Belief Induced by the Partial Knowledge of the Probabilities

Philippe Smets
IRIDIA
Université Libre de Bruxelles
Brussels-Belgium

Abstract :

We construct the belief function that quantifies the agent' beliefs about which event of Ω will occurred when he knows that the event is selected by a chance set-up and that the probability function associated to the chance set up is only partially known.

Keywords: belief function, upper and lower probabilities.

1. INTRODUCTION.

1) The use of belief functions to quantify degrees of belief is muddled by problems that result from the confusion between belief and lower probabilities (or between plausibility and upper probabilities). Beliefs can be induced by many types of information. In this paper, we consider only one very special case: beliefs induced on a frame of discernment Ω when the elements of Ω will be selected by a random process. It seems reasonable to defend the idea that the belief of an event should be numerically equal to the probability of that event. This principle is called the Hacking Frequency Principle (Hacking 1965).

But there are cases where the probability function that governs the random process is not exactly known. This lack of knowledge can be encountered when probabilities are partially defined or when data are missing. As an example, suppose an urn where there are 100 balls. Its composition is not exactly known. All that is known is that there are between 30 and 40 black balls, between 10 and 50 white balls, and the other are red. What is your belief that the next randomly selected ball will be black? Suppose you have selected 50 balls at random with replacement and you have observed 15 black balls, 20 white, 10 reds and 5 'not black'. What is your belief now that there are between 35 and 37 black balls? What is your belief now that the next randomly selected ball will be black? These are the problems we solve in this paper.

In this paper, we accept that beliefs are quantified by belief functions, as described in the transferable belief model (Smets 1990b, Smets and Kennes 1994). The transferable belief model is a model for quantified beliefs developed independently of any underlying probabilistic model. It is neither Dempster's model nor its today versions (Shafer, 1990, Kohlas, 1994). It is not a model based on inner measures (Halpern and Fagin, 1990).

What we study here is just a special case of belief function. We study the belief induced by the knowledge of the existence of an objective chance set up that generates random events according to a probability function, probability function that happens to be only partially known to us.

2) Suppose a frame of discernment Ω, i.e., a set of mutually exclusive and exhaustive events such as one and only one of them is true (we accept the close world assumption (Smets 1988)). Suppose the true element will be selected by a chance process. Let $P:2^{\Omega}\rightarrow[0,1]$ be the probability function over Ω where P(A) for $A\subseteq\Omega$ quantifies the probability (chance) that the selected element is in A. We accept that this probability measure is "objective". The problem is to assess Your degree of belief. You denotes the agent who hold the beliefs. Your beliefs are quantified by a belief function $bel:2^{\Omega}\rightarrow[0,1]$, about the fact that the selected element is in A, given You only have some partial knowledge about the value of P.

Should You know P, then by Hacking Frequency Principle (1965) Your degree of belief bel(A) for each $A\subseteq\Omega$ should be equal to P(A):

If You know that $P(A) = p_A\ \forall A\subseteq\Omega$

then $bel(A) = p_A\ \forall A\subseteq\Omega$

In that case bel is a probability function over Ω. But remember that bel and P do not have the same meaning, they only share the same values. P quantifies the probability (chance) of the events in Ω, bel quantifies the belief over Ω induced in You by the knowledge of the value of the probabilities. P exists independently of me, bel cannot exist if You do not exist.

Let $\mathbb{P}_{\Omega}$ be the set of probability functions over Ω. Suppose that You know only that the probability function P that governs the random process over Ω is an element of a subset P of $\mathbb{P}_{\Omega}$. The problem is to determine Your belief about Ω given You know only that P is an element of P (but You do not know which one).

In many cases, P is uniquely defined by its upper and lower probabilities functions P* and P_* where:

$$P_*(A) = \min \{ P(A) : P \in \mathcal{P} \}$$
$$P^*(A) = \max \{ P(A) : P \in \mathcal{P} \} = 1 - P_*(\bar{A})$$

or $\mathcal{P} = \{P : P \in \mathbb{P}_\Omega, P_*(A) \le P(A) \le P^*(A), \forall A \subseteq \Omega\}$.

Just as P and bel characterize different concepts, P_* and bel characterize also different concepts, even when P_* is mathematically a belief function. The function bel concerns Your belief over Ω. The function P_* gives the lowest possible values for the probability of the events in Ω compatible with what You know.

This knowledge that $P \in \mathcal{P} \subseteq \mathbb{P}_\Omega$ is translated into a belief $bel_{\mathbb{P}_\Omega}$ over $\mathbb{P}_\Omega$.[1] That belief only supports $\mathcal{P}$, i.e., its basic belief masses are:

$$m_{\mathbb{P}_\Omega}(Q) = 1 \quad \text{if } Q = \mathcal{P}$$
$$= 0 \quad \text{otherwise}$$

Given Your belief over $\mathbb{P}_\Omega$, can You build Your belief over Ω. In this paper, we will show how to build such a belief function.

Classical material about belief functions and the transferable belief model can be found in Shafer (1976), Smets (1988) and Smets and Kennes (1994).

2. IMPACT OF HACKING FREQUENCY PRINCIPLE.

The general frame consists of :
- Ω: the finite set of possible elementary events ω_i, i=1, 2...n, (the outcomes of the stochastic experiment),
- $\mathbb{P}_\Omega$: the set of probability functions P over Ω,
- $\mathbb{B}_{\mathbb{P}_\Omega}$: the set of belief functions over $\mathbb{P}_\Omega$.

Let N = {1, 2...n}. Let $W = \mathbb{P}_\Omega \times \Omega$. All subsets A of W can be represented as the finite union of the intersection of A with each of the elementary events ω_i:

$$\forall A \subseteq W, \qquad A = \bigcup_{i \in N} (A_i , \omega_i) \qquad (2.1)$$

where $A_i = proj(A \cap cyl(\omega_i)) \subseteq \mathbb{P}_\Omega$, cyl(X) is the cylindrical extension of X on W where X denotes a subset of Ω (or $\mathbb{P}_\Omega$), and proj(B) is the projection of $B \subseteq W$ on Ω (or $\mathbb{P}_\Omega$) (context makes it clear which domain and which range are involved).

The major problem solved in this paper is the construction of the belief function bel_W on W that would result if You were in a state of total ignorance about the value of P. If You have some prior belief $bel_{\mathbb{P}_\Omega}$ about the value of P, the belief bel_W over W would be combined with the vacuous extension of $bel_{\mathbb{P}_\Omega}$ on W by the application of Dempster's rule of combination. We will treat essentially the case where You only know that $P \in \mathcal{P}$ where $\mathcal{P}$ is a subset of $\mathbb{P}_\Omega$, i.e., when $\mathcal{P}$ is the only focal element and Your belief over $\mathbb{P}_\Omega$ can be represented by the basic belief assignment with $m_{\mathbb{P}_\Omega}(\mathcal{P}) = 1$. Generalization for a finite (or countable) numbers of focal elements is immediate. Further generalization is more delicate.

Let $\mathbb{B}_W$ be the set of belief functions over W. What is their nature? We are going to construct the equivalent of the basic belief masses (bbm) on $\mathbb{B}_W$. We say equivalent as W is not a finite space and the concept of basic belief masses has to be extended in order to cope with the structure of W. The bbm will become some sort of 'densities'. For simplicity sake, they are also denoted by $m_W : 2^W \to [0,1]$. The value $bel_W(A)$ is defined as the 'integral' of the m_W values given to the non empty subsets of A. It happens that in the case considered in this paper, m_W is a real density for which classical integrals are well defined. We call the m_W function a basic belief density (bbd) to enhance its particular nature. Those subsets A of W such that $m_W(A) > 0$ are called the focal elements of m_W.

The first constraint about m_W results from Hacking Frequency Principle. Suppose You know the values $P(\omega_i)$ of the objective probability function P on Ω for every $\omega_i \in \Omega$ (what is translated by $\mathcal{P} = \{P\}$). Let $bel_\Omega^{\{P\}}$ denotes Your belief over Ω when You know that $\mathcal{P} = \{P\}$.[2] By Hacking Frequency Principle, the value $bel_\Omega^{\{P\}}(X)$ for any subset X of Ω is numerically equal to the probability P(X) given to X.

By construction, $bel_\Omega^{\{P\}}$ results from the marginalizing of $bel_W^{\{P\}}$ over Ω:

$$bel_\Omega^{\{P\}}(X) = bel_W^{\{P\}}(cyl(X)) \qquad \forall X \subseteq \Omega$$

Hacking Frequency Principle implies the next requirement.

Requirement 1:
If You know that $\mathcal{P} = \{P\}$

$$\text{then } bel_W^{\{P\}}(cyl(X)) = P(X) \quad \forall X \subseteq \Omega \qquad (2.2)$$

Let $\pi(P) = cyl(\{P\}) \subseteq W$, then $\pi(P) = \bigcup_{i \in N} (\{P(\omega_i)\}, \omega_i)$ by 2.1. Thus $bel_W^{\{P\}}(cyl(X)) = bel_W(cyl(X) | \pi(P))$. The second term is just the result of the conditioning of bel_W on $cyl(\{P\})$, what is achieved by the application of Dempster's rule of conditioning. Hence the bbd $m_W(A)$, $A \subseteq W$ is transferred to $A \cap \pi(P)$. Let $A = \bigcup_{i \in N} (A_i, \omega_i) \subseteq W$, then m(A) is transferred to $A \cap \pi(P) = \bigcup_{i \in N} (A_i \cap \{P(\omega_i)\}, \omega_i)$.

---

[1] subscripts of m and bel denote their domain.

[2] Superscripts of bem and m denote Your knowledge about P, i.e., the focal element of $bel_{\mathbb{P}_\Omega}$.

The result of this conditioning on $\pi(P)$ is a probability function. Hence $bel_W^{\{P\}}$ must be a Bayesian belief function, i.e., only singletons can be focal elements and $bel_W^{\{P\}}(W) = 1$. The singletons of W have the form $(\{P(\omega_i)\}, \omega_i)$ for $i \in \{1,...n\}$, $P \in \mathbb{P}_\Omega$. Hence m_W must satisfy:

$$m_W(\bigcup_{i\in N}(A_i, \omega_i)) = 0 \quad \text{if } A_i \cap A_j \neq \emptyset \text{ for some } i \neq j \in N$$
$$= 0 \quad \text{if } \bigcup_{i\in N} A_i \neq \mathbb{P}_\Omega$$
$$\geq 0 \quad \text{otherwise.} \qquad (2.3)$$

The impact of Hacking Frequency Principle, translated by 2.3, is very strong. It implies that the focal elements of m_W can be represented as $\bigcup_{i\in N}(A_i, \omega_i)$ where the A_i, i=1,...n, are non empty elements of a partition of $\mathbb{P}_\Omega$.

3. THE CASE WHERE $|\Omega| = 2$.

We study now the case where $|\Omega| = 2$. Let $\Omega = \{S, F\}$ where S and F denote Success and Failure, respectively.

Figure 1: Structure of the domain of m_W and one example of bbd centered on a when $|\Omega| = 2$.

Let $A \subseteq W$ be a focal element of m_W, then $A = (\alpha, S) \cup (\overline{\alpha}, F)$ where $\alpha \subseteq [0, 1]$ and $\overline{\alpha}$ is the complement of α relative to [0, 1]. For simplicity sake, the bbd $m_W((\alpha, S) \cup (\overline{\alpha}, F))$ is written as $m_W(\alpha)$.

W can be graphically represented by two [0, 1] intervals where the upper [0, 1] interval is the intersection of W with S, and the lower one is the intersection of W with F (see figure 1). Every focal element of m_W is made of a set of mutually exclusive and exhaustive intervals that are either in the S domain or the F domain. By convention, intervals are defined as closed to the left and open to the right, except when 1 is the right limit, in which case the interval is also closed to the right.

We introduce an extra assumption.

Requirement 2:
If $|\Omega|=2$, $bel_{\mathbb{P}_\Omega}([a,b) \cup [c,d) \mid S) = bel_{\mathbb{P}_\Omega}([a,d) \mid S)$

for every $0 \leq a \leq b \leq c \leq d \leq 1$.

It is equivalent to assuming that $m_W(\alpha)$ is null except if $\alpha = [a, 1]$. The origin of the assumption is to be found in the meaning of the bbd. The bbd $m_W(\alpha)$ for $\alpha \subseteq [0, 1]$ is that part of belief (a density here) that supports the fact that $P(S) \in \alpha$ (and $P(F) \in \overline{\alpha}$). Suppose we condition m_W on S. Each bbd $m_W(\alpha)$ is transferred to $(\alpha, S) \subseteq W$. Requirement 2 means that if after conditioning on S a bbd supports $P(S) = x \in [0,1]$, it also supports every value in [0, 1] larger than x. Observing a success could support $P(S) = .3$, but that support should then also be given to $P(S) = .4$ etc....

This assumption means that each focal element is a step function that starts from $(\{0\}, F)$, jumps from the F domain to the S domain at some a in [0,1], and ends at $(\{1\}, S)$ (see figure 1).

Finally, if we apply again the Hacking Frequency Principle, we obtain after conditioning on $P = \{P\}$ with $P(S) = p$, $P(F) = 1-p$:

$$bel_\Omega^{\{P\}}(S) = p = \int_0^p m_W([x,1])\, dx.$$

The second equality results from the fact that only those bbd that jump before p will touch $(\{p\}, S)$ and $bel_\Omega^{\{P\}}(S)$ is equal to the integral of those bbd that touch $(\{p\}, S)$. Derivating both terms on p implies that:

$$m_W([p,1]) = 1 \qquad \forall p \in [0,1].$$

In conclusion we have derived the bbd on $_W$ when $|\Omega| = 2$.

Some properties can be easily derived.

1) Suppose the agent knows that $P = \{P: a \leq P(S) \leq b, 0 \leq a < b \leq 1\}$. We condition m_W on the cylindrical extension of [a, b]. The bbd $m_W(A)$ for $A \subseteq W$ is transferred to $A \cap cyl([a, b])$. $bel_W^P(S)$ is the integral of all the bbd that touch only S after conditioning on cyl([a, b]), i.e., those bbd that jump to S before a:

$$bel_W^P(S) = \int_0^a m_W([x,1])\, dx = a$$

Similarly $pl_W^P(S)$ is the integral of all the bbd that touch S, i.e., that jump to S before b:

$$pl_W^P(S) = \int_0^b m_W([x,1])\, dx = b$$

This result should not be extrapolated blindly to higher dimensions (see section 4).

2) The case $|\Omega| = 2$ can be nicely represented by figure 2 (Smets, 1978). Each point in the triangle corresponds to one interval of [0,1]. In general, if positive bbd are given

only to intervals, we assign the bbd given to [a,b] to the point (a,b) of the triangle. Then:

$$\text{bel}([a,b]) = \int_a^b \int_x^b m([x,y])\, dy\, dx$$

$$\text{pl}([a,b]) = \int_0^b \int_{a\vee x}^1 m([x,y])\, dy\, dx$$

$$q([a,b]) = \int_0^a \int_b^1 m([x,y])\, dy\, dx$$

$$m([a,b]) = -\frac{\partial^2 \text{bel}([a,b])}{\partial a\, \partial b} = -\frac{\partial^2 q([a,b])}{\partial a\, \partial b}$$

The result of the application of Dempster's rule of combination is given by multiplying the commonality functions.

Figure 2: Parametric representation on beliefs on [0,1] when the focal elements are intervals. The shaded areas are those on which integration is performed in order to compute bel([a,b] (a triangle) , q([a,b]) (a rectangle) and pl([a,b]) (a rectangle with right lower corner truncated).

In the present case ($|\Omega|=2$) the non-null bbd of m_W obtained after conditioning on S are given to the intervals [a,1], hence they cluster on the upper horizontal line. Those obtained after conditioning on F are given to the intervals [0,b], hence they cluster on the left vertical line.

Suppose You perform n independent experiments and observe r successes, s failures where $r + s \leq n$ (the difference n - (r + s) is the number of experiments for which the outcome is not available). The commonality function induced on $\mathbb{P}_\Omega = [0, 1]$

- by a success is: $q_{\mathbb{P}_\Omega}([a,b] \mid S) = a$
- by a failure is: $q_{\mathbb{P}_\Omega}([a,b] \mid F) = 1\text{-}b$
- by a 'S∪F' is: $q_{\mathbb{P}_\Omega}([a,b] \mid S\cup F) = 1$

The belief function induced by 'S∪F' is the vacuous belief function that reflect the state of total ignorance in which You are after just learning the tautology 'S∪F'. Hence we can just as well drop all 'vacuous' results and assume n = r+s.

The commonality function induced by r successes and s failures in n independent (Bernoullian) trials is obtained by multiplying the corresponding commonality functions. Hence:

$$q_{\mathbb{P}_\Omega}([a,b] \mid r, s) \approx a^r\, (1\text{-}b)^s$$

In that case, by derivating $q_{\mathbb{P}_\Omega}([a,b] \mid r, s)$ and appropriate normalization, we get:

$$m_{\mathbb{P}_\Omega}([a,b] \mid r, s) = \frac{\Gamma(r+s+1)}{\Gamma(r)\,\Gamma(s)}\, a^{r-1}\, (1\text{-}b)^{s-1}$$

where Γ is the gamma function.

When $n\to\infty$, $r\to np$, $s\to n(1\text{-}p)$ (hence $p = \lim \frac{r}{r+s}$), the limit of $m([a,b] \mid r, s)$ tends to 0 except for a dirac function at p. In that case $\text{bel}(A \mid r,s) = 1$ if $p\in A$ and 0 otherwise. After accumulating an infinite number of information, You will be in a state of 'total certainty', of 'knowledge' about the value of P(S).

3) Suppose You want to compute the belief that the next outcome is a success (or a failure) given You have already observed r successes and s failures in n independent trials. We use $m([a,b] \mid r, s)$ as the a priori belief over [0, 1]. Dempster's rule of combination $m_{12} = m_1 \oplus m_2$ can be represented as (Dubois and Prade, 1986, Smets, 1993a):

$$m_{12}(A) = \sum_{B\subseteq\Omega} m_1(A \mid B)\, m_2(B)$$

$$\text{bel}_{12}(A) = \sum_{B\subseteq\Omega} \text{bel}_1(A \mid B)\, m_2(B)$$

where $m_1(A \mid B)$ and $\text{bel}_1(A \mid B)$ are unnormalized conditional basic belief masses and belief functions. Generalizing this relation in the present context and denoting $\text{bel}_\Omega^{\{P:\, P(S)\in\{a,b]\}}(S)$ by $\text{bel}_\Omega(S \mid P(S)\in[a,b])$ (which value equals a), one obtains:

$$\text{bel}_\Omega(S|r,s) = \int_0^1 \int_a^1 \text{bel}_\Omega(S|P(S)\in[a,b])\, m_{\mathbb{P}_\Omega}([a,b]|r,s)\, db\, da$$

$$= \int_0^1 \int_a^1 a\, \frac{\Gamma(r+s+1)}{\Gamma(r)\,\Gamma(s)}\, a^{r-1}\, (1\text{-}b)^{s-1}\, db\, da$$

So: $$\text{bel}_\Omega(S \mid r, s) = \frac{r}{r+s+1}$$

$$\text{bel}_\Omega(F \mid r, s) = \frac{s}{r+s+1}$$

and $$m_\Omega(S\cup F \mid r,s) = \frac{1}{r+s+1}.$$

This result shows that the observed proportion is an excellent approximation of bel_Ω if r+s is not too small.

4. CASE WITH $|\Omega| = 3$.

Suppose $|\Omega| = 3$ where $\Omega = \{A, B, C\}$. $\mathbb{P}_\Omega$ can be represented by an equilateral triangle where each point corresponds to an element of $\mathbb{P}_\Omega$. The three heights are equal to the three probabilities P(A), P(B) and P(C). The height of such a triangle is 1 and the length of its side is equal to $\sqrt{4/3}$.

By requirement 1, we know that the focal elements of m_W can be represented by:

$$(P_A, A) \cup (P_B, B) \cup (P_C, C)$$

where (P_A, P_B, P_C) are the elements of a partition of $\mathbb{P}_\Omega$.

In order to specify the form of the subsets P_X, $X \in \{A, B, C\}$, we consider the conditioning of m_W on the set $P_L \subseteq \mathbb{P}_\Omega$ where

$$P_L = \{P: P = (p_A, p_B, p_C): p_B = b_0 + \frac{1 - a_1 - b_0}{a_1} p_A, \quad p_C = 1 - p_A - p_B\}. \qquad (4.1)$$

where $b_0, a_1 \in [0, 1]$, $b_0 < 1-a_1$.

This set P_L corresponds to the subset of $\mathbb{P}_\Omega$ where $P(A) \in [0, a_1]$ and P(B) and P(C) are linearly related to P(A). Requirement 3 states that, after conditioning m_W on P_L, the bbd so obtained on the space P_L is identical to those obtained when $|\Omega|=2$ (indeed every element of the new subdomain is characterized by P(A) as when $|\Omega|=2$). Therefore after further conditioning on A, the focal elements on P_L should be of the form of intervals $[a, a_1]$ (see figure 3). This requirement is sufficient in order to derive the structure of the focal elements of m_W.

Requirement 3:
If $|\Omega| = 3$, for every P_A, there exists an $\alpha \in [0, a_1]$ such that the projection of P_A on P_L is the interval $[\alpha, a_1]$.

Requirement 3 can identically be defined as:

Requirement 3':
If $|\Omega| = 3$, $m_W(X \mid P_L) \geq 0$ if it exists an $\alpha \in [0, a_1]$ and $X = \{P: P \in P_L, \ p_A \geq \alpha\}$.
$= 0$ otherwise.

Each requirement implies that the limits between P_A and P_B must be a straight line passing through the corner where P(C) = 1 and that crosses the opposite side of the triangle (and similarly for the other limits). Every focal elements of m_W can be labeled by an element $q = (q_A, q_B, q_C) \in \mathbb{P}_\Omega$. The focal element labeled by q is the set

$$P(q) = (P_A(q), A) \cup (P_B(q), B) \cup (P_C(q), C) \qquad (4.2)$$

where:

$$P_A(q) = \{P: P = (p_A, p_B, p_C) \in \mathbb{P}_\Omega, \quad p_A \geq \max(p_B \frac{q_A}{q_B}, p_C \frac{q_A}{q_C})$$

and similarly for $P_B(q)$ and $P_C(q)$ where the A,B,C-indexes are symmetrically exchanged. The graphical representation of $P_A(q)$ is the upper corner of the $\mathbb{P}_\Omega$ triangle that includes all points in $\mathbb{P}_\Omega$ between the upper corner and the two straight lines drawn from the two other corners through q. $bel_W(X \mid A)$ for $X \subseteq \mathbb{P}_\Omega$ will be the 'integral' of all the bbd given to the focal elements $P(q)$ such that $X \subseteq P_A(q)$.

Figure 4 shows the structure of the partition so generated.

The value of $bel_\Omega(A \mid P = \{(a, b, c)\})$ is the 'integral' of m_W taken over all q in the triangle which corners are (0, 0, 1), (0, 1, 0) and (a, b, c). Reapplying the Hacking Frequency Principle we have:

$$bel_\Omega(A \mid P = \{(a, b, c)\}) = a \qquad (4.3)$$

Figure 3: Explanation of Requirement 3. Left figure: the $\mathbb{P}_\Omega$ space with $|\Omega| = 3$, and the P_L domain. Middle figure: a bbd that satisfies Requirement 3. Right figure: a bbd that does not satisfy Requirement 3

It can then be proved that the only function m_W symmetric in the three arguments of q that satisfies (4.3) for every $(a, b, c) \in \mathbb{P}_\Omega$ is the function $m_W(P(q)) = \sqrt{3}$ for every $q \in \mathbb{P}_\Omega$.

Figure 4: Structure of the domain of m_W and one example of bbd labeled q when $|\Omega| = 3$. The shaded area is $P_A(q)$.

Some properties derived from this solution are detailed.

1) If P = {(.5, 0, .5), (.5, .5, 0)}, then $m_\Omega(A) = 1/3$, $m_\Omega(B) = m_\Omega(C) = 0$, $m_\Omega(A\cup B) = m_\Omega(A\cup C) = 1/6$, $m_\Omega(B\cup C) = 1/3$, $m_\Omega(A\cup B\cup C) = 0$. This result merits some reflection. One might be surprised that even though P(A) = .5 is exactly known, one does not have $bel_\Omega(A) = .5$. If the frame had been A versus $\overline{A}$, the critic would have been appropriate, except that in such a frame we just have the required results. The difference observed here reflects the fact that there are three elements. What is nice is that the pignistic probability induced in this case is such that BetP(A) = .5 (the pignistic transformation is detailed in next section).

2) If P = {(.5, b, c): b+c = .5}, then $m_\Omega(A) = 1/3$, $m_\Omega(B) = m_\Omega(C) = 0$, $m_\Omega(A\cup B) = m_\Omega(A\cup C) = 1/6$, $m_\Omega(B\cup C) = 1/4$, $m_\Omega(A\cup B\cup C) = 1/12$. The same remarks hold as for the case 1, but BetP(A) = .5 as it should.

3) If P = {(a, b, c): a≤.5, b≤.5, c≤.5}, then $m_\Omega(A) = m_\Omega(B) = m_\Omega(C) = 0$, $m_\Omega(A\cup B) = m_\Omega(A\cup C) = m_\Omega(B\cup C) = .25$, $m_\Omega(A\cup B\cup C) = .25$.

4) Suppose You know that P is characterized by a lower probability function P_* on Ω. Let P* be the upper probability function dual of P_*, i.e., $P^*(X) = 1 - P_*(\overline{X})$ for $X\subseteq\Omega$. Let $a = P_*(A)$, $b = P_*(B)$, $c = P_*(C)$, $A = P^*(A)$, $B = P^*(B)$, $C = P^*(C)$. The belief on Ω induced by the set P of probability distributions P on Ω compatible with the upper and lower probabilities (i.e., $\forall A\subseteq\Omega$, $P_*(A)\le P(A)\le P^*(A)$) is given by:

Figure 5: Domain of PΩ when $|\Omega| = 3$. The hexagon represents the set P of probability functions compatible with a given lower probability function. The values of $m_\Omega(A \mid P)$ and $m_\Omega(A\cup B \mid P)$ are the shaded surfaces. $m_\Omega(A\cup B\cup C \mid P)$ is the surface of the hexagon plus the three small left over triangles fixed on its side.

$$bel_\Omega(A) = \frac{a}{a+B+C}$$
$$bel_\Omega(B) = \frac{b}{A+b+C}$$
$$bel_\Omega(C) = \frac{c}{A+B+c}$$
$$bel_\Omega(A\cup B) = (1-C)^2 + C\,(a+b)$$
$$bel_\Omega(A\cup C) = (1-B)^2 + B\,(a+c)$$
$$bel_\Omega(B\cup C) = (1-A)^2 + A\,(b+c)$$
$$bel_\Omega(A\cup B\cup C) = 1$$

These results are obtained by computing the various surfaces described in figure 5.

It is worth noticing that bel is not equal to P_*, even when P_* is a belief function. Why should they? The transferable belief model never requires that the belief function that quantities our belief should be the lower envelop of a set of probability function.

5. PIGNISTIC PROBABILITY.

In Smets (1990a, 1993b) and Smets and Kennes (1994), we have shown how to build the appropriate probability function BetP, called the pignistic probability function, from a belief function when a decision must be made. We have shown that the only 'rational' transformation, called the pignistic transformation, must satisfy the following rule when the betting frame Ω is finite. Let m be the basic

belief assignment quantifying the agent's beliefs over Ω. For ω∈Ω,

$$\mathrm{BetP}(\omega) = \sum_{A:\omega\in A\subseteq\Omega} \frac{1}{|A|}\, m(A)$$

where |A| is the number of elements of Ω in A. Any other probability function would lead to irrationality in the betting behavior of the agent. Its extension to continuous cases is easy to realize if the bbd are really densities, in which case sums become classical integrals.

We study how decision should be made when beliefs are induced by a set of probabilities, i.e., how to derive the appropriate pignistic probability from the initial belief induced by the knowledge that $P\in \mathcal{P}\subseteq\mathbb{P}_\Omega$. The choice of the appropriate betting frame is important. We could think to build the belief function over Ω that quantifies our belief over Ω and apply the pignistic transformation to such a belief function over Ω using Ω as the betting frame. But this is an erroneous strategy as the betting frame is not Ω but W. The beliefs induced by $P\in\mathcal{P}$ is a belief over W, the belief derived on Ω is only the result of the marginalization of the first one on Ω.

Using W as the betting frame, we apply the pignistic transformation to Bel_W. For $X\subseteq\mathbb{P}_\Omega$ let S(X) be the surface of X. Suppose the agent who wants to bet on Ω knows only that $P\in\mathcal{P}$. The pignistic transformation implies that the bbd $m_W(\mathcal{P}(q))$ given to $\mathcal{P}(q)$ (see (4.2)) be equally distributed among the elements of $\mathcal{P}$:

$$\mathrm{BetP}_\Omega(A) = \int_{q\in\mathbb{P}_\Omega} \frac{S(\mathcal{P}_A(q)\cap\mathcal{P})}{S(\mathcal{P})}\, dq.$$

Interchanging the order of integration, one gets that

$$\mathrm{BetP}_\Omega(A) = \frac{1}{S(\mathcal{P})} \int_{P\in\mathcal{P}} \int_{q\in\mathbb{P}_\Omega} I(\mathcal{P}_A(q)\cap\{P\})\, dq\, dP$$

where $I(X) = \begin{cases} 1 & \text{if } X\neq\emptyset, \\ 0 & \text{otherwise.}\end{cases}$

One has: $\int_{q\in\mathbb{P}_\Omega} I(\mathcal{P}_A(q)\cap\{P\})\, dq = P(A)$,

hence:

$$\mathrm{BetP}_\Omega(A) = \frac{1}{S(\mathcal{P})} \int_{P\in\mathcal{P}} P(A)\, dP.$$

The pignistic probability $\mathrm{BetP}_\Omega(A)$ so derived is equivalent to the probability one would derive by assuming an equi a priori density over $\mathbb{P}_\Omega$, conditioning it on $\mathcal{P}$, and computing the expected probability of P(A).

In particular, when $|\Omega| = 2$ and $\mathcal{P} = [a, b]$, the result is:

$$\mathrm{BetP}(S) = \frac{1 + a - b}{2}.$$

These are quite natural results. $\mathrm{BetP}_\Omega(. \mid \mathcal{P})$ indeed happens to be the center of gravity of $\mathcal{P}$, but its derivation does not result from the use of an equi a priori density over $\mathbb{P}_\Omega$. It just happens that both approaches lead to the same results: 1) the equi a priori density over $\mathbb{P}_\Omega$ and 2) the application of the pignistic transformation combined with the evaluation of $\mathrm{BetP}_\Omega(X \mid \mathcal{P})$ as $\mathrm{BetP}_W(\mathrm{cyl}(X) \mid \mathcal{P})$ for $X\subseteq\Omega$, where BetP_W is the pignistic probability obtained from $\mathrm{bel}_W(. \mid \mathcal{P})$ over the betting frame W.

6. CONCLUSIONS:

1) Generalization to $|\Omega| > 3$ is conceptually easy, but very laborious when solutions must be written down. Nothing new comes out of it. In practice, computation will not been based on the explicit equations, but on some Monte Carlo method.

2) Generalization of the procedure can be achieved if one has a non-degenerated belief function on $\mathbb{P}_\Omega$ if there are only a finite number of subsets of $\mathbb{P}_\Omega$ that receive positive basic belief masses (more general cases are not considered here). Let $\{\mathcal{P}_i: i= 1, 2...n\}$ be the set of focal elements of $\mathrm{bel}_{\mathbb{P}_\Omega}$ with their basic belief masses $m_{\mathbb{P}_\Omega}(\mathcal{P}_i)$. For each focal element $\mathcal{P}_i$, we derive $\mathrm{bel}_W(. \mid P\in\mathcal{P}_i)$ over W. The belief function bel_W over Ω induced $\{(\mathcal{P}_i, m_{\mathbb{P}_\Omega}(\mathcal{P}_i)): i= 1, 2...n\}$ is:

$$\forall A\subseteq W \qquad \mathrm{bel}_W(A) = \sum_{i=1}^{n} \mathrm{bel}_W(A \mid P\in\mathcal{P}_i)\, m_{\mathbb{P}_\Omega}(\mathcal{P}_i)$$

3) Suppose two pieces of evidence that say that $P\in\mathcal{P}_1$ and $P\in\mathcal{P}_2$, respectively. The combination of these two pieces of evidence leads to the knowledge $P\in\mathcal{P}_1\cap\mathcal{P}_2$.

One could build bel_1 on W as the belief function induced by the knowledge that $P\in\mathcal{P}_1$. Identically, one could build bel_2 on W as the belief function induced by the knowledge that $P\in\mathcal{P}_2$. One could then be tempted, erroneously in fact, to combine bel_1 and bel_2 into $\mathrm{bel}_1\oplus\mathrm{bel}_2$ by Dempster's rule of combination.

One could also build bel_{12} on W as the belief function induced by the knowledge that $P\in\mathcal{P}_1\cap\mathcal{P}_2$. In general $\mathrm{bel}_{12} \neq \mathrm{bel}_1\oplus\mathrm{bel}_2$. Only bel_{12} is correct. Indeed Dempster's rule of combination is applicable iff both pieces of evidence are distinct, and distinctness is not satisfied in the present context because of the existence of a unique underlying probability function on Ω that create a link between the two pieces of evidence.

4) In conclusion, the knowledge that the probability function P over Ω belongs to some subset $\mathcal{P}$ of $\mathbb{P}_\Omega$ permits the construction of a belief function bel over $\mathbb{P}_\Omega$ x Ω and over Ω. It must be enhanced that in general the belief function bel_Ω induced over Ω by a lower probability function P_* will not satisfy $\mathrm{bel}_\Omega = P_*$ even if P_* happens to be a belief function. By showing what is

the belief induced by a lower probability, we hope we have been able to show the fundamental difference between the upper and lower probabilities model and the transferable belief model (see also Smets, 1987, Smets and Kennes, 1994, Halpern and Fagin, 1990).

Acknowledgment: Research work has been partly supported by the Action de Recherches Concertées BELON funded by a grant from the Communauté Française de Belgique and the ESPRIT III, Basic research Action 6156 (DRUMS II) funded by a grant from the Commission of the European Communities.

Bibliography:

DUBOIS D. and PRADE H. (1986) On the unicity of Dempster rule of combination. Int. J. Intelligent Systems, 1:133-142.
HACKING I. (1965) Logic of statistical inference. Cambridge University Press, Cambridge, U.K.
HALPERN J.Y. and FAGIN R. (1990) Two views of beleif: beleif as generalized probability and belief as evidence. submitted for publication.
KOHLAS J. and MONNEY P.A (1994) Representation of Evidence by Hints. in Advances in the Dempster-Shafer Theory of Evidence. Yager R.R., Kacprzyk J. and Fedrizzi M., eds, Wiley, New York, pg. 473-492.
SHAFER G. (1976) A mathematical theory of evidence. Princeton Univ. Press. Princeton, NJ.
SHAFER G. (1990) Perspectives in the theory and practice of belief functions. Intern. J. Approx. Reasoning, 4:323-362.
SMETS Ph. (1978) Un modèle mathématico-statistique simulant le processus du diagnostic médical. Doctoral dissertation, Université Libre de Bruxelles, Bruxelles, (Available through University Microfilm International, 30-32 Mortimer Street, London W1N 7RA, thesis 80-70,003)
SMETS P. (1987) Upper and lower probability functions versus belief functions. Proc. International Symposium on Fuzzy Systems and Knowledge Engineering, Guangzhou, China, July 10-16, pg 17-21.
SMETS Ph. (1988) Belief functions. in SMETS Ph, MAMDANI A., DUBOIS D. and PRADE H. ed. Non standard logics for automated reasoning. Academic Press, London p 253-286.
SMETS Ph. (1990a) Construucting the pignistic probability function in a context of uncertainty. Uncertainty in Artificial Intelligence 5, Henrion M., Shachter R.D., Kanal L.N. and Lemmer J.F. eds, North Holland, Amsterdam, , 29-40.
SMETS Ph. (1990b) The combination of evidence in the transferable belief model. IEEE-Pattern analysis and Machine Intelligence, 12:447-458.
SMETS P. (1993a) Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem. Int. J. Approximate Reasoning 9:1-35.
SMETS P. (1993b) An axiomatic justifiaction for the use of belief function to quantify beliefs. IJCAI'93 (Inter. Joint Conf. on AI), San Mateo, Ca, pg. 598-603.
SMETS Ph. and KENNES (1994) The transferable belief model. Artificial Intelligence, 66:191-234.