New Semantics for Quantitative Possibility Theory.

Didier Dubois, Henri Prade and Philippe Smets*

IRIT, Université Paul Sabatier,
118 route de Narbonne, 31062, Toulouse, cedex 4, France
dubois@irit.fr, prade@irit.fr
and
IRIDIA Université Libre de Bruxelles
50 av. Roosevelt, CP 194-6, 1050 Bruxelles, Belgium
psmets@ulb.ac.be

Abstract

New semantics for numerical values given to possibility measures are provided. For epistemic possibilities, the new approach is based on the semantics of the transferable belief model, itself based on betting odds interpreted in a less drastic way than what subjective probabilities presupposes. It is shown that the least informative among the belief structures that are compatible with prescribed betting rates is nested, i.e. corresponds to a possibility measure. It is also proved that the idempotent conjunctive combination of two possibility measures corresponds to the hyper-cautious conjunctive combination of the belief functions induced by the possibility measures. This view differs from the subjective semantics first proposed by Giles and relying on upper and lower probability induced by non-exchangeable bets. For objective possibility degrees, the semantics is based on the most informative possibilistic approximation of a probability measure derived from a histogram. The motivation for this semantics is its capability to extend a well-known kind of confidence intervals around the mode of a distribution to a fuzzy confidence interval. We show how the idempotent disjunctive combination of possibility functions is related to the convex mixture of probability distributions.

1 Introduction

Quantitative possibility theory has been proposed as a numerical model which could represent quantified uncertainty (Zadeh, 1978; Dubois & Prade, 1988; Dubois, Nguyen, & Prade, 2000). In order to sustain this claim, it is necessary to examine the representation power of possibility theory regarding uncertainty in both objective and subjective contexts. In the objective context, quantitative possibility can be devised as an approximation of upper and lower frequentist probabilities, due to the presence of incomplete statistical observations (Dubois & Prade, 1986a; Gebhardt & Kruse, 1993). In the subjective context, quantitative possibility theory somehow competes with the probabilistic model in its personalistic or Bayesian views and with the transferable belief model (TBM) (Smets & Kennes, 1994; Smets, 1997, 1998), both of which also intend to represent degrees of belief. A major issue when developing formal models that represent psychological quantities (belief is such an object) is to produce an operational definition of what these degrees are supposed to quantify. Such an operational definition, and the assessment methods that can be derived from it, provide a meaning, a semantics, to the .7 encountered in statements like 'my degree of belief is .7'. Such an operational definition has been produced long ago by the Bayesians. They claim that any state of incomplete knowledge of an agent can be modeled by a single probability distribution on the appropriate referential, and that the probabilities can be revealed by a betting experiment in which the agent provides betting odds under an exchangeable bet assumption. A similar setting exists for imprecise probabilities (Walley, 1991), relaxing the assumption of exchangeable bets, and more recently for the TBM as well (Smets & Kennes, 1994; Smets, 2001), introducing several betting frames corresponding to various partitions of the referential. In that sense, the numerical values encountered in these three models are well defined.

Quantitative possibility theory (QPT) did not have such a wealth of operational definitions so far, despite an early proposal by Giles (Giles, 1982) in the setting of upper and lower probabilities, recently taken over by De Cooman and colleagues (Walley, 1997; De Cooman & Aeyels, 1999). One way to avoid the measurement problem is to develop a *qualitative* epistemic possibility theory where only ordering relations are used (Dubois & Prade, 1998).

Nevertheless QPT seems to be a theory worth exploring as well, and rejecting it because of the current lack of convincing semantics would be unfortunate. The recent revival of a form of subjectivist QPT due to De Cooman and colleagues, and the development of possibilistic networks based on incomplete statistical data (Gebhardt & Kruse, 1998) suggests on the contrary that it is fruitful to investigate various operational semantics for possibility theory. This is due to several reasons: first possibility theory is a special case of most existing non additive uncertainty theories, be they numerical or not. Hence progress in one of these theories usually has impact in possibility theory. Another major reason is that possibility theory is very simple, certainly the simplest competitor for probability theory. Hence it can be used as useful approximate representation by other theories. A last reason is that previous works have suggested strong links between possibility theory and non-Bayesian statistics, especially the use of likelihood functions without prior (Smets, 1982), and confidence intervals. It is not absurd to think that, in the future, possibility theory may contribute to unify and shed some light on some aspects of non-Bayesian statistics.

The aim of this paper is to propose two new semantics for possibility theory: a subjectivist semantics and an objectivist one. We use the term 'subjectivist' to mean that we consider the concepts of beliefs (how much we believe) and betting behaviors (how much would we pay to enter into a game) without regard to the possible random nature and repeatability of the events. We use the term 'objectivist' to mean that we consider data generated by random processes where repetition is natural, and where histograms can summarize the data. The distinction is somehow similar to the one made between the personal and the frequential interpretations of probabilities. It also reflects that in the 'subjectivist' case, we start from a betting behavior, whereas in the 'objectivist' case we start from a histogram.

The subjective semantics differ from the upper and lower probabilistic setting of the subjective possibility proposed by Giles and followers, without questioning its merit. Instead of making the bets non-exchangeable, we assume that the exchangeable betting rates only imperfectly reflect the agent's beliefs. The objectivist semantics suggests a flexible extension of particular confidence intervals.

Moreover we show that the basic combination rules in possibility theory, minimum and maximum, can be interpreted in the proposed settings: the former using a minimal commitment assumption in the subjectivist setting; the latter using an information preservation principle in the frequentist setting.

This paper provides an overview of these semantics. Detailed theorems and proofs can be found in the long version of this paper, which pursues an investigation started in (Smets, 2000). Up-to-date presentations of the TBM and possibility theory can be found in (Smets, 1998; Dubois & Prade, 1998), respectively.

2 Subjectivist semantics

2.1 The transferable belief model and bets.

For long, it had been realized that possibility functions are mathematically identical to consonant plausibility functions (Shafer, 1976), so using the semantics of the TBM for producing a semantics for quantitative epistemic possibility theory is an obvious approach, even if not explored in depth so far. This link had already been realized long ago. What was missing was to show that the analogy goes further.

Suppose You (hereafter You is the agent who holds the beliefs) consider what beliefs You should adopt on what is the actual value of a variable ranging on a set Ω, called the frame of discernment. You have decided that Your beliefs should be those produced by a fully reliable source. It is assumed that such beliefs can be represented by a belief function . A belief function can be mathematically described as a finite random set which has a very specific interpretation. The so-called basic belief mass assigned to each set is understood as the weight given to the fact that all You may know from the source is that the value of the variable of interest lies somewhere in that set. Sets with positive mass are thus pieces of incomplete information, unspecific evidence, not only sets in the usual sense of a conjunction of elements. Sets here are disjunctions of elements, one of which is the true one. A belief function (resp: plausibility function) is a set-function that assigns to each event (subset of the 'frame of discernment') the sum of the masses given to its subsets (resp: to the subsets that intersect it). It evaluates to what extent the event is implied by (resp. consistent with) the available evidence. When the sets with positive mass are nested, the plausibility function is called a possibility measure, and can be characterized, just like probability, by an assignment of weights to singletons, called a possibility distribution.

Should You know the beliefs of the source, they would be Yours. Unfortunately, it occurs that You only know the value of the 'pignistic' probabilities the source would use to bet on the actual value of Ω (Smets, 1990; Smets & Kennes, 1994). The pignistic probability induced by a belief function is built by defining a uniform probability on each set of positive mass, and performing the convex mixture of these probabilities according to the mass function. In terms of game theory it corresponds to the Shapley value of a game; in terms of upper and lower probabilities it is the center of gravity of the set of probabilities dominating the belief function. The knowledge of the values of the probabilities allocated to the elements of Ω is not sufficient to construct a unique underlying belief function. Many belief functions can induce the same probabilities. For instance, uniform betting rates on Ω either correspond to complete ignorance on the values of the variable, or to the knowledge that the variable is random and uniformly distributed. So all You know about the belief function that represents the source's beliefs is that it belongs to the set of beliefs that induce the supplied pignistic probabilities. Under this scheme, we do not question the exchangeability of bets, as done by Walley, Giles and others. What we question is the assumption of a one-to-one correspondence between betting rates produced by the source, and the actual beliefs entertained by the source. Betting rates do not tell if the uncertainty of the source results from the perceived randomness of the phenomenon under study or from a simple lack of information about it.

Since several belief functions, lead to the same betting rates, You have to select one that most plausibility reflects the actual states of belief of the source. A cautious approach is to obey a 'least commitment principle' that states that You should never presuppose more beliefs than justified. Then, You can select the 'least committed' element in the family of belief functions compatible with the pignistic probability function prescribed by the obtained betting rates. The first result of this paper is that the least committed belief function is consonant, that is, the corresponding plausibility function is a possibility function. This possibility function is the unique one in the set of belief functions having a prescribed pignistic probability, because the pignistic transformation is a bijection between possibilities and probabilities. So this possibility function turns out to be the least committed belief function whose pignistic transformation is equal to the pignistic probabilities supplied by the source.

More formally let $m(A)$ be the basic belief mass allocated to subset A. The function m is called a basic belief assignment (bba). The sum of these masses across all events is 1. The degree of belief $bel(A)$ is defined for all $A \subseteq \Omega$, by:

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B)$$

The degree of plausibility $pl(A)$ is:

$$pl(A) = \sum_{B \cap A \neq \emptyset} m(B) = bel(\Omega) - bel(\overline{A}).$$

In order to enhance the fact that we work with non-normalized belief functions ($m(\emptyset)$ can be positive), we use the notation *bel* and *pl*, whereas Shafer uses the notation *Bel* and *Pl*. Another useful function that is also in one to one correspondence with any of m, *bel* and *pl* is the commonality function q such that:

$$q(A) = \sum_{B: A \subseteq B} m(B).$$

2.2 Consonant belief functions

A belief function is said to be *consonant* iff its focal elements are nested (Shafer, 1976, pg 219). By extension, we will speak of consonant basic belief assignments, commonality functions, plausibility functions, ...

Theorem 2.1 ***Consonant belief functions.*** *(Shafer, 1976, Theorem 10.1, pg 220) Let m be a bba on Ω. Then the following assertions are all equivalent:*

1. *m is consonant.*
2. $bel(A \cap B) = \min(bel(A), bel(B)),\ \forall A, B \subseteq \Omega.$
3. $pl(A \cup B) = \max(pl(A), pl(B)),\ \forall A, B \subseteq \Omega.$
4. $pl(A) = \max_{\omega \in A} pl(\omega)$*, for all non empty* $A \subseteq \Omega$.
5. $q(A) = \min_{\omega \in A} pl(\omega) = \min_{\omega \in A} q(\omega)$*, for all non empty* $A \subseteq \Omega$.

Items 3 and 4 shows that consonant plausibility and belief functions are possibility and necessity functions, usually denoted by Π and N respectively, while the $pl(\omega)$'s define a possibility distribution, that contains all the necessary information for building the other set-functions. The fact that we work with unnormalized bba's does not affect these properties, being understood that we never require that possibility and necessity functions be normalized. The difference between $\Pi(\Omega)$ or $pl(\Omega)$ and 1, that equals $m(\emptyset)$, represents the amount of conflict between the pieces of evidence that were used to build these functions.

2.3 Least commitment

So far, what 'least committed' means has not been explained, and refers to the capability of comparing belief functions by their informational contents. Dubois and Prade (1987) have made three proposals to order belief functions according to the 'specificity', or precision of the beliefs they represent. Let m_1 and m_2 be two bba's on Ω. The statement that m_1 contains at least as much information as, is at least as precise as m_2 is denoted $m_1 \sqsubseteq_x m_2$ corresponding to some x-ordering where we vary the subscript x. Then m_2 is said to be x-less committed than m_1. The proposed information orderings are:

- *pl-ordering.* If $pl_1(A) \leq pl_2(A)$ for all $A \subseteq \Omega$, we write $m_1 \sqsubseteq_{pl} m_2$
- *q-ordering.* If $q_1(A) \leq q_2(A)$ for all $A \subseteq \Omega$, we write $m_1 \sqsubseteq_q m_2$
- *s-ordering.* If m_1 is a specialization of m_2, we write $m_1 \sqsubseteq_s m_2$

where pl denotes the plausibility function and q denotes the commonality function.

The idea behind the pl-ordering is that a belief function is all the more specific as the intervals ranging from the belief degree to the plausibility degree of each event are small. Due to the duality between plausibility and belief, $pl_1(A) \leq pl_2(A)$ for all $A \subseteq \Omega$ is indeed equivalent, for normalized bba's, to $bel_1(A) \geq bel_2(A)$ for all $A \subseteq \Omega$, where bel denotes the belief function.

The idea behind the q-ordering is maybe less obvious. The commonality function of an event reflects the amount of support this event may received from its supersets. So, $q(A)$ represents the portion of belief that may eventually be assigned to A. The more amount of belief remains unassigned, i.e. the bigger the focal elements having a high mass assignment, the higher the commonality degrees and the less informative is the belief function. In particular, if $m(\Omega) = 1$, then $q(A) = 1$ for all non empty sets. More generally, to consider $m(\Omega)$ as a rough measure of uninformativeness of a belief function seems reasonable. Suppose now we know that the actual world belongs to $A \subseteq \Omega$. Then $m[A](A)$ obtained by conditioning m with Dempster's rule of conditioning becomes the 'conditional measure of uninformativeness' in context A. It just happens that $m[A](A) = q(A)$, so the commonality function is the set of conditional measure of uninformativeness, and the fact that a measure of information content turns out to be a function of the q's becomes very natural.

The concept of specialization (s-ordering) (Dubois & Prade, 1986b; Yager, 1986) is at the core of the transferable belief model (Klawonn & Smets, 1992). The intuitive idea is that the smaller the focal elements, the more focused are the beliefs. Let $m_Y^\Omega[BK]$ be the basic belief assignment that represents Your belief on Ω given the background knowledge (BK) accumulated by You. The impact of a new piece of evidence Ev induces a change in Your beliefs characterized by a redistribution of the basic belief masses of $m_Y^\Omega[BK]$ such that $m_Y^\Omega[BK](A)$ is reallocated to the subsets of A. In a colloquial way, we say that 'the masses flow down'. The new belief function is said to be a specialization of the former one. More generally, m_2 is a specialization of m_1 if every mass $m_1(A)$ is reallocated to subsets of A in m_2. See (Dubois & Prade, 1986b) for the technical definition.

When the belief functions are consonant, all these comparisons become equivalent and reduce to fuzzy set inclusion between the membership functions representing the corresponding possibility distributions. Formally for $x = pl, q, s$, $m_1 \sqsubseteq_x m_2$ if and only if $pl_1(\omega) \geq pl_2(\omega)$ for all $\omega \in \Omega$. We say that pl_1 is less specific than pl_2.

Dubois and Prade (1986b) prove that :

- $m_1 \sqsubseteq_s m_2$ implies $m_1 \sqsubseteq_{pl} m_2$ and $m_1 \sqsubseteq_q m_2$, but the converse is not true, and
- $m_1 \sqsubseteq_{pl} m_2$ and $m_1 \sqsubseteq_q m_2$ do not imply each other.

2.4 Pignistic probabilities

Suppose a bba m^Ω that quantifies Your beliefs on Ω. When a decision must be made that depends on the actual value ω_0 where $\omega_0 \in \Omega$, You must construct a probability function in order to make the optimal decision, i.e., the one that maximizes the expected utility. This is achieved by the pignistic transformation. Its nature and its justification are defined in (Smets, 1990; Smets & Kennes, 1994; Smets, 1998).

Let F be the *betting frame*, the set of 'atoms' on which stakes will be allocated. Bets can then only be built on the elements of the power set of that frame. Let $BetP^F$ denote the pignistic probability function You will use to bet of the alternatives in F. $BetP^F$ is a function of F and m^Ω,

$$BetP^F = \Gamma(m^\Omega, F).$$

Smets (1990) has shown that the only transformation from m^Ω to $BetP^F$ that satisfies some rationality requirements is the so-called pignistic transformation that satisfies:

$$BetP^F(f) = \sum_{A: f \in A \subseteq F} \frac{m^F(A)}{|A|(1 - m^F(\emptyset))}, \qquad \forall f \in F \tag{1}$$

where $|A|$ is the number of elements of F in A, and m^F is the bba induced by m^Ω on F, (we have assumed that F is compatible with Ω (Shafer, 1976, pg. 114 *et seq.*)).

It is easy to show that the function $BetP^F$ is indeed a probability function and the pignistic transformation of a probability function is the probability function itself. We call it pignistic in order to avoid the confusion that would consist in interpreting $BetP^F$ as a measure representing Your beliefs on F. From the mathematical point of view, the pignistic probability of *bel* is its Shapley value in a game-theoretic setting.

The result showing that the least committed set of beliefs yielding a prescribed pignistic probability can be represented by a possibility function, has been formally obtained in two ways, depending on how belief functions are compared in terms of information contents. Comparing the belief functions having a prescribed pignistic probability, it can be proved that the least informed one in the sense of the q-ordering is a possibility function. The belief functions having a prescribed pignistic probability are called isopignistic. The following theorem has been obtained:

Theorem 2.2 *Let $BetP^\Omega$ be a pignistic probability function defined on Ω with the elements ω_i of Ω so labeled that :*

$$BetP^\Omega(\omega_1) \geq BetP^\Omega(\omega_2) \geq \ldots \geq BetP^\Omega(\omega_n)$$

where $n = |\Omega|$. Let $\mathfrak{B}isoP(BetP^\Omega)$ be the set of isopignistic belief functions induced by $BetP^\Omega$. The q-least committed element in $\mathfrak{B}isoP(BetP^\Omega)$ is the consonant belief function of mass $\widehat{m}$ whose only focal elements are the subsets $A_i = \{\omega_1, \omega_2, \ldots, \omega_i\}$ and:

$$\widehat{m}(A) = |A| \cdot (BetP^\Omega(\omega_i) - BetP^\Omega(\omega_{i+1}))$$

where $BetP^\Omega(\omega_{n+1})$ is 0 by definition.

Proof. Let $\hat{q}$ and $\widehat{BetP}$ be the commonality function and the pignistic probability function related to $\widehat{m}$. The pignistic transformation of $\hat{m}$ being $BetP$, $\hat{m}$ belongs to $\mathfrak{B}isoP(BetP^\Omega)$. Let $m \in \mathfrak{B}isoP(BetP^\Omega)$ be a bba also defined on Ω with q and $BetP$ its related commonality function and pignistic probability function. Let $A_i = \{\omega_1, \omega_2 \ldots, \omega_i\}$ for $i = 1, \ldots, n$. In order to prove that $\hat{q}$ is the q-least committed element in $\mathfrak{B}isoP(BetP^\Omega)$, we need only to show that $q \in \mathfrak{B}isoP(BetP^\Omega)$ implies $q(A) \leq \hat{q}(A)$, $\forall A \subseteq \Omega$.

Consider A_n. Suppose $q(A_n) > \hat{q}(A_n)$. It is equivalent to $m(A_n) > \widehat{m}(A_n)$, in which case $BetP(\omega_n) > \widehat{BetP}(\omega_n)$, contrary to the hypothesis. So

- either $m(A_n) = \widehat{m}(A_n)$ and for all $A \subseteq A_n$ with $\omega_n \in A$, $m(A) = 0$ so that we keep $BetP(\omega_n) = \widehat{BetP}(\omega_n)$.
- or $m(A_n) < \widehat{m}(A_n)$ in which case $q(A_n) < \hat{q}(A_n)$, and thus q is not q-less committed that $\hat{q}$.

Therefore the first alternative is true.

Consider A_{n-1}. We repeat the same reasoning replacing n by $n-1$ and get $m(A_{n-1}) = \widehat{m}(A_{n-1})$ and for all $A \subseteq A_{n-1}$ with $\omega_{n-1} \in A$, $m(A) = 0$. Iterating for $n-2, n-3 \ldots$, we prove the theorem. □

The probability-possibility transformation described by the theorem was independently proposed by Dubois and Prade (1982; 1983) a long time ago, using a very different rationale. The other informational orderings do not lead to unique least informed solutions. However a scalar index for comparing belief functions in terms of specificity has been proposed in (Dubois & Prade, 1985). The idea is based on the fact that the level of imprecision of a set used to represent a piece of incomplete knowledge is its cardinality (or the logarithm thereof). A belief function is formally a random set, and the degree of imprecision of belief function is simply its expected cardinality (or expected logarithm of the cardinality). Let

$$I(m) = \sum_{A \subseteq F} |A| \cdot m(A)$$

Comparing isopignistic belief functions in terms of expected cardinality, the same result as above obtains:

Theorem 2.3 *The belief function of maximal expected cardinality $I(Bel)$ among isopignistic belief functions induced by $BetP^\Omega$ is the unique possibility function having this pignistic probability.*

Proof. (sketch). Consider a belief function Bel with mass function m. Denote $Bet(Bel) = BetP^\Omega = p$ its pignistic probability. Conversely given a probability distribution on Ω, Let $\pi = Bet^{-1}(p)$ be the possibility distribution whose pignistic transformation yields the probability distribution p. Overall $\pi = Bet^{-1}(Bet(Bel))$. It is easy to see that $I(\pi) = \sum_{\omega \in \Omega} \pi(\omega)$ and $\pi(\omega) = \sum_{\omega' \in \Omega} min(p(\omega), p(\omega'))$ (See (Dubois & Prade, 1983)). The technical part of the proof consists in proving that $I(Bet^{-1}(Bet(Bel)) \geq I(Bel)$. Since each of these expressions is a linear function of the masses $m(A)$, we prove that for each focal element A, the coefficient of $m(A)$ in $I(Bel)$ is not greater than the corresponding coefficient in $I(Bet^{-1}(Bet(Bel))$. □

3 The minimum rule

The story goes on. Suppose we collect the pignistic probabilities about the actual value of Ω from two sources. From these two pignistic probabilities,You build two consonant plausibility function's, i.e., the two possibility functions induced by the observed betting rates as presented above. How to conjunctively combine the data collected from the two sources? Do we have to redo the whole betting procedure or can we get the result directly by combining the two possibility functions? We will show in this section that indeed the last idea is correct.

In possibility theory, there exists such a combination rule that performs the conjunction of two possibility functions. Let π_1 and π_2 be two possibility distributions on Ω that we want to combine conjunctively into a new possibility function π_{12}. The most classical conjunctive combination rule to build π_{12} consists in using the minimum rule: $\pi_{12}(\omega) = \min(\pi_1(\omega), \pi_2(\omega))$ for all $\omega \in \Omega$ and its related possibility measure is given by $\Pi_{12}(A) = \max_{\omega \in A \subseteq \Omega} \pi_{12}(\omega)$. Could it be applied in the present context? We will show here that it is indeed the case.

We must first avoid a classical trap. In belief function theory, the conjunctive rule of combination for the bba' produced by two distinct pieces of evidence is Dempster's rule of combination. It is well known that Dempster's rule of combination applied to two consonant plausibility functions does not produce a consonant plausibility function. So Dempster's rule of combination does not seem adequate to combine possibility measures. It seems thus that the analogy between consonant plausibility functions and possibility functions collapses here. This is not the case. Dempster's rule of combination requires that the involved pieces of evidence are 'distinct', a concept analogous to independence in random set theory. All we have here are the betting behaviors of the two sources, and 'distinctness' of the sources cannot be assumed.

In fact, other combination rules exist in the TBM, based on some kind of cautious approach and where 'possible correlations' between the involved belief functions are considered. How to combine two bba's conjunctively, when you cannot assume they are produced by two 'distinct' pieces of information? You may assume that the result of the combination must be a specialization of each of them (since the result of the combination should be a belief function at least as informative as the ones You start with). As said above, a specialization of a bba m_1 is a transformation of m_1 into a new bba m_2, both on the same frame of discernment, such that every mass $m_1(A)$ given by the first bba to a subset A of its frame is split and reallocated to the subsets of A so as to form the second bba.

So consider all belief functions that are specialization of *both* initial possibility functions derived from the pignistic probabilitys produced by the two sources. In that family, apply again the 'cautious' approach and select as Your belief the least committed element of that family in the sense of specialization, which is the stronger notion of information comparison. The main result is that this procedure again yields a consonant plausibility function and it turns out to be exactly the result obtained within possibility theory when using the minimum rule.

More formally: Let $\mathcal{SP}(m_1)$ and $\mathcal{SP}(m_2)$ be the set of specializations of m_1 and m_2, respectively. The result of the conjunctive combination of m_1 and m_2 should then belong to $\mathcal{SP}(m_1) \cap \mathcal{SP}(m_2)$. That family is never empty as long as the result of Dempster rule of combination, itself a specialization, always belongs to it. As far as no information on the correlation between the sources is available, the cautious attitude would consist in representing Your beliefs by $m_{1 \oslash 2}$, the least committed element in $\mathcal{SP}(m_1) \cap \mathcal{SP}(m_2)$, in the sense of the s-ordering. Unfortunately such a minimum does not always exist, except when m_1 and m_2 are both consonant. In that case , we consider what would be $m_{1 \oslash 2}$, and prove that it is equal to the result found when conjunctively combining two possibility functions by the minimum. Note that if Π_1 and Π_2 are two possibility functions with q_1 and q_2 their related commonality functions, the commonality function q_{12} of their minimum-based conjunctive combination Π_{12} satisfies : $q_{12}(A) = \min(q_1(A), q_2(A))$ for all $A \subseteq \Omega$.

Theorem 3.1 *Let m_1 and m_2 be two consonant belief functions on Ω with q_1 and q_2 their corresponding commonality functions. Let $\mathcal{SP}_1$ and $\mathcal{SP}_2$ be the set of specializations of m_1 and m_2, respectively. Let $q_{12}(A) = \min(q_1(A), q_2(A))$ for all $A \subseteq \Omega$, and m_{12} its corresponding bba. Then $m_{12} = m_{1 \oslash 2} = \min\{m : m \in \mathcal{SP}(m_1) \cap \mathcal{SP}(m_2)\}$ in the sense of s-ordering, and this minimally specific element is unique.*

Proof. If m_{12} is a specialization of m_1, then, for all $A \subseteq \Omega$, $q_{12}(A) \leq q_1(A)$. So $q_{12}(A) \leq \min(q_1(A), q_2(A))$. Thus the s-least specific element is obtained when we have $q_{12}(A) = \min(q_1(A), q_2(A))$, $\forall A \subseteq \Omega$. The only problem left is to prove that this computed q_{12} is indeed a commonality function, i.e., that its related basic belief masses are non negative. For it, we only need to prove that $q_{12}(A) = \min_{\omega \in A} q_{12}(\omega)$. m_1 and m_2 being consonant, we have $q_i(A) = \min_{\omega \in A} q_i(\omega)$ for $i = 1, 2$. So

$$\begin{aligned} q_{12}(A) &= \min(q_1(A), q_2(A)) \\ &= \min(\min_{\omega \in A} q_1(\omega), \min_{\omega \in A} q_2(\omega)) \\ &= \min_{\omega \in A}(\min(q_1(\omega), q_2(\omega)) \\ &= \min_{\omega \in A}(q_{12}(\omega)). \end{aligned}$$

So q_{12} is consonant, in which case m_{12} is a belief function. □

We call the last combination the hyper-cautious conjunctive combination rule.

So the direct combination approach developed in possibility theory and the one derived using the TBM detour are the same (see Figure 1). This result restores the coherence between the two models, and thus using the TBM operational definition to explain the meaning of the possibility values is perfectly valid and appropriate. The fact that applying the same approach to the combination of general belief functions does not yield a unique least specialized belief function stresses the difficulty of characterizing a feasible idempotent combination rule for non-consonant belief functions, a problem that remains open.

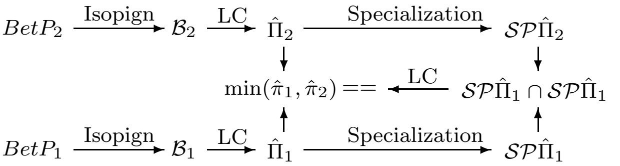

Figure 1: Epistemic possibilities. Isopign = finding the set of belief functions that share the same pignistic probabilities. LC = least committed.

4 Objectivist semantics

Since possibility measures are special cases of plausibility functions, they are also, at the mathematical level envelopes of special families of probability functions (See, e.g. (Dubois & Prade, 1992; De Cooman & Aeyels, 2000). Let Π be a possibility measure and $\mathfrak{P}(\Pi)$ be the set of probability functions dominated by Π.

Suppose a probability function P is obtained via some statistical experiment. This probability function is a very rich piece of information, if the number of statistical experiments supporting it is high enough. Suppose that for some reason one wishes to use a possibilistic representation of this information, maybe because we just need an approximation of it, or because we want to compute a linear convex combination of them without knowing the weights (see section 5). The possibility measures Π that are candidates for representing P must clearly be such that $P \in \mathfrak{P}(\Pi)$. We shall say that Π covers P. Again there are many possibility measures obeying this constraint. It again makes sense to use an informational principle to pick the best Π induced by P. However the situation is very different from the subjectivist setting. In the latter the pignistic probability is just what is revealed about the epistemic state of the agent by the betting experiments. So a principle of cautiousness prevails in order to be faithful to the incompleteness of the information. In the objectivist setting of precisely observed identical and independent, and sufficiently numerous experiments, P is supposed to represent all the information that can ever be captured. Moving from a probabilistic to a possibilistic setting means losing information since we only get (special) probability bounds.

So the natural informational principle for picking a reasonable possibility distribution representing P is to preserve as much information as possible, hence picking the most informative possibility distribution (in the sense of any x-ordering above) in $\mathbf{\Pi}(P) = \{\Pi : P \in \mathfrak{P}(\Pi)\}$, that is, in the sense of fuzzy set inclusion (i.e., by taking the possibility function that is pointwise minimal). It has been proved that generally this maximally informed possibility distribution exists and is unique. Namely this is the case when P defines a total order of a finite referential. It is also true for "bell-shaped" unimodal distributions on the real line. When there are elements of equal probability, unicity is recovered if, due to symmetry, we also enforce equal possibility of these elements. See details in (Dubois, Prade, & Sandri, 1993; Lasserre, Mauris, & Foulloy, 1998).

In the case of unimodal continuous pdfs, the level cuts of the obtained possibility distribution π, namely $A_\alpha = \{\omega : \pi(\omega) \geq \alpha\}$, are closely related to some kinds of 'confidence intervals' (with confidence levels α) around the mode of the distribution. Namely, let p be a unimodal continuous density. Let $I_\alpha = \{r, p(r) \geq \alpha\}$. Let l_α be the length of this interval. It is easy to prove that I_α is the interval of highest probability among all those whose length is l_α. I_α is conversely the most precise interval representative of p with a given confidence level equal to $P(I_\alpha)$. The most specific possibility distribution π containing P is such that $\pi(\inf I_\alpha) = \pi(\sup I_\alpha) = 1 - P(I_\alpha)$. So the cut at level α of π that is, $A_\alpha = \{r, \pi(r) \geq \alpha\}$ is the confidence interval around the mode of p with confidence level $1 - \alpha$.

These setting equips quantitative possibility theory with a reasonable objectivist semantics. Note that this semantics can be considered as indirect, as it relies on the frequentist view of probability, and is built on top of it. It differs from the direct frequentist semantics that considers a possibility distribution as the one-point coverage mapping of a random set describing imprecise statistical data.

However our approach provides a convenient representation of a family of confidence intervals used by statisticians when they try to estimate a parameter from a sample. So-doing, they trade-off randomness for incomplete knowledge with a view to get an easier-to-interpret and ready for use representation of information. It is as if they move to a (crude) possibilistic representation. One of the weaknesses of this procedure is the arbitrariness of the confidence level used to pick the proper interval (usually 0.95, but why not another figure close to 1?). Our approach captures the whole family of confidence intervals by means of fuzzy intervals ((Dubois, Kerre, Mesiar, & Prade, 2000)) thus postponing the decision on the confidence level.

5 The maximum rule

Again the story can be pursued considering the fusion of two probability distributions P_1 and P_2 coming from two statistical experiments pertaining to the same phenomena. If the fusion takes place on the data, it is usually enough to add the two sets of data, and derive the corresponding probability. It comes down to a linear convex combination of P_1 and P_2 whose weights reflect the relative amount of data of each source.

However if the original data sets are lost and only P_1 and P_2 are available, the relative weights of the data sets are unknown. The probability distribution resulting from merging the two data sets is of the form $\alpha P_1 + (1-\alpha)P_2$ where α is unknown. It gives a family of probability distributions F and the question is to find the most informative possibility distribution Π such that F is included in $\mathfrak{P}(\Pi)$ using the above principle of information preservation. Let Π_1 and Π_2 be the possibility measures encoding the most informative confidence sets respectively associated to P_1 and P_2. Then $\Pi_1 \geq P_1$ and $\Pi_2 \geq P_2$, eventwise. Now it is obvious that

$$\max(\Pi_1, \Pi_2) \geq \max(P_1, P_2) \geq \alpha P_1 + (1-\alpha)P_2$$

It turns out that the set function $\Pi^{12} = \max(\Pi_1, \Pi_2)$ is also a possibility measure with possibility distribution $\max(\pi_1, \pi_2)$. So $\Pi^{12} = \max(\Pi_1, \Pi_2)$ encodes a family of probability measures that contains $\alpha P_1 + (1-\alpha)P_2$ for any α in the unit interval. However there are events A, B such that $P_1(A) = \Pi_1(A)$, and $P_2(B) = \Pi_2(B)$, basically the complements of the confidence sets. So $\Pi^{12} = \max(\Pi_1, \Pi_2)$ is actually the valid upperbound, i.e. it covers all the convex mixtures of P_1 and P_2. Now let Π_α be the most informative possibility measure covering $\alpha P_1 + (1-\alpha)P_2$. Obviously, the intersection over α of all sets of possibility measures less specific than Π_α has $\sup_\alpha \Pi_\alpha$ as a lower bound and it is the most specific possibility measure covering all the convex mixtures of P_1 and P_2. However it is clearly less than or equally specific as Π^{12}. Hence it is equal to it. It is thus proved that the most informative possibility distribution covering all the convex mixtures of P_1 and P_2 can be obtained as the idempotent disjunctive combination of the possibility measures Π_1 and Π_2 obtained from P_1 and P_2. Hence this setting justifies the maximum combination rule of possibility theory (see Figure 2).

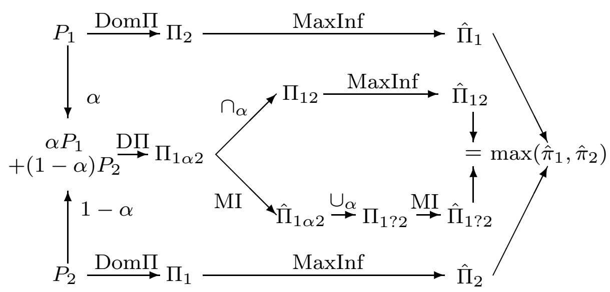

Figure 2: Objective possibilities. DΠ = DomΠ = dominating possibility measures. MI = MaxInf = maximally informative possibility measure. $\cap_\alpha$ = intersection over all α in [0,1]. $\cup_\alpha$ = union over all α in [0,1]. Other symbols as in text.

6 Conclusion

This paper studies two operational settings for the measurement of degrees of possibility. In the first one, Quantitative Epistemic Possibility theory can be viewed as a very cautious application of the TBM. It uses the operational definition of the TBM as an operational definition of the values of the possibility function whereby the betting rates provided by an agent only partially reflect beliefs. In a frequentist setting, a possibility measure can be induced from frequency observations as a consonant family of certain confidence sets. These operational settings shed light on well-known idempotent combination rules of possibility theory. The minimum and maximum rules are justified, one in each setting, based on opposite information principles. The link between possibility functions and membership functions of linguistic terms had been pointed out from the start by Zadeh (1978) and is discussed by Walley and De Cooman (2000) in the upper/lower probability setting. Here we indirectly provide other semantics for fuzzy set theory through quantitative possibility theory, based either on standard behavioral methods of subjective probability or as an extension of standard statistical practice. In both cases a probability measure is replaced by a possibility measure. Some may find it debatable or pointless. However, in the subjective setting, it is based on the claim that taking a subjective probability as a faithful representation of an agent's beliefs is far too optimistic, considering that people usually have incomplete knowledge. Our possibilistic representation is arguably more faithful because more cautious since explicitly accounting for this situation. In the objective setting, moving from the probabilistic to the possibilistic representation corresponds to an apparent loss of information and looks all the more debatable. However, a confidence interval alone corresponds to a more drastic loss of information with respect to the original provability distribution. Our proposal makes confidence intervals more expressive by refraining from choosing a confidence level. Note that the obtained transformation is again bijective so that the probability measure can be entirely recovered from the possibility measure.

References

De Cooman, G., & Aeyels, D. (1999). Supremum-preserving upper probabilities. *Information Sciences*, *118*, 173–212.

De Cooman, G., & Aeyels, D. (2000). A random set description of a possibility measure and its natural extension. *IEEE Trans on Systems, Man and Cybernetics*, *30*, 124–131.

Dubois, D., Kerre, E., Mesiar, R., & Prade, H. (2000). Fuzzy interval analysis. In D. Dubois & H. Prade (Eds.), *Fundamentals of fuzzy sets* (pp. 483–581). Kluwer , Boston, Mass.

Dubois, D., Nguyen, H., & Prade, H. (2000). Possibility theory, probability theory and fuzzy sets: misunderstandings, bridges and gaps. In D. Dubois & H. Prade (Eds.), *The handbook of fuzzy sets series* (pp. 344–438). Kluwer Academic Publishers, Boston.

Dubois, D., & Prade, H. (1983). Unfair coins and necessity measures: a possibilistic interpretation of histograms. *Fuzzy Sets and Systems*, *10*, 15–20.

Dubois, D., & Prade, H. (1985). A note on measures of specificity for fuzzy sets. *International Journal of General Systems*, *10*(4), 279-283.

Dubois, D., & Prade, H. (1986a). Fuzzy sets and statistical data. *Europ. J. Operations Research*, *25*, 345–356.

Dubois, D., & Prade, H. (1986b). A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets. *International Journal of General Systems*, *12*, 193-226.

Dubois, D., & Prade, H. (1987). The principle of minimum specificity as a basis for evidential reasoning. In B. Bouchon & R. R. Yager (Eds.), *Uncertainty in knowledge-based systems* (pp. 75–84). Springer Verlag, Berlin.

Dubois, D., & Prade, H. (1988). *Possibility theory : an approach to computerized processing of uncertainty.* New York: Plenum Press.

Dubois, D., & Prade, H. (1992). When upper probabilities are possibility measures. *Fuzzy Sets and Systems*, *49*, 65–74.

Dubois, D., & Prade, H. (1998). Possibility theory: qualitative and quantitative aspects. In D. M. Gabbay & P. Smets (Eds.), *Handbook of defeasible reasoning and uncertainty management systems, vol. 1* (Vol. 1, pp. 169–226). Kluwer, Doordrecht, The Netherlands.

Dubois, D., Prade, H., & Sandri, S. A. (1993). On possibility/probability transformations. In R. Lowen (Ed.), *Fuzzy logic: State of the art* (pp. 103–112). Kluwer Academic Publ.

Dubois, D., & Pradre, H. (1982). On several representations of an uncertain body of evidence. In M. M. Gupta & E. Sanchez (Eds.), *Fuzzy information and decision processes* (pp. 167–181). North Holland, Amsterdam.

Gabbay, D. M., & Smets, P. (Eds.). (1998). *Handbook of defeasible reasoning and uncertainty management systems, vol. 1.* Kluwer, Doordrecht, The Netherlands.

Gebhardt, J., & Kruse, R. (1993). The context model - an intergating view of vagueness and uncertainty. *Int. J. Approximate Reasoning*, *9*, 283–314.

Gebhardt, J., & Kruse, R. (1998). Parallel combination of information sources. In D. M. Gabbay & P. Smets (Eds.), *Handbook of defeasible reasoning and uncertainty management systems, vol. 3* (Vol. 3, pp. 393–440). Kluwer, Doordrecht, The Netherlands.

Giles, R. (1982). Foundation for a possibility theory. In M. M. Gupta & E. Sanchez (Eds.), *Fuzzy information and decision processes* (pp. 83–195). North Holland, Amsterdam.

Gupta, M. M., & Sanchez, E. (Eds.). (1982). *Fuzzy information and decision processes.* North Holland, Amsterdam.

Klawonn, F., & Smets, P. (1992). The dynammic of belief in the transferable belief model and specialization- generalization matrices. In D. Dubois, M. P. Wellman, B. D'Ambrosio, & P. Smets (Eds.), *Uncertainty in artificial intelligence 92* (pp. 130–137). Morgan Kaufman, San Mateo, Ca.

Lasserre, V., Mauris, G., & Foulloy, L. (1998). A simple modelisation of measurement uncertainty: The truncated triangular possibility distribution. In IPMU-98 (Ed.), *Information processing and management of uncertainty* (pp. 10–17).

Shafer, G. (1976). *A mathematical theory of evidence.* Princeton Univ. Press. Princeton, NJ.

Smets, P. (1982). Possibilistic inference from statistical data. In A. Ballester, D. Cardus, & E. Trillas (Eds.), *Second world conference on mathematics at the service of man* (pp. 611–613). Universidad Politecnica de Las Palmas, Spain.

Smets, P. (1990). Constructing the pignistic probability function in a context of uncertainty. In M. Henrion, R. D. Shachter, L. N. Kanal, & J. F. Lemmer (Eds.), *Uncertainty in artificial intelligence 5* (pp. 29–40). North Holland, Amsteram.

Smets, P. (1997). The normative representation of quantified beliefs by belief functions. *Artificial Intelligence*, *92*, 229–242.

Smets, P. (1998). The transferable belief model for quantified belief representation. In D. M. Gabbay & P. Smets (Eds.), *Handbook of defeasible reasoning and uncertainty management systems, vol. 1* (Vol. 1, pp. 267–301). Kluwer, Doordrecht, The Netherlands.

Smets, P. (2000). Quantified possibility theory seen as an hyper cautious transferable belief model. In *Rencontres francophones sur les logiques floues et ses applications. LFA 2000, La Rochelle, France* (pp. 343–353). Toulouse, France: Cepadues-Editions.

Smets, P. (2001). Decision making in a context where uncertainty is represented by belief functions. In R. P. Srivastava (Ed.), *Belief functions in business decisions.* Physica-Verlag.

Smets, P., & Kennes, R. (1994). The transferable belief model. *Artificial Intelligence*, *66*, 191–234.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities.* Chapman and Hall, London.

Walley, P. (1997). Statistical inferences based on a second-order possibility distribution. *International Journal of General Systems*, *26*, 337–383.

Walley, P., & De Cooman, G. (2000). A behavioral model for linguistic uncertainty. In P. Wang (Ed.), *Computing with words.* Wiley, New York.

Yager, R. R. (1986). Arithmetic and other operations on Dempster-Shafer structures. *Int. J. Man-Machines Studies*, *25*, 357–366.

Zadeh, L. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, *1*, 3–28.