calculi, and shows that the rigorous application of probability theory overcome these deficiencies.

Besides, the emergence of fuzzy rule-based systems in process control problems led AI researchers, that criticized MYCIN-like systems, to reject fuzzy logic on the same grounds of dubious compositionality assumptions. For instance, Elkan [45] questioned its well-foundedness and cast serious doubts on the reasons of its success, arguing that "fuzzy logic mathematically collapses to two-valued logic". This claim was in fact due to the use of too strong a notion of logical equivalence which is valid in two-valued logic, but which has nothing to do with many-valued logics nor fuzzy set theory. Assuming a fully compositional *many-valued* calculus on a *Boolean* algebra of propositions, the logical system indeed collapses to two-valued logic, as recalled later on in this paper. But fuzzy logic is *not* Boolean. Furthermore, Elkan [45] does not mention the important distinction between degrees of truth and degrees of uncertainty, between the problem of handling many-valued propositions (which is what fuzzy set connectives do) and the problem of exploiting incomplete knowledge (which is the purpose of possibility theory). He never mentions possibility theory, but wonders at the success of fuzzy logic controllers in the face of the trivialization result. However, there is no treatment of uncertainty in fuzzy controllers. Elkan's trivialization result kills truth-functional belief propagation systems (for the nth time), and certainly does not harm fuzzy set theory nor the interpolation device at work in fuzzy controllers.

3. The uncertainty calculus of propositional logic

Propositional logic is the simplest logic. It is always described as the logic of all-or-nothing entities. It is a calculus of Boolean variables conventionally taking values in {0, 1}, where 0 is called false and 1 is called true. However, in the scope of Artificial Intelligence, a set K of well-formed Boolean formulae is understood as a set of propositions believed by an intelligent agent, what is often called a *belief base*. This set represents the current knowledge of the agent, i.e., all that is known by this agent. The belief base enables questions pertaining to the state of affairs to be answered. Such questions are often put in the form "is proposition p true?". However, it is important to notice that, in the context of knowledge representation, such questions mean: "Is the agent *sure* that p is true on the basis of the belief base K?"

3.1. Boolean possibility theory

Answering questions in this elementary framework comes down to proving propositions from K taken as a set of axioms, under the laws and inference rules of propositional logic. When a proposition p follows from K, it means that the agent believes p. If the beliefs of the agent match the actual states of affairs, this implies that p is true (in the usual sense), or, more precisely, that the truth-value of p is 1. However, the deduction procedure of classical logic does not directly compute the (truth) value of p. It rather computes a (Boolean) degree of belief about p. Namely, if p is provable from K, then the agent believes p and the degree of belief in p is 1; if p is not provable from K, then the agent does not believe p and the degree of belief in p is 0. Generally, an agent has but limited knowledge, and it implies that K is not complete. That is, for some propositions p, it may happen that neither p nor its negation $\neg p$ are provable from K. In epistemic terms, it means that the agent believes neither p nor $\neg p$. This possibility confirms that the question-answering scheme of classical logic does not compute truth-values of propositions directly. In particular, if the agent does not believe p, it does NOT imply that (s)he believes $\neg p$. The set of all believed propositions (entailed by a belief base K) is called a belief set [48].

The above analysis makes it clear that

(i) propositional logic, used as a tool for knowledge representation, embeds a calculus of uncertainty (belief) that does not reduce to a truth assignment procedure,

(ii) uncertainty in propositional logic is *ternary* and not binary: either a proposition p is believed, or its negation is believed, or neither of them are believed.

Let $N(p) \in \{0, 1\}$ denote the (Boolean) degree of belief of p. By convention, let $N(p) = 1$ when K proves p, and 0 otherwise. The state of belief of the agent regarding p is captured by the pair $(N(p), N(\neg p))$. Namely,

- $(N(p), N(\neg p)) = (1, 0)$ iff the agent believes p,
- $(N(p), N(\neg p)) = (0, 1)$ iff the agent believes $\neg p$,

- $(N(p), N(\neg p)) = (0, 0)$ iff the agent believes neither p nor its negation.

The third belief state corresponds to a fully ignorant agent, who is totally uncertain about p. The fourth value of the pair $(N(p), N(\neg p)) = (1, 1)$ is obtained only if the belief base K is inconsistent, since in the presence of a contradiction in K, everything is entailed. Most uncertainty theories, as well as belief revision theories, assume that belief bases, regardless of the way they are represented, are consistent, which forbids this fourth situation. Nevertheless there exists a specific literature about inconsistency-tolerant reasoning in the propositional setting (see [8,10] for surveys).

Now let us turn to the issue of compositionality. A major feature of Boolean logic lies in the possibility of computing the truth-values of propositions from the truth-values of their components. Is it possible to compute the above two-valued degrees of belief in the same way? Not for all connectives. First, the value of $N(p)$ does not fully determine $N(\neg p)$, due to the ternary nature of uncertainty in this setting. The quantity $\Pi(p) = 1 - N(\neg p)$ represents to what extent the agent considers p as possible. Indeed, $\Pi(p) = 0$ is equivalent to $N(\neg p) = 1$, that is, the agent believes $\neg p$, hence considers p as impossible. When $\Pi(p) = 1$, either K implies p or it implies neither p nor $\neg p$. It precisely depicts "p is possible" as "p is consistent with the belief base K". The state of ignorance about p is equivalently described by the pair $(\Pi(p), \Pi(\neg p)) = (1, 1)$, which says that for the agent both p and $\neg p$ are possible.

Similarly, it is not possible to compute the degree of belief of a disjunction from the degrees of belief in its components. When $N(p) = 0$ and $N(q) = 0$, the value of $N(p \vee q)$ can be 0 or 1; this is obvious as it expresses the well-known fact that when K implies neither p nor q, it may sometimes imply $p \vee q$, but sometimes not. If $p = q$ and $N(p) = 0$ then $N(p \vee q) = 0$, clearly. Another important particular case is when $q = \neg p$ and the agent is ignorant about p. Nevertheless (mind that our setting is propositional logic), the agent must believe $p \vee \neg p$ since it is a tautology, hence $N(p \vee q) = 1$, although $N(p) = N(\neg p) = 0$. We stand here in the crux of the confusion pervading the literature on the validity or not of the law of excluded middle. In a *compositional* three-valued logic, the case $(N(p), N(\neg p)) = (0, 0)$ is encoded as, hence confused with, a third truth-value. Actually, the belief function N is compositional with respect to conjunction only. Namely, in $\{0, 1\}$:

$$N(p \wedge q) = \min\bigl(N(p), N(q)\bigr). \tag{1}$$

It expresses in a concise way the well-known fact that a belief base implies a conjunction of propositions if and only if it implies each conjunct. Dually the Boolean possibility function Π is compositional for disjunction only:

$$\Pi(p \vee q) = \max\bigl(\Pi(p), \Pi(q)\bigr). \tag{2}$$

Generally, the problem under study is described by means of a language, which is a set of state variables. They are Boolean ones for classical logic. The Cartesian product of the ranges of variables is called "frame of discernment" by Shafer [76] because the description language conditions our ability to distinguish between actual states of affairs. This frame of discernment is denoted Ω in the following. Each element of Ω is a conjunction of Boolean values called interpretations of the language. It clusters states of affairs indistinguishable according to the language. Nevertheless, such elements will also be abusively called "states of affairs", although they stand for imperfect descriptions thereof (using Boolean variables in logic, real variables in interval analysis).

Fixing an interpretation, the truth-value (1 or 0) of any formula in the language in the corresponding state of affairs can be computed. To each proposition p is assigned a subset $[p]$ of Ω containing the interpretations where p is true. These interpretations are called the models of p. Define the set function Π from 2^Ω to $\{0, 1\}$ by letting for any subset $A \subseteq \Omega$, $\Pi(A) = \Pi(p)$ whenever $[p] = A$. This definition is consistent since, in classical logic, logically equivalent propositions have the same set of models and are equally believed. Then the set function Π is indeed a (two-valued) possibility measure in the sense of Zadeh [85]. The dual set function built from the belief function N in the same way is a necessity measure [35].

A belief base K can be replaced by the set E of interpretations which make all propositions in K true. E is the set of models of K. The models of K represent possible states of affairs, one of which is the actual one. The characteristic function of E can thus be called a possibility distribution. A proposition p is surely true if $E \subseteq [p]$, surely false if $E \subseteq [p]^c$ (the complement of $[p]$) and p is totally uncertain if both $E \cap [p] \neq \emptyset$, $E \cap [p]^c \neq \emptyset$. This is again the crude trichotomy mentioned above in the presence of incomplete knowledge. And the basic definitions of possibility theory are retrieved: $\Pi(A) = 1$ iff $E \cap A \neq \emptyset$; $N(A) = 1$ iff $E \subseteq A$. The two modes of representation, belief bases and sets of possible states of affairs, are equivalent because propositional calculus is sound and complete.

3.2. *Is interval analysis based on a 3-valued logic?*

As an example of application of this simple calculus, consider the case of interval analysis. It has often been considered as a natural example where a three-valued logic can be used. The reason for this claim is found when comparing the value of a real-valued quantity x to a threshold θ, when the value of x is only known to belong to an interval $[a, b]$. In this example, the available knowledge is modeled by the set $E = [a, b]$ and the proposition $p =$ "$x > \theta$" has set of models $[p] = (\theta, +\infty)$. Then three situations are possible:

(i) $a > \theta$: then $x > \theta$ is true;

(ii) $b < \theta$: then $x > \theta$ is false;

(iii) $a \leqslant \theta \leqslant b$.

It is tempting to interpret the last situation as "$x > \theta$ is neither true nor false", and to assign a third truth-value between 0 and 1 to the proposition $p =$ "$x > \theta$". In the light of the above explanation, it is clear that this view is erroneous. First, $x \in [a, b]$ describes an incomplete state of knowledge, here in terms of a set of possible values of x instead of a belief base. Second, asserting $a > \theta$ is equivalent to saying that the agent, who knows $x \in [a, b]$, *believes* that $x > \theta$ is true. It only implies that $x > \theta$ is true. But the converse (i.e., $x > \theta$ implies $a > \theta$) is false since the case when $a > \theta$ does not cover all cases where $x > \theta$ is actually true (namely the cases, ignored by the agent, when $x \in (\theta, b]$). It is clear that there does not exist any state of affairs where $x > \theta$ is neither true nor false, since the actual value of x is either greater than θ or not. Situation (iii) just expresses that $x > \theta$ is neither believed nor disbelieved by the agent since both $[a, b] \cap [p] \neq \emptyset$, $[a, b] \cap [p]^c \neq \emptyset$. Interval analysis does not fit with a three-valued logic, it requires the same elementary uncertainty theory as propositional calculus: Boolean possibility theory. The three cases should thus be interpreted as follows:

(i) $a > \theta$, then $(N(p), N(\neg p)) = (1, 0)$: the agent believes $x > \theta$ is true,

(ii) $b < \theta$ then $(N(p), N(\neg p)) = (0, 1)$: the agent believes $x > \theta$ is false,

(iii) $a \leqslant x \leqslant b$, then $(N(p), N(\neg p)) = (0, 0)$: the agent is ignorant.

3.3. Does probability theory extend classical logic?

None of the two set-functions N and Π coincides with a probability function, except in the trivial case when the belief base K is complete, which implies that it possesses a single model. Then $\Pi = N$ is a two-valued probability measure. One state is then known to be the true one. This is also the only case when a probability measure can be two-valued. Besides, suppose the three epistemic states that describe the uncertainty calculus of propositional logic are mapped to a three-valued chain $\{0, \alpha, 1\}$ where $0 < \alpha < 1$. The corresponding set-function g is such that $g(A) = 1$ if A is believed, $g(A) = 0$ if A is disbelieved, and $g(A) = \alpha$, in case of ignorance. This function is almost never a probability function [43]. It can be so only if the set of models of K contains only two elements (then $\alpha = 1/2$). However, if K is incomplete and possesses more than two models, the function g is not a probability measure and is compositional for negation only. Namely, using a uniform probability distribution on the set E of models of K to express ignorance leads to assigning distinct probabilities to at least two supposedly totally uncertain propositions p and q. This is counterintuitive and not in agreement with the uncertainty calculus embedded in propositional logics (since one then finds for instance that p is more probable than q). Hence, probability theory is not a faithful representation of incomplete knowledge in the sense of classical logic.

In the literature, degrees of probabilities are viewed as an extension of binary truth-values. This is because, in probability theory, the statement that the probability of a proposition is 1 (respectively 0) is *equivalent* to stating that this proposition is actually true (respectively false), provided that the probabilistic knowledge of the agent is correct. Hence, contrary to the case of the previous section, a probabilistic agent which does not believe p at all is forced to believe $\neg p$. Of course, it is obvious that compositionality is not valid in this setting. Nevertheless, many prominent authors did consider probabilistic logic as a non-compositional variant of many-valued logics. But very early, when many-valued logics came to light, some scholars in the foundations of probability became aware that probabilities differ from what logicians call truth-values. De Finetti [24], witnessing the emergence of many-valued logics (especially the works of Lukasiewicz), pointed out that uncertainty, or partial belief, as captured by probability, is a meta-concept with respect to truth degrees. It does not contradict the fact that a proposition is usually understood as a binary notion. On the contrary, the notion of partial truth, as put forward by Lukasiewicz [64], leads to changing the very notion of proposition, which ceases to be an all-or-nothing matter. Indeed, the definition of a proposition is a matter of convention. To quote De Finetti (our translation from the French):

"Propositions are assigned two values, true or false, and no other, not because there "exists" an a priori truth called "excluded middle law", but because we call "propositions" logical entities built in such a way that only a yes/no answer is possible ... A logic, similar to the usual one, but leaving room for three or more [truth] values, cannot aim but at compressing several ordinary propositions into a single many-valued logical entity, which may well turn out to be very useful ... ".

In this statement, one can perceive an idea that fuzzy sets would later formalize, namely a collection of sets (level-cuts) representing a single proposition. It also puts it very clearly that many-valued logics (hence, fuzzy sets describing the meaning of sentences), deal with many-valuedness in a logical format, not with uncertainty or probability. On the contrary, uncertainty pertains to the beliefs held by an agent, who is not totally sure whether a proposition of interest is true or false, without questioning the fact that ultimately this proposition cannot be but true or false. To quote De Finetti [24] again:

"Even if, in itself, a proposition cannot be but true or false, it may occur that a given person does not know the answer, at least at a given moment. Hence for this person, there is a third attitude in front of a proposition. This third attitude does not correspond to a third truth-value distinct from yes or no, but to the doubt between the yes and the no. (As people, who, due to incomplete or indecipherable information, appear as of "unknown sex" in a given statistics. They do not constitute a third sex. They only form the group of people whose sex is unknown.)"

Probabilistic logic, contrary to multiple-valued logic, is not a substitute to Boolean logic. It is probability theory built *on top* of Boolean logic. Degrees of probability lie at the meta-level with respect to Boolean propositions. However this point is not always clearly made by the forerunners of many-valued logics. Carnap [15] also points out the difference in nature between truth-values and probability values (hence degrees thereof), precisely because "true" (respectively false) is not synonymous to "known to be true" (respectively known to be false), or in other words, "verified" (respectively falsified). He criticizes Reichenbach on his claim that probability values should supersede the two usual truth-values. Quoting Carnap:

> "A given sentence is often neither verified nor falsified; nevertheless it is either true or false, whether anybody knows it or not. In this way an inadvertent confusion of "true" and "verified" may lead to doubts about the validity of the principle of

excluded middle ... I agree with Reichenbach that a concept referring to an absolute and unobtainable maximum should be replaced by a concept referring to a high degree in a continuous scale. However, what is superseded by "highly probable" ... is the concept of "confirmed to the highest degree or verified", and not the concept "true"."

4. Partial truth vs. uncertainty

A standard analogical example that points out the difference between degrees of truth and degrees of uncertainty is that of a bottle (e.g., [12]). In terms of binary truth-values, a bottle is viewed as full or empty. If one accounts for the quantity of liquid in the bottle, one may say the bottle is "half-full" for instance. Under this way of speaking, "full" becomes a fuzzy predicate and the degree of truth of "the bottle is full" reflects the amount of liquid in the bottle. The situation is quite different when expressing our ignorance about whether the bottle is either full or empty (given that we know only one of the two situations is the true one). To say that the probability that the bottle is full is $1/2$ does not mean that the bottle is half full. Degrees of uncertainty are clearly a higher level notion, higher than degrees of truth.

This distinction between degrees of truth and degrees of uncertainty that, at least, goes back to De Finetti [24], seems to have been almost completely forgotten by Artificial Intelligence. The confusion pervading the relationship between truth and uncertainty in the expert systems literature is apparently due to the lack of a dedicated paradigm for interpreting partial truth, and grades of uncertainty in a single framework. This is what we try to provide here.

First, we must do away with the commonsense view of truth *as the compatibility between a statement and reality*. This naïve definition of truth is often criticized by philosophers (see, e.g., Gochet's discussion of Dubois and Prade [38]). In the scope of information systems, the debatable word "reality" must be changed into "what is known about reality by an agent" and interpreting the latter as "the possibly incomplete description of some actual state of facts as stored in a data base". Moreover, instead of computing the degree of truth of a proposition, one can only evaluate the degree of belief that this proposition holds, which is really what estimating its conformity with the description of what is known of the actual state of facts comes down to. Only when the knowledge about the state of facts is complete can we obtain degrees of truth from degrees of belief about truth.

Besides, the word "truth" is perhaps too loaded with philosophical presuppositions and controversies, and should not be taken for granted in Artificial Intelligence. In this particular framework, a truth assignment comes down to assigning values 1 or 0 to Boolean (propositional) variables. Computing the degree of truth of a proposition means computing the value of the Boolean function that models the proposition. Moving to many-valued logics means going from Boolean variables to non Boolean variables, usually ranging on a common totally ordered scale. This has nothing to do with uncertainty handling.

With these remarks in mind, belief evaluation comes down to a semantic matching procedure. This point of view is in accordance with the test-score semantics of Zadeh [87] for natural languages. We start again with, on the one hand, the epistemic state of an agent, and, on the other hand, a question to be answered, namely to what extent the agent believes that a proposition p is true. As pointed out in previous publications (noticeably [38]) several situations can be encountered.

4.1. Complete knowledge

If the agent precisely knows everything about the state of the world, (s)he has full belief about what is true and what is not for any proposition p of interest. In Boolean logic, a proposition is true or false. If the actual state of facts is known and encoded as a single element ω of a set of interpretations, the agent is then sure that p is true if $\omega \in [p]$, and false otherwise.

In the setting of multiple-valued logics, the convention prescribing that a proposition is true or false is changed. A more refined range is used for the function that represents the meaning of a proposition. This is usual in natural language when words are modeled by fuzzy sets. For instance, the compatibility of "tall" in the phrase "a tall man", with some individual of a given height, is often graded: the man can be judged *not quite* tall, *somewhat* tall, *rather* tall, *very* tall, etc. Changing the usual true/false convention leads to a new concept of proposition whose compatibility with a given state of facts is a matter of degree, and can be measured on an ordered scale L that is no longer $\{0, 1\}$, but the unit interval for instance. Many-valuedness reflects linguistic hedges such as "somewhat", "rather", "very", etc. This type of convention leads to identifying a "fuzzy proposition" p with *a fuzzy set* of possible states of affairs; the degree of membership of a state of affairs to this fuzzy set evaluates the degree of fit between the proposition and the state of facts it refers to. This degree of fit $\tau_\omega(p) \in L$ is called degree of truth of proposition p in the interpretation ω (state of affairs). Many-valued logics, especially truth-functional ones, provide compositional calculi of degrees of truth, including degrees between "true" and "false" (see [50,53]).

In the many-valued framework, a question such as "is proposition p believed to be true (or false)?" makes no sense strictly speaking. Indeed, it should be clear which element of the scale L the word "true" refers to. If it systematically refers to the top element of L (fully true), then the shades of truth are not used. One should ask questions such as "is proposition p believed to be true to level α (exactly, or at least, or at most)?" However these questions have little intuitive appeal (e.g., who could answer the question: "is Paul tall to degree 0.8?"). Nevertheless, in the complete knowledge case, questions like "to what extent is p true ?" can be addressed for comparison purposes, by producing the corresponding degree of truth, that is, the value of p (viewed as a non-Boolean function).

4.2. Incomplete knowledge and Boolean queries

The full Boolean case has been addressed in section 3. The simplest representation of a state of incomplete knowledge is a subset E of states of affairs, one of which is the actual one. The set E may be, in the Boolean case, the set of models of a belief base, or yet an interval, or a Cartesian product thereof. Only two-valued belief degrees pertaining to a proposition p and its negation can be computed as in section 3.1, in the framework of possibility theory.

The knowledge of the agent could alternatively be captured by a single probability measure P on the frame of discernment. Then, degrees of belief in any proposition are probabilities $P([p])$. However, the knowledge is not incomplete stricto sensu in this situation, since, as seen above, it does not cover the Boolean incomplete knowledge situation. Probabilistic knowledge is an extension of the *complete* knowledge case to many-valued belief degrees, where the characteristic property of complete belief bases (K implies p or implies $\neg p$, for all p) is generalized by the property $P([p]) + P([\neg p]) = 1$.

A genuine extension of the Boolean incomplete knowledge situation to many-valued degrees of belief is obtained when the set E representing the agent knowledge is a fuzzy set, whose membership function then represents a possibility distribution π over the frame of discernment. The set E of possible states of fact is then merely ordered in terms of plausibility, normality and the like. $\pi(\omega)$ reflects to what extent the state of affairs is unsurprising, expected as normal [42]. For simplicity, the range V of π is the unit interval, but this is a matter of convention. Any totally ordered chain V would do. Moreover, the consistency of the state of knowledge is ensured if $\pi(\omega) = 1$ for some ω (normalization). It means that at least one interpretation is completely possible, i.e., is considered as a normal state of affairs.

The overlapping between E and the ordinary set of models of a proposition p will be a matter of degree. This degree of overlapping is the *degree of possibility* defined by

$$\Pi(p) = \max\{\pi(\omega), \omega \in [p]\}, \tag{3}$$

which generalizes the Boolean degree of possibility introduced in section 3. The rationale behind this evaluation is that the plausibility of p is evaluated in the most normal situation where p is true. It assumes that the agent presupposes that the actual state of affairs is always as normal as can be. $\Pi(p)$ is also the degree of consistency between the state of knowledge of the agent and the proposition p, and describes to what extent p is possibly true.

Dually, the degree of belief in p reflects the lack of plausibility of the negation of p, and is captured by a many-valued necessity function:

$$N(p) = \min\{1 - \pi(\omega), \omega \notin [p]\} = 1 - \Pi(\neg p). \tag{4}$$

It evaluates a degree of inclusion of the fuzzy set E in the set of models of p, and does correspond to an evaluation of the extent to which the state of knowledge implies p. $N(p)$ evaluates to what extent p is certainly true. $N(p) = 1$ corresponds to absolute belief in p (p is true in all not impossible states of affairs however implausible), and entails its truth. $N(p) > 0$ represents the case when p is expected to be true, meaning that p is true in all normal states of affairs (ω, such that $\pi(\omega) = 1$).

The set-function Π (respectively N) is decomposable with respect to arbitrary disjunctions (respectively conjunctions): properties (1) and (2) still hold. This framework leaves room for the three epistemic attitudes of the agent regarding p, depicted in section 3, and especially total ignorance ($N(p) = N(\neg p) = 0$). Possibility theory is thus a genuine simple extension of the uncertainty theory embedded in the propositional calculus.

4.3. *Believing a fuzzy proposition in the face of incomplete knowledge*

Suppose the epistemic state of the agent is represented by a many-valued possibility distribution π on Ω and the set $[p]$ of models of the proposition p is fuzzy. In that case the truth state of a proposition may altogether be a matter of degree and may be ill-known. There are two orthogonal scales involved in the picture: a truth scale L which contains membership grades $\mu_{[p]}(\omega)$, and a plausibility scale V which contains possibility degrees $\pi(\omega)$ of states of affairs ω according to the agent. For any candidate truth-value α of proposition p, the degree of possibility that "p is true to degree α" can be computed as

$$\tau_\pi(\alpha) = \sup\bigl\{\pi(\omega)\colon\ \mu_{[p]}(\omega) = \alpha\bigr\}. \tag{5}$$

This fuzzy set τ_π of more or less possible truth-values forms a so-called *fuzzy truth-value* [86]. A fuzzy truth-value is more than a graded truth-value. It combines the ideas of partial truth and of uncertainty about truth and is a more complex notion than the usual notion of truth-value. The complete answer to a query evaluating to what extent an agent believes a fuzzy proposition p, when his state of knowledge is described by a possibility π is thus captured by a possibility distribution over truth-values. This possibly hard-to-interpret answer can be summarized by the possibility and necessity of the many-valued proposition p (when $V = L$):

$$\Pi(p) = \sup_{\omega\in\Omega} \min\bigl(\mu_{[p]}(\omega), \pi(\omega)\bigr),$$
$$N(p) = \inf_{\omega\in\Omega} \max\bigl(\mu_{[p]}(\omega), 1 - \pi(\omega)\bigr).$$

These are standard fuzzy pattern matching indices in fuzzy information systems [16]. They are easily retrieved from the fuzzy truth-value as follows (assuming $V = L = [0, 1]$), after Baldwin and Pilsworth [5]:

$$\Pi(p) = \sup_{t\in[0,1]} \min\bigl(t, \tau_\pi(t)\bigr),$$
$$N(p) = \inf_{t\in[0,1]} \max\bigl(t, 1 - \tau_\pi(t)\bigr).$$

To see that this case encompasses all other situations, notice the following:

(i) If p is a Boolean proposition, then $\tau_\pi(\alpha) = 0$ if $\alpha \neq 0, 1$. Moreover, $\tau_\pi(1) = \Pi(p)$ and $\tau_\pi(0) = \Pi(\neg p) = 1 - N(p)$.

(ii) Suppose π is the characteristic function of a classical set of models E of a belief base K. Then $\tau_\pi(\alpha) = 1$ if $\exists\omega \in E$ such that $\mu_{[p]}(\omega) = \alpha$. So τ_π is a classical subset of truth-values, upper bounded by the value $\sup\{\mu_{[p]}(\omega),\ \omega \in E\} = \Pi(p)$ and lower bounded by the value $\inf\{\mu_{[p]}(\omega),\ \omega \in E\} = N(p)$. Note that here $\Pi(p)$ and $N(p)$ are possibility and necessity of a *fuzzy* event, based on a *binary* possibility distribution).

(iii) If both π and p are two-valued, then $L = \{0, 1\} = V$. Then $\tau_\pi = \{1\}$, if $E \subseteq [p]$, $\tau_\pi = \{0\}$ if $E \cap [p] = \emptyset$ and $\tau_\pi = \{0, 1\}$ if $E \cap [p] \neq \emptyset$ and $E \cap [\neg p] \neq \emptyset$. Again, the three-valued description of uncertainty.

Zadeh [83] claims that fuzzy logic is a logic with such a type of fuzzy truth-values, rather than a mere many-valued logic. The terminology "fuzzy truth-value" is somewhat misleading, since a fuzzy truth-value is a fuzzy set of truth-values, and not a single graded truth-value. The move from a many-valued logic to a fuzzy-valued logic in the sense above is very different from the move from Boolean logic to many-valued logics (where the truth set {0, 1} is turned into a more general totally ordered set). It does not consist in changing a many-valued truth set L into another truth set made of fuzzy sets. Namely, in the Zadehian fuzzy logic with "fuzzy truth-values", the fuzzy propositions are still evaluated on L; a fuzzy proposition is, in a given state of affairs, still true to a precise level in L. A fuzzy truth-value, in this setting is not a truth-value in the usual sense, it is a possibility distribution on truth-values, a description of the state of belief of an agent regarding a proposition p, as clearly indicated in the Boolean situation (iii) above, where τ_π proves to be a non-empty *subset* of {0, 1}.

So, the move proposed by Zadeh towards fuzzy truth-values is to start computing an agent's belief in a fuzzy proposition on the basis of a possibility distribution describing the agent's knowledge. This move is the exact counterpart, for many-valued logics, to the attempt we made to lay bare the uncertainty theory in propositional calculus, which yields possibility theory. We did it first by considering a set of formulas as a belief base, thus using logic as a tool for representing knowledge (rather than considering it as metamathematics). What Zadeh proposed, by offering fuzzy logic with fuzzy truth-values, is to add possibility theory to many-valued logics, thus making them more suitable for knowledge representation and reasoning about knowledge and belief. To support this claim, note that the various existing many-valued logic systems (see the extensive presentation by Hajek [53]) never had much impact on AI research, in the community of knowledge representation. In our opinion, this is because a many-valued logic without the underlying belief evaluation tools is of little use for reasoning about knowledge. However, Zadeh never produced the syntactic machinery for uncertain reasoning in fuzzy logic. This also created some misunderstandings with logicians, for whom a logic without a syntactic machinery is no logic (see [52], for instance). Possibilistic logic [32,33] is one step towards such a syntactic tool capturing fuzzy logic in the sense advocated by Zadeh [83], although it only extends the uncertainty calculus of propositional logic to the many-valued case, while keeping binary propositional variables (see section 6).

Changing the possibility distribution π into a probability measure P on the frame of discernment, the fuzzy truth-value τ_π becomes a probability distribution on the truth set (the set of membership grades), namely $P(\alpha) = \sum_\omega \{P(\omega)\colon \mu_{[p]}(\omega) = \alpha\}$, in finite settings. This probability measure is again not a truth-value.

5. The impossibility of compositionality in weighted logics of uncertainty

An important consequence of the above distinction between degrees of truth and degrees of belief is that degrees of belief bearing on classical propositions cannot be compositional for all connectives. The notion of compositional degree of belief is mathematically inconsistent, in general. This result holds when the information that is used to evaluate propositions is incomplete. By contrast, truth-values of propositions (fuzzy or not) are compositional when these truth-values can be precisely evaluated (i.e., under complete information). Sets of fuzzy propositions (interpreted as families of fuzzy subsets) are no longer Boolean algebras but form weaker structures compatible with the unit interval. For instance, using max, min, $1 - (\cdot)$ for expressing disjunction, conjunction and negation of fuzzy propositions equips sets of such propositions with a distributive lattice structure that is compatible with the unit interval (it is a De Morgan Algebra, containing a distributive lattice); this structure is the only one where all laws of Boolean algebra hold except the laws of non-contradiction and of excluded middle [6]. Under complete information, truth-values in the unit interval can be compositional. Sometimes, arguments against fuzzy set theory rely on the impossibility of compositionality (e.g., [45,81]). Usually these arguments are based on the wrong assumption that the algebra of propositions to be evaluated is Boolean, and what are called "degrees of fuzziness", or truth-values, are actually degrees of belief (see the discussions of Elkan's paper [45]). Note that *fuzzy* truth-values (introduced in the previous section) are not truth-functional, generally. This section summarizes existing trivialization results pertaining to the compositionality of measures of confidence attached to Boolean propositions, points out some open problems, and draws some conclusions pertaining to the use of truth-functional logics. It enables some mistakes and misunderstandings to be corrected concerning possibility theory.

5.1. The assumption of fully compositional beliefs

Let $\mathbf{B}(\wedge, \vee, \neg)$ be a Boolean algebra, assumed finite for simplicity and equipped with the usual entailment ordering $\models$. $\mathbf{B}$ represents, for instance, a classical language quotiented by the equivalence relation between formulas, or may equivalently represent the set of subsets of a finite referential $\Omega = \{\omega_1, \ldots, \omega_n\}$, i.e., $2^\Omega = \mathbf{B}$. The top of $\mathbf{B}$ is denoted $\mathbf{T}$ and $\perp$ its bottom. Elements of $\mathbf{B}$ are denoted $p, q, r, \ldots$, and correspond to events $A, B, C, \ldots$ when $2^\Omega = \mathbf{B}$. The entailment ordering $p \models q$ corresponds to the inclusion $[p] \subseteq [q]$.

A confidence measure g can thus be viewed as a function $\mathbf{B} \to V$, where V is a totally ordered scale with a maximum 1 and a minimum element 0, and such that, $\forall p, q \in \mathbf{B}$,

$$g(\bot) = 0, \qquad g(\mathbf{T}) = 1,$$
$$p \models q \Rightarrow g(p) \leqslant g(q) \quad \text{(monotonicity)}.$$

Examples of confidence measures are probability measures, possibility and necessity measures, belief functions, etc. The full compositionality assumption reads as follows: g is *fully compositional* if and only if there exists a function $\nu : V \to V$, and two binary operations $*$, $\oplus$ on V such that

$$g(\neg p) = \nu\bigl(g(p)\bigr), \tag{6}$$
$$g(p \vee q) = g(p) \oplus g(q), \tag{7}$$
$$g(p \wedge q) = g(p) * g(q). \tag{8}$$

The following results should be noticed.

Fact 1. The negation function ν is order-reversing and involutive. In particular, $\nu(1) = 0$.

This is due to entailment preservation. Namely, if $p \models q$ then $\neg q \models \neg p$. Let $\alpha = g(p)$ and $\beta = g(q)$. It follows that if $\alpha \leqslant \beta$ then $\nu(\beta) \leqslant \nu(\alpha)$. Similarly, $\nu(\nu(\alpha)) = g(\neg\neg p) = g(p) = \alpha$. Since ν is decreasing in the wide sense and involutive, it is bijective, hence order-reversing. The identity $\nu(1) = 0$ is due to $g(\bot) = \nu(g(\mathbf{T}))$.

Fact 2. $\alpha \oplus \beta = \nu(\nu(\alpha) * \nu(\beta))$, $\forall \alpha, \beta \in V$.

This is an obvious consequence of De Morgan laws.

Fact 3. $g(p \vee q) = \max(g(p), g(q))$; $g(p \wedge q) = \min(g(p), g(q))$.

Indeed, due to monotonicity, $g(p \vee q) = g(p) \oplus g(q) \geqslant \max(g(p), g(q))$. Now suppose that $p \models q$, so that $g(q) \geqslant g(p)$. Then $g(p \vee q) = g(p) \oplus g(q) = g(q)$: the idempotence of $\oplus$ is enforced by the one of the Boolean disjunction. The proof is similar for conjunction.

Fact 4. $\forall p$, $g(p) = 0$ or $g(\neg p) = 1$.

To see it, just let $q = \neg p$ in fact 3 which then reads

$$\max\bigl(g(p), g(\neg p)\bigr) = 1, \qquad \min\bigl(g(p), g(\neg p)\bigr) = 0.$$

Fact 5. $g(p) = 0 \Rightarrow g(\neg p) = 1$.

Due to $\nu(1) = 0$ and fact 4. Alternatively, assume $g(p) = g(\neg p) = 1$. Then, fact 3 yields $0 = g(\bot) = g(p \wedge \neg p) = \min(g(p), g(\neg p)) = 1$. A contradiction.

Hence, fact 5 follows from fact 3.

Remark. The above proofs should be understood as properties which hold, fixing the connectives $\nu(\cdot)$, $\oplus$, $*$, and *varying the confidence function* g. This is especially important if we choose $V = [0, 1]$, as we have assumed a finite setting. If we fix g, then the choice of connectives in [0, 1] is much wider, as the attained range of g is a finite subset of the unit interval. For instance, in fact 1, ν is requested to be involutive only on a finite subset of [0, 1]. There are many values $\alpha \in [0, 1]$ which are not of the form $g(p)$ for a fixed g. On these values, there is no constraint on the choice of the functions representing the connectives. However, for any p, q, α, β , there exists a confidence function g such that $\alpha = g(p)$ and $\beta = g(q)$. On this issue, see Paris [69, chap. 3] and Halpern's [54] discussions of a theorem of Cox [18] proving the unicity of probability measures from first principles on a finite set of events.

These results prove the main part of the following theorem, that is, they force any fully compositional confidence function to be deterministic:

Theorem 1. The family of confidence functions fully compositional with respect to conjunction, disjunction and negation on a Boolean algebra coincides with the set of binary truth-assignments. Moreover, any 2 properties among (6)–(8) imply the third one.

Proof. Equations (7) and (8) imply (6) once it is noticed that fact 3 results from (7) and (8), and alone implies facts 4 and 5, that in turn imply (6). □

For instance, if $\mathbf{B} = 2^{\Omega}$, and g is a fully compositional confidence measure on events, it means that there is an element ω in Ω such that for any subset A of Ω, $g(A) = 1$ if A contains ω, and 0 otherwise. In other words, there is no uncertainty and the actual state of affairs is known to be ω.

This result is basically proved independently and with different notations, in [38, 81] and indirectly [45]. It is also a direct consequence of the well-known fact in mathematics that a non-trivial Boolean algebra that is linearly ordered has only two elements, since V has the same structure as $\mathbf{B}$, due to (6)–(8). Compositional belief seems to have been so popular that Paris [69, chapter 5] devotes a whole chapter of his book to it. However, he also points out some of the limitations of such an assumption. In [9], the same author and colleagues even propose a least-squares-based evaluation of the extent to which a fully compositional belief calculus (they improperly call a "fuzzy logic") approximates probability theory.

5.2. *Almost preserving compositionality*

As said above, measures of confidence defined on a Boolean algebra and taking values in a chain V containing more than two elements cannot be fully compositional with respect to all the logical connectives, just because we cannot equip V with a structure of Boolean algebra. Since full compositionality cannot hold, one may think of being as compositional as possible. The case used to trivialize (7) and (8) is when $q = \neg p$. Hence the definition:

Definition. g is *almost compositional* if and only if (6) holds, and (7), (8) always hold except when $q = \neg p$, i.e., we assume $g(p \wedge \neg p) = 0,\ g(p \vee \neg p) = 1$.

This type of set-function was actually introduced by Schwartz [74] who proposed a logic of likelihood, using a self-dual confidence function (called likelihood) governed by the following laws, for all p, q

$$g(\neg p) = \nu\big(g(p)\big), \tag{9}$$

$$g(p \vee q) = \max\big(g(p), g(q)\big), \quad \text{except if } p = \neg q, \tag{10}$$

$$g(p \wedge q) = \min\big(g(p), g(q)\big), \quad \text{except if } p = \neg q. \tag{11}$$

The connectives in (9) and (10) are enforced by the idempotence of Boolean conjunctions and disjunctions, as shown in the proof of theorem 1 (fact 3); moreover, (9) and (10) imply (11). Such a measure of confidence g is as compositional as possible to avoid trivialization by reduction to a binary truth-assignment. Moreover, only operations with a qualitative flavor are used to combine the likelihood degrees. Only a totally ordered set equipped with an order reversing involution is required as a likelihood scale. In the following we investigate what is the power of expressiveness of these measures of likelihood, in the finite case.

For simplicity, let $\mathbf{B} = 2^{\Omega}$, and the elements $\omega_1, \ldots, \omega_n$ of Ω be the atoms of the Boolean algebra. For $i = 1, \ldots, n$, denote $g_i = g(\omega_i) \in V$ the values of the uncertainty distribution of g on Ω throughout the section. The next trivialization result is described by the following.

Theorem 2. Suppose a function g from $\mathbf{B} = 2^{\Omega}$ to a totally ordered scale V belongs to the family of almost compositional confidence functions. Assuming $g_1 \geqslant g_2 \geqslant \cdots \geqslant g_n$ for the distribution of g without loss of generality, it holds that:

(i) $g(\mathbf{B}) \subset V$ has at most four elements,

(ii) if $g_1 = 1$, then $\forall j \neq 1,\ g_j = 0$, and g is a binary truth-assignment,

(iii) if $\exists \alpha,\ 0 < \alpha < 1,\ g_1 = \alpha$, then $\alpha \geqslant \nu(\alpha)$, and $\forall i \neq 1,\ g_i = \nu(\alpha)$.

Proof. First, note that

$$\forall i, \quad g_i = \nu\big(g\big(\Omega - \{\omega_i\}\big)\big) = \nu\Big(\max_{j \neq i} g_j\Big) = \min_{j \neq i} \nu(g_j).$$

So, if $\exists i,\ g_i = 1$, then $g(\Omega - \{\omega_i\}) = 0$, and then $\forall j \neq i,\ g_j = 0$. It corresponds to the *deterministic* case. Let us suppose that $\exists i,\ g_i = \alpha \in (0, 1)$. Let us suppose, without loss of generality, that $\alpha = g_1 \geqslant g_2 \geqslant \cdots \geqslant g_n$. Then, applying (9) and (10):

$$\forall j \neq 1, \quad g_j \leqslant \max_{k \neq 1} g_k = g\big(\Omega - \{\omega_1\}\big) = \nu(g_1) = \nu(\alpha),$$
$$g_2 = \nu\big(g\big(\Omega - \{\omega_2\}\big)\big) = \nu\Big(\max_{j \neq 2} g_j\Big) = \nu(g_1) = \nu(\alpha).$$

Since g_1 is the maximal level, it follows that $\alpha \geqslant \nu(\alpha)$. Similarly we have:

$$g_3 = \nu\big(\max(g_1, g_2, g_4, \ldots, g_n)\big) = \nu(\alpha), \quad \ldots, \quad g_n = \nu(\alpha).$$

Thus if $g_1 < 1$, we can only have $1 > g_1 \geqslant g_2 = g_3 = \cdots = g_n = \nu(g_1) > 0$. So we can describe only one of the following:

- either a pseudo-deterministic situation where $\exists i,\ g_i = \alpha$, and $\forall j \neq i,\ g_j = \nu(\alpha) < \alpha$,
- or a case of total uncertainty described by $\forall i,\ g_i = \alpha = \nu(\alpha)$ (provided that V contains such an element). □

This theorem trivializes the almost compositionality assumption: indeed, situation (ii) is the case of no uncertainty, and $V = \{0, 1\}$. Situation (iii) is almost deterministic as soon as $\lambda = \alpha > \nu(\alpha) = \upsilon$ and $V = \{0, \upsilon, \lambda, 1\}$. Then ω_1 is a *likely* situation (λ) and ω_i is *unlikely* (υ) otherwise (but never impossible). Then $g(A) = \lambda$ if $\omega_1 \in A \neq \emptyset$; $g(A) = \upsilon$ if $\omega_1 \notin A \neq \emptyset$, $g(\Omega) = 1$, $g(\emptyset) = 0$. There exists *one* alternative ω_1 which, without being necessarily completely certain, appears to be more likely than the others which are considered as having a smaller, undifferentiated level of likelihood. Lastly, if $\alpha = \nu(\alpha)$, then α is the median element of $V = \{0, \alpha, 1\}$ and g expresses complete uncertainty: $g(A) = \alpha\ \forall A \neq \emptyset, \Omega$. This representation of belief does not really need the unit interval since at most a 4-element totally ordered set $\{0, \lambda, \upsilon, 1\}$ is needed. This proposal corresponds to the most elementary logic of likelihood which can be imagined. This framework, called the "simplified English probabilistic logic" by Aleliunas [3], is not sufficiently expressive for any practical use.[1] In the following, a confidence measure satisfying (9)–(11) is called a *simplified linguistic probability* function (SLP).

One could try and escape the above trivialization result by dropping assumption (9) (negation-compositionality). g is said to be *weakly almost compositional* iff (10) and (11) hold. A trivialization result again follows:

Theorem 3. Suppose a function g from $\mathbf{B} = 2^{\Omega}$ to a totally ordered scale V belongs to the family of weakly almost compositional confidence functions. Assuming $g_1 \geqslant g_2 \geqslant \cdots \geqslant g_n$ for the distribution of g, without loss of generality, it holds that:

(i) $g(\mathbf{B})$ has at most 4 elements $\{0 < \upsilon < \lambda < 1\}$,

[1] In more recent works, Schwartz [75] uses axioms (9)–(11), only when p and q are distinct independent literals, thus avoiding trivialization.

(ii) if $g_1 = g_2 \neq 1$, then $\forall i,\ g_i = \alpha \in V$,

(iii) if $1 \neq g_1 > g_2$ then $\forall i > 2,\ g_1 = \lambda > g_i = g_2 = \upsilon$.

Proof. Using the above notations, note that, applying (11) to $\{\omega_i\} = \bigcap_{j \neq i} \Omega - \{\omega_\mathrm{j}\}$

$$g_i = \min_{j \neq i} g\big(\Omega - \{\omega_j\}\big),$$

where

$$g\big(\Omega - \{\omega_j\}\big) = \begin{cases} g_1 & \text{if } j \neq 1 \text{ applying (10),} \\ g_2 & \text{if } j = 1. \end{cases}$$

Hence for $i \geqslant 2,\ g_i = \min(g_1, g_2) = g_2$. □

So we are almost in the same situation as previously: theorem 3 allows the situations $g_i = \alpha \in V$ for any α, while this was allowed only if $\alpha = \nu(\alpha)$ by theorem 2. However, the gain in expressivity is not significant.

Comments and open problem

Since full compositionality of confidence functions cannot hold without trivialization, and especially only one of properties (6)–(8) may hold at a time, each of these three assumptions lead to one family of set-functions which are fully compositional with respect to one connective:

(i) $g(p) = \nu(g(\neg p))$ alone: it is too weak to allow for a convenient use of this set function; however these self-dual set-functions, that are compositional with respect to negation are called participation measures [78]. They include probability measures.

(ii) $g(p \vee q) = \max(g(p), g(q))$ alone: g is a possibility measure and it leads to possibilistic logic [32,33], not fuzzy logic; generally, only $g(p \wedge q) \leqslant \min(g(p), g(q))$ holds.

(iii) $g(p \wedge q) = \min(g(p), g(q))$ alone: g is a necessity measure, obtained from possibility measures by duality: $g^*(p) = \nu(g(\neg p))$ is a necessity measure iff and only if g is a possibility measure. It again leads to possibilistic logic.

Restricting compositionality with respect to disjunction to when $p \wedge q = \bot$, the choice of operation $\oplus$ as maximum is no longer compulsory. For instance, decomposable measures are obtained [36,68,80]. Then, along with negation compositionality (6), this assumption of course no longer leads to trivialization, as shown by the existence of probability measures, when $\oplus$ is the bounded sum ($\alpha \oplus \beta = \min(1, \alpha + \beta)$) and $V = [0, 1]$; with grades of probability, $P(p) = 1 - P(\neg p)$ but $P(p \vee q) = P(p) + P(q)$ only for mutually exclusive events, and $P(p \wedge q) = P(p) \cdot P(q)$ only in situations of stochastic independence. Probability theory cannot be cast in a straightforward manner into a truth-functional many-valued logic whose truth set would be [0, 1].

An open problem is the study of probability-like measures, compositional with respect to negation, taking values in a finite totally ordered set. Of course, this study should be based on finite monotonic semi-groups [47]. This type of qualitative uncertainty theory has been envisaged by Aleliunas [2–4].

It is interesting to see whether SLP functions induce a comparative probability ordering on events. Namely a comparative probability ordering $\geqslant$ is a complete and transitive relation among events $A \subseteq \Omega$, such that $A \geqslant \emptyset$, $\forall A \subseteq \Omega$, and $\geqslant$ satisfies the *pre-additivity* axiom [46]:

$$\forall A_1, \quad \text{if } A_1 \cap (A_2 \cup A_3) = \emptyset, \quad \text{then } A_2 > A_3 \Leftrightarrow A_1 \cup A_2 > A_1 \cup A_3, \tag{12}$$

where $A_1 > A_2$ means $A_1 \geqslant A_2$ and not $(A_2 \geqslant A_1)$. Any non-degenerate SLP function g classifies the events in Ω into 4 classes of level $0 < \upsilon < \lambda < 1$, respectively. Namely, $\exists \omega_1$ such that the class of level λ is $\{A \neq \Omega,\ \omega_1 \in A\}$, the class of level υ is $\{A \neq \emptyset,\ \omega_1 \notin A\}$. The class of level 1 is $\{\Omega\}$ and the one of level 0 is $\{\emptyset\}$. Particularly, we have, for $A_1 \neq A_2$, $A_1 > A_2$ if and only if $A_1 = \Omega$ or $A_2 = \emptyset$ or ($\omega_1 \in A_1$ and $\omega_1 \notin A_2$). Let us consider whether (12) holds.

Theorem 4. Any non-degenerate SLP function only satisfies the following relaxation of pre-additivity: $\forall A_1, A_2, A_3$ such that (i) either $A_2 = \Omega$, or (ii) $A_1 \cup A_2 \neq \Omega$, $A_1 \cap (A_2 \cup A_3) = \emptyset$, $A_3 \neq \emptyset$: $A_2 > A_3 \Leftrightarrow A_1 \cup A_2 > A_1 \cup A_3$. If $A_1 \cup A_2 = \Omega$, or $A_3 = \emptyset$, pre-additivity may fail.

Proof. If $A_2 = \Omega$ then $A_1 = \emptyset$ and (12) is trivial.

From now on $A_1 \neq \emptyset$; hence $A_2 \neq \Omega$. If $A_3 \neq \emptyset$ then assume $A_2 > A_3$. It means $\omega_1 \in A_2$, $\omega_1 \notin A_3$. Since $A_1 \cap A_2 = \emptyset$, $\omega_1 \notin A_1$. Hence $\omega_1 \notin A_1 \cup A_3$ and $A_1 \cup A_2 > A_1 \cup A_3$.

Conversely, by assumption $\Omega \neq A_1 \cup A_2 > A_1 \cup A_3$. Clearly, $A_1 \cup A_3 \neq \emptyset$; we have $\omega_1 \in A_1 \cup A_2$, $\omega_1 \notin A_1 \cup A_3$. Hence, $\omega_1 \notin A_1$, and $\omega_1 \in A_2 - A_1$. Hence, $A_2 > A_3$.

However, consider the case $A_1 \cup A_2 = \Omega > A_1 \cup A_3$; then since $A_1 \cap (A_2 \cup A_3) = \emptyset$, it follows that $A_3 \subseteq A_2$. If $\omega_1 \in A_2 - A_3$ then $g(A_1 \cup A_2) = 1 > g(A_1 \cup A_3) = \upsilon$ and $g(A_2) = \lambda > g(A_3) = \upsilon$. If $\omega_1 \in A_3$ we have $g(A_2) = g(A_3) = \lambda$ and $g(A_1 \cup A_2) = 1 > g(A_1 \cup A_3) = \lambda$. Hence, (12) fails when $A_1 \cup A_2 = \Omega$.

When $A_3 = \emptyset$ then (12) fails too, if $\omega_1 \in A_1$, since then $g(A_2) > g(A_3)$ but $g(A_1 \cup A_2) = g(A_1 \cup A_3) = \lambda$, generally. $\square$

As a consequence, the SLP function *almost* satisfies the axioms of a comparative probability relation.

5.3. Consequences of the trivialization: possibility theory vs. fuzzy logic

Several authors (e.g., [17,51,63,71,88] presupposed that the following properties of a confidence function were characteristic of possibility measures[2]:

$$g(\neg p) = 1 - g(p), \qquad g(p \vee q) = \max\bigl(g(p), g(q)\bigr),$$
$$g(p \wedge q) = \min\bigl(g(p), g(q)\bigr).$$

The above results show that it cannot be so, without trivializing the confidence function. In fact, these connectives are formally those of Kleene logic. This is a recurring mistake that has created a latent confusion between two distinct calculi: (i) possibility theory on a Boolean lattice of propositions, and (ii) fuzzy set degrees of membership, which can be interpreted as degrees of truth with respect to fuzzy propositions forming a complete pseudo-complemented, non-Boolean, distributive lattice [41]. This state of facts has generated many a misunderstanding both within the field of fuzzy set theory and also among fuzzy set opponents ([17], for instance).

In possibility theory, a possibility measure, usually denoted Π, is a function from 2^{Ω} to the chain V, characterized in the finite case by the *single* compositionality axiom [85]:

$$\Pi(p \vee q) = \max\bigl(\Pi(p), \Pi(q)\bigr), \tag{13}$$

where p and q are classical *propositions* which can be true or false only (no intermediary degrees of truth). The underlying structure for such propositions is a Boolean algebra. As recalled in the previous section, $\Pi(p)$ estimates the possibility of p being true, not a truth-value. It is generally assumed that classical tautologies have possibility 1 and classical contradictions have possibility 0. The epistemic status of p can be estimated through the *pair* of values $\Pi(p)$, and $\Pi(\neg p)$ which is the possibility of p being false. Equivalently we can use the pair $(\Pi(p), N(p))$ where $N(p)$ is the certainty or necessity of p being true, defined by duality between possibility and necessity, i.e., on the unit interval [35,37]:

$$N(p) = \nu\bigl(\Pi(\neg p)\bigr), \tag{14}$$

where ν is the order-reversing function on V. It must be noticed $\nu(\Pi(\neg p))$ is not equal to $\Pi(p)$ (contrary to what is found in some literature) but rather defines a new measure N whose characteristic property, the min-decomposability with respect to conjunction, follows directly from (13) and (14), namely the property

$$N(p \wedge q) = \min\bigl(N(p), N(q)\bigr). \tag{15}$$

Any possibility measure Π on a finite set Ω is characterized by its possibility distribution π, a function from Ω to V, and Π is recovered by equation (13). So a possibility measure has a representation whose complexity is linear with the number of elements

[2] Paris [69] avoids this pitfall, although he deals with possibility theory in his chapter on truth-functional beliefs.

in Ω. The possibility distribution π is the membership function of a fuzzy set F (the one of possible state of affairs), i.e., $\pi = \mu_F$. However, in properties (13)–(15) only this fuzzy set is implicitly involved, while when fuzzy sets connectives are defined as usual:

$$\mu_{F^c} = 1 - \mu_F, \qquad \mu_{F\cup G} = \max(\mu_F, \mu_G), \qquad \mu_{F\cap G} = \min(\mu_F, \mu_G),$$

more that one fuzzy set is involved. In general, *only* the following inequalities hold in possibility theory, as shown in the Boolean case in section 2.

$$N(p \vee q) \geqslant \max\bigl(N(p), N(q)\bigr),$$
$$\Pi(p \wedge q) \leqslant \min\bigl(\Pi(p), \Pi(q)\bigr).$$

Indeed, let $q \equiv \neg p$, for instance; we always have $N(p \vee q) = N(p \vee \neg p) = N(\mathbf{T}) = 1$ and $\Pi(p \wedge q) = \Pi(p \wedge \neg p) = \Pi(\bot) = 0$ (where $\mathbf{T}$ and $\bot$ denote tautology and contradiction, respectively), while $\Pi(p)$ and $\Pi(\neg p)$ remain free provided they respect the constraint

$$\max\bigl(\Pi(p), \Pi(\neg p)\bigr) = 1, \tag{16}$$

a direct consequence of (13) and $\Pi(\mathbf{T}) = 1$ (since $p \vee \neg p \equiv \mathbf{T}$). We may have $\Pi(p) = \Pi(\neg p) = 1$. Note that (16) can be equivalently written, using (13):

$$N(p) > 0 \Rightarrow \Pi(p) = 1.$$

The equality $\Pi(p \wedge q) = \min(\Pi(p), \Pi(q))$ holds for two given propositions p and q if and only if $\Pi(p \wedge q) \geqslant \min(\Pi(\neg p \wedge q), \ \Pi(p \wedge \neg q))$ [28]. Moreover the equality $\Pi(X \wedge Y) = \min(\Pi(X), \Pi(Y))$ holds for $X \in \{p, \neg p\}, \ Y \in \{q, \neg q\}$, if and only if X and Y are non-interactive Boolean variables, as pointed out by Zadeh [84]. It means that

(i) X and Y are logically independent (none of the four possible situations $p \wedge q, \ \neg p \wedge q, \ p \wedge \neg q, \ \neg p \wedge \neg q$ is a contradiction), and

(ii) $\exists (X, Y) \in \{p, \neg p\} \times \{q, \neg q\}$ such that $\Pi(X \wedge Y) \geqslant \max(\Pi(\neg X \wedge Y), \Pi(X \wedge \neg Y))$ and $\Pi(\neg X \wedge \neg Y) = \min(\Pi(\neg X \wedge Y), \Pi(X \wedge \neg Y))$.

It implies that the joint possibility $\Pi(X \wedge Y)$ can at most take 3-values $1 \geqslant \alpha \geqslant \beta$ since it is determined by the marginals $\Pi(p), \Pi(\neg p), \Pi(q), \Pi(\neg q)$ with $\max(\Pi(p), \Pi(\neg p)) = 1, \ \max(\Pi(q), \Pi(\neg q)) = 1$. See [28,29] for an extensive study of conditioning and event independence in possibility theory, and [20] as well as [7] for the study of independence between possibilistic variables. The consequences of enforcing truth-functionality in possibility theory are also studied by De Cooman [22] who shows that it results in a conservative approximation closely related to Kleene systems. Mathematical aspects of qualitative possibility theory are studied in De Cooman [21].

The situation of possibility measures is similar to the one of probability measures which are compositional for negation only, but where $\mathrm{Prob}(p \wedge q)$ does not depend only on $\mathrm{Prob}(p)$ and $\mathrm{Prob}(q)$ generally, nor does $\mathrm{Prob}(p \vee q)$ depend on $\mathrm{Prob}(p), \ \mathrm{Prob}(q)$ only, except for special situations. $\mathrm{Prob}(p \wedge q) = \mathrm{Prob}(p) \cdot \mathrm{Prob}(q)$ requires that p and q be stochastically independent.

Remark. Possibility measures have been shown to be the numerical counterpart of so-called qualitative possibility relations $\trianglerighteq$ on sets of events (where $A_1 \trianglerighteq A_2$ reads "A_1 is at least as possible as A_2"). Possibility relations $\trianglerighteq$ are supposed to be reflexive, complete ($A_1 \trianglerighteq A_2$ or $A_2 \trianglerighteq A_1$), transitive, non-trivial ($\Omega \triangleright \emptyset$), such that $\forall A,\ A \trianglerighteq \emptyset$ (impossibility of contradiction) and to satisfy the characteristic axiom [27,39]:

$$\forall A_1, \quad \text{if } A_2 \trianglerighteq A_3 \quad \text{then } A_1 \cup A_2 \trianglerighteq A_1 \cup A_3. \tag{17}$$

A full-fledged representation theorem holds. Namely, for any possibility relation $\trianglerighteq$, there exists a possibility measure Π such that

$$\forall A_1,\ \forall A_2, \quad \Pi(A_1) \geqslant \Pi(A_2) \quad \text{if and only if } A_1 \trianglerighteq A_2,$$

and conversely any possibility measure induces a possibility relation. For comparative necessity the above axiom (17) is changed by substituting $\cap$ to $\cup$. This result shows the qualitative nature of possibility and necessity measures.

In contrast to possibility theory, whose extensive study is recent, the fully compositional calculus behind fuzzy set theory, defined for multiple-valued logical variables using the De Morgan triple (max, min, $1 - \cdot$), has already been extensively studied in the seventies. On the problem of checking whether two fuzzy logic expressions are equivalent, see [62,72]. The latter prove that fuzzy logic is isomorphic to a 3-valued logic, i.e., two fuzzy propositions are equivalent in [0, 1] iff they are such on {0, 1/2, 1}. Syntactic aspects of this fuzzy logic are studied by Lee [61]. The fact that some classical logic equivalences no longer hold is not surprising since the involved algebra is not Boolean. However, the properties of a complete pseudo-complemented distributive lattice are enough to guarantee that for any formula under a conjunctive normal form, there exists a fuzzy logic equivalent disjunctive normal form as shown by Greenberg [51]. There are a lot of older works on fuzzy normal forms, noticeably Davio and Thayse [19], Kandel [56] and Mukaidono [66]. For more details, the reader is invited to consult the following monographs and surveys: Kandel and Lee [57], Dubois and Prade [35, pp. 151–158], Mukaidono [66] Dubois, Lang and Prade [31]. More general De Morgan structures are studied by Gehrke et al. [49]. See Klement and Navarra [59] for logical axiomatizations of general truth-functional calculi of this type. This literature differs from the residuated structures studied in the many-valued logics described in Hajek [53] in which the negation is the by-product of the implication ($\neg p = p \to \bot$) and is not involutive (see also the survey by Gottwald [50]).

6. The case for possibilistic logic

Since possibility theory is a genuine and simple extension of the uncertainty theory embedded in the propositional calculus, one may wonder whether there is a type of logic that enables states of knowledge such as possibility distributions to be expressed at a syntactical level. The answer is positive, and this is possibilistic logic [32,33].

A possibilistic belief base is a weakly ordered set of propositions $(K, \geqslant)$, which can be encoded as well as a set of pairs $\{(p_i, \alpha_i),\ i = 1, n\}$ where p_i is a Boolean proposition and α_i is a weight in a totally ordered scale ([0, 1] for simplicity). The corresponding fuzzy set E of interpretations induced by K is retrieved as follows.

First, each weighted formula (p_i, α_i) is interpreted as the constraint $N(p_i) \geqslant \alpha_i$, expressing that p_i is believed to degree at least α_i. The corresponding fuzzy set E_i of interpretations is defined by its membership function π_i such that

$$\pi_i(\omega) = \begin{cases} 1 & \text{if } \omega \text{ is a model of } p_i, \\ 1 - \alpha_i & \text{if } \omega \text{ is not.} \end{cases}$$

Namely, the higher the degree of belief in p_i, the lower the plausibility of counter-models of p_i. Note that when two fuzzy sets E and F representing states of knowledge are such that $E \subseteq F$, then F is less complete (= informative) than E. The possibility distribution π_i is the least informative (least committed) possibility distribution such that $N(p_i) \geqslant \alpha_i$ holds. Then the possibility distribution representing the fuzzy set of models of $(K, \geqslant)$ is simply obtained, using a fuzzy set intersection:

$$\pi = \min_{i=1,\ldots,n} \pi\pi_i. \tag{18}$$

K is logically consistent if and only if π is normalized. Moreover, the inequality $N(p) \geqslant \alpha$ holds if and only if the weighted proposition (p, α) can be syntactically derived from $\{(p_i, \alpha_i),\ i = 1, \ldots, n\}$ using the inference machinery of possibilistic logic. Indeed possibilistic logic is sound and complete with respect to possibility theory [33].

Possibilistic logic bears some similarity with so called "signed logics". These many-valued weighted logics handle pairs (p, I) consisting of a many-valued proposition and a set of truth-values I [55]. The meaning of (p, I) is that the truth-value of p is constrained to lie in the set $I \subseteq L$. These semantics radically differs from the kind of intuition developed in possibilistic logic. Especially, while p is a many-valued proposition in signed logics, it is clear that the pair (p, I) can only be true or false, according to whether the truth-value of p lies in I or not. In some sense, attaching a weight α to a fuzzy proposition p, viewed as an interval $[\alpha, 1]$, makes (p, I) Boolean. On the contrary, in possibilistic logic, attaching a weight to a Boolean proposition makes it many-valued. However, both the sign I in (p, I) and the weight α in (p, α) can be viewed as special cases of fuzzy truth-values in the sense of Zadeh, since a fuzzy truth-value is a fuzzy set of truth-values: while the sign I is a usual subset of the many-valued truth set, the weight α of (p, α) in possibilistic logic encodes a fuzzy set of $\{0, 1\}$ with membership function $\mu(1) = \Pi(p) = 1$ and $\mu(0) = \Pi(\neg p) = 1 - \alpha$ as explained in section 4.3.

An interesting issue is the comparison of possibilistic logic and the max–min De Morgan logic on the unit interval (e.g., [14,59,61]). Indeed, despite their basic difference (the former is a non-fully compositional logic of uncertainty, and the other is a many-valued logic based on a non-Boolean structure) they have many similarities, both heavily relying on the use of maximum and minimum operations, and both admitting the same family of tautologies, if a tautology is understood as a proposition whose truth-value is always at least 1/2. The two weighed logics share the same syntax. However, the extended resolution principle:

$$\text{from } (p \vee r, \alpha) \text{ and } (q \vee \neg r, \beta) \text{ deduce } \big(p \vee q,\ \min(\alpha, \beta)\big),$$

which is valid without any restriction in possibilistic logic, is no longer applicable in Lee's [61] fuzzy resolution logic when $\min(\alpha, \beta) \leqslant 1/2$ (see [31]). A deeper comparison is a matter of further research.

Actually, there is a way of accommodating non compositional uncertainty calculi in the framework of many-valued logics. It consists in interpreting a pair formed by a Boolean proposition and its degree of confidence as a many-valued proposition and treating this pair as a single syntactic entity to be combined by logical connectives. For instance, possibilistic logic can be embedded within Gödel many-valued logic. This program is carried out by Boldrin and Sossai [11], for instance. In Gödel logic (see [50, 53]), the conjunction is based on minimum, the implication is induced via residuation

$$\alpha \to \beta = \sup\{\lambda, \min(\lambda, \alpha) \leqslant \beta\} = \begin{cases} 1 & \text{if } \alpha \leqslant \beta, \\ \beta & \text{otherwise,} \end{cases}$$

the negation is different from the one in fuzzy set theory (it is non-involutive):

$$\sim p = p \to \bot \quad \text{such that } v(\sim p) = \begin{cases} 1 & \text{if } v(p) = 0, \\ 0 & \text{otherwise,} \end{cases}$$

and the semantic entailment is Zadeh's fuzzy set inclusion. Possibilistic logic only uses the Gödel logic conjunction, as shown by the semantics of fuzzy logic: a possibilistic formula (p, α) has a fuzzy set of models and can thus be viewed as a many-valued literal. Putting several possibilistic formulas in a belief base comes down to consider their conjunction $\wedge_{i=1,n}(p_i, \alpha_i)$, which at the semantic level is achieved by the minimum operation between their fuzzy sets of models like in (18). So possibilistic logic can be viewed as a fragment of Gödel logic. It is not clear, however what the negation $(p, \alpha) \to \bot$ means nor what it is useful for.

Remark. Other (non-logical) frameworks exist that justify the representation of epistemic states by possibility distributions, especially:

- Zadeh's [86] approximate reasoning methodology which tries to translate sets of sentences in natural languages into soft constraints on a frame of discernment,
- fuzzy interval analysis [30] where uncertainty is represented by the possibilistic counterpart of random variables.

More generally, the epistemic states of an agent can as well be described by random sets, or families of probability functions. In the case of random sets, propositions are evaluated using belief and plausibility functions of Shafer [76]. In the case of probability families, propositions are evaluated by means of upper and lower probability bounds. Note that weighted belief bases $K = \{(p_i, \alpha_i),\ i = 1, n\}$ where (p_i, α_i) is interpreted in a probabilistic way (e.g., Nilsson [67]) as $P(p_i) = \alpha_i$ lead to a semantic view of epistemic states in terms of families of probability functions. Indeed, most of the time K induces more than one probability function on the frame of discernment (and sometimes none, if the probability assignments are inconsistent). Deriving the probability (or bounds thereof) of a given proposition p on the basis of a probabilistic belief base K is the central problem addressed by De Finetti [25] in his approach to probability theory (see also [1]). In this approach, the state of total ignorance about p is captured when non-informative probability bounds (0 and 1) are obtained for $P(p)$, in full agreement with propositional calculus, and Boolean possibility theory.

7. Concluding remarks

The intended purpose of this paper is to emphasize the distinction between the treatment of gradual (or vague) propositions in the presence of complete information and the handling of uncertainty for propositions which are either true or false. Non-binary truth-assignment functions of fuzzy propositions can be handled in a fully truth-functional way like in the case of usual propositions. This is, for instance, the case for most applications in fuzzy control where controller inputs are precise, and fuzzy rules are evaluated on the basis of complete information. In the case of uncertainty handling, possibility theory offers a qualitative framework for handling partial belief which can be cast in a logical formalism. Possibility theory, as probability theory and other uncertainty calculi is *not fully compositional* with respect to all connectives. Adopting a compositional uncertainty calculus by mimicking truth-tables of many-valued logics results in an important loss of expressiveness, reducing belief measures to two-valued truth-assignment functions. This paper points out the intrinsic lack of expressiveness of representations of uncertainty that try to take advantage of compositionality as far as possible. It is shown here that, not only the full compositionality of any uncertainty calculus is not possible, but retaining this property as much as mathematical consistency allows, only leads to a very crude, almost deterministic representation of belief.

Another consequence of the lack of compositionality of uncertainty calculi is that many-valued logics, in so far as they are truth-functional, cannot be directly used to represent graded belief. As such, their truth-tables are insufficient for reasoning about knowledge. An uncertainty representation must be added on top of these many-valued logics, for instance intervals of truth-values, attached to formulas. But the bounds of these intervals will not be fully truth-functional. We agree with Elkan [45] on the point that what he calls "truth-functional fuzzy logic" is not adapted to a proper handling of uncertainty in knowledge-based systems. But our agreement is not based on an alleged self-inconsistency of fuzzy set theory leading to a collapse. It is based on the fact that fuzzy set theory primarily offers a calculus of truth-values, and not directly a tool for reasoning under uncertainty. Fuzzy set theory may have to do with uncertainty modeling when membership functions are reinterpreted as possibility distributions, just like sets can be used for the modeling of incomplete information. The uncertainty theory based on sets has been described in section 3. It is ternary, but should not be confused with a three-valued compositional logic. Similarly, the simplest uncertainty theory based on fuzzy sets is possibility theory. Possibility theory can be viewed either as a qualitative approach to uncertainty, or, in the numerical case, one approach among other non-Bayesian uncertainty calculi that rely on probability bounds [23,40].

References

[1] E.W. Adams and H.P. Levine, On the uncertainties transmitted from premises to conclusions in deductive inferences, Synthese 30 (1975) 429–460.

[2] R. Aleliunas, Models of reasoning based on formal deductive probability theories, Draft report, University of Waterloo, Canada (1986).

[3] R. Aleliunas, A new normative theory of probabilistic logic, in: *Knowledge Representation and Defeasible Reasoning*, eds. H.E. Kyburg, Jr., et al. (Kluwer Academic, Dordrecht, 1990) pp. 387–403.

[4] R. Aleliunas, A summary of a new normative theory of probabilistic logic, in: *Uncertainty in AI*, Vol. 4, eds. R.D. Shachter et al. (North-Holland, Amsterdam, 1990) pp. 199–206.

[5] J.F. Baldwin and B.W. Pilsworth, Fuzzy truth definition of possibility measure for decision classification, Int. J. of Man–Machine Studies 11 (1979) 447–463.

[6] R.E. Bellman and M. Giertz, On the analytic formalism of the theory of fuzzy sets, Information Science 5 (1973) 149–157.

[7] N. Ben Amor, S. Benferhat, D. Dubois, H. Geffner and H. Prade, Independence in qualitative uncertainty frameworks, in: *Proc. of 7th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR 2000)*, Breckenridge (2000) pp. 235–246.

[8] S. Benferhat, D. Dubois and H. Prade, Some syntactic approaches to the handling of inconsistent knowledge bases: a comparative study, Studia Logica 58 (1997) 17–45.

[9] A.D.C. Bennett, J.B. Paris and A. Venkovska, A new criterion for comparing fuzzy logics for uncertain reasoning, J. Logic, Language, Information 9 (2000) 31–63.

[10] P. Besnard and A. Hunter, eds., *Reasoning with Actual and Potential Contradictions*, Handbook of Defeasible Reasoning and Uncertainty Management Systems, Vol. 2, eds. D.M. Gabbay and P. Smets (Kluwer Academic, Dordrecht, 1998).

[11] L. Boldrin and C. Sossai, Truth-functionality and measure-based logics, in: *Fuzzy Sets, Logics and Reasoning about Knowledge*, eds. D. Dubois, H. Prade and E.P. Klement (Kluwer Academic, Dordrecht, 1999) pp. 351–380.

[12] J.C. Bezdek and S. Pal, *Fuzzy Models for Pattern Recognition* (IEEE Press, New York, 1996) chapter 1.

[13] B.G. Buchanan and E.H. Shortliffe, *Rule-Based Expert Systems* (Addison-Wesley, Reading, MA, 1984).

[14] D. Butnariu, E.P. Klement and S. Zofrany, On triangular norm-based propositional fuzzy logics, Fuzzy Sets and Systems 69 (1995) 241–255.

[15] R. Carnap, Two concepts of probability, Philosophy and Phenomenological Research 5 (1945) 513–532.

[16] M. Cayrol, H. Farreny and H. Prade, Fuzzy pattern matching, Kybernetes 11 (1982) 103–116.

[17] P. Cheeseman, Probabilistic versus fuzzy reasoning, in: *Uncertainty in Artificial Intelligence 1*, eds. L. Kanal and J. Lemmer (North-Holland, Amsterdam, 1988) pp. 85–102.

[18] R. Cox, Probability, frequency and reasonable expectation, American J. Phys. 14 (1946) 1–13.

[19] M. Davio and A. Thayse, Representation of fuzzy functions, Philips Res. Rep. 29 (1973) 93–106.

[20] L.M. De Campos and J.F. Huete, Independence concepts in possibility theory, Part I: Fuzzy Sets and Systems 103 (1999) 127–152; Part II: Fuzzy Sets and Systems 103 (1999) 487–505.

[21] G. De Cooman, Possibility theory – Part I: Measure- and integral-theoretics groundwork; Part II: Conditional possibility; Part III: Possibilistic independence, Int. J. of General Systems 25(4) (1997) 291–371.

[22] G. De Cooman, From possibilistic information to Kleene's strong multi-valued logics, in: *Fuzzy Sets, Logics and Reasoning about Knowledge*, eds. D. Dubois, H. Prade and E.P. Klement (Kluwer Academic, Dordrecht, 1999) pp. 315–323.
[23] G. De Cooman and D. Aeyels, Supremum-preserving upper probabilities, Information Sciences 118 (1999) 173–212.
[24] B. De Finetti, La logique de la probabilité, in: *Actes Congrès Int. de Philos. Scient.*, Paris 1935 (Hermann et Cie Editions, Paris, 1936) pp. IV1–IV9.
[25] B. De Finetti, *Theory of Probability – A Critical Introductory Treatment*, Vols. 1, 2 (Wiley, Chichester, 1974).
[26] M. De Glas, A few critical remarks on possibility theory, Research report 88/23, LAFORIA, University Paris VI (1988).
[27] D. Dubois, Belief structures, possibility theory and decomposable confidence measures on finite sets, Comput. Artif. Intell. (Bratislava) 5(5) (1986) 403–416.
[28] D. Dubois, L. Fariñas del Cerro, A. Herzig and H. Prade, An ordinal view of independence with application to plausible reasoning, in: *Proc. of the 10th Conf. on Uncertainty in Artificial Intelligence* eds. R. Lopez de Mantaras and D. Poole, Seattle, WA (July 29–31, 1994) pp. 195–203.
[29] D. Dubois, L. Fariñas del Cerro, A. Herzig and H. Prade, Qualitative relevance and independence: A roadmap, in: *Proc. of the 15h Int. Joint Conf. on Artificial Intelligence (IJCAI'97)*, Nagoya, Japan (August 23–29, 1997) pp. 62–67. Extended version "A roadmap of qualitative independence", in: *Fuzzy Sets, Logics and Reasoning about Knowledge*, eds. D. Dubois, H. Prade and E.P. Klement (Kluwer Academic, Dordrecht, 1999) pp. 325–350.
[30] D. Dubois, E. Kerre, R. Mesiar and H. Prade, Fuzzy interval analysis, in: *Fundamentals of Fuzzy Sets* eds. D. Dubois and H. Prade (Kluwer Academic, Dordrecht, 2000) pp. 483–582.
[31] D. Dubois, J. Lang and H. Prade, Fuzzy sets in approximate reasoning – Part 2: Logical approaches, Fuzzy Sets and Systems 40 (1991) 203–244.
[32] D. Dubois, J. Lang and H. Prade, Automated reasoning using possibilistic logic: Semantics, belief revision, and variable certainty, IEEE Trans. on Knowledge and Data Engineering 6 (1994) 64–71.
[33] D. Dubois, J. Lang and H. Prade, Possibilistic logic, in: *Handbook of Logic in Artificial Intelligence and Logic Programming*, Vol. 3 eds. D.M. Gabbay, C.J. Hogger, J.A. Robinson and D. Nute (Oxford University Press, Oxford, 1994) pp. 439–513.
[34] D. Dubois, W. Ostasiewicz and H. Prade, Fuzzy sets: History and basic notions, in: *Fundamentals of Fuzzy Sets*, eds. D. Dubois and H. Prade (Kluwer Academic, Dordrecht, 2000) pp. 21–124.
[35] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications* (Academic Press, New York, 1980).
[36] D. Dubois and H. Prade, A class of fuzzy measures based on triangular norms – A general framework for the combination of uncertain information, Int. J. of General Systems 8(1) (1982) 43–61.
[37] D. Dubois and H. Prade, *Possibility Theory* (Plenum Press, New York, 1988).
[38] D. Dubois and H. Prade, An introduction to possibilistic and fuzzy logics (with discussions), in: *Non Standard Logics for Automated Reasoning*, eds. P. Smets et al. (Academic Press, New York, 1988) pp. 287–315, 321–326.
[39] D. Dubois and H. Prade, Epistemic entrenchment and possibilistic logic, Artificial Intelligence 50 (1991) 223–239.
[40] D. Dubois and H. Prade, When upper probabilities are possibility measures, Fuzzy Sets and Systems 49 (1992) 65–74.
[41] D. Dubois and H. Prade, Possibility theory is not fully compositional! – A comment on a short note by H.J. Greenberg, Fuzzy Sets and Systems 95 (1998) 131–134.
[42] D. Dubois and H. Prade, Possibility theory: qualitative and quantitative aspects, in: *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, Vol. 1, eds. D.M. Gabbay and P. Smets (Kluwer Academic, Dordrecht, 1998) pp. 169–226.
[43] D. Dubois, H. Prade and P. Smets, Representing partial ignorance, IEEE Trans. on Systems, Man and Cybernetics 26 (1996) 361–377.

[44] R. Duda, J. Gaschnig and P. Hart, Model design in the Prospector consultant system for mineral exploration, in: *Expert Systems in the Microelectronic Age*, ed. D. Michie (Edinburgh University Press, 1981) pp. 153–167.

[45] Ch. Elkan, The paradoxical success of fuzzy logic, in *Proc. AAAI'93*, Washington, DC, (July 11–15, 1993) pp. 698–703. Extended version (with discussions), IEEE Expert 9(4) (1994) 2–49.

[46] T. Fine, *Theories of Probability* (Academic Press, New York, 1973).

[47] L. Fuchs, *Partially Ordered Algebraic Systems* (Pergamon Press, Oxford, 1963).

[48] P. Gärdenfors, *Knowledge in Flux* (MIT Press, Cambridge, MA, 1988).

[49] M. Gehrke, C. Walker and E. Walker, De Morgan systems on the unit interval, Int. J. of Intelligent Systems 11 (1996) 733–750.

[50] S. Gottwald, Many-valued logic and fuzzy set theory, in: *Mathematics of Fuzzy Sets: Logic, Topology and Measure Theory*, eds. U. Hoehle and S. Rodabaugh, The Handbooks of Fuzzy Sets Series (Kluwer Academic, Dordrecht, 1999) pp. 5–90.

[51] H.J. Greenberg, Possibilities of logically equivalent expressions, Fuzzy Sets and Systems 86 (1997) 249–250.

[52] S. Haack, Do we need "fuzzy logic"?, Int. J. of Man–Machine Studies 11 (1979) 437–445.

[53] P. Hajek, *The Metamathematics of Fuzzy Logics* (Kluwer Academic, Dordrecht, 1998).

[54] J. Halpern, A counterexample to theorems of Cox and Fine, J. Artificial Intelligence Res. 110 (1999) 67–85.

[55] R. Hähnle, *Automated Deduction in Multiple-Valued Logics* (Oxford Univ. Press, Oxford, UK, 1994).

[56] A. Kandel, On minimization of fuzzy functions, IEEE Trans. on Computers 22 (1973) 826–832.

[57] A. Kandel and A. Lee, *Fuzzy Switching and Automata* (Crane Russak, New York, 1979).

[58] S.C. Kleene, *Introduction to Metamathematics* (North-Holland, Amsterdam, 1952).

[59] E.P. Klement and M. Navarra, Propositional fuzzy logics based on Frank t-norms: A comparison, in: *Fuzzy Sets, Logics and Reasoning about Knowledge*, eds. D. Dubois, H. Prade and E.P. Klement (Kluwer Academic, Dordrecht, 1999) pp. 25–47.

[60] M. Laviolette and J.W. Seaman, Jr., The efficacy of fuzzy representations of uncertainty, IEEE Trans. on Fuzzy Systems 2 (1994) 4–15.

[61] R.C.T. Lee, Fuzzy logic and the resolution principle, J. of the ACM 19 (1972) 109–119.

[62] R.T.C. Lee and C.L. Chang, Some properties of fuzzy logic, Information and Control 19 (1971) 417–431.

[63] L. Lesmo, L. Saitta and P. Torasso, Evidence combination in expert systems, Int. J. Man–Machine Studies 22 (1985) 307–326.

[64] J. Lukasiewicz, Philosophical remarks on many-valued systems of propositional logic, (1930) reprinted in: *Selected Works*, ed. Borkowski, Studies in Logic and the Foundations of Mathematics (North-Holland, Amsterdam, 1970) pp. 153–179.

[65] G. Moisil, La logique des concepts nuancés, in: *Essais sur les Logiques Non Chrysippiennes* (Editions Acad. Repub. Soc. Roum, Bucharest, 1972) pp. 57–163.

[66] M. Mukaidono, The representation and minimization of fuzzy functions, in: *The Analysis of Fuzzy Information*, Vol. 1, ed. J. Bezdek (CRC Press, Boca Raton, FL, 1987) pp. 213–229.

[67] N.J. Nilsson, Probabilistic logic, Artificial Intelligence 28 (1986) 71–87.

[68] E. Pap, *Null-Additive Set Functions* (Kluwer Academic, Dordrecht, 1994).

[69] J. Paris, *The Uncertain's Reasoner's Companion* (Cambridge University Press, 1994).

[70] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (Morgan Kaufmann, San Mateo, CA, 1988).

[71] J. Pearl, Integrating probability and logic, in: *Readings in Uncertainty for Artificial Intelligence*, eds. J. Pearl and G. Shafer (Morgan Kaufmann, San Francisco, 1990) pp. 677–679.

[72] F. Preparata and R.T. Yeh, Continuously-valued logic, J. Comput. & Syst. Sci. 6 (1972) 397–418.

[73] H. Reichenbach, *The Theory of Probability* (University of California Press, 1949).

[74] D.G. Schwartz, A min–max semantics for fuzzy likelihood, in: *Proc. of 1st IEEE Int. Conf. on Fuzzy Systems*, San Diego, CA (March 8–12, 1992) pp. 1393–1398.

[75] D.G. Schwartz, Layman's probability theory: a calculus for reasoning with linguistic likelihood, Information Sciences 126 (2000) 71–82.
[76] G. Shafer, *A Mathematical Theory of Evidence* (Princeton University Press, Princeton, NJ, 1976).
[77] P. Smets and R. Kennes, The transferable belief model, Artificial Intelligence 66 (1994) 191–234.
[78] D. Tsichritsis, Participation measures, J. Math. Analysis & Appl. 36 (1971) 60–72.
[79] P. Walley, *Statistical Reasoning with Imprecise Probabilities* (Chapman and Hall, 1991).
[80] S. Weber, $\perp$-decomposable measures and integrals for Archimedean t-conorms $\perp$, J. of Math. Anal. and Appl. 101 (1984) 114–138.
[81] T. Weston, Approximate truth, J. Philos. Logic 16 (1987) 203–227.
[82] L.A. Zadeh, Fuzzy sets, Information and Control 8 (1965) 338–353.
[83] L.A. Zadeh, Fuzzy logic and approximate reasoning (in memory of Grigore Moisil), Synthese 30 (1975) 407–428.
[84] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning, Information Sciences, Part 1: 8 (1975) 199–249, Part 2: 8 (1975) 301–357, Part 3: 9 (1975) 43–80.
[85] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1 (1978) 3–28.
[86] L.A. Zadeh, A theory of approximate reasoning, in: *Machine Intelligence*, Vol. 9, eds. J.E. Hayes, D. Michie and L.I. Mikulich (Elsevier, New York, 1979) pp. 149–194.
[87] L.A. Zadeh, Test score semantics for natural languages and meaning representation via PRUF, in: *Empirical Semantics*, Vol. 1, ed. B.B. Rieger (Bochum, Brockmeyer, 1982) pp. 281–349.
[88] H.-J. Zimmermann, *Fuzzy Set Theory and its Applications* (Kluwer Academic, Norwell, MA, 3rd ed., 1996).