A Definition of Subjective Possibility

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

**IRIDIA Université Libre de Bruxelles

50 av. Roosevelt, CP 194-6, 1050 Bruxelles, Belgium

psmets@ulb.ac.be

(to appear in « Badania Operacyjne i Decyzije (Operational Research and Decisions) »

Abstract□Based on the setting of exchangeable bets, this paper proposes a subjectivist view of numerical possibility theory. It relies on the assumption that when an agent constructs a probability measure by assigning prices to lotteries, this probability measure is actually induced by a belief function representing the agent's actual state of knowledge. We also assume that going from an underlying belief function to an elicited probability measure is achieved by means of the so-called pignistic transformation (mathematically equivalent to the Shapley value). We pose and solve the problem of finding the least informative belief function having a given pignistic probability. We prove that it is unique and consonant, thus induced by a possibility distribution. This result exploits a simple informational index, in agreement with partial orderings between belief functions, comparing their information content. The obtained possibility distribution is subjective in the same sense as in the subjectivist school in probability theory. However, we claim that it is the least biased representation of the agent's state of knowledge compatible with the observed betting behaviour.

1. Introduction

Quantitative possibility theory was proposed as an approach to the representation of linguistic imprecision (Zadeh, 1978) and then as a theory of uncertainty of its own (Dubois and Prade 1988, 2002; Dubois et al. 2000). In order to sustain this claim, operational semantics are requested. In the subjectivist context, quantitative possibility theory competes with probability theory in its subjectivist or Bayesian views and with the Transferable Belief Model (Smets and Kennes 1994; Smets 1998), both of which also intend to represent degrees of belief. The term *subjectivist* means that we consider probability, and other numerical set-functions proposed for the representation of uncertainty, as tools for quantifying an agent's beliefs in events without regard to the possible random nature and repeatability of the events. An operational definition, and the assessment methods that can be derived from it, provides a meaning to the value.7 encountered in statements like "my degree of belief is.7". Bayesians claim that any state of incomplete knowledge of an agent can (and should) be modelled by a single probability distribution on the appropriate referential, and that degrees of belief coincide with probabilities that can be revealed by observing the betting behaviour of the agent (how much would the agent pay to enter into a game). In such a betting experiment, 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 Transferable Belief Model as well (Smets, 1997), introducing several betting frames corresponding to various partitions of the referential. In that sense, numerical values encountered in these three theories are well-defined.

Quantitative possibility theory seems to be worth exploring as well from this standpoint. Rejecting it because of the current lack of convincing semantics would be unfortunate, simply because it entertains close formal relationships between other theories: possibility measures are consonant Shafer plausibility measure, and thus encode special families of probability functions. Since possibility theory is a special case of most existing non-additive uncertainty theories, be they numerical or not, progress in one of these theories usually has impact in possibility theory. The recent revival along the lines of Walley's imprecise probabilities, by De Cooman and Aeyels (1999), of a form of subjectivist possibility theory initiated by Giles (1982), and the development of possibilistic networks based on incomplete statistical data (Borgelt and Kruse, 2003) also suggest that it is fruitful to investigate various operational semantics for possibility theory. Another major reason for studying possibility theory is that it is very simple, certainly the simplest challenger for probability theory, especially in the form of fuzzy numbers, a mathematical model extensively used by Chanas in his works (e.g. Chanas and Kuchta, 1998; Chanas and Zielinski, 2001) as well as many other scholars in fuzzy optimization.

The aim of this paper is to propose subjectivist semantics for numerical possibility theory based on exchangeable bets. Such subjectivist semantics differ from the upper and lower probabilistic setting proposed by Giles (1982), Walley (1991) and followers, without questioning its merit. This school interprets the maximal acceptable buying price of a lottery ticket pertaining to the occurrence of an event as its lower probability, and the minimal sale price of the same lottery ticket as its upper probability, both prices being distinct. Here, we assume exchangeable bets, just like the Bayesian School, but we consider that betting rates only partially reflect an agent's beliefs. In other words, betting rates produce a unique probability distribution but they are induced by the agent's beliefs without being in one-to-one correspondence with them. For instance, an agent may assign equal probabilities to the facets of a die, either because the fairness of the die has been experimentally validated, or, by symmetry, just because this agent does not know if the die is biased or not. Clearly, beliefs entertained by the agent in both situations are very distinct (Dubois and Prade, 1990). In this paper we assume that beliefs are more naturally modelled by means of a belief function, thus leaving room for incomplete knowledge (Dubois et al, 1996).

In previous works, Smets (1990) axiomatically argued that there exists a natural transformation of a belief function into a (so-called pignistic) probability function such that if the agent's beliefs are modelled by the former, his betting rates are captured by the latter. He called it the pignistic transformation. It was previously suggested by Dubois and Prade (1982) in the setting of belief functions, and formally coincides with the Shapley value in game theory (Shapley, 1953); see Dubois and Prade (2002). Denneberg and Grabisch (1999) have generalized it to so-called interaction weights attached to all subsets (not only to singletons). Moreover in the case of possibility distributions, corresponding to consonant plausibility functions, the transformation is one-to-one. In general, however, distinct belief functions may correspond to the same pignistic probability. The pignistic transformation has been proposed by several authors in yet a different context. Kaufmann (1980) and Yager (1982) proposed a scheme for the random simulation of a finite fuzzy set: picking a membership grade at random in the unit interval, and them randomly picking a value of the variable in the corresponding cut of the fuzzy set. In the continuous setting, Chanas and Nowakowski (1988) proposed a more general probabilistic interpretation of fuzzy intervals based on a similar interpretation.

This paper formalizes and solves the following problem: given a subjective pignistic probability distribution p provided by an agent under the form of betting rates, find a suitable least committed belief function whose pignistic transform is p. Such a belief function is a cautious representation of the agent's belief, assuming minimal statistical knowledge. For instance, if the agent supplies a uniform probability, it is assumed by default that the agent has no information. In that case, an unbiased representation is the vacuous belief function, or equivalently, the uniform possibility distribution, thus reversing Laplace's principle of indifference.

The main result of the paper is that the least committed belief function with prescribed pignistic transform is unique and consonant, that is, it can be modelled as a possibility distribution. This result was already announced by the authors (Dubois et al. 2001), but its proof is still unpublished. Since the pignistic transformation is one-to-one for possibility distributions, this result also provides the converse transform with a natural interpretation, first suggested with a different rationale by Dubois and Prade (1983). This result also sheds light on the probabilistic interpretation of fuzzy numbers suggested by Chanas and colleagues in his work.

2. Belief functions

Consider beliefs held by an agent on what is the actual value of a variable ranging on a set Ω, called the frame of discernment. It is assumed that such beliefs can be represented by a belief function. A belief function can be mathematically defined from a (generally finite) random set that has a very specific interpretation. A so-called basic belief mass $m(A)$ is assigned to each subset A of Ω, such that $m(A) \geq 0, \forall A \subseteq \Omega$; moreover:

$$\sum_{A \subseteq \Omega} m(A) = 1.$$

The degree $m(A)$ is understood as the weight given to the fact that all the agent knows is that the value of the variable of interest lies somewhere in set A, and nothing else. In other words, the probability allocation $m(A)$ is potentially shared between elements of A, but remains suspended for lack of knowledge. A set E such that $m(E) > 0$ is called a focal set. In the absence of conflicting information it is generally assumed that $m(\emptyset) = 0$. This is what is assumed in the following. A belief function *Bel* as well as a plausibility function *Pl*, attached to each event (or each proposition of interest) can be bijectively associated with the basic mass function m (Shafer, 1976). They are defined by

$$Bel(A) = \sum_{\emptyset \neq E \subseteq A} m(E), \qquad Pl(A) = 1 - Bel(A^c) = \sum_{E: E \cap A \neq \emptyset} m(E),$$

where A^c is the complement of A. The belief function evaluates to what extent events are logically implied by the available evidence. The plausibility function evaluates to what extent events are consistent with the available evidence. A companion set-function, called commonality, and denoted by Q, is defined by reversing the direction of inclusion in the belief function expression:

$$Q(A) = \sum_{A \subseteq E} m(E).$$

$Q(A)$ is the share of belief totally unassigned and free to potentially support any proposition in the context where the agent accepts that A holds true[1]. It can been argued that $Q(A)$ is a measure of guaranteed plausibility of A because it clearly provides a lower bound of the plausibility of *each* element in A.

The function Pl restricted to singletons, induced by a mass function m is called its contour function (Shafer, 1976), and is denoted π_m, defined by $\pi_m(\omega) = Pl(\{\omega\})$. When the focal sets are nested, the plausibility function is called a possibility measure (Zadeh, 1978), and can be characterized, just like probability, by its contour function, then called a possibility distribution π. In such a situation, the primitive object can be the possibility distribution, and each of the functions *m, Pl, Bel,* can be reconstructed from it, noticing that (Dubois and Prade, 1982)

$$Pl(A) = \max_{\omega \in A} \pi(\omega) \tag{1}$$

---

[1] When conditioning a mass function on event A, the mass m(E) of each focal set E is allocated to the subset $A \cap E$. The overall (possibly subnormal) mass finally allocated to a subset C of A is denoted $m(C \mid A)$. Then $Q(A)$ coincides with the mass $m(A \mid A)$ assigned to set A before normalizing. So, up to normalization, $Q(A)$ is a measure of unassigned belief in the context where the agent accepts that A holds true.

The set function *Pl* is then often denoted Π. If $\Omega = \{\omega_1, \ldots, \omega_n\}$, and letting $\pi_i = \pi(\omega_i)$, such that $1 = \pi_1 \geq \pi_2 \geq \ldots \geq \pi_n \geq \pi_{n+1} = 0$, then the mass function generating Π is denoted m_π such that

$$m_\pi(A) = \pi_i - \pi_{i+1} \text{ if } A = \{\omega_1, \ldots, \omega_i\} \quad (2)$$

$$(= 0, \text{ otherwise}).$$

If the mass function *m* is not consonant the contour function is not enough to recover it as in (2) since *m* then needs up to $2^{\mathrm{card}(\Omega)}$ terms to be determined from π_m where *card* stands for cardinality.

3. The pignistic transformation

It is assumed that the actual beliefs of the agent can be faithfully modelled by a mass function on Ω. A probability measure induced by a mass function can be built by defining a uniform probability on each set with positive mass, and performing the convex mixture of these probabilities according to the mass function. This transformation, which, as pointed out earlier, recurrently appears in various contexts since the fifties, was called the pignistic transformation by Smets (1990). Let *m* be a mass function from 2^Ω to $[0, 1]$. The pignistic transform of *m* is a probability distribution BetP = *Pig(m)* such that :

$$BetP(\omega) = \sum_{A: \omega \in A} m(A) / card(A) \quad (3)$$

where *card(A)* is the cardinality of *A*. It could be viewed as an extension of Laplace indifference principle, according to which equally possible outcomes have equal probability. It looks like a weighted form thereof, since, by symmetry, each focal set is then interpreted as a uniform probability. According to (Smets, 1991), the agent's beliefs cannot be directly assessed. All that can be known are the values of the "pignistic" probabilities the agent would use to bet on the frame Ω. Only the probability distribution *BetP*, not the belief function accounting for the agent's beliefs, is obtained by eliciting an agent's betting rates on the frame Ω (Smets, 2002).

The pignistic probability depends on the chosen betting frame. Changing Ω into one of its refinements, thus modifying the granularity, a different probability is obtained. It has been proved that for any event *A*, the minimal (resp. maximal) value of $BetP(A) = \sum_{\omega \in A} BetP(\omega)$ over all possible changes of granularity yields back *Bel(A)* (resp. *Pl(A)*) (Wilson, 1993). So, the interval *[Bel(A), Pl(A)]* contains all possible values of the pignistic probability of *A*, across all betting frames. This is related to the fact that all probability functions *P* dominating the belief function *Bel* induced by *m* (that is $P \geq Bel$) can be generated by changing each focal set *E* into a probability distribution $p(\cdot \mid E)$ with support *E*. Namely:

$$p(\omega) = \sum_E p(\omega \mid E) \cdot m(E).$$

In Bayesian terms, this is an application of the total probability theorem where $p(\omega \mid E)$ is the (subjective) probability of ω when all that is known is the piece of evidence E, and $m(E)$ is the probability of knowing this piece of evidence only. So, in terms of upper and lower probabilities, *BetP* is the centre of gravity of the set of probabilities dominating the belief function (Dubois et al., 1993). In terms of game theory, it corresponds to the Shapley value of a game.

In the special case of consonant belief functions, the pignistic transformation can be expressed in terms of the possibility distribution π such that $1 = \pi_1 \geq \pi_2 \geq \ldots \geq \pi_n \geq \pi_{n+1} = 0$ as follows, letting $p_i = BetP(\omega_i)$:

$$p_i = \sum_{j = i, \ldots, n} (\pi_j - \pi_{j+1})/j \quad \forall i = 1, \ldots, n$$

It can be checked that $p_1 \geq p_2 \geq \ldots \geq p_n$ and that the transformation is bijective between probabilities and possibilities. Its converse Pig^{-1} was independently suggested by Dubois and Prade (1983). It reconstructs the possibility distribution as follows

$$\pi_i = \sum_{j = 1, \ldots, n} min(p_i, p_j), \quad \forall i = 1 \ldots, n \tag{4}$$

and we write $\pi = Pig^{-1}(BetP)$. Note that another probability-possibility transformation exists, of the form (Dubois and Prade, 1982; Delgado and Moral 1987):

$$\pi_i = \sum_{j = i, \ldots n} p_j, \quad \forall i = 1, \ldots, n \tag{5}$$

The latter transformation of a probability distribution p yields the most specific (= restrictive) possibility distribution such that $\Pi(A) \geq P(A)$. When p stems from validated statistical data, one may argue that this transformation yields the most legitimate possibilistic representation of P (Dubois et al, 2001) since p represents a complete model of the studied random phenomenon and (5) yields the most specific possibility distribution respecting the ordering of elements of Ω induced by p, in the sense that $\sum_{j = 1, \ldots, n} \pi_j$ is minimal (minimal cardinality of the fuzzy set with membership grades π_j). However, in the subjective probability case, it is questionable whether the expert possesses a complete model, even if the betting framework enforces it. If the parameter under concern is random, the agent may have only partial knowledge about it. If the parameter is not random (just ill-known), a complete model should come down to knowing its precise value. Hence the optimal (maximally specific) transformation (5) does not convincingly apply to subjective probabilities.

4. The most cautious belief function inducing a subjective probability

The knowledge of the values of the probability p allocated to the elements of Ω by the agent is not sufficient to reconstruct a unique underlying belief function whose pignistic transform is p. Many belief functions induce the same pignistic probability distribution. As already said, 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 that is known about the mass function that represents the agent's beliefs is that it belongs to the ones that induce the available subjective probability. Under this scheme, we do not question the exchangeability of bets, as done by Walley (1991), Giles (1982) and others. What we question is the assumption of a one-to-one correspondence between the betting rates produced by the agent, and the actual beliefs entertained by this agent. Betting rates do not tell if the uncertainty of the agent results from the perceived randomness of the phenomenon under study or from a simple lack of information about it.

Since several mass functions may lead to the same betting rates, one has to select the least committed among these mass functions, as the one that, by default, reflects the actual state of belief of the agent. The belief functions whose pignistic transform is p are called *isopignistic* belief functions and form the set $\mathcal{IP}(p)$. A cautious approach among isopignistic belief functions is to obey a "least commitment principle". It states that one should never presuppose more beliefs than justified. Then, one should select the least committed element, that is, the least informed one, in the family of isopignistic belief functions corresponding to the pignistic probability function prescribed by the obtained betting rates.

4.1 Informational comparison of belief functions

There are several methods to compare belief functions in terms of their informational contents. Some informational indices extend the probabilistic notion of entropy. Other ones generalize the notion of cardinality of a set representing incomplete knowledge, yet other ones combine both (see the recent survey by Klir and Smith (2001), for instance). Besides, three partial orderings comparing the information content of two belief functions in terms of specificity have been proposed by Yager (1985) and Dubois and Prade (1986).

A first natural specificity ordering of belief functions compares intervals limited by belief and plausibility. Namely the interval $[Bel(A), Pl(A)]$ is all the wider as the information concerning A is scarce. So, a partial information order on the set of belief functions over Ω can be defined as follows: Bel_1 is at least as *precise* as Bel_2 if and only if $[Bel_1(A), Pl_1(A)] \subseteq [Bel_2(A), Pl_2(A)], \forall A \subseteq \Omega;$ it corresponds to an inclusion relation between sets of probabilities dominating Bel_1 and Bel_2. In fact, this ordering can be defined equivalently and more simply as $Pl_1(A) \leq Pl_2(A), \forall A \subseteq \Omega$ due to the duality between Bel and Pl.

Interestingly, this partial ordering does not imply any relationship between the commonality functions Q_1 and Q_2 (see Dubois and Prade, 1986 and the counterexample below). Another partial informational ordering between belief functions has thus been defined by comparing the commonality functions: Bel_1 is at least as Q-informed as Bel_2 if and only if $Q_1(A) \leq Q_2(A)$, $\forall A \subseteq \Omega$. This direction of inequality is natural since it ensures that for singletons, $Pl_1(\{\omega\}) \leq Pl_2(\{\omega\})$, due the identity of Pl and Q functions on singletons.

A third partial informational ordering can be described directly from the mass functions m_1 and m_2. The idea is that Bel_1 is at least as informed as Bel_2 whenever it is possible to turn m_2 into m_1 by consistently reassigning each weight $m_2(E)$ to subsets of E that are focal sets of m_1 (possibly splitting the masses among them). It is called the specialization ordering. Namely, m_1 is more specialized than m_2 if and only if there is a stochastic matrix W whose rows correspond to focal sets of m_1 and columns to focal sets of m_2, such that $m_1 = W \cdot m_2$. Here, mass functions are encoded as vectors and entry w_{ij} reflects the proportion of the mass $m_2(E_j)$ allocated to focal set F_i of m_1, with the condition that F_i must be a subset of E_j for w_{ij} to be positive.

This third ordering is more demanding than the other ones and implies them. But the Q-informativeness and the precision orderings are not comparable.

Example 1. Suppose $\Omega = \{\omega_1, \omega_2, \omega_3\}$, $E = \{\omega_1, \omega_2\}$, $F = \{\omega_1, \omega_3\}$, $\alpha \in (0.5, 1)$. Consider the mass function $m(E) = \alpha$, $m(F) = 1 - \alpha$, and the possibility measure π such that $\pi(\omega_1) = 1$, $\pi(\omega_2) = \alpha$, $\pi(\omega_3) = 1 - \alpha$. It is clear that $Pl(\{\omega\}) = \pi(\omega)$, $\forall \omega$; the mass function associated to π by (2) is $m_\pi(\{\omega_1\}) = 1 - \alpha$, $m_\pi(E) = 2\alpha - 1$, $m_\pi(\Omega) = 1 - \alpha$. It is obvious that none of the two mass functions m and m_π is a specialization of the other since m_π has a focal element contained in none of E or F, and a focal element containing none of them. Now it is obvious that m is at the same time less precise and more Q-informed than m_π. Indeed, $Pl(A) \geq \Pi(A)$, $\forall A$, and $Pl(\{\omega_2, \omega_3\}) = 1 > \Pi(\{\omega_2, \omega_3\}) = \alpha$. However $Q_\pi(A) \geq Q(A)$, $\forall A$, and $Q_\pi(\{\omega_2, \omega_3\}) = 1 - \alpha > Q(\{\omega_2, \omega_3\}) = 0$.

In view of this situation, the interpretation of the Q-informativeness is somewhat problematic. Nevertheless, all three orderings coincide for possibility measures and come down to the possibilistic ordering of specificity on singletons (Yager, 1983; Dubois and Prade, 1988): π_1 is at least as informed as π_2 if and only if $\pi_1 \leq \pi_2$).

While the merit of such partial informational orderings is to lay bare the meaning of the comparison, they often lead to non-comparability. Indeed, one may try to define the least debatable representation of an agent's belief as a minimally informative isopignistic mass function according to one of these orderings. Unfortunately, unicity may easily fail for these least informative mass functions, as the corresponding optimization problem comes down to vector-maximization.

4.2 Using expected cardinality

An easier problem is to maximize an information index. A natural measure of non-commitment of a belief function is the average of the cardinalities of its focal elements, weighted by the mass function m:

$$I(m) = \sum_{A \subseteq \Omega} m(A) \cdot card(A)$$

It is the simplest imprecision measure. It is easy to see that $I(m)$ is the cardinality of the fuzzy set whose membership function coincides with the contour function (Dubois and Jaulent, 1987), namely, $I(m) = \sum_{\omega \in \Omega} \pi_m(\omega)$.

It is clear that this index is compatible with the specialization ordering (hence with the two other informational orderings), namely that if m_1 is more specialized than m_2 then $I(m_1) \leq I(m_2)$.

We define the least biased belief representation, for an agent supplying a pignistic probability p, as the belief function whose mass m^* maximizes $I(m)$ among isopignistic belief functions whose pignistic transform according to eqn.(4) is p. The following result is now established.

Theorem 1: The unique mass function which maximizes $I(m)$ under the constraint $Pig(m) = p$ exists and is consonant. It is the possibility distribution π defined by the converse of the pignistic transformation applied to p (restricted to possibility measures, as defined by (4)).

The proof of this result is based on the following lemma:

Lemma: For any belief function with mass function m, $I(Pig^{-1}(Pig(m)) \geq I(m)$, and $I(Pig^{-1}(Pig(m)) = I(m)$ only if m is consonant.

Proof: Consider $p = Pig(m)$, such that $p_1 \geq p_2 \geq \ldots \geq p_n$ and $\pi = Pig^{-1}(p)$ such that $1 = \pi_1 \geq \pi_2 \geq \ldots \geq \pi_n \geq \pi_{n+1} = 0$. It can be checked that

$$I(Pig^{-1}(p)) = \sum_{i=1,\ldots n} \pi_i$$
$$= \sum_{i=1,\ldots n} \sum_{j=1,\ldots n} min(p_i, p_j)$$
$$= \sum_{i=1,\ldots n} (2i - 1) \cdot p_i$$

($I(Pig^{-1}(p))$ is the sum of entries in the $n \times n$ matrix with coefficients $min(p_i, p_j)$. There is only one entry containing p_1, 3 entries containing p_2, etc.).

Now, since $p_i = \sum_{E:\, \omega_i \in E} m(E)/card(E)$, it remains to be shown that

$$\sum_{i=1,\ldots n} (2i - 1) \cdot \sum_{E:\, \omega_i \in E} m(E)/card(E) \geq \sum_{i=1,\ldots n} \sum_{E:\, \omega_i \in E} m(E).$$

Subtracting the right-hand side from the left-hand side, and factoring $m(E)$, it is enough to prove that the multiplicative coefficient of $m(E)$ is positive, that is, denoting by χ_E the indicator function of E:

$c(E) = \sum_{i = 1, \dots, n} (2i - 1) \cdot \mu_E(\omega_i)/ card(E) - \sum_{i = 1, \dots, n} \mu_E(\omega_i) \geq 0$.

Let E be a subset with k elements of the form $\{\omega_{i_1}, \dots, \omega_{i_k}\}$ such that $p(\{\omega_{i_1}\}) \geq p(\{\omega_{i_2}\}) \geq \dots \geq p(\{\omega_{i_k}\})$. Then:

$c(E) = [(2i_1-1)/k + (2i_2-1)/k + \dots + (2i_k-1)/k] - k$

It is minimal for $i_j = j$ for all $j = 1, \dots k$. Hence

$c(E) \geq [(2-1)/k + (4-1)/k + \dots + (2k-1)/k] - k = (2/k) \cdot (\sum_{j = 1, \dots k} j) - k - 1 = 0.$

It is clear that if $c(E) > 0$ for any E, then $I(Pig^{-1}(Pig(m)) > I(m)$ as soon as $m(E) > 0$. But since $i_j \geq j$ by construction, the only way of having $c(E) > 0$ for some set E is to have $i_j > j$ for some i, that is E is not of the form $\{\omega_1, \dots, \omega_k\}$ for some k. But m is not consonant as soon as $m(E) > 0$ for such a set E (as $Pig(m) = p$ and the only consonant m in $\mathcal{IP}(p)$ is π). Hence as soon as m is not consonant in $\mathcal{IP}(p)$, $I(Pig^{-1}(p)) > I(m)$.

Proof of theorem 1: Since $I(Pig^{-1}(Pig(m)) \geq I(m)$ from the lemma, and Pig is a bijection on possibility measures, the consonant belief function associated to $Pig^{-1}(Pig(m))$ is not more cardinality-specific than the belief function induced by m. Conversely fixing the probability distribution p, and choosing any non-consonant m in $\mathcal{IP}(p)$, $I(Pig^{-1}(p)) > I(m)$. It follows that the consonant mass function associated to $Pig^{-1}(p)$ is the unique maximum of $I(m)$.

3.2 Comparing commonalities

Smets (2000) suggested that the least specific isopignistic belief function according to the commonality ordering is also $Pig^{-1}(Pig(m))$. This ordering is less intuitive than the specialization ordering and the inclusion of *Bel-Pl* intervals. However, there is indeed a unique minimally Q-informative belief function in $\mathcal{IP}(p)$, and it is precisely the one found by maximizing $I(m)$. In order to show it, we first prove that, for ensuring comparability in the sense of the Q-informativeness ordering between a consonant belief function and a belief function, it is enough to rely on singletons:

Lemma 2: Consider a belief function with mass function m and a possibility distribution π with respective commonality functions Q and Q_π. Then $Q_\pi(A) \geq Q(A), \forall A \subseteq \Omega$ if and only if $\pi(\omega) \geq Pl(\{\omega\}), \forall \omega \in \Omega$.

Proof: It is obviously enough to prove the "if" part since $Q(\{\omega\}) = Pl(\{\omega\})$. Besides, note that for possibility measures $Q_\pi(A) = \min_{\omega \in A} \pi(\omega)$. Now assume $\pi(\omega) \geq Pl(\{\omega\}), \forall \omega \in \Omega$. Then: $Q_\pi(A) = \min_{\omega \in A} \pi(\omega) = \pi(\omega_*) \geq Pl(\{\omega_*\}) \geq Q(A)$ since function Q is antimonotonic with respect to inclusion.

Theorem 2: The unique consonant mass function in $\mathcal{IP}(p)$ (induced by the possibility distribution defined by (4)), is minimally Q-informative.

As previously we need one more lemma.

Lemma 3 : Consider a belief function with mass function m, $p = Pig(m)$, and $\pi = Pig^{-1}(p)$. Then $\pi \geq \pi_m$ i.e. π is not more specific than the contour function of m.

Proof : Consider $p = Pig(m)$, such that $p_1 \geq p_2 \geq \ldots \geq p_n$ and $\pi = Pig^{-1}(p)$ such that $1 = \pi_1 \geq \pi_2 \geq \ldots \geq \pi_n \geq \pi_{n+1} = 0$. Now $\pi_k = \pi_k(\omega)$ is defined in terms of m as

$$\pi_k = k \cdot p_k + \sum_{j = k+1, \ldots n} p_j$$

$$= k \cdot \sum_{E: \omega_k \in E} m(E)/card(E) + \sum_{j = k+1, \ldots n} \sum_{E: \omega_j \in E} m(E)/card(E)$$

We must show that this expression is not less than $\sum_{E: \omega_k \in E} m(E) = Pl(\{\omega_k\}) = \pi_m(\omega_k)$. To this end we proceed focal set by focal set, with fixed cardinality. Denote by $c(E)$ the multiplicative coefficient of $m(E)$ in the expression of π_k, namely, denoting by μ_E the indicator function of E:

$$c(E) = k \cdot \mu_E(\omega_k)/card(E) + \sum_{j = k+1, \ldots n} \mu_E(\omega_j)/card(E)$$

Let us show that $c(E) \geq 1$ whenever $\omega_k \in E$ (otherwise $m(E)$ does not contribute to $\pi_m(\omega)$).

First, assume $card(E) = n$. It means that $E = \Omega$. The coefficient $c(\Omega)$ of $m(\Omega)$ is $(k/n + (n - k)/n) = 1$ since all terms in the second summand of the expression of $c(E)$ are present.

Now, assume $card(E) = i > k$. There are at least $i - k$ terms in the second summand of the expression of $c(E)$. Then $c(E) \geq (k/i + (i - k)/i) = 1$.

Assume $card(E) = i \leq k$. Then the second summand of the expression of $c(E)$ may be zero since E may fail to contain any ω_j for $j > k$. It is no problem since then $c(E) \geq k/i \geq 1$ by assumption.

Proof of theorem 2: Based on lemma 3, we know that $\pi \geq \pi_m$ for $\pi = Pig^{-1}(Pig(m))$. Due to lemma 3 it implies that π is not more Q-informative that m. fixing $p = Pig(m)$, this property holds for all belief functions in $\mathcal{IP}(p)$, and $\pi \in \mathcal{IP}(p)$, by construction. Hence π is not more Q-informative than any belief function in $\mathcal{IP}(p)$.

Note that Lemma 3 is stronger than Lemma 1. It clearly implies it since Lemma 1 compares the sum $\sum_{i = 1, \ldots n} \pi_i$ to the sum $\sum_{i = 1, \ldots n} \pi_m(\omega_i)$. However the proof of Lemma 1 is more direct. Moreover Lemma 2 shows that when comparing mass functions in terms of commonality, one of them being consonant, commonality functions play no particular role. Only contour functions matter. So, the optimality of the possibility measure in $\mathcal{IP}(p)$ is really in the sense of the pointwise comparison, in the fuzzy set inclusion sense, of the plausibility functions on singletons, i.e. the contour functions. In particular, Lemma 3 implies that the consonant mass function $Pig^{-1}(p)$ is certainly minimally precise in the sense of the comparison of *Bel-Pl* intervals, in $\mathcal{IP}(p)$.

Let us now turn to the issue of unicity of the least informative mass function in the sense of the pointwise comparison of contour functions. The unicity problem can be stated as follows: given a possibility distribution π on Ω, whose pignistic transform is a probability distribution $p = Pig(\pi)$, is there another (non-consonant) mass function $m \neq m_\pi$ such that $p = Pig(m)$ and $\pi = \pi_m$?

Theorem 3 : The one and only mass function m, such that $p = Pig(m)$ and $\pi = \pi_m$, where $\pi = Pig^{-1}(p)$, is the one underlying the possibility distribution π.

Proof: Fix the probability distribution p such that $p_1 \geq p_2 \geq \ldots \geq p_n$ and $\pi = Pig^{-1}(p)$ such that $1 = \pi_1 \geq \pi_2 \geq \ldots \geq \pi_n \geq \pi_{n+1} = 0$. From Lemma 3, the condition $\pi = \pi_m$ must be enforced. The mass function m must then satisfy the following constraints:

$\sum_{E: \omega_k \in E} m(E)/card(E) = p_k$ for $k = 1, \ldots, n$;

$\sum_{E: \omega_k \in E} m(E) = \pi_k$ for $k = 1, \ldots, n$, where $\pi_k = k \cdot p_k + \sum_{j = k+1, \ldots n} p_j$;

$\sum_E m(E) = 1$.

Note that since $\pi_1 = 1$, $m(E) = 0$ as soon as $\omega_1 \notin E$. Now, for $k = n$, it holds that $\pi_n = n \cdot p_n$ so that $\sum_{E: \omega_n \in E} m(E) = n(\sum_{E: \omega_n \in E} m(E)/card(E))$. It reads

$\sum_{E: \omega_n \in E} m(E)(1 - n/card(E)) = 0$, hence $m(E) = 0$ whenever $\omega_n \in E$, $Card(E) < n$.

So, all such masses $m(E)$ in the pair of equations number $k = n$ are zero except $m(\Omega) = n \cdot p_n$.

Suppose all masses $m(E) = 0$ whenever $\omega_j \in E$, $Card(E) < j$ in the pairs of equations $j = k + 1, \ldots n$, except $m(\Omega) = n \cdot p_n$, and $m(\{\omega_1, \ldots, \omega_j\}) = j(p_j - p_{j+1})$. Consider the pair of equations number k. It comes:

$\sum_{E: \omega_k \in E} m(E) = k \cdot \sum_{E: \omega_k \in E} m(E)/card(E) + \sum_{j = k+1, \ldots n} \sum_{E: \omega_j \in E} m(E)/card(E)$

Subtracting the right-hand side from the left-hand side, consider the coefficients of the remaining focal sets: If $\omega_j \in E$ and $j > k$, then $E = \{\omega_1, \ldots, \omega_j\}$, the coefficient of $m(E)$ is $1 - (k/j + (j - k)/j) = 0$. If $E = \{\omega_1, \ldots, \omega_k\}$, the coefficient is $1 - k/k = 0$. If $E \subset \{\omega_1, \ldots, \omega_k\}$, the coefficient is $1 - k/card(E) \neq 0$. Hence $m(E) = 0$. Hence the mass $m(\{\omega_1, \ldots, \omega_k\})$ can be completely determined as the unique solution to the equation:

$\sum_{j = k, \ldots n} m(\{\omega_1, \ldots, \omega_j\}) = k \cdot pk + \sum_{j = k+1, \ldots n} p_j$

since all $m(\{\omega_1, \ldots, \omega_j\})$, for $j > k$ are determined in the previous steps. Overall only subsets of the form $E = \{\omega_1, \ldots, \omega_j\}$, $k = 1, \ldots, n$ may receive positive mass if the mass function has pignistic transform p and contour function $\pi = Pig^{-1}(p)$. Hence, m is consonant, and because there is only one consonant mass function in $\mathcal{IP}(p)$, it precisely yields the one underlying $Pig^{-1}(p)$.

Putting together Theorems 2 and 3, the minimally Q-informative mass function with pignistic probability p exists, is unique and is consonant. It is actually the mass function having the least specific (i.e. pointwisely maximal in Ω) contour function, hence also least precise in the sense of the comparison of *Bel-Pl* intervals restricted to singletons. It suggests that most of the time, a unique least precise non-consonant mass function in $\mathcal{IP}(p)$ in the sense of the comparison of *Bel-Pl* intervals for all events will not exist. Indeed if $m \in \mathcal{IP}(p)$ is a least precise mass function different from the one inducing $\pi = Pig^{-1}(p)$, then $\pi(\omega) > \pi_m(\omega)$, for some $\omega \in \Omega$, due to the unicity result in Theorem 3. Since m is among minimally precise ones, it must also hold that $Pl(A) > \Pi(A)$ for some non-singleton event A. So m and m_π are not comparable. That this non-unicity situation does occur can be checked from Example 1.

Example 1 (continued). Assume $\alpha = 1/2$. So, $m(\{\omega_1, \omega_2\}) = m(\{\omega_1, \omega_3\}) = 1/2$. The pignistic probability p induced by m is clearly: $p(\omega_1) = 1/2$, $p(\omega_2) = 1/4$, $p(\omega_3) = 1/4$; $\pi = Pig^{-1}(p)$ is $\pi(\omega_1) = 1$, $\pi(\omega_2) = 3/4$, $\pi(\omega_3) = 3/4$. The contour function of m is $\pi_m(\omega_1) = 1$, $\pi_m(\omega_2) = 1/2$, $\pi_m(\omega_3) = 1/2$. It is more specific than $Pig^{-1}(p)$ as expected. Note that Pl($\{\omega_1, \omega_3\}) = 1$, while

$\pi(\{\omega_1, \omega_3\}) = 3/4$. Hence m and $Pig^{-1}(p)$ are not comparable in the sense of the precision ordering; they are both minimally precise in $\mathcal{IP}(p)$.

5. Conclusion

The main result of this paper is that, on finite sets, the least committed mass function among the ones which share the same pignistic transform, is unique and consonant, that is, the corresponding plausibility function is a possibility function. This possibility function is the unique one in the set of plausibility functions having this prescribed pignistic probability, because the pignistic transformation is a bijection between possibilities and probabilities. So this possibility function corresponds to the least committed mass function whose transform is equal to the subjective probability supplied by an agent. It suggests a new justification to a probability-possibility transform previously suggested by two of the authors.

This result provides an operational basis for defining subjective possibility degrees, hence the membership function of (discrete) fuzzy numbers. It tentatively addresses objections raised by Bayesian subjectivists against the use of fuzzy numbers and numerical possibility theory in decision-making and uncertainty modelling tasks. Interestingly, our approach refutes neither the Bayesian operational setting (unlike Walley(1999) and De Cooman and Aeyels (1999)) nor the use of standard expected utility for decisions (since the pignistic probability can be used for making decisions). It only questions the interpretation of betting rates as full-fledged degrees of belief. Bayesians may then claim that our approach makes no contribution, since the underlying possibility distribution is not used for selecting decisions. However the proposed subjective possibility approach, just like the Transferable Belief Model, does differ from the Bayesian approach in a dynamic environment. In our non-classical setting, when an event is known to have occurred, the revision of information takes place by modifying the possibility distribution underlying the pignistic probability, not this probability directly. It means that the new probability distribution obtained from the agent is no longer assumed to coincide with the result of conditioning the original pignistic probability, but that the agent would bet again based on a different frame supporting the revised knowledge (see e.g. (Dubois et al. 1996), (Smets 2002), on this matter).

In order to fully bridge the gap between the above results and the probabilistic interpretations of fuzzy numbers after Chanas and Nowakowski (1988), the next step is to extend the result of this paper to the infinite case, using continuous belief functions whose focal sets are closed intervals. This is a topic for further research.

References

Borgelt C. Kruse R. (2003) Learning graphical possibilistic models from data. *IEEE trans. On Fuzzy Systems*, 11, 159-171.

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

Chanas S., Kuchta D., (1998) "Discrete fuzzy optimization", in *Fuzzy Sets in Decision Analysis, Operations Research and Statistics*, R. Slowinski (Ed.), p. 249-277, The Handbooks of Fuzzy Sets Series, Kluwer Acad. Publ., Boston, 1998.

Chanas, S. Zielinski P. (2001) Critical path analysis in the network with fuzzy activity times, *Fuzzy Sets and Systems*, 122 (2), 195-204.

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

Delgado M. and Moral S. (1987). On the concept of possibility-probability consistency, *Fuzzy Sets and Systems*, **21**, 311-318.

Denneberg D., Grabisch M. (1999): Interaction transform of set-functions over a finite set. *Information Sciences* 121(1-2): 149-170.

Dubois D. and Jaulent M.C. (1987) A general approach to parameter evaluation in fuzzy digital pictures", Pattern Recognition Letters, 6, 251-259.

Dubois D. and Prade H. (1982) On several representations of an uncertain body of evidence, in: *Fuzzy Information and Decision Processes* (M.M. Gupta, E. Sanchez, eds.), North-Holland, 167-181.

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

Dubois D. and Prade H. (1986) A set-theoretic view of belief functions. *Int. J. of General Systems*, 12(3), 193-226.

Dubois D. and Prade H. (1988). *Possibility Theory*, Plenum Press, New York.

Dubois D. and Prade H. (1990) : Coping with uncertain knowledge - In defense of possibility and evidence theories, *Computers and Artificial Intelligence* (Bratislava), 9(2), 115-144.

Dubois D. and Prade H. (2002) Quantitative possibility theory and its probabilistic connections. In□ *Soft Methods in Probability, Statistics and Data Analysis*. (Grzegorzewski, P. et al., Eds). Physica Verlag, Heidelberg - Germany , 3-26 ,.

Dubois D., Nguyen H. T., Prade H. (2000) Possibility theory, probability and fuzzy sets: misunderstandings, bridges and gaps. In: *Fundamentals of Fuzzy Sets,* (Dubois, D. Prade,H., Eds.), Kluwer , Boston, Mass., The Handbooks of Fuzzy Sets Series , 343-438.

Dubois D., Prade H., Sandri S.A. (1993) On possibility/probability transformations. In : *Fuzzy Logic: State of the Art* (R. Lowen, M. Lowen, eds.), Kluwer Academic Publ., 103-112.

Dubois D., Prade H., Smets P. (1996): Representing partial ignorance. *IEEE Trans. on Systems, Man and Cybernetics*, 26(3), 1996, 361-377.

Dubois D. and Prade H. Smets P. (2001) New semantics for quantitative possibility theory. *Proc.of the 6th.European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty* (ECSQARU 2001, Toulouse, France). Springer-Verlag, p. 410-421.

Giles R. (1982). Foundations for a theory of possibility, *Fuzzy Information and Decision Processes* (Gupta M.M. and Sanchez E., eds.), North-Holland, 183-195.

Kaufmann A (1980). La simulation des ensembles flous, *CNRS Round Table on Fuzzy Sets*, Lyon, France (unpublished proceedings).

Klir G. and Smith R.M. (2001) On measuring uncertainty and uncertainty-based information□ recent developments. *Annals Math. and AI.* 32, 5-34.

Shapley, L. S. (1953) A value for n-person games. In Kuhn and Tucker, eds., *Contributions to the Theory of Games*, II, Princeton University Press, 307-317.

Shafer G. (1976). *A Mathematical Theory of Evidence*, Princeton University Press, Princeton.

Smets P. (1990). Constructing the pignistic probability function in a context of uncertainty, *Uncertainty in Artificial Intelligence 5* (Henrion M. *et al.*, Eds.), North-Holland, Amsterdam, 29-39.

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

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. *Handbook of Defeasible Reasoning and Uncertainty Management Systems, vol. 1.* (D.M. Gabbay, P. Smets, eds) Kluwer, Doordrecht, The Netherlands, 267—301.

Smets, P. (2000) Quantified possibility theory seen as an hypercautious transferable belief model. *Rencontres Francophones sur les Logiques Floues et ses Applications* (LFA 2000, La Rochelle, France), Cepadues-Editions, Toulouse, France, pp.343--353.

Smets, P. (2002) Decision making in a context where uncertainty is represented by belief functions. In: Belief Functions in Business Decisions (R.P. Srivastava□T.J. Mock, Eds.). Physica-Verlag. Heldelberg, 17-61.

Walley P. (1991). *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall.

Wilson N. (1993) : Decision-Making with Belief Functions and Pignistic Probabilities. In□ Symbolic and Quantitative Approaches to Reasoning and Uncertainty (Proc. of the Europ. Conf. ECSQARU'93, Granada, Spain) (M. Clarke, R. Kruse, S. Moral, eds.), Lecture Notes in Computer Science, Vol. 747, Springer Verlag, Berlin,: 364-371

Yager R.R. (1982) Level sets for membership evaluation of fuzzy subsets. In : *Fuzzy Sets and Possibility Theory : Recent Developments* (R.R. Yager, ed.), Pergamon Press, Oxford, 90-97.

Yager R.R. (1983). An introduction to applications of possibility theory, *Human Systems Management*, **3**, 246-269.

Yager R.R. (1985) The entailment principle for Dempster-Shafer granules. Tech. Report MI-512, Iona College, New Rochelle, N.Y.

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