# Resolving misunderstandings about belief functions.

## In Response to J. Pearl's Criticisms in "Reasoning with Belief Functions: an Analysis of Compatibility"

Philippe SMETS[1]
IRIDIA-Université Libre de Bruxelles
50 av. F. Roosevelt, 1050 Brussels, Belgium

Judea Pearl has made an excellent presentation of many of the errors that pervade the belief function (BF) literature, some being put forward, even though Pearl knows their answer, for the mere purpose of provoking reactions. So it is a good opportunity for me to clarify the situation. I shall start with a summary of my interpretation of the use of BF to quantify someone's belief (called the transferable belief model) and follow on with a systematic refutation of Pearl's criticisms. Sections and relations numbers are preceded by a P or an S referring to Pearl's or my own presentation, respectively.

## S.1. The Transferable Belief Model (TBM)

The transferable belief model (Smets 1988) is my interpretation of Dempster-Shafer's model. It fits in essentially with Shafer's initial presentation (Shafer 1976) except inasmuch as it rejects explicitly all connections with any probabilistic model. It is based on the following assumptions:

1) our degree of belief is quantified by a number between 0 and 1
2) there exists a **two-level structure**:
    - a **credal** level where beliefs are entertained and
    - a **pignistic** level where beliefs are used to make decisions
3) beliefs at the credal level are quantified by **belief functions**.
4) beliefs at the pignistic level are quantified by **probability functions**.
5) the credal level precedes the pignistic level in that at any time beliefs are entertained (and updated) at the credal level, the pignistic level appearing only when a decision must be made.
6) when a **decision** must be made, beliefs at the credal level are transformed into beliefs at the pignistic level, i.e. there exists a transformation, called the **pignistic transformation**, from belief functions to probability functions. (Smets 1989)

---

[1] The following text presents research results of the Belgian National incentive-program for fundamental research in artificial intelligence initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. Scientific responsibility for this text is assumed by its author. Research has been partly supported by the DRUMS projects which is funded by the European Community as part of the ESPRIT Basic Research Project 3085.

To describe the TBM, one starts with $\Omega$, a non-empty finite set called the **frame of discernment.** $\Omega$ is equipped with a Boolean algebra $\Re$ of some of its subsets. Every element of $\Re$ is called a proposition. The singletons of $\Omega$ are called the elementary propositions. At most, one elementary proposition of $\Omega$ is "true". The singletons of $\Omega$ can be seen as possible answers to a certain question, at most one answer being correct. A proposition of $\Re$ is true if it contains the element of $\Omega$ that characterizes the true singleton of $\Omega$.

A **credal state** on $\Omega$ is a description of our subjective, personal judgment that the propositions $A\in\Re$ are true. It results from known, possibly inconclusive pieces of evidence that induce partial beliefs on the propositions of $\Re$. It is an epistemic construct as it is relative to our knowledge.

The **transferable belief model** postulates that the impact of a piece of evidence consists in allocating parts of an initial unitary amount of belief among the propositions of $\Re$. For $A\in\Re$, m(A) is a part of our belief that supports A, i.e. that the 'truth' is in A, and that, due to lack of information, does not support any strict subproposition of A. The m's are called the **basic belief masses** (bbm).

If further evidence becomes available and implies that the proposition $\neg$B is false, then the mass m(A) initially allocated to A is transferred to A&B. Hence, the name of transferable belief model. This transfer of belief corresponds to **Dempster's rule of conditioning** (except for the normalization factor).

## S.2. Probabilities: to be or not to be.

An important but often neglected distinction must be made between the following two types of **probabilistic reasoning** : (P.1.3)

*Type 1.* **There exists a probability function** P: $\Re\rightarrow[0, 1]$ where $\Re$ is the finite boolean algebra of propositions on which we rest our beliefs. But the values of P(A) for $A\in\Re$ are only known to belong to some intervals whose limits are called the upper ($P^*(A)$) and lower ($P_*(A)$) probabilities. For $\forall A\in\Re$, one has:

$$P_*(A) \leq P(A) \leq P^*(A) \qquad \text{(S1)}$$

This is essentially the **upper and lower probabilities** (ULP) model.

**Dempster** (1967) created a similar model. He postulated the existence of a probability function $P_X$ on some space X and a one-to-many mapping M from X to Y. The probabilities $P_Y$ induced on Y by $P_X$ through M are such that the lower (upper) probabilities function is a belief (plausibility) function. Up to here, Dempster's model (which I call the **PXMY model**) is merely a particular form of ULP model[2].

In both the ULP and the PXMY contexts, the interval $[P_Y^*(A), P_{Y*}(A)]$ and its length for $A \subseteq Y$ reflects the imprecision of our knowledge of $P_Y(A)$.

*Type 2.* **No probability function is postulated to exist** on $\Re$. The degree of belief given to $A \in \Re$, at the credal level, is quantified by a point-valued function bel(A). Neither additivity nor Cox's axioms are postulated. The TBM postulates the primitive concept of "part of belief", the basic belief masses (bbm)[3]. Such a model corresponds closely to the one described in Shafer's book.

But in more recent publications, Shafer became less radical. He accepts the existence of some underlying probability measure, but of a very different nature than that of the ULP models.[4] He claims that the **meaning** of a piece of evidence is random. He speaks about the probability that the evidence means A, not the probability that A is true. His model is essentially a PXMY model (see his translator paradigm, Shafer and Tversky 1986)

In type 2 context where underlying probability functions are not postulated, equations like:

$$\mathrm{bel}(A) \leq P(A) \leq \mathrm{pl}(A) \qquad \text{(S2)}$$

are highly misleading. It is true that given a bel and its associated plausibility function pl, there exists a set of compatible probability functions P that satisfy (S2), but one must resist the temptation to give any interpretation to those probability functions other than to be mathematical objects without relevance to our quantified belief. Only in type 1 context does the P function in (S2) have a perfectly well-defined relation with our belief.

Pearl's definition of **incomplete information** (P.2 §1) is really unusual. By incomplete knowledge, I mean knowledge that is insufficient to reduce the set of possible worlds into a single element. Should I know "everything", I would know in which world I am. The lack of complete knowledge is responsible for the fact that I can only express my beliefs about which world prevails. Wether that belief should be quantified by a probability function or a belief function is obviously THE crucial question.

Strat's (1989) statement quoted by Pearl is indeed unfortunate as it leads one to believe that BF should be used for "**unavailable probabilities**". If "unavailable" means that

---

[2] Mathematically lower probabilities functions are capacities of order 2. In the PXMY model the induced lower probabilities functions are capacities of order infinite; see Choquet (1953).

[3] Note that undefined concepts appear also in plain probability theory: the word "probability" itself is also a primitive concept !

[4] Shafer's ideas - but not the solution - are similar to those defended in the Evidentiary Value Model (Gardenfors et al. 1983)

"it exists but we don't know its value", then ULP are to be used. If "unavailable" means "it doesn't exist", then Strat is correct. Pearl's discussion and comments result from the inappropriate use of BF for typical ULP problems, a confusion unfortunately shared by many users of BF.

The lesson to be derived from this distinction is that the TBM is not a model for **poorly known probabilities**. For this case the appropriate model is the ULP model - maybe under the PXMY form, if one can explain the origin of its components. The TBM is a model that uses whatever is available, not a model to be used when some probabilities are missing. The reason for this formulation is that people will usually interpret the idea of "missing probabilities" not as *absent* probabilities but as *existing* probabilities whose values are missing. The correct statement should read more like: "the TBM is a model in which we are ignorant of the existence of probabilities" (whereas the ULP models are models in which we know of the existence of probabilities but are ignorant of their values).

## S.3. The static and dynamic components.

To study any model for quantified belief, one must consider not only its **static component** (how are beliefs allocated?) but also its **dynamic component** (how are beliefs updated?). In the TBM, the static component corresponds to the basic belief masses allocation and the dynamic component to
1) the **transfer** of those basic belief masses among the propositions and
2) the **combination** of the beliefs induced by several pieces of evidence.

It must be emphasised that any **comparison** of the TBM with others models must consider both components. Far too often, authors concentrate on the static component and discover many relations between the TBM and upper and lower probabilities (ULP) models, inner and outer measures (Fagin and Halpern 1989), random sets (Nguyen 1978), probabilities of provability (Pearl 1988) (P.1.4.§2), probabilities of necessity (Ruspini 1986)) etc. But these authors usually do not explain or justify the dynamic component, i.e. how updating (conditioning) is to be handled (except in some cases by defining conditioning as a special case of combination). So I feel these partial comparisons are incomplete, especially since all these interpretations lead to different updating rules.

### Updating - Combination: which comes first?.

**Shafer** presents his model by introducing successively 1) the static component (the belief functions), 2) the combination process (Dempster's rule of combination) and 3) the updating process (Dempster's rule of conditioning) as a particular case of the combination process. We present the **TBM** by introducing successively 1) the static component (the basic belief masses) 2) the updating process (the mass transfer) and 3) the combination process that we derive from the previous processes by requiring essentially the compositionality of the combination (i.e. $bel_1 \oplus bel_2$ is only a function of $bel_1$ and $bel_2$) (Smets 1990a). I think this order of presentation of the TBM is more natural.

Updating by a true fact (P.1.3) is the most fundamental dynamic component. Later, one introduces the concept of updating by uncertain facts (as with Jeffrey's rule of conditioning, Jeffrey (1983), Shafer (1980)) and finally the concept of combining symmetrically two BF induced by two "distinct" pieces of evidence. That the two updating processes turn out to be mathematically special forms of Dempster's rule of combination is not required by the model.

To show why I feel the updating process precedes the combination process, let us examine the equivalent processes within the **probabilistic framework**. Remember that if both BF $bel_1$ and $bel_2$ are probability functions $P_1$ and $P_2$, then Dempster's rule of combination can be reduced to $m_{12}(x) = P_{12}(x) = \alpha\, P_1(x)\, P_2(x)$ for $x \in \Omega$ if 1) there is an equi a priori probability on $\Omega$ and 2) the two pieces of evidence that induced $P_1$ and $P_2$ are conditionally independent for each $x \in \Omega$. If one tries to define the conditioning as a special case of combination, the conditional probability $P(x|X)$ is no longer defined as $P(x)/P(X)$ for $x \in X$. It is defined as the result of the combination (with the equi prior and the conditional independence being postulated) of the probability function P with a probability function P' such that $P'(x) = 1/|X|$ for $x \in X$ and 0 otherwise. Can such a new definition of the conditional probability be regarded as acceptable? It is mathematicaly correct but looks faintly surrealistic. To claim that conditioning with BF is a special form of combination is one and the same thing and that is why I am not in favor of that kind of an approach.

## S.4. What is TBM not ?

The TBM is not a model based on **random sets** or on **upper and lower probabilities**. They share the same static components, not the dynamic ones. With random sets, conditioning should be performed by the application of the geometric rules (Suppes and Zanotti, 1977), except in very twisted cases (Smets 1990b). With upper and lower probabilities, the G-rule[5] is usually appropriate.

A very important element of the TBM is its total disconnection with any concept of **randomness** or **additive probabilities**. Whenever randomness (or decision) is involved, additive probabilities are of course perfectly appropriate, but this context is not

[5] G-rule of conditioning is what Pearl calls FH conditioning (see P.3.3, equation P19). Its description can be found in Dempster (1967). I called it the G-rule as I traced the work on upper and lower probabilities to Good's papers. (see also Planchet, 1989)

the one covered by the TBM. In the TBM context (belief entertained outside any context of randomness or decision process), the use of additive probabilities to quantify our beliefs can be justified by axioms like those of Cox (1946). Cox's axiom 1 states that the belief of the negation of a proposition should be a function of the belief of the proposition

$$\text{belief}(\neg A) = f(\text{belief}(A))$$

Cox's axiom is not postulated in the TBM… and therefore the additive probabilities that derive from it are not required in the TBM.

I also dissociate the TBM from the **PXMY** model because Dempster's rule of conditioning can be criticized in the latter case. Why indeed should one apply Dempster's rule of conditioning in a PXMY context? (Levi 1983). When the truth is known to be mapped into a subset Z of Y, Dempster adapts the M mapping accordingly but does not update the initial probability distribution P on X by taking the new fact Z into account. I developed the TBM as a model totally unlinked to any underlying probability model just to avoid any such criticism. When I build a BF on some space Y, I never postulate the existence of some space X endowed with a probability function and of a one-to-many mappings from X to Y.

Beware nevertheless that the Bayesian analysis is always a particular case of the TBM analysis (P.5 §1). The conclusions reached by a Bayesian analysis require more information than those reached by a TBM analysis. The fact is that in order to perform the Bayesian analysis, many probabilities must be specified on top of those used by the TBM analysis and of course these **extra specifications** lead to results different from those reached by a TBM analysis that does not use them. If these extra specifications had also been introduced in the TBM analysis, both analyses would have led to the same results. In fact, all results derived from a Bayesian analysis can also be obtained from a TBM analysis if the same information is used in both analyses.

## S.5. Open or closed worlds?

In Smets (1988) I explain the difference between the open and closed world assumptions. Reconsider the definition of Ω in S.1. We postulate either that **at most one** elementary proposition of Ω is true (open world assumption) or that **one and only one** elementary proposition of Ω is true (closed world assumption). The **normalization** results from the closed world assumption. Shafer postulates the closed world assumption. We generalize the problem, accept that both contexts could exist and select the appropriate assumption according to the problem under scrutiny. (P.1.3.3)

Pearl could easily have generalized his description of Dempster's rule of combination in P.1.4. by dropping the statement "given that the two pieces of evidence are non-contradictory" (§3) or "to condition beliefs on the evidence being non contradictory" (§4). Distinguishing between the open and closed world contexts would have solved the "counter examples" in P.1.4 §6 and §7 for which the paradoxical results are due to the normalization factor. His criticisms on the adequacy of Dempster's rule of conditioning concerns the normalization problem more than the rule itself. Under an open-world assumption, there is no normalization when applying both Dempster's rules.

## S.6. Belief about unknown probabilities.

The three examples in P.2.1 and the remarks in P.2.2 §5 are irrelevant to BF, they are ULP problems. They correspond to the type 1 context as explained in S.2. To grasp the origin of the confusion, let us analyse example 1, the other two being solved identically.

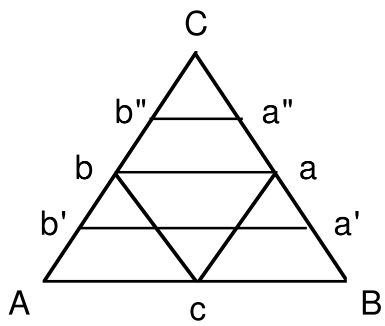


**Figure 1:** Simplex ABC.

In **example 1** of P.2.1, $\Omega$ is the simplex ABC (figure 1) characterized by all three elements vectors $(x_1, x_2, x_3)$ where $x_i \geq 0$ and $\sum_{i=1}^{3} x_i = 1$. ($x_i$ represents the unknown $P(E_i)$). The three pieces of evidence are $x_i \in [0, .5]$, i=1,2,3. For instance the first piece of evidence induces a belief function with basic belief masses $m_1(AbaB) = 1$ (see figure 1). The second and third pieces of evidence induce the bbm $m_2(CacA) = 1$ and $m_3(BcbC) = 1$. Their combination by Dempster's rule of combination leads to $m_{123}(abc) = 1$. So any point in the triangle abc is a solution to the problem raised in example 1. Similar results are derived if one considers the three inequalities of equation P7 as a single piece of evidence that just says the solution is in the triangle abc.

The object language of the probabilities deals with the events $E_i$. The object language of our belief deals with the vectors of probabilities (points in the simplex ABC). The belief statements do not concern the events $E_i$. An open question is how to build your beliefs on $E_i$ given your beliefs on their prtobabilities. The metalanguage (our belief statements) expresses our knowledge about which "object" (vector of probabilities) is true. Example 1 (and the other two) confuses the two levels.

The introduction of bbm $m_i$ is of course not very useful in these examples. So in order to show what the difference between the object language and the metalanguage really is, I shall generalize the problem raised in example 1 by considering e.g. that the bbm $m_1$ are:
$m_1(Ab'a'B) = .3$ $m_1(AbaB) = .4$ $m_1(Ab''a''B) = .2$ $m_1(ABC) = .1$
which translates e.g.

I believe at level .3 that $P(E_1) \leq .25$
I believe at level .7 that $P(E_1) \leq .50$
I believe at level .9 that $P(E_1) \leq .75$
I believe at level .6 that $.25 \leq P(E_1) \leq .75$ …

If similar bbm had been obtained for $m_2$ and $m_3$, then the full strength of the TBM could be realized. Many errors, like those of Pearl's three examples, derive from the fact that each BF is characterized by one focal element and the expression $bel(P(E_1) \leq .5) = 1$ is translated into $bel(E_1) = .5$. A belief about a probability of an event is not identical to a belief about the event. In these examples, the BF are analogous to the meta-probabilities (probabilities of order 2)

As to the remark after equation P8 in P.2.1, I agree that the function defined by P7 is not a BF, but then, I never claimed that lower probability functions are belief functions. They are unrelated. $P_*$ expresses one thing, bel another.

## S.7. The unspoiling of the sandwich.

The sandwich principle raised in P.2.2 example 4 and P.3.2 can be stated as

$$bel(A) \in [\, bel(A|B) , bel(A|\neg B) \,] \qquad (S3)$$

It is an important issue that must be solved as S3 is not required with the TBM. I shall present an example in which S3 seems not to be required, proceed to explain the relation between bel(A), bel(A|B) and bel(A|¬B) as deduced from the Generalized Bayes Theorem, then why S3 is not required within the TBM, and, finally, answer a few specific questions raised by Pearl.

### S.7.1. Examples against the sandwich principle.

Suppose a given person X that can be either male (M) or female (F), young (Y) or old. I try to assess my belief that X is young.
Consider the following pieces of evidence.

$E_1$: A not fully reliable witness $W_1$ claims he saw X and says: "X is not an old man". By 'not fully reliable' I mean that maybe $W_1$ saw X and is telling the truth about X, or maybe he did not see X and saw somebody else.

$E_2$: Another not fully reliable witness $W_2$ says : "X is not an old woman".

M: I know that X is a man

F: I know that X is a woman

For simplicity's sake, let us accept that the reliabilities of both witnesses are the same and our prior belief about X's age is independent of our knowledge about X's sex.

How would I compare $bel_{12}(Y)$ and $bel_{1M}(Y)$ where :

$bel_{12}(Y)$ is the belief that X is young given $E_1$ and $E_2$.

$bel_{1M}(Y)$ is the belief that X is young given $E_1$ and M.

Which of the following relations is more natural?

$$bel_{12}(Y) = bel_{1M}(Y) \qquad \text{(S4)}$$

$$bel_{12}(Y) < bel_{1M}(Y) \qquad \text{(S5)}$$

$bel_{1M}(Y)$ reflects the fact that X is male (known for sure) and $W_1$ was reliable when he said that "X is not an old man". So $bel_{1M}(Y)$ is related to the reliability given to $W_1$. $bel_{12}(Y)$ results from the combination of the pieces of evidence provided by the two witnesses. If both witnesses are reliable, then X is young. Otherwise I do not know. So $bel_{12}(Y)$ is related to the fact that both $W_1$ and $W_2$ are reliable.

I personally feel that (S4) is not required, and (S5) is more appropriate. But once (S5) is accepted, it is proved below that:

$$bel_{12}(Y) < \min ( bel_{12M}(Y) , bel_{12F}(Y) ) \qquad \text{(S6)}$$

where $bel_{12M}(Y)$ and $bel_{12F}(Y)$ are respectively the beliefs that X is young given the pieces of evidence $E_1$, $E_2$ and M or F. They correspond in fact to the conditional beliefs that X is young given X is male or female (and $E_1$ and $E_2$).

So the inequality S6 contradicts the "sandwich" principle described in S3 (P.3.2 §2) as S6 is equivalent to:

$$bel_{12}(Y) < \min ( bel_{12}(Y|M) , bel_{12}(Y|F) )$$

***Proof that (S5) implies (S6).***

The symmetry of the reliabilities implies that $bel_{1M} = bel_{2F}$, so (S5) becomes $bel_{12}(Y) < bel_{2F}(Y)$. Once M is known, $E_2$ becomes a tautology, so $bel_{1M} = bel_{12M}$. Identically $bel_{2F} = bel_{12F}$. Replacing $bel_{1M}$ and $bel_{2F}$ in (S5) one gets (S6). QED

In conclusion, the sandwich principle might seem natural though I think my example casts some doubt on any such naturalness. If you accept the sandwich principle, you must accept S4. (In fact S4 results from a Bayesian analysis, whereas S5 results from a TBM analysis.)

As another example, suppose you have 3 potential killers, A, B or C. Each can use a gun or a knife. I shall select one of them but you will not know how I select the killer. The killer selects his weapon by a random process with p(gun) = .2 and p(knife)=.8. Each of A,B,C has his own personal random device, the random devices are unrelated, and it just so happens they share the same probabilities (the .2 values).

Suppose you are Bayesian and must express your 'belief' that the killer will use a gun. The BF solution gives bel(gun) = .2x.2x.2 = .008. What about you? Would you defend .2? But this only applies if I select the killer with a random device (in which case the BF also gives .2 as a result). But I never said I would use a random device, I might be a very hostile player, and cheat whenever I can.

So before assigning the killing to A, B and C, I will order each of them to note, on a piece of paper, the weapon they would use if selected. They use their personal random device and note the name of the weapon they would use if selected. Suppose I am an hostile player. In that case, I would do whatever I could to show your prediction about the weapon used is wrong. You express your belief about the gun, so if I can, I will assign the killing to somebody who wrote 'knife' on his paper (I just look at the papers). So what is the 'chance' I cannot beat you? What is the probability that whatever I do you win. Answer: the chance that the 3 killers wrote 'gun', i.e. .008.

So your .008 belief for gun corresponds to the strength of your belief that whatever selection procedure I use - even the most hostile - you will win if you bet on 'gun'.

So you could interpret bel(x) as the probability that you are sure to win whatever Mother Nature (however hostile) will do (and pl(x) = 1 - .8x.8x.8 then corresponds to what a most friendly Nature would do). The betting behaviour proposed by Stratt (1989) is in fact based on these ideas.

### S.7.2. Marginal beliefs based on conditional beliefs.

Let $\Omega_A$ and $\Omega_B$ be two finite frames of discernment. Let $\Re_A$ be an algebra on $\Omega_A$. Let $\mathfrak{B} = \{B_i: i=1,\ldots n\}$ where the $B_i$ are the elementary propositions of $\Omega_B$. Suppose one knows $bel(A \mid B_i)$ for each $A\in\Re_A$ and each $B_i\in\mathfrak{B}$. Suppose the conditional BF are distinct. Let $\Re_B$ be an algebra on $\Omega_B$. Given some prior belief $bel_0$ on $\Re_B$ (with bbm $m_0$), then the marginal $bel(A) = bel(A\mid \bigcup_{i=1}^{n} B_i)$ defined on $\Re_A$ is:

bel(A) = \i\su(B∈ $\Re_B$,, { $m_0(B)$ )\i\pr($B_i\in\mathfrak{B}$, $B_i\subseteq B$,, $bel(A \mid B_i)$ }) (S7)

(This relation is deduced from the so-called conditional embedding underlying the **Generalized Bayes Theorem**, see Smets (1978, 1979, 1981, 1986, 1987), Shafer (1982)).

In Pearl's example 4, I consider that I am in a state of total ignorance as regards $\Omega_B$, hence by S7

$$bel(A) = bel(A \mid B)\ bel(A \mid \neg B) \le \min(\ bel(A \mid B)\ ,\ bel(A \mid \neg B)\ )$$

which contradicts the sandwich principle.

### S.7.3. Why is the sandwich principle not required?

Let cyl(A), cyl(B) and cyl(¬B) be the cylindric extensions of $A \in \Re_A$ , B and $\neg B \in \Re_B$ on the cartesian product $\Omega = \Omega_A \times \Omega_B$. Let $bel_\Omega$ be the BF defined on Ω that could generate the various bel(A), bel(A | B) and bel(A | ¬B) by marginalization or conditioning.

Then
$$\begin{aligned} bel(A) &= bel_\Omega(cyl(A) \mid cyl(B) \vee cyl(\neg B)) \\ bel(A \mid B) &= bel_\Omega(cyl(A) \mid cyl(B)) \\ bel(A \mid \neg B) &= bel_\Omega(cyl(A) \mid cyl(\neg B)) \end{aligned}$$

All these bel concern the same event on Ω (cyl(A)) but the last two result from some updating on some true facts (cyl(B) and cyl(¬B)). In the latter two cases we know more than in the first. It seems reasonable to accept that to some context the more we know the stronger our belief could be[6]. Each piece of evidence (the one inducing the updating on cyl(B) and the one inducing the updating on cyl(¬B)) induces some information about cyl(A). That the resulting beliefs on A (bel(A | B) and bel(A | ¬B)) are increased in both cases (min (bel(A | B) , bel(A | ¬B)) ≥ bel(A)) becomes perfectly acceptable. In particular, when we know nothing our belief is vacuous and it will increase by the accumulation of new pieces of evidence. So our belief on cyl(A) increases also once we know that B is true (or that ¬B is true).

In fact, suppose one knows there is no *a priori* whatsoever on $\mathfrak{B}$, then (S7) becomes:

$$bel(A) = \prod_{i=1}^{n} bel(A \mid B_i) \qquad \text{(S8)}$$

Hence with large n, bel(A) tends in general towards a vacuous belief function, which I consider to be perfectly reasonable. Should I have some probabilistic *a priori* on $\mathfrak{B}$, then bel(A) would be some weighted average of the bel(A | $B_i$) (as required by the sandwich principle). But when the belief on $\mathfrak{B}$ is vacuous, the principle does not hold. That bel(A) tends to 0 reflects the fact that cyl(A) is hardly supported when n is large.

### S.7.4. Further considerations.

1) The conclusions at the end of P.2.2.§2 are wrong. I agree that **conditional beliefs** bel(A | $B_i$) for A in $\Omega_A$ within each context $B_i \in \mathfrak{B}$ are usually the real foundation for expressing **expert opinions** and that there is no prior over the space $\Re_B$. Our Generalized Bayes Theorem was developed with exactly that idea in mind. Pearl's

[6] This argument does not mean that bel(A) will always increase as new pieces of evidence become known, as the latter could also be in contradiction with A.

criticism is inappropriate. It results, of course, from his erroneous requirement of relation S3.

**2) The pork eating problem.** (P.2.3. reasoning by cases §2)
Again see relation S8 for the origin of the .63 result. Furthermore, one should be very careful when subdividing a category (the Jews) into sub-categories. One must remember that relation (S8) must be respected. Pearl's data .7 imply that there is an *a priori* bbm m on the four Jews sub-categories (OJ, CJ, RJ, NAJ) among the Jews such that:

$$.70 = .999 \cdot m(OJ) + .8 \cdot m(CJ) + .4 \cdot m(RJ) + .2 \cdot m(NAJ) + \\ + .999 \cdot .8 \cdot m(OJ \vee CJ) + \ldots + .999 \cdot .8 \cdot .4 \cdot .2 \cdot m(OJ \vee CJ \vee RJ \vee NAJ)$$

How to derive m is still an open question. Should we apply the least specific principle as advocated by Dubois and Prade, or minimize some measure of information (Smets, 1983).

## S.8. De re and de dicto conditionals.

Pearl (P.2.2 §2 and §3, and P.2.3) raises the problem of the quantified conditional, already studied by Harper et al. (1981). In fact most problems come from the ambiguity of the proposition:

"I believe that A implies B at level p " (S9)

Should (S9) be translated into:

| | | |
|---|---|---|
| if A then (belief(B) = p) | $A \supset (bel_A(B) = p)$ | de re interpretation |
| belief(if A then B) = p | $bel(A \supset B) = p$ | de dicto interpretation |

I call these two interpretations the **de re** and **de dicto** interpretations (Smets et al., 1990).

The implication $A \supset B$ is translated as a **material implication** $\neg A \vee B$. $bel_A(B)$ denotes the belief given to B in a context where A is true. It corresponds to $bel(B \mid A)$.

With BF one has:

$$bel_A(B) = bel(B \mid A) = \frac{bel(\neg A \vee B) - bel(\neg A)}{1 - bel(\neg A)} = bel(\neg A \vee B) = bel(A \supset B)$$

if $bel(\neg A) = 0$, an accepted requirement as it only translates the idea that the knowledge of an implication should not induce any a priori belief on the antecedent domain. **The de**

**re and de dicto interpretations lead to the same results** - a nice property that resolves most (if not all) conflicts related to the concept of probability of conditionals versus conditional probabilities (Lewis, 1976).

A particular case is raised in **Pearl's Footnote 5:** If all we know is bel(If B then A) = p, then the use of bel(B $\supset$ A) = bel($\neg B \vee A$) = p leads to the results derived by the deconditionalization process (called conditional embedding in Shafer 1982) (see relation S7 in S.7.2 with m(A|B) = p).

Using deconditionalization (see S7), one can solve the case where I know that:

I believe at level $p_i$ that $B_i$ implies $A_i$ for i = 1, 2…n,

with $B_i \in \mathfrak{B}$ and $A_i \in \mathfrak{R}_A$ (see P.2.2). Each uncertain implication is translated into the simple support functions with $m(\neg B_i \vee A_i) = p_i$, and combined via Dempster's rule of combination.

Suppose that $B \in \mathfrak{R}_B$ is known to be true. Then from S7 the BF induced on $A \in \mathfrak{R}_A$ by "B is true" and the set of uncertain implications is

$$\mathrm{bel}(A \mid B) = \prod_{B_i \in B} \mathrm{bel}(A \mid Bi) = \prod_{i:B_i \in B} c_i\, p_i \qquad \text{(S10)}$$

with $c_i$ = 1 if $A_i \subseteq A$, 0 otherwise.

A pity these formulae (S7 and its special case S10) never became very popular. In actual fact, they were completely ignored. Let us hope the present papers will help to promote them.

## S.9. Constraints or combinations.

I would like to introduce another problem concerning the meaning to be given to the available information on our belief (P.2.3).

*Situation 1:* With probability theory, one usually postulates the existence of some underlying probability function on some space and all pieces of evidence correspond to some constraints on this probability function (e.g. I know P(A) and P(B) but not P(A&B) as in P.2.1 example 2).

*Situation 2:* With BF, it is usually postulated that each piece of evidence induces a BF on some space, that they are distinct (a concept which incidently still requires a formalized definition) and that they must be combined through Dempster's rule of combination.

The two situations are totally different.

The first situation can also be encountered with BF even though I do not think it is often relevant. This is what Pearl tries to solve in P.2.1, examples 1 and 2. In such partly defined belief functions, one can always try to apply the **minimum specificity** approach of Dubois and Prade (1987), i.e. find the least specific BF that satisfies the given constraints where $bel_1$ is less specific than $bel_2$ if $bel_1(A) \leq bel_2(A)$ $\forall A$ in $\Omega$.

The second situation can also be encountered with probability function … but it raises the problem of how to **aggregate probabilities** (Genest and Zidek, 1986).

For instance, in (P.2.3.reasoning by cases §1) one must distinguish between the fact that Pearl's equations P11 are

1) either two constraints on the same underlying BF: in that case $bel(B \mid A) = .9$ and $bel(B \mid \neg A) = .7$ and the least specific underlying BF is such that $m(B) = .6$, $m(A \vee B) = .1$ and $m(\neg A \vee B) = .3$.)
2) or produced by two distinct pieces of evidence, in which case one will obtain $bel(B) = .63$ by the application of (S8). That solution does not run counter to common sense as explained in S.7.

**S.10. Tweety splashed by the Sprinkler while eating his sandwich.** (P.2.3.chaining and example 5)

Tweety and the Sprinkler examples deserve a full TBM analysis in order to avoid the erroneous conclusions derived from shallow analysis.

**Tweety is an atypical bird.**

The TBM analysis of the **Tweety problem** is the following :
Let $\mathfrak{B} = \{TB, NTB, NB\}$ where TB = typical bird, NTB = non typical bird and NB = non-bird. Let $\mathcal{F} = \{F, \neg F\}$ where F = "it flies". The data are presented in table 1. They translate:

In general typical birds fly.
In general non-typical birds do not fly.
In general non-birds do not fly.

| | TB | NTB | NB |
|---|---|---|---|
| F | $\alpha$ | 0 | 0 |
| ¬ F | 0 | $\beta$ | $\gamma$ |

**Table 1:** The bbm allocated to F and ¬F given TB, NTB and NB.

Furthermore, "in general birds are typical birds" ($bel(TB \mid \neg NB) = \delta$). $\delta$ is the bbm given to $TB \cup NB$ that is transferred to TB by the conditioning on $\neg NB$.

*case 1:* Let it be known that Tweety is a bird. The application of (S7) implies that

$$\begin{aligned} bel(F \mid TB \vee NTB) &= m(TB \mid TB \vee NTB)\ bel(F \mid TB) \\ &\quad + m(NTB \mid TB \vee NTB)\ bel(F \mid NTB) \\ &\quad + m(TB \vee NTB \mid TB \vee NTB)\ bel(F \mid TB)\ bel(F \mid NTB) \\ &= \delta\, \alpha \end{aligned}$$

$$bel(\neg F \mid TB \vee NTB) = 0$$

*case 2:* Let it be known that Tweety is a penguin and all penguins are NTB, then by (S7)

$$bel(F \mid NTB) = 0$$
$$bel(\neg F \mid NTB) = \beta$$

and Tweety is still a bird as it should ($bel(NTB) = 1$, so $bel(TB \vee NTB) = 1$). Who said that BF is essentially a theory of monotonic logic? (see P.2.2 §3)

*case 3:* Let it be known that Titi is a typical bird. Thus by S7, $bel(F|TB) = \alpha$, $bel(\neg F|TB) = 0$ and $bel(B) = 1$. Hence our belief that typical birds fly is larger than our belief that birds fly… which is required (but not always derived in the penguin analysis).

The TBM analysis of the Tweety example (and **rules subject to exceptions**) deserves a fuller treatment (see Smets and Hsia, 1990). As shown in case 2, the facts "Tweety is a penguin" and "Tweety is a bird" lead to $bel(F) = 0$ and $bel(\neg F) = \beta$, not to a vacuous BF as claimed by Pearl (P.5 §6).

**Sprinkler is on and it does not rain.**

The **sprinkler problem** is indeed a tricky one, but so is it with many probabilistic problems (see Székely G., 1986).

Let S = "Sprinkler was on" $\mathcal{S} = \{S, \neg S\}$
R = "It rained" $\mathcal{R} = \{R, \neg R\}$
W = "Ground is wet" $\mathcal{W} = \{W, \neg W\}$

$bel_1(S \supset W) = 1$
$bel_2(R \supset W) = 1$
$bel_3(W \supset S \vee R) = 1$
$bel_4(\neg S) = \alpha$

The BF express that the antecedents "Sprinkler was on" ($bel_1$) or "It rained" ($bel_2$) imply the consequent "Ground is wet" and no other causes could be invoked ($bel_3$). $bel_4$ expresses that usually sprinklers are not on.

The vacuous extension of the four BF on the product space $\mathcal{S} x \mathcal{R} x \mathcal{W}$ and their combination induces the BF whose bbm are

$$m((\neg W,\neg S,\neg R) \vee (W,\neg S,R)) = \alpha$$
$$m((\neg W,\neg S,\neg R) \vee (W,S,\neg R) \vee (W,\neg S,R) \vee (W,S,R)) = 1 - \alpha$$

By conditioning on S, $\alpha$ is transferred to Ø, and 1 - $\alpha$ is transferred to (W,S). So after normalization, bel(S) = 1, bel(W) = 1 and bel(R) = 0 as it should.

By conditioning on W, $\alpha$ is transferred to $(W,\neg S,R)$ and $1-\alpha$ to $(W,S,\neg R) \vee (W,\neg S,R) \vee (W,S,R)$. So bel(W) = 1 and bel(R) = $\alpha$, which translates the idea that "Ground is wet" suggests "It rained".

The paradoxes described by Pearl do not exist when BF are used correctly.

**The Peter, Paul and Mary sandwich** (P.3.2. example 5)

Consider the product space CxS where C ={H, T} and S = {h, t} (translation: Coin, Head, Tail, Sandwich, ham, turkey). Let the four elements of CxS be a = (H,h), b = (H,t), c = (T,h), d = (T,t). The problem is to build a belief function on CxS. We know that Pr(H) = Pr(T) = bel(H) = bel(T) = .5. (I think that you should always accept Hacking's frequency principle: if Pr(x) = p then bel(x) = p.). So on CxS,

$$bel(a \cup b) = bel(c \cup d) = .5.$$

Next what are bel(H|h) and bel(H|t). The first remark is that they must satisfy the constraint:

$$\text{'}bel(H|h) = x\text{'} \equiv \text{'}bel(H|t) = x\text{'}$$

whatever x (here x = .5). The only way to build a belief function on CxS such that this constraint is always satisfied is by allocating the basic belief masses on CxS such that $m(a \cup b) = m(c \cup d) = .5$ . It satisfies all constraints (the one relative to $bel(a \cup b)$... and the very strong "correltation" ('bel(H|h) = x' $\equiv$ 'bel(H|t) = x').

What is the chance that Paul will win $1000. To win (W) is equivalent to obtaining $a \cup d$. If Paul knows that the sandwich is h, then his chance of winning is .5. (idem if turkey). (one has bel(W|h) = bel(H|h) = Pr(H|h) = .5, etc...)

If Paul does not know which sandwich was prepared, then bel(W) = $bel(a \cup d)$ = 0 = $bel(\neg W)$ = $bel(b \cup c)$. But if Paul must bet on W, which of course is his real problem (the

\$1000 attest to it), then Paul builds his pignistic probabilities on CxS by applying my pignistic transformation: The .5 basic belief mass given to a∪b is split equally between a and b (idem for the other .5) . So each singleton of CxS receives a BetP of .25, and BetP(W) = BetP(a) + BetP(d) = .5, as requested by Pearl.

**Other paradoxes.**

Most paradoxes raised by Pearl in P.2.3.contraposition are paradoxes related to the translation of IF…THEN by material implications. **They exist in logic** and are unrelated to the use of BF.

I consider it logically correct to state "If you are hungry, you haven't eaten cakes". That it is not a useful statement is irrelevant. One should be careful not to confuse **truth and usefulness**.

With regard to the bird, in a world with **"only" birds**, the translation of "typical birds fly" has to be done as with Tweety. Two categories need to be created: Typical birds (those who usually fly) and Non-typical birds (those who usually do not fly). The paradox then disappears.

With the **kind person**, if all I know are the two rules given by Pearl, then I cannot understand Pearl's conclusion. I feel it is perfectly reasonable to believe that Joe is unkind (with strength $m^2$). Note that the same conclusion (and Pearl's dissatisfaction) would appear in plain logic. Given the two rules, F will imply ¬K. To avoid such a conclusion - as seems to be Pearl's intention - there must be further rules. If someone would only tell me what they are I would include them in the TBM analysis. Dissatisfaction with BF models is often due to the fact that only some rules are considered, which then lead to unsatisfactory conclusions. But the reason why these conclusions are unsatisfactory is that there are extra constraints we would like to see fulfilled… but did not include in our initial analysis. Introduce them in the TBM analysis, and the results will be satisfactory.

## S.11. The three prisoners problem. (P.3.1)

The **three prisoners problem** is a nice problem that deserves a TBM analysis. Let $A_1$, $A_2$ and $A_3$ be the three prisoners. Everybody knows that two prisoners are going to be executed. I ($A_1$) call the guard and ask him to give me the name of one of the other two prisoners that is going to be executed. He answers $A_3$. Should I change my belief that I will be saved given I know $A_3$ is going to be executed?

Before I ask the guard, my belief that the saved prisoner S was $A_1$ was: bel(S=$A_i$) = 1/3. It was based on the accepted idea that the judge would randomly select (with probability 1/3) the prisoner that would be saved. Two situations could then be considered.

*Context 1:* The guard looked over the judge's shoulder, saw the result of the selection, came to my cell and told me "$A_3$ is not saved". It is equivalent to saying "$A_3$ was not selected": conditioning on $\neg A_3$ seems to be realistic. Hence my belief that I would be saved increased to 1/2.

$$\text{bel}(S=A_1 \mid S\neq A_3) = 1/2$$

This is how the solution disliked by Pearl is arrived at. But it does not correspond to the context that concerns Pearl.

*Context 2:* The context that Pearl uses is the following: I called the guard, told him I knew one of $A_2$ or $A_3$ had to die, that knowing the name of one of those who would be executed would not change my belief that I would be saved, so that the guard could answer my question. And so he did. Why did he agree ? Because the guard had made sure that answering the question would not change my belief about being saved. To reach this goal, all he had to do was to make it clear to me that he would select with probability 1/2 between $A_2$ and $A_3$ in case I was going to be saved. In such a case, the TBM analysis is reduced to a plain probabilistic analysis. One gets:

$$\text{bel}(S=A_1 \mid \text{guard accept to answer and say } S\neq A_3) = 1/3.$$

The TBM and probability solutions are identical, which is normal as the probability solution corresponds to the TBM solution whenever beliefs happen to be additive.

## S.12. Decision making.

Pearl (P.1.4.§5) raises the problem of **decision based on BF**. A frequent error consists in interpreting bel(A) as the maximal rate at which one would be willing to bet on event A. This interpretation concerns lower probabilities, not belief functions. In both Shafer and the TBM interpretation bel provides only the strength of our belief that A is true. It does not directly claim the probability that A is true, as A may be true by chance over and above the necessity implied by bel(A). Where **betting context** is concerned, let me refer to the paper (Smets 89) in which I explain what p**ignistic probability** is, i.e. the probability derived from my belief and used in order to bet "coherently". Beliefs quantified by BF can be used as a basis for **rational decision making**. Once decision is involved, our BF is transformed into an additive measure (a probability function) on the set of alternatives, and decisions are then based on these probabilities (and the utilities involved).

**Tu Quoque Judea**

The translation of the sandwich principle into decision theory cannot be used as a criticism against BF. The fact is the principle is violated even within probability theory as shown by the **Newcomb paradoxes** (Eells, 1982, Gärdenfors and Sahlin, 1988).

As an example of such a paradox, take the story of **King David and Bath-Cheba**. (2 Samuel 11.2-4). King David fell in love with Bath-Cheba, Urie's wife. He could easily have sent his guards to get her but worried about the consequences of such an act: the pleasure of Bath-Cheba's company had to be weighed against the risk that this iniquitous act would cause a revolt,which for him would be disastrous. So, he was going to abstain when the grand vizier Seyab explained to him that revolts result from lack of charisma, not from iniquitous acts. Either a King has charisma and revolts are unlikely or he does not have charisma and revolts are likely. That he has committed iniquitous acts is irrelevant in both contexts. So why not enjoy Bath-Cheba's presence? Where does the paradox come from?

**Origin of the paradox.**

The fact is King David was well-versed in statistics and decision theory, so before deciding, he collected data from a hundred previous kings, on their charisma status (C), on whether they had committed iniquitous acts (B) and whether there had been a revolt (R).

| | C | | ¬C | |
|---|---|---|---|---|
| | B | ¬B | B | ¬B |
| R | 2 | 3 | 24 | 16 |
| ¬R | 18 | 27 | 6 | 4 |
| | 20 | 30 | 30 | 20 |

**Table 2:** Distribution of 100 kings according to charisma (C), the committing of iniquitous acts (B) and occurrence of revolt (R).

| | B | ¬B |
|---|---|---|
| R | 26 | 19 |
| ¬R | 24 | 31 |

**Table 3:** Distribution of 100 kings according to the committing of iniquitous acts (B) and occurrence of revolt (R).

The utility of each act and consequences are given in table 4.

| u | B | ¬ B |
|---|---|---|
| R | $-\beta+\varepsilon$ | $-\beta$ |
| ¬ R | $\varepsilon$ | 0 |

**Table 4:** Utility of each act (B or ¬B) and consequence (R or ¬R). ε is the (positive) utility that results from Bath-Cheba's company, -β is the (strongly negative) utility that results from revolt (β>>ε)

King David's first computation is based on the probabilities given in table 3 where he does not consider the charisma status. In that case, the expected utilities are

$$u(B) = \varepsilon - \beta \cdot \frac{26}{50} \ll u(\neg B) = -\beta \cdot \frac{19}{50}$$

As β is much larger than ε, King David should not send his guards to get Bath-Cheba.

But if King David considers his charisma status (therefore the probabilities given in table 2), one gets:
- if King David has charisma

$$u(B) = \varepsilon - \beta \cdot \frac{2}{20} > u(\neg B) = -\beta \cdot \frac{3}{30}$$

- if King David does not have charisma

$$u(B) = \varepsilon - \beta \cdot \frac{24}{30} > u(\neg B) = -\beta \cdot \frac{16}{20}$$

So in both contexts, King David should send his guards and get Bath-Cheba. What are we to make of this?

The **sandwich principle** questioned once already in S.7 does not stand up to criticism within strict probability theory.

## S.13. Conclusion.

I think I have answered all of Pearl's criticisms against BF and therefore also against the TBM.

Pearl complains about the lack of effective procedure for deciding whether a problem can be represented by BF. He is right, but I would also like to have an effective procedure for deciding whether a problem can be represented by a probability function (or possibility functions, or fuzzy sets theory etc)

I think it is equally important to expect a procedure that could correctly describe a BF problem (see the Tweety, the Sprinkler and Pork Eating solutions). It is so awfully easy to be wrong. But this is, unfortunately, not specific for BF. Just consider the case where you collect the conditional probabilities on A given B, on B given C and on C given A. Careless manipulations could easily lead to absurdity as many constraints underlying the three sets of conditional probabilities might be violated.

The comparison between BF and likelihood is not so enlightening. In fact a more interesting relation exists between likelihood and **possibility theory** (Smets, 1982).

I wish to apologize for possibly being somewhat harsh in some of my replies. I was very happy that Pearl's paper provides me with an opportunity to show how the TBM can resist assaults of that kind. Pearl's criticism, unfortunately, reflects the status of BF literature. I think the foundation of any models for quantifying belief is a very important and all too often neglected topic. To use BF and Dempster's rules blindly can be very dangerous indeed. Their use must be justified by the nature of the problem to which they are applied, not by fad or ad hoc argument.

## Bibliography.


CHOQUET G. (1953) Theory of capacities. Ann. Inst. Fourier, Grenoble, 5:131-295.

COX R.T. (1946) Probability, frequency and reasonable expectation. Amer.J.Phys. 14:1-13.

DEMPSTER A.P. (1967) Upper and lower probabilities induced by a multivalued mapping. Annals of Math. Statistis 38:325-339.

DUBOIS D. and PRADE H. (1987) The principle of minimum specificity as a basis for evidential reasoning. in: Uncertainty in knowledge-based systems, BOUCHON B. and YAGER R. eds, Springer Verlag, Berlin, pg.75-84.

EELLS E. (1982) Ratioanl decision and causality. Cambridge Univ. Press, Cambridge, G.B.

FAGIN R. and HALPERN J.Y. (1989) Unceertainty, belief and probability. Proc.Intl. Joint Conf. in AI, IJCAI89, Detroit, 1161-1167.

GARDENFORS P., HANSSON B. and SAHLIN N.E. (1983) Evidentiary value: philosophical, judicial and psychological aspects of a theory. C.W.K. Gleerups, Lund, Sweden.

GARDENFORS P. and SAHLIN N.E. (1988) Decision, probability and utility. Cambridge Univ. Press, Cambridge, G.B.

GENEST C. and ZIDEK J.V. (1986) Combining probability distributions: a critical and annotated bibliography. Statistical Sc. 1:114-148.

HARPER W.L., STALNOCKER R. and PEARCE G. (1981) Ifs. Reidel, Dordrecht.

JEFFREY R.C. (1983) The logic of decision. 2nd Ed. Univ. Chicago Press, Chicago.

LEVI I. (1983) Consonance, dissonance end evidentiary mechanisms. in GARDENFORS P., HANSSON B. and SAHLIN N.E. (eds) Evidentiary value: philosophical, judicial and psychological aspects of a theory. C.W.K. Gleerups, Lund, Sweden. pg. 27-43.

LEWIS D. (1976) Probabilities of conditionals and conditional probabilities. Philosophical Review 85: 297-315.

NGUYEN H.T. (1978) On random sets and belief functions. J. Math. Anal. and Applic. 65:539-542.

PEARL J. (1988) Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann Pub. San Mateo, Ca, USA.

PLANCHET B. (1989) Credibility and conditioning. J. Theor. Probabil. 2:289-299.

RUSPINI E.H. (1986) The logical foundations of evidential reasoning. Technical note 408, SRI International, Menlo Park, Ca.

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

SHAFER G. (1980) Jeffrey's rule of conditioning. Personal communic.

SHAFER G. (1982) Belief functions and parametric models. J. Roy. Statist. Soc. B44:322-352.

SHAFER G. and TVERSKY A. (1985) Languages and designs for probability. Cognitive Sc. 9:309-339.

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. (1979) Modèle quantitatif du diagnostic médical. Bulletin de l'Académie Royale de Médecine de Belgique 134 : 330-343.

SMETS Ph. (1981) Medical Diagnosis : Fuzzy Sets and Degrees of Belief. Fuzzy Sets and systems, 5 : 259-266.

SMETS Ph. (1982) Possibilistic Inference from Statistical Data. In : Second World Conference on Mathematics at the Service of Man. A. Ballester, D. Cardus and E. Trillas eds. Universidad Politecnica de Las Palmas, pp 611-613.

SMETS Ph. (1983) Information Content of an Evidence. Int. J. Man Machine Studies, 19 : 33-43.

SMETS Ph. (1986) Bayes' theorem generalized for belief functions. Proc. ECAI-86,vol II. 169-171.

SMETS Ph. (1987) Belief functions and generalized Bayes theorem. Second IFSA Congress, Tokyo, Japan, July 20-25. pg. 404-407.

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 pg 253-286.

SMETS Ph. (1989) Construucting the pignistic probability function in a context of uncertainty. Proc. Fifth Workshop on Uncertainty in AI, Windsor, Canada, 319-326. also in: Uncertainty in Artificial Itelligence 5, Henrion M., Shachter R.D., Kanal L.N. and Lemmer J.F. ed., North Holland, Amsterdam 1990, 29-39.

SMETS Ph. (1990a) Belief functions and their combinations.IEEE Trans. PAMI, 12: 447-458..

SMETS Ph. (1990b) The transferable belief model and random sets. IPMU-90, pg. 93-95.

SMETS Ph., CLARKE M.R.B., and FROIDEVAUX C. (1990) Uncertainty, Conditional and Non Monotonicity. To appear in Int. J. Non Standard Logic.

SMETS Ph. and HSIA Y.T. (1990) Default reasoning and the transferable belief model. 6th Int. Conf. on Uncertainty in AI. Boston.

STRAT T.M. (1989) Making decisions with belief functions. Proc. 5th Workshop on Uncertainty in IA. Windsor, 351-360.

SUPPES P. and ZANOTTI M. (1977) On using random relations to generate upper and lower probabilities. Synthesis 36:427-440.

SZEKELY G.J. (1986) Paradoxes in probability theory and mathematical statistics. Reidel, Dordrecht., The Netherlands.

## Acknowledgements.

The author is indebted to Y.T. Hsia, Robert Kennes, Bruno Marchal and Alessandro Saffiotti for their useful comments.