Decision Making in a Context where Uncertainty is Represented by Belief Functions.

Philippe Smets
IRIDIA
Université Libre de Bruxelles
50 av. Roosevelt, CP 194-6, 1050 Bruxelles, Belgium
psmets@ulb.ac.be
http://iridia.ulb.ac.be/~psmets

October 6, 2000

Abstract

A quantified model to represent uncertainty is incomplete if its use in a decision environment is not explained. When belief functions were first introduced to represent quantified uncertainty, no associated decision model was proposed. Since then, it became clear that the belief functions meaning is multiple. The models based on belief functions could be understood as an upper and lower probabilities model, as the hint model, as the transferable belief model and as a probability model extended to modal propositions. These models are mathematically identical at the static level, their behaviors diverge at their dynamic level (under conditioning and/or revision). For decision making, some authors defend that decisions must be based on expected utilities, in which case a probability function must be determined. When uncertainty is represented by belief functions, the choice of the appropriate probability function must be explained and justified. This probability function does not represent a state of belief, it is just the additive measure needed to compute the expected utilities. Other models of decision making when beliefs are represented by belief functions have also been suggested, some of which are discussed here.

Contents

1 Introduction.

Uncertainty induces beliefs, i.e. graded dispositions that guide our behavior. Translated within a normative approach, this leads usually to the construction of a model to represent quantified beliefs that is linked directly to 'rational' agent behavior described within decision contexts (DeGroot, 1970). It has been argued that decisions are 'rational' only if we use a probability measure over the various possible states of the nature and compute with it the expected utility of each possible act, the optimal act being the one that maximizes these expected utilities (DeGroot, 1970; Savage, 1954). Accordingly, whenever a decision must be made, the decider's beliefs must induce a probability function on the set of possible outcomes. Nevertheless the fact that beliefs can only be observed through our decisions does not necessarily mean that beliefs cannot be entertained without any revealing behavior manifestations (Smith & Jones, 1986 p.147). The idea that entertained beliefs and beliefs in a decision context can be distinguished is defended in the transferable belief model (Smets & Kennes, 1994). When these two types of beliefs are distinguished the classical justification for the use of probability functions to represent quantified beliefs does not apply for what concerns the entertained beliefs. Of course it still applies for what concerns decisions themselves.

New normative models to represent quantified belief have been proposed recently, in particular those based on belief functions. These models are often criticized for supposedly not providing a method for rational decision making. This paper reviews solutions proposed in some of the models where uncertainty is represented by belief functions. Four types of such models have been described.

1. Some authors consider that the belief function is the lower envelope of a family of probability functions (section 4.1). Some of them assume the existence of some underlying but imprecisely known probability function (Walley, 1987; Giles, 1982). Others just assume that beliefs are represented by the family of probability functions itself, not explicitly assuming some underlying partially unknown probability function (Kyburg, 1987b; Levi, 1980; Smith, 1961; Voorbraak, 1993). In both cases, a set of expected utilities can be computed for each decision, but these sets usually do not induce an unambiguous order among the set of possible acts. To resolve that ambiguity, Schmeidler (1989) proposes the use of the lower expectation to make decisions. A generalization of it can be found in Jaffray (1988)) and Strat (1990b, 1990a) who, inspired by Hurwicz (Luce & Raiffa, 1957), advocate the use of a weighted average of the upper and the lower expectations (Section 6).

2. In the transferable belief model (Smets & Kennes, 1994) (section 4.3), we have proposed and justified the use of the so-called pignistic probabilities as the appropriate probability function to be used when decision is required (Smets, 1990; Smets & Kennes, 1994; Smets, 1993c) (section 5). In a similar framework, Appriou suggested another solution based on the most plausible singletons (Appriou, 1991) (section 5.9).

3. In Dempster's original work (Dempster, 1967) and in the hint model

of Kohlas and Monney (Kohlas & Monney, 1995) (section 4.2) the problem of decision making is not essential, and no particular solution seems to have been defended. Nevertheless (Schaller, 1991) studies the use of the hint models for multiple criteria decision making. The same remark applies to the models where a belief is understood as the probability of knowing (Ruspini, 1986, 1987)) or of proving (Pearl, 1988) (section 4.4).

In this paper, we review some of the solutions proposed for decision making when the uncertainty is represented by a belief function. We consider the case of one decider and one criteria. Before proceeding to decision making itself, we present the mathematical background on belief functions (Section 2) and probability theory (Section 3) useful in this paper and some of the models that lead to a belief function (Section 4). It is essential that these models are not confused. They might look similar. Indeed when considering just the static representation of uncertainty, they all end up with a belief function. But once the dynamic of the model is considered, the differences show up.

Sections 5 and 6 present some of the solutions proposed to make rational decisions in the transferable belief model and in the upper and lower probabilities frameworks, respectively. We focus on the theoretical aspects of decision making under uncertainty represented by belief functions. Applications can be found in the other papers published in this monograph.

2 Mathematical background on belief functions.

In this section, we summarize the mathematical knowledge about belief functions needed for this presentation. Details can be found in (Shafer, 1976a; Smets, 1988, 1998b).

2.1 Belief functions.

Let Ω be a finite space, called the frame of discernment. The elements of the set Ω are called 'worlds'. One world corresponds to the actual world; it is denoted hereafter ω_0. There is an agent, denoted You (De Finetti, 1974), who does not know which world is the actual world and who can only express the strength of his/her opinion (called degree of belief) that the actual world belongs to this or that subset of Ω.

That ω_0 belongs to Ω is assumed under the so-called 'closed world assumption', whereas under the 'open world assumption' ω_0 might not be an element of Ω (Smets, 1988, 1998b). The closed world assumption is assumed in probability theory and it justifies the axiom $P(\Omega) = 1$. In the transferable belief model, we acknowledge that You may have built Ω so that ω_0 does not belong to Ω.

On Ω, we define a function $bel : 2^\Omega \rightarrow [0, 1]$ called a 'belief function' (Shafer, 1976a). For $A \subseteq \Omega$, $bel(A)$ denotes the strength of Your belief that $\omega_0 \in A$. Mathematically, bel is a Choquet capacity monotone of infinite order . It satisfies the following inequalities:

$$
\begin{aligned}
& 1)\ bel(\emptyset) = 0 \\
& 2)\ \forall n \geq 1, \forall A_1, A_2, ...A_n \subseteq \Omega \\
bel(A_1 \cup A_2 \cup ...A_n) \quad \geq & \quad \sum_i bel(A_i) - \sum_{i>j} bel(A_i \cap A_j) \\
& \quad .... - (-1)^n bel(A_1 \cap A_2 \cap ...A_n).
\end{aligned}
$$

Under the closed world assumption, one adds $bel(\Omega) = 1$. Under the open world assumption, one adds only $bel(\Omega) \leq 1$.

Other useful functions have been defined, like the basic belief assignment m, the plausibility function pl, the implicability functions b, the commonality function q, the weight function w ... (Smets, 1998b). Their usefulness depends on the context. They are in one-to-one relations with bel.

The basic belief assignment (bba) related to a belief function bel is the function $m : 2^\Omega \to [0, 1]$ defined as:

$$
\begin{aligned}
m(A) &= \sum_{B \subseteq A} (-1)^{|A|-|B|} bel(B), \quad \forall A \subseteq \Omega, A \neq \emptyset, \\
m(\emptyset) &= 1 - bel(\Omega).
\end{aligned}
$$

Mathematically, m is the Möbius transform of bel. The value $m(A)$ for $A \subseteq \Omega$ is called the basic belief mass (bbm) given to A. It may happen that $m(\emptyset) > 0$. It reflects either that we are working under an open world assumption or that there exists some kind of contradiction in the belief state (Smets, 1992b).

The relation from m to bel is given by:

$$
bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B), \ \forall A \subseteq \Omega.
$$

The plausibility function $pl : 2^\Omega \to [0, 1]$ is defined as:

$$
pl(A) = bel(\Omega) - bel(\overline{A}), \ \forall A \subseteq \Omega.
$$

Shafer (1976a) accepts the closed world assumption, what can be expressed by saying that bel is 'normalized'. In that case, $bel(\Omega) = 1$, $pl(\Omega) = 1$ and $m(\emptyset) = 0$. In the transferable belief model, we do not require such a normalization (Smets, 1992b). We use the notation bel and pl, whereas Shafer uses the notation Bel and Pl. The difference should enhance that the functions bel and pl are not necessarily normalized.

The meaning of bel and of the other functions varies according to the meaning the authors give to $bel(A)$. In this presentation, we will discuss the upper and lower probabilities model, the hint model, the transferable belief model and the probability model extended to modal propositions (section 4).

2.2 Notation.

In order to homogenize the notation, we use the next conventions that we have found convenient, even though it might seems cumbersome in some cases. The full notation for *bel* and its related functions is:

$$bel_{Y,t}^{\Omega,\Re}[EC_{Y,t}](\omega_0 \in A) = x.$$

It denotes that the degree of belief held by the agent Y (shortcut for You) at time t that the actual world ω_0 belongs to the set A of worlds is equal to x, where A is a subset of the frame of discernment Ω and $A \in \Re$ where $\Re$ is a Boolean algebra of subsets of Ω. The belief is based on the evidential corpus $EC_{Y,t}$ held by Y at t, where $EC_{Y,t}$ represents all what agent Y knows at t. In practice many indices can be omitted for simplicity sake. Usually $\Re$ is 2^Ω, the power set of Ω. When $\Re$ is not explicitly stated, it means that *bel* is defined on 2^Ω. '$\omega_0 \in A$' is often denoted as 'A'. Y, t and/or Ω are omitted when the values of the missing elements are clearly defined from the context. So $bel^\Omega[E](A)$ or even $bel(A)$ are often used.

Note that $bel_{Y,t}^{\Omega,\Re}[EC_{Y,t}]$ (and its simplified forms) denotes the belief function, and can be understood as a finite vector of length $|\Re|$, which components are the values of $bel_{Y,t}^{\Omega,\Re}[EC_{Y,t}](A)$ for every $A \in \Re$.

In the above notation, *bel* can be replaced by any of m, pl, q, b, etc... The indices should made it clear what the links are. So $m_{Y,t}^{\Omega,\Re}[EC_{Y,t}]$ and $pl_{Y,t}^{\Omega,\Re}[EC_{Y,t}]$ are the bba and the plausibility function related to $bel_{Y,t}^{\Omega,\Re}[EC_{Y,t}]$.

2.3 Set of probability functions compatible with a belief function.

Given any normalized belief function bel^Ω on Ω, one can always define a family $\Pi(bel^\Omega)$ of probability functions P^Ω on Ω that satisfy any of the next three definitions:

$$\begin{aligned}
\Pi(bel^\Omega) &= \{P^\Omega : bel^\Omega(A) \leq P^\Omega(A),\ \forall A \subseteq \Omega\} \\
\Pi(bel^\Omega) &= \{P^\Omega : P^\Omega(A) \leq pl^\Omega(A),\ \forall A \subseteq \Omega\} \\
\Pi(bel^\Omega) &= \{P^\Omega : bel^\Omega(A) \leq P^\Omega(A) \leq pl^\Omega(A),\ \forall A \subseteq \Omega\}
\end{aligned}$$

The three definitions are equivalent. $\Pi(bel^\Omega)$ and bel^Ω are in one-to-one correspondence. We call $\Pi(bel^\Omega)$ the set of probability functions compatible with bel^Ω.

2.4 The belief function kinematic.

Let Ω be a frame of discernment and bel^Ω be a belief function on Ω. For $E \subseteq \Omega$, let Ev_E be the evidence that states that the actual world is not in $\overline{E}$. Suppose You learn Ev_E for sure. You must then revise Your beliefs accordingly. This is achieved by Dempster's rule of conditioning, i.e.:

$$bel^\Omega[Ev_E](A) = bel^\Omega(A \cup \overline{E}) - bel^\Omega(\overline{E}) \tag{1}$$

We use the strange expression 'the actual world is not in $\overline{E}$' instead of the more classical one 'the actual world is in E' as far as we don't assume initially that the actual world is an element of Ω (see the open world assumption, (section 2.1)).

Suppose now that we have two pieces of evidence Ev_1 and Ev_2 that bear on Ω. Let $bel^\Omega[Ev_1]$ and $bel^\Omega[Ev_2]$ be the belief functions derived from each piece of evidence taken individually. If these pieces of evidence are 'distinct', than we can compute the belief $bel^\Omega[Ev_1, Ev_2]$ induced by their conjunction on Ω. Its related bba satisfies:

$$m^\Omega[Ev_1, Ev_2](A) = \sum_{B,C:B\cap C=A} m^\Omega[Ev_1](B)m^\Omega[Ev_2](C).$$

This rule can also be written as:

$$m^\Omega[Ev_1, Ev_2](A) = \sum_{B,C\subseteq\overline{A}:B\cap C=\emptyset} m^\Omega[Ev_1](A\cup B)m^\Omega[Ev_2](A\cup C).$$

Another very useful rewriting of the same equation is:

$$f^\Omega[Ev_1, Ev_2](A) = \sum_{B\subseteq\Omega} f^\Omega[Ev_1, Ev_B](A)m^\Omega[Ev_2](B),$$

where $f \in \{m, b, bel, pl, q\}$ and $bel^\Omega[Ev_1, Ev_B]$ is the belief function obtained by the application to $bel^\Omega[Ev_1]$ of Dempster's rule of conditioning on Ev_B (relation 1).

The definition of 'distinctness' is detailed in (Smets, 1992a, 1998b). It translates the fact that the sources who produce the two belief functions based on Ev_1 and Ev_2 , respectively, are 'distinct', 'unrelated', 'independent'. No precise definition of these terms are presented here, as they will hardly be needed in this paper.

3 Mathematical background on probability theory.

3.1 An operational definition of P.

The classical definition of a subjective probability is based on an analysis of rational betting behavior. The (subjective) probability of a proposition is usually characterized as the value of the opportunity to gain a unit value if the proposition is true (Ramsey, 1964). More formally, one variant of the operational definition of a subjective probability is the following:

Definition 3.1 ***Operational definition of subjective probabilities.*** *Consider a finite space Ω, a game on the betting frame Ω, a player and a banker. We have '$P^\Omega_{You,t}(A) = x$' iff You consider at time t and for any $M > 0$ that the* *player must pay $\$xM$ to the banker to enter a game where the player wins $\$M$ from the banker if the actual world belongs to A and \$0 otherwise, and You are ready to be any of the player or the banker.*

We insist on the fact that You are not allowed to 'run away' from the game. You must accept to be either the banker or the player, this being settled after You have assessed the value of x. The present definition is based on 'forced bets'. The case where You could 'run away' is analyzed in the upper and lower probabilities framework (section 6).

3.2 The assessment of P.

In order to assess the value of a subjective probability, one can consider the following method. Let a finite space Ω and $A \subseteq \Omega$. Consider two bets. In bet 1, You bet on A versus $\overline{A}$ where You gain $\$x$ if A is true, and \$0 otherwise (with $\$x$ being any reasonable prize like \$100). In bet 2, You have an urn with a proportion p of Black balls. You bet on Black versus not-Black where You gain $\$x$ if the randomly selected ball (where every ball has the same chance to be selected) is Black, and \$0 otherwise. Which bet do You prefer?

- If You prefer bet 1, it means that $P(A) > p$.
- If You prefer bet 2, it means that $P(A) < p$.
- If You are indifferent between the two bets, it means that $P(A) = p$.

By varying p, one can (in theory) always find a state of indifference between the two bets. So one can assess the value of $P(A)$.

In practice, this method is not good to assess probabilities and more elaborated methods have been developed by psychometricians. Nevertheless many of the methods they developed are ingenious variants of the one we just described.

3.3 Decision making.

Suppose a decision maker wants to select an optimal act among an exhaustive set of possible acts, denoted A. The outcome of each act depends on the exact value of the actual world ω_0. Let Ω denote the set of possible worlds, the actual one being assumed to belong to Ω (the closed world assumption is assumed). Savage (1954) has proposed a set of rationality axioms that justify the use of an additive measure $P^\Omega : 2^\Omega \rightarrow [0,1]$ and a utility function $u : A \times \Omega \rightarrow \text{Reals}$. The optimal act is the one that maximizes the expected utility $\overline{u}(a)$ where:

$$\overline{u}(a) = \sum_{\omega \in \Omega} u(a,\omega) P^\Omega(\omega).$$

The naturalness of this set of axioms, even though often criticized (Rivett, 1975), has nevertheless stood the test of time. We will accept them hereafter.

That the additive measure P^Ω represents *stricto sensu* Your beliefs is not required in Savage's approach. In fact, Savage only proves the existence of the additive measure, its interpretation being left aside. Nevertheless most authors take it for granted that this additive measure represents Your beliefs. This is one of the major assumptions relaxed in the transferable belief model.

3.4 Dutch books.

The reason why beliefs must be represented by a probability function is often based on the Dutch book argument. A Dutch book is a set of bets that would lead to a sure loss whatever is the value of the actual world. The only way to avoid being vulnerable to a Dutch book, i.e., to be certain to always avoid facing sure loss, consists in representing Your beliefs by a probability function. In that case, Your beliefs are said to be 'coherent'.

Just to illustrate how the rule $P(\overline{A}) = 1 - P(A)$ is derived, consider the next two bets on A versus $\overline{A}$. In bet 1, player wins if A, and the prize You propose to enter the game is \$.3. In bet 2, player wins if $\overline{A}$, and the prize You propose to enter the game is \$.6. Then I decide that You will be the banker in both bets. In that case the player will pay You \$.9 to enter the two games. Whatever occurs, A or $\overline{A}$, the banker, thus You, will have to pay \$1 to the player. So You are doomed to loose \$.1 whatever occurs, a behavior normally considered as irrational. If the sum of the prizes proposed to enter the game is larger than \$1, then You will be the player, and doomed to loose money whatever occurs. The only way to avoid such sure losses is achieved by using prizes that add to \$1, hence the corresponding axiom in probability theory.

Similar bets can be organized to derive the additivity axioms of probability measures.

Bayes rule of conditioning, i.e., $P(A|B) = \frac{P(A \cap B)}{P(B)}$, is derived through the consideration of bets where the player wins \$1 if $A \cap B$ occurs, \$0 if $\overline{A} \cap B$ occurs, and the bet is canceled (the player gets back the money paid to enter the game) if $\overline{B}$ occurs.

It is essential to realize that the Dutch book argument does not tell what the values of these probabilities must be. It only states that if You bet at .7 on Heads in a coin tossing experiment, than You must bet at .3 on Tails to avoid to be victim of a Dutch book. The Dutch book argument does not require that the probability of Heads is .5 (or any other value). The latter is a problem of reasonableness that has nothing to do with the Dutch book argument (Kyburg, 1988)). One can be widely unreasonable without sinning against either logic or probability. This might seem surprising and even unpalatable to some probabilists, but no rationality arguments can be proposed that force You to bet at .5 on Heads.

Note that Savage's approach and the Dutch book argument consider only forced bets. You are not allowed to refuse the proposed bets. *A* more general approach that cope with unforced bets has been considered in (Smith, 1961; Giles, 1982; Walley, 1991). They concern the upper and lower probability theory (see Section 6).

3.5 A note on conditioning and belief revision.

In probability theory, the changes in belief that result from conditioning and belief revision are usually not distinguished. Nevertheless, Hacking (1988) uses these two forms of changes in belief to distinguish between personalist and Bayesian probabilists, the latter being a special case of the former.

A personalist probabilist considers that $Prob^H[f](h)$, what is classically denoted by $Prob_f(h)$, is the number representing Your personal probability for $h \subseteq H$, when You know f. Note furthermore that one assumes not only that fact f is true, but also that You know that fact f is true.

Probability given facts, denoted $P[f]$ is not to be confused with conditional probability, denoted $P(.|e)$ where e is the conditioning event. Using our full notation, conditional probability is defined as,

$$Prob^H[f](h|e) = \frac{Prob^H[f](he)}{Prob^H[f](e)}$$

for positive denominators, where f is the background facts known to You and e is the conditioning event. Conditional probabilities indicate how confident You would be, knowing only f, if You knew e as well.

When it comes to the problem of avoiding a Dutch Book, the fact f cannot be reconsidered. It is fixed and known, so no bet is to be considered if f were not the case. For what concerns e, the situation is different. We are considering bets where You could win or loose when e is the case, and bet is canceled if e is not the case. So when e is concerned, we must be careful on how the probabilities are allocated, as if they do not satisfy probability theory, then a Dutch Book could be built against the bettor. Therefore, conditioning must be performed according to the Bayesian rule. But as described, conditioning consist in considering 'hypothetical events', i.e., events that have not yet occurred but might.

The distinction between probability given facts (factual revision) and conditional probabilities (hypothetical conditioning) is usually not made. In the TBM, such a distinction is made.

Note: As an example consider the Peter, Paul and Marry Saga (see section 5.5.1, we suggest to the reader who is not familiar with the Saga to skip this note and to come back to this note once he/she has reached the section presenting the Saga). Let f denote Your knowledge about the killer before learning about Peter's alibi, and let $P[f]$ be Your pignistic probability on {Peter, Paul, Marry}. In the Saga, we have $P[f]$(Peter) = 1/4, $P[f]$(Paul) = 1/4, $P[f]$(Marry) = 1/2. Suppose You want to bet on Paul versus Marry given the hypothetical (conditioning) event that Peter is not the killer. Probabilities would be $P[f]$(Paul|not Peter) = 1/3, $P[f]$(Marry|not Peter) = 2/3. You must use these probabilities in order to avoid a Dutch Book. Indeed, You must be ready to face bets that are canceled if Peter is the killer or if Peter is not the killer, as well as bets that are canceled if Paul is or is not the killer, etc... Once You learn about Peter's alibi, the bets that would be canceled if Peter is the killer cannot be considered anymore. In that case the arguments based on cancellable bets that lead to Bayes' rule do not apply anymore when it comes to the fact 'Peter is not the killer'. You 'revise' Your probabilities and get $P[f,\text{not Peter}](\text{Paul}) = 1/2$, what is different from the previous $P[f](\text{Paul}|\text{not Peter})$. In the TBM, revision and conditioning are different. Revision is performed by Dempster's rule of conditioning whereas conditioning is done by Bayes rule applied to the pignistic probabilities. Too bad the word 'conditioning' was used to describe Dempster's rule, but the tradition is very strong. It would have been better to call it 'Dempster's rule of revision'.□

Suppose Your beliefs on H are represented by the probability function P^H. We propose the next definition:

1. the 'factual revision on f' is the process that transforms P^H into $P^H[f]$,
2. the 'hypothetical conditioning on e' is the process that transforms P^H into $P^H(.|e)$.

They correspond to the 'given facts' and the 'conditioning' processes, respectively. The value of $P^H(.|e)$ is obtained by Bayes' rule of conditioning. The value of $P^H[f]$, as such, is still to be defined. For that purpose and staying within the probability context, Hacking introduces the next assumption, called the 'dynamic assumption':

$$Prob^H[f,e](h) = Prob^H[f](h|e) = \frac{Prob^H[f](he)}{Prob^H[f](e)}$$

i.e., You equate Your conditional probabilities to Your probabilities given facts.

A Bayesian probabilist is then defined as a personalist probabilist who accepts the dynamic assumption. For a Bayesian, the distinction between hypothetical conditioning and factual revision is not important. Jeffrey (1988) and Teller (1976) produce arguments based on so-called dynamic Dutch books in order to justify the adequacy of the dynamic assumption.

In the TBM, the dynamic assumption is not assumed, therefore revisions and conditioning are not identical. Nevertheless the TBM resists to the dynamic Dutch books as shown in (Smets, 1993c) (see also section 5.8).

4 Various interpretations attached to a belief function.

The expression 'Dempster-Shafer theory' is confusing because it covers several models. Usually the expression 'Dempster-Shafer theory' concerns mathematical models dealing with uncertainty. It corresponds to a 'package' made of at least three components: a belief function and two rules called Dempster's rule of conditioning and Dempster's rule of combination. Unfortunately often authors omit to state exactly what the belief function is supposed to quantify, and most confusion encountered in the literature about belief functions results from the fact people shift between definitions.

4.1 The upper and lower probability model.

The strength of the agent's opinion is quantified by a probability function P defined on Ω. This function P is only known to belong to a family of probability functions defined on Ω. This family is denoted Π. Then P_* is defined as the lower envelope and P^* as the upper envelope of Π.

$$\begin{aligned} P_*(A) &= \min\{P(A) : P \in \Pi\}, \\ P^*(A) &= \max\{P(A) : P \in \Pi\}. \end{aligned}$$

The lower envelope does not always characterizes the Π family. We will restrain ourselves to the classical case where $P_*(A) = \inf\{P(A) : P \in \Pi\}$. In that case, there is a one-to-one correspondence between P_* and Π, where Π is then equal to the set of probability functions compatible with P_* (section 2.3):

$$\Pi = \{P : P_*(A) \leq P(A) \leq P^*(A),\ \forall A \subseteq \Omega\}.$$

The constraints on P^* are in fact redundant as $P^*(A) = P_*(\Omega) - P_*(\overline{A})$. They can thus be neglected.

In general, the lower envelope P_* of a convex family of probability functions is not a belief function (not even capacities monotone of order 2). Nevertheless, for some convex families, the lower envelope happens to be a belief function, and this is the kind of families considered by those using the interpretation of *bel* as a lower probability.

In that model, for every $A \subseteq \Omega$, there exists a value $P(A)$ that quantifies the strength of the opinions held by the agent that the actual world belongs to A. But the function P is not known, and we only know that the function P belongs to a convex family Π the lower envelope of which is a (normalized) belief function.

In this context, the dynamic assumption is accepted (section 3.5). So conditioning and revision become synonymous. The appropriate rule to represent conditioning is the natural extension rule ((Walley, 1991; Jaffray, 1992; Fagin & Halpern, 1991). Suppose the conditioning event is A, i.e., You know for sure that $\omega_0 \in A$. Each P in Π is conditioned on A by Bayes rule. This conditioning procedure induces a new family of (conditional) probability functions, denoted Π_A, where:

$$\Pi_A = \{P[A] : P[A](X) = P(X \cap A)/P(A), \forall X \subseteq \Omega, P \in \Pi\}.$$

This family is convex. When the lower envelope of Π is a belief function, then the lower envelope of Π_A, denoted $bel[A]$, is also a belief function (Jaffray, 1992; Fagin & Halpern, 1991). It satisfies:

$$bel[A](X) = \frac{bel(X \cap A)}{bel(X \cap A) + pl(\overline{B} \cap A)}.$$

Jaffray (1992) presents the relation between the basic belief assignment related to *bel* and the basic belief assignment related to $bel[A]$.

In that upper and lower probability context, Dempster's rule of conditioning can be derived in very special cases (Moral & de Campos, 1990; Gilboa & Schmeidler, 1992)). Suppose the conditioning event is A, and we apply Bayes conditioning rule only to the probability functions in Π that satisfy $P(A) = pl(A)$. Then the lower envelope of the family Π^*_A of conditional probability functions so built is equal to the belief function obtained by applying Dempster's rule of conditioning to *bel*, the lower envelope of Π.

$$\Pi^*_A = \{P[A] : P[A] \in \Pi_A, P(A) = \max\{P'(A) : P' \in \Pi\}\},$$

$$\min\{P[A](X) : P[A] \in \Pi^*_A\} = bel(X \cup \overline{A}) - bel(\overline{A}),$$

where

$$bel(X) = \min\{P(X) : P \in \Pi\}.$$

Nevertheless, this restricted conditioning process hardly fits with the general idea of conditioning and that it would be the adequate way to condition in the upper and lower probability context is still an open question. In fact, Kyburg (1987a) and Voorbraak (1991) both criticize Dempster-Shafer theory as inadequate because Dempster's rule of conditioning leads to incorrect results for conditioning. But their implicit interpretation of belief functions was the upper and lower probabilities model, a context where indeed Dempster's rule of conditioning should not be applied (or at least its use should then be clearly justified).

In the upper and lower probability context, Dempster's rule of combination is not used as it does not seem to correspond to any realistic process.

In fact Kyburg and Voorbraak do not consider the explicit existence of some unknown probability function. They just assume that a family Π of probability functions characterizes the agent beliefs and they defend that the conditioning process be defined by the natural extension rule.

That *bel*, defined as the lower envelope of Π, represents 'degrees of belief', as admitted by these authors, cannot be answered as the concept of a 'degree of belief' is hardly uniquely defined.

4.2 Dempster's model and the hint model of Kohlas and Monney.

Dempster (1967, 1968, 1972) studied belief functions while trying to solve the problem of fiducial inference. Dempster's approach assumes two finite spaces X and Y, a probability measure P^X on X, and a one-to-many mapping $\Gamma : X \to 2^Y$ from X to the power set of Y. So P^X induces 'random sets' on Y, and all that can be stated about the probability $P^Y(y_0 \in A)$ that the actual (but unknown) value y_0 of Y is in $A \subseteq Y$ is that

$$P^Y(y_0 \in A) \in [bel^Y(A), pl^Y(A)]$$

where

$$
\begin{aligned}
bel^Y(A) &= \sum_{\Gamma(x)\subseteq A} P^X(x), \\
pl^Y(A) &= \sum_{\Gamma(x)\cap A\neq\emptyset} P^X(x).
\end{aligned}
$$

.

It is mathematically true that for any belief function bel^Y on Y, one can describe a space X, a probability measure P^X on X and a mapping $\Gamma : X \to 2^Y$ that together generate the belief function bel^Y. If the X domain, the P^X and the Γ mapping have a meaningful interpretation, then the model can be appropriate.

It can be shown that the model can be reduced to the previous upper and lower probability model. Suppose we knew the values of the conditional probability functions $P^Y[x]$ defined on Y for every $x \in X$. Then:

$$P^Y(A) = \sum_{x\in X} P^Y[x](A)P^X(x).$$

The conditional probability functions happen to be unknown, except for the constraint that $P^Y[x](y) = 0$ if $y \notin \Gamma(x)$. Let Π be the family of all conditional probability functions on Y that satisfy the last constraints. Then:

$$bel^Y(A) = \min\{P^Y(A) = \sum_{x\in X} P^Y[x](A)P^X(x) : P^Y[x] \in \Pi\}$$

So in a certain sense, there is a mathematical link between Dempster's model and the upper and lower probability model. That this link corresponds to some reality depends on the application to which the model is applied.

The hint model of Kohlas and Monney (1994) is a full-grown example of Dempster's model. The x's are assumptions, P^X is a probability measure that expresses the agent's beliefs about which assumption x holds, and Y is a space of hypotheses. The Γ mapping expresses the fact that assumption x is a support for the set of hypotheses $\Gamma(x)$.

Kohlas and Monney assume Dempster's original structure (X, P^X, Γ, Y) where X, Y, P^X and Γ are as defined above. They assume the existence of a question whose answer is unknown. The set Y is the set of possible answers to the question. One and only one element of Y is the correct answer to the question. The goal is to make assertions about the answer in the light of the available information. We assume that this information allows for several different interpretations, depending on some unknown circumstances. These interpretations are regrouped into the set X and there is exactly one correct interpretation. Not all interpretations are equally likely and the known probability measure P^X on X reflects our information in that respect. Furthermore, if the interpretation x∈X is the correct one, then the answer is known to be in the subset $\Gamma(x) \subseteq Y$. The structure (X, P^X, Γ, Y) is called a hint. An interpretation $x \in X$ supports the hypothesis $H \subseteq Y$ if $\Gamma(x) \subseteq H$ because in that case the answer is necessarily in H. The degree of support of H, denoted $sp(H)$, is defined as the probability of all supporting interpretations of H (Kohlas & Monney, 1994, page vi), with a few adaptations). The sp function is the bel function studied in this paper.

The hint theory corresponds to Dempster's original approach. They call their measure a degree of support, instead of belief, to avoid personal, subjective connotation, but degrees of support and degrees of belief are mathematically equivalent and conceptually very close. In the hint theory, the primitive concept is the hint from which degrees of supports are deduced, whereas, in the transferable belief model, the primitive concept is the degree of belief.

Conditioning in the hint model must be handled with care (Smets, 1993b). One can:

1. condition on the fact that the actual value x_0 of X is in $A \subseteq X$, in which case P^X is conditioned on A according to Bayes' rule. Then:

$$bel^Y[x_0 \in A](H) = \frac{bel^Y(\Gamma(A) \cap H)}{bel^Y(\Gamma(A)}, \ \forall H \subseteq Y.$$

This rule is known as the geometrical rule of conditioning (Shafer, 1976b; Suppes & Zanotti, 1977).

2. condition on the fact that the actual value y_0 of Y is in $H' \subseteq Y$. Then:

$$pl^Y[H'](H) = \frac{pl^Y(H \cap H')}{pl^Y(H')},$$

i.e., Dempster's rule of conditioning.

Decision process in the hint model is not an essential element of the theory which is more focussed on reasoning under uncertainty than deciding. Apparently no specific solutions seems to have been advanced in such a context, except for the work of Schaller (1991). It is not obvious which, if any, of the pignistic transformation or the upper and lower expectation solution is applicable. We will not further study how decision should be made in that model. We refer to Schaller (1991).

4.3 The transferable belief model of Smets.

The TBM (for transferable belief model) provides a model for the representation of quantified beliefs. The value $bel(A)$ represents the agent's belief that the actual world belongs to $A \subseteq \Omega$. No concept of probability measure underlies the description of the TBM. A study of the rationality properties that should be satisfied by a function which purpose is to quantify someone's beliefs leads to the use of belief functions (Smets, 1997, 1993d). These axiomatic studies lead also to the derivation of Dempster's rule of conditioning. From this construction, we have derived (and often justified) many other concepts like the conjunctive combination rule (that is essentially equal to Dempster's rule of combination), the disjunctive rule of combination, the specialization concept, the least commitment principle, the cautious combination rule, the measure of information content, the concept of doxastic independence, ...

The core of the TBM is essentially inspired by what is described in Shafer's book (note that some of Shafer's later papers enhance other interpretations).

The TBM is a model to represent quantified beliefs free from any probability connotation. It is based on the assumption that beliefs manifest themselves at two mental levels: the credal level where beliefs are entertained and the pignistic level where beliefs are used to make decisions.

Usually these two levels are not distinguished and probability functions are used to quantify beliefs at both levels. The justification for the use of probability functions is usually linked to "rational" behavior to be held by an ideal agent involved in some decision contexts (Ramsey, 1964; Savage, 1954; DeGroot, 1970). This result is accepted here, except that these probability functions quantify the uncertainty only when a decision is really involved.

In the TBM, we assume that the pignistic and the credal levels are distinct which implies that the justification for using probability functions at the credal level does not hold anymore (Dubois, Prade, & Smets, 1996). At the credal level, beliefs are represented by a belief function; at the pignistic level, this belief function induces a probability function that is used to make decision. This probability function should not be understood as representing Your beliefs, it is nothing but the additive measure needed to make decision, i.e., to compute the expected utilities. Of course this probability function is directly induced by the belief function representing Your belief at the credal level. The link between the two levels is achieved by the pignistic transformation that transforms a belief function into a probability function. Its nature and justification is detailed in section 5.

In the TBM, the basic belief assignment receives a natural interpretation. For $A \subseteq \Omega$, $m(A)$ is that part of Your belief that supports A, i.e., that the actual world ω_0 belongs to A, and that, due to lack of information, does not support any strict subset of A.

If some further pieces of evidence become available to You and You accept them as valid, and if their only impact bearing on Ω is that they imply that the actual world ω_0 does not belong to $\overline{B}$, then the mass $m(A)$ initially allocated to A is transferred to $A \cap B$. Indeed, some of Your belief (quantified by $m(A)$) was allocated to A, and now You accept that $\omega_0 \notin \overline{B}$, so that mass $m(A)$ is transferred to $A \cap B$ (hence the name of the model). The resulting new basic belief assignment is the one obtained by the application of Dempster's rule of conditioning .

The degree of belief $bel(A)$ quantifies the total amount of justified specific support given to A. It is obtained by summing all basic belief masses given to subsets $X \subseteq \Omega$ with $X \subseteq A$ (and $X \neq \emptyset$). Indeed a part of belief that supports that the actual world ω_0 is in B also supports that ω_0 is in A whenever $B \subseteq A$. So for all $A \subseteq \Omega$,

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

We say justified because we include in $bel(A)$ only the basic belief masses given to subsets of A. For instance, consider two distinct elements x and y of

Ω. The basic belief mass $m(\{x, y\})$ given to $\{x, y\}$ could support x if further information indicates this. However given the available information the basic belief mass can only be given to $\{x, y\}$. We say specific because the basic belief mass $m(\emptyset)$ is not included in $bel(A)$ as it is given to the subset $\emptyset$ that supports not only A but also $\overline{A}$.

The degree of plausibility $pl(A)$ for $A \subseteq \Omega$ quantifies the maximum amount of potential specific support that could be given to A. It is obtained by adding all the basic belief masses given to subsets X compatible with A, i.e., such that $X \cap A \neq \emptyset$:

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

We say potential because the basic belief masses included in $pl(A)$ could be transferred to non-empty subsets of A if new information could justify such a transfer. It would be the case if we learn that $\overline{A}$ is impossible.

The plausibility function pl is just another way of representing the information contained in bel and could be as well forgotten, except it often provides a mathematically convenient alternative representation of the beliefs.

4.4 Probability functions extended to modal propositions.

Ruspini (1986, 1987) has proposed to consider $bel^{\Omega}_{Y,t}(A)$ as the probability that the agent Y knows at time t that A holds. Pearl (1988) proposed to understand it as the probability that A is provable. In both cases, we assume a set of worlds to which probabilities are attached. We can write

$$bel^{\Omega}(A) = P(\Box A) = P(\{w : w \in \Omega, w \models \Box A\}).$$

where the $\Box$ operator denotes the modal operator which meaning, here, can be seen as 'knowing' or 'proving'.

The static analysis of the $P(\Box A)$ function where A is a non-modal proposition shows that it is indeed a belief function over the set of non-modal propositions (see also Tsiporkova *et al.* (1999b, 1999, 1999a, ?)). The dynamic analysis allows to derive both the geometrical rule of conditioning and Dempster's rule of conditioning, the first after conditioning on a subsets of worlds that entail a given proposition, the second by a minimal chance of the accessibility relation described in Kripke semantic for modal logic. We are not aware of any particular solution proposed within such frameworks to make decisions. They will not be further studied in this paper.

5 Decision making in the TBM framework.

5.1 The origin of the pignistic probability function.

Let Ω be a finite set of worlds, and ω_0 denote the actual world. Let bel^{Ω} denote Your belief about the actual value of ω_0. When a decision must be made that depends on ω_0, You construct a probability function P^Ω on Ω in order to select the optimal decision, i.e., the one that maximizes the expected utility. We assume that P^Ω is a function of the belief function bel^Ω. It translates the saying that beliefs guide our actions. Hence one must transform bel^Ω into a probability function that will be used for selecting the best decision.

Let F be the betting frame, i.e., the frame on which decisions must be made. It is the set of alternatives in Ω on which we must build the probability function. Here, the betting frame F will be equal to Ω, but later on (section 5.4), we will see that F and Ω can be different.

Let $BetP^F$ denote the probability function on F needed for selecting the optimal decision (Bet is for betting and P for probability). The transformation between bel^Ω and $BetP^F$ that we will derive will be called the pignistic transformation. It is denoted by Γ_F. So

$$BetP^F = \Gamma_F(bel^\Omega).$$

We call $BetP^F$ a pignistic probability to insist on the fact that it is only a probability measure used to make decisions and not a probability function that represents Your beliefs. Of course $BetP^F$ is mathematically a classical probability measure on F.

The structure of the pignistic transformation is derived from the rationality requirement that underlies the following scenario.

Example 5.1. Buying Your friend's drink. Suppose You have two friends, Glen (G) and Judea (J). You know they will toss a fair coin and the winner will visit You tonight. You want to buy the drink Your friend would like to have tonight: coke, wine or beer. You can only buy one drink. Let $D = \{\text{coke, wine, beer}\}$ and $F = D$.

Let $bel^D[G]$ quantify Your belief about the drink Glen will ask for, should he come. Given $bel^D[G]$, You build the pignistic probability $BetP^D[G]$ about the drink Glen will ask by applying the (still to be defined) pignistic transformation. You build in the same way the pignistic probability $BetP^D[J]$ based on $bel^D[J]$, Your belief about the drink Judea will ask for, should he come. The two pignistic probability distributions $BetP^D[G]$ and $BetP^D[J]$ are the conditional probability distributions about the drink that will be asked for given Glen or Judea comes. The pignistic probability distributions $BetP^D$ about the drink that Your visitor will ask for is then:

$$BetP^D(d) = .5BetP^D[G](d) + .5BetP^D[J](d),\ \forall d \in D.$$

You will use the pignistic probability function $BetP^D$ to decide which drink to buy.

But You might as well reconsider the whole problem and first compute Your belief bel^D about the drink Your visitor (V) would like to have. We have shown (Smets, 1997) that bel^D is given by:

$$bel^D(d) = .5bel^D[G](d) + .5bel^D[J](d),\ \forall d \subseteq D.$$

Given bel^D, You could then build the pignistic probability $BetP^D$ You should use to decide which drink to buy. It seems reasonable to assume that both solutions must be equal. This requirement is called the linearity assumption and formally defined below. In such a case, the pignistic transformation is uniquely defined. □

5.2 Deriving the pignistic transformation.

Formally, we require that the pignistic transformation satisfies the following assumptions.

Proposition 5.1 ***Credal-Pignistic Link.*** *Let F be a finite set and let bel^F be a belief function defined on F. Let $BetP^F$ be a probability function on F. For all $\omega \in F$,*

$$BetP^F(\omega) = \Gamma_F(bel^\Omega)(\omega).$$

Axiom 5.1 translates the idea that our beliefs guide our behaviors. The function Γ_F is called the pignistic transformation. Evaluation of $BetP^F(A)$ for $A \subseteq F$ is obtained by adding the probabilities $BetP^F(\omega)$ for $\omega \in A$.

Proposition 5.2 ***Linearity.*** *Let bel_1 and bel_2 be two belief functions on the frame of discernment F. Let Γ_F be the pignistic transformation that transforms a belief function over F into a probability function $BetP^F$ over F. Then Γ_F satisfies, for any $\alpha \in [0,1]$,*

$$\Gamma_F(\alpha bel_1 + (1-\alpha)bel_2) = \alpha\Gamma_F(bel_1) + (1-\alpha)\Gamma_F(bel_2)$$

The origin of the linearity requirement was explained in section 5.1. Technical assumptions must be added that are hardly arguable:

Proposition 5.3 ***Efficiency.*** $BetP^F(\Omega) = 1$.

Proposition 5.4 ***Anonymity.*** *Let R be a permutation function from F to F. The pignistic probability given to the image of $A \subseteq F$ after permutation of the elements of F is the same as the pignistic probability given to A before applying the permutation:*

$$BetP^{F*}(R(A)) = BetP^F(A), \ \forall A \subseteq F,$$

where $BetP^{F*}$ is the pignistic probability function on $F*$ after applying the permutation function.

Proposition 5.5 ***Impossible Event.*** *The pignistic probability of an impossible event is zero.*

Proposition 5.6 ***Projectivity.*** *If bel^F happens to be a probability function P defined on F, then $\Gamma_F(P) = P$.*

Proposition 5.3 tells that the pignistic probabilities given to the elements of F add to one. Proposition 5.4 states that renaming the elements of Ω does not change the pignistic probabilities. Proposition 5.5 tells that the pignistic probabilities given to a subset is not changed when one adds to it any element known to be impossible. Proposition 5.6 recognizes that if someone's belief is already described by a probability function, then the pignistic probabilities and the degrees of belief are numerically equal.

Under these assumptions, it is possible to derive uniquely Γ_F. The proof can be found in Shapley (1953).

Theorem 5.1 ***Pignistic Transformation Theorem.*** *Let bel^F be a belief function on space F and m^F its related bba. Let $BetP^F = \Gamma_F(bel^F)$. The only solution $BetP^F$ that satisfies propositions 5.1 to 5.6 is:*

$$BetP^F(\omega) = \sum_{A \subseteq F, \omega \in A} \frac{1}{|A|} \frac{m^F(A)}{(1 - m^F(\emptyset))}, \ \forall \omega \in F, \tag{2}$$

where $|A|$ is the number of elements of Ω in A, and

$$BetP^F(A) = \sum_{\omega \in A} BetP^F(w), \ \forall A \subseteq F.$$

.

It is easy to show that the function $BetP^F$ obtained from (2) is indeed a probability function.

Relation (2) can be also expressed as a function of bel^F instead of m^F. It becomes:

$$BetP^F(\omega) = \sum_{A \subseteq \overline{\omega}} \frac{|A|! \ (|\Omega| - |A| - 1)!}{|\Omega|!} \ \frac{bel^F(w \cup A) - bel^F(A)}{bel^F(\Omega)}.$$

As such, this equation is hardly useful but it enhances that the pignistic transformation produces a probability function whenever bel^F is monotone for inclusion (hence a capacity monotone of order 1).

Example 5.2. A pignistic transformation. Let $F = \{a, b, c\}$, and $m^F(\{a\}) = .3$, $m^F(\{a, b\}) = .2$, $m^F(\{b, c\}) = .2$, and $m^F(\{a, b, c\}) = .3$. Then $BetP^F(a) = .3 + .2/2 + .3/3 = .5$, $BetP^F(b) = .2/2 + .2/2 + .3/3 = .3$ and $BetP^F(c) = .2/2 + .3/3 = .2$. We also have, a.o., $BetP^F(\{a, b\}) = .5 + .3 = .8$. □

The pignistic probability so derived fits with common sense. It consists in distributing every basic belief mass equally among the elements that belong to its focal element. In (Smets & Kennes, 1994), we had called this property the generalized insufficient reason principle. It was a very unfortunate name as some readers thought that the pignistic transformation is *justified* by a generalization of the insufficient reason principle. It is not the case. The insufficient reason principle, as such, is not an acceptable rationality principle as it is responsible for most contradictions encountered in probability theory. The reason why the TBM escapes from these contradictions is that the betting frame F is established before applying the transformation. All contradictions encountered when applying the insufficient reason principle result from the fact the betting frame (or its equivalent) can vary, what leads to contradictions. Had it been fixed, the insufficient reason principle would not be prone to contradictions and would become a quite valid principle.

5.2.1 Historical notes.

1. In a context similar to ours, Shapley (1953) derived the same relation (2). The model he derived was later called the 'transferable utility model' (Roth, 1988) whereas, unaware of it but amazingly, we called our model the 'transferable belief model'.

2. The solution derived from the pignistic transformation was already proposed in (Dubois & Pradre, 1982; Williams, 1982) as 'natural' solutions but without justification.

5.3 The case of two independent frames.

Let bel^{Ω} be a belief function on a frame of discernment Ω. Let Ω' be a coarsening of Ω, i.e., the elements of Ω' are the elements of a partition of Ω. Let $bel^{\Omega\downarrow\Omega'}$ denote the belief function defined on Ω' such that:

$$bel^{\Omega\downarrow\Omega'}(A) = bel^{\Omega}(A),\ \forall A \subseteq \Omega`.$$

Marginalization is a special case of coarsening: for example let $\Omega = X \times Y$ and $\Omega` = X$.

Suppose a belief function $bel^{X\times Y}$ on the frame of discernment $X \times Y$ which bba $m^{X\times Y}$ satisfies:

$$\begin{aligned} m^{X\times Y}(w) &= m^{X\times Y\downarrow X}(x) m^{X\times Y\downarrow Y}(y),\ \text{if } w = x \times y \text{ for } x \subseteq X, y \subseteq Y, \\ &= 0, \text{otherwise}. \end{aligned}$$

In that case, X and Y are said to be non-interactive under $bel^{X\times Y}$. In probability theory, non-interactivity and stochastic independence are equivalent. In the TBM, this equivalence is more delicate (Ben Yaghlane, Smets, & Mellouli, 2000).

Consider the pignistic probabilities on the spaces $X \times Y$, and those on X and on Y. They will be related as if the two variables were stochastically independent.

Theorem 5.2 *Let X and Y be two non-interactive variables under $bel^{X\times Y}$. Then, $\forall x \in X, \forall y \in Y$:*

$$\Gamma_{X\times Y}(bel^{X\times Y})(x, y) = \Gamma_X(bel^{X\times Y\downarrow X})(x)\Gamma_Y(bel^{X\times Y\downarrow Y})(y),$$

or equivalently:

$$BetP^{X\times Y}(x,y) = BetP^X(x)BetP(^Y(y).$$

This theorem states that the pignistic probabilities computed on the elements of the $X \times Y$ space are equal to the product of the pignistic probabilities computed on their marginals. This is of course a rephrasing of the major property encountered in probability theory under stochastic independence. This result is important, as its absence would have raised serious doubts about the validity of the pignistic transformation.

It was hoped that this constraint would be only satisfied by the pignistic transformation, in which case we would have obtained a second justification for its use. But this is not the case. Indeed, if the probability needed at the pignistic level was defined on the elements of F as proportional to the plausibility given to these elements, then the last relation of theorem 5.2 would be also satisfied.

5.4 The betting frame.

The pignistic transformation depends on the structure of the frame on which the decision must be made. One must first define the 'betting frame' F, i.e., the set of atoms on which stakes will be allocated. The granularity of this frame F is defined so that a stake could be given to each element of F independently of the stakes given to the other elements of F. Suppose one starts with a belief function bel^Ω on a frame Ω. If the stakes given to elements A and B of Ω must necessarily be always equal, both A and B belong to the same granule of the betting frame F. For instance, suppose You want to bet on the fact that Michel lives in Paris or in Brussels. Let M = 'Michel lives in Paris' and T = 'it rains now in Tokyo'. Let $\Omega = \{M\&T,\ M\&\neg T,\ \neg M\&T,\ \neg M\&\neg T\}$. There are this four worlds in Ω. Nevertheless, the stakes put on the worlds $M\&T$ and $M\&\neg T$ are by 'necessity' equal as we bet only on Michel's home town. So $M\&T$ and $M\&\neg T$ are in the same granule of the betting frame, and F has only two elements: $F = \{M,\ \neg M\}$.

The betting frame F is organized so that the granules of Ω are the elements of F. The problem is then to define bel^F on F given bel^Ω. This is achieved by applying a sequence of coarsenings and/or refinements on Ω. The only constraints that must hold between F and Ω is that the frames F and Ω are compatible, i.e., F and Ω have a common refinement (Shafer, 1976a page 114). The pignistic probability $BetP^F$ is then built from the belief function bel^F so derived on F from bel^Ω. Therefore we can write $BetP^F = \Gamma_F(bel^\Omega)$ where it is understood that the Γ_F operator not only transforms a belief function into a probability function but it also transforms bel^Ω into bel^F (provided F and Ω are compatible, a constraint usually satisfied).

5.4.1 Betting under total ignorance.

To show the power of our approach, let us consider the next examples based on total ignorance and which solution in probability theory is often considered as disturbing.

Example 5.3. Betting and total ignorance. Consider a guard in a huge power plant. On the emergency panel, alarms A_1 and A_2 are both on. The guard never heard about these two alarms, they were hidden in a remote place. He takes the instruction book and discovers that alarm A_1 is on iff circuit C is in state C_1 or C_2 and that alarm A_2 is on iff circuit D is in state D_1, D_2 or D_3. He never heard about these C and D circuits. Therefore, his beliefs on the C circuit will be characterized by a 'vacuous' belief function on space $\Omega_C = \{C_1, C_2\}$, i.e., a belief function which bba satisfies $m^{\Omega_C}(\Omega_C) = 1$ (this particular belief function is the one that represents the state of total ignorance, as the only supported subset is the whole space Ω itself). By the application of (2) his pignistic probabilities will be given by

$$BetP^{\Omega_C}(C_1) = BetP^{\Omega_C}(C_2) = 1/2.$$

Similarly for the D circuit, the guard's belief on space $\Omega_D = \{D_1, D_2, D_3\}$ will be vacuous, i.e., $m^{\Omega_D}(\Omega_D) = 1$, and the pignistic probabilities are

$$BetP^{\Omega_D}(D_1) = BetP^{\Omega_D}(D_2) = BetP^{\Omega_D}(D_3) = 1/3.$$

Now, by reading the next page on the manual, the guard discovers that circuits C and D are so made that whenever circuit C is in state C_1, circuit D is in state D_1 and vice-versa. So he learns that C_1 and D_1 are equivalent (given what the guard knows) and that C_2 and (D_2 or D_3) are also equivalent as C is either C_1 or C_2 and D is either D_1 or D_2 or D_3. In the TBM, this information does not modify his belief about which circuit is broken.

If the guard had been a trained Bayesian, he would have assigned value for $P^{\Omega_C}(C_1)$ and $P^{\Omega_D}(D_1)$ (given the lack of any information, they would probably be 1/2 and 1/3, but any value could be used). Once he learns about the equivalence between $C_1$and D_1, he must adapt his probabilities as they must give the same probabilities to C_1 and D_1. Which set of probabilities is he going to update: P^{Ω_C} or P^{Ω_D}, and why?, especially since it must be remembered that he has no knowledge whatsoever about what the circuits are. In a probabilistic approach, the difficulty raised by this type of example results from the requirement that equivalent propositions should receive identical beliefs, and therefore identical probabilities.

Within the transferable belief model, the only requirement is that equivalent propositions should receive equal beliefs (it is satisfied as $bel^{\Omega_C}(C_1) = bel^{\Omega_D}(D_1) = 0$). Pignistic probabilities depend not only on these beliefs but also on the structure of the betting frame. The difference between $BetP^{\Omega_C}(C_1)$ and $BetP^{\Omega_D}(D_1)$ reflects the difference between the two betting frames. □

The fact the TBM can cope easily with such states of ignorance results from the partial dissociation between the credal and the pignistic levels. Bayesians do not consider such a distinction and therefore work in a more limited framework, hence the difficulty they encounter in the present example.

5.4.2 Undefined betting frame.

We consider now the problem where the betting frame is ill defined. Suppose bel^Ω is a belief function on a frame Ω and let F be a betting frame compatible with Ω. Let $BetP^F$ be the pignistic probability obtained by applying the pignistic transformation Γ_F to bel^Ω:

$$BetP^F = \Gamma_F(bel^\Omega).$$

Suppose another betting frame F' compatible with Ω. We build $BetP^{F'} = \Gamma_{F'}(bel^\Omega)$ in a way similar to the one used in the previous case. Consider now the family F_A made of all the betting frames F_1, $F_2 \ldots$ compatible with Ω and such that $A \subseteq F_i$ for all i. Compute $BetP^{F_i}(A)$ for all $F_i \in F_A$. Wilson (1993) shows that the minimum of $BetP^{F_i}(A)$ taken over the $F_i \in F_A$ is equal to $bel^\Omega(A)$:

$$\min\{BetP^{F_i}(A) : F_i \in F_A\} = bel^\Omega(A).$$

and that the set of pignistic probabilities $BetP^{F_i}$ that can be obtained from bel^Ω by varying the betting frame F_i is equal to the set $\Pi(bel^\Omega)$ of probability functions compatible with bel^Ω (see section 2.3).

So whatever the betting frame F compatible with Ω such that $A \subseteq F$,

$$BetP^F(A) \geq bel^\Omega(A), \ \forall A \subseteq \Omega.$$

Suppose You ignore what is the appropriate betting frame F. You nevertheless know that, for all $A \subseteq \Omega$, the lowest bound of $BetP^F(A)$ is $bel^\Omega(A)$. Therefore $bel^\Omega(A)$ can be understood as the lowest pignistic probability one could give to A when the betting frame is not fixed (Giles, 1982).

This set $\Pi(bel^\Omega)$ of probability functions compatible with a belief function bel^Ω gets a meaning within the TBM thanks to this result. It is the set of pignistic probability functions defined on betting frames F compatible with Ω that could be induced by bel^Ω when varying the betting frame. Its definition follows from bel^Ω, not the reverse as assumed by the authors who understand bel^Ω as the lower envelope of some class of probability functions. In the TBM, we derive $\Pi(bel^\Omega)$ from bel^Ω, not bel^Ω from $\Pi(bel^\Omega)$.

When betting must be done and the betting frame is totally unknown, all You know is that $BetP^?(A) \in [bel^\Omega(A), pl^\Omega(A)]$, where the ? superscript indicates that the betting frame is unknown. Such a case can be handled by using the procedure developed by Jaffray (1988) and Strat (1990b, 1990a), and detailed in section 6. In fact, if You only know that the betting frame belongs to a given subset F' of the set of all possible betting frames (and not necessarily the set of all possible betting frames), then You can compute the upper and lower bounds of $BetP^?(A)$ as

$$\begin{aligned} BetP_*(A) &= \min\{BetP^F(A) : F \in F'\} \\ BetP^*(A) &= \max\{BetP^F(A) : F \in F'\} \end{aligned}$$

Ordering these intervals is of course not obvious in general. Procedure based on weighted average of the limits of the intervals, like in the Jaffray - Strat solution, is a natural solution. Nevertheless we feel that the present problem is quite artificial (where F' comes from?), and we will not further consider it. Only the case where F' is the set of all possible betting frames is realistic, and it is discussed in section 6.

5.5 The impact of the two-level model.

In order to show that the introduction of the two-level mental model is not just an intellectual game, we present an example where the results will be different if one takes the two-level approach as advocated in the transferable belief model or a one-level model like in probability theory.

5.5.1 The Peter, Paul and Marry Saga.

Big Boss has decided that Mr. Jones must be murdered by one of the three people present in his waiting room and whose names are Peter, Paul and Marry. Big Boss has decided that the killer on duty will be selected by a throw of a dice: if it is an even number, the killer will be female, if it is an odd number, the killer will be male. You, the judge, know that Mr. Jones has been murdered and who was in the waiting room. You know about the dice throwing, but You do not know what the outcome was and who was actually selected. You are also ignorant as to how Big Boss would have decided between Peter and Paul in the case of an odd number being observed. Given the available information at time t_0, Your odds for betting on the sex of the killer would be 1 to 1 for male versus female.

At time $t_1 > t_0$, You learn that if Big Boss had not selected Peter, then Peter would necessarily have gone to the police station at the time of the killing in order to have a perfect alibi. Peter indeed went to the police station, so he is not the killer. The question is how You would bet now on male versus female: should Your odds be 1 to 1 (as in the transferable belief model) or 1 to 2 (as in the most natural Bayesian model).

Note that the alibi evidence makes 'Peter is not the killer' and 'Peter has a perfect alibi' equivalent. The more classical evidence 'Peter has a perfect alibi' would only imply P('Peter is not the killer' | 'Peter has a perfect alibi') = 1. But P('Peter has a perfect alibi' | 'Peter is not the killer') would be undefined and would then give rise to further discussion, which would be useless for our purpose. In this presentation, the latter probability is also 1.

5.5.2 The transferable belief model solution.

Let k be the killer. The information about the waiting room and the dice throwing pattern induces the following basic belief assignment $m_{t_0}^{\Omega}$:

$$k \in \Omega = \{\text{Peter, Paul, Mary}\},$$

$$
\begin{aligned}
m_{t_0}^{\Omega}(\{\text{Mary}\}) &= .5 \\
m_{t_0}^{\Omega}(\{\text{Peter, Paul}\}\ ) &= .5
\end{aligned}
$$

The basic belief mass .5 given to {Peter, Paul} corresponds to that part of belief that supports ‘Peter or Paul’, could possibly support each of them, but given the lack of further information, cannot be divided more specifically between Peter and Paul.

Let $BetP_{t_0}^{\Omega}$ be the pignistic probability obtained by applying the pignistic transformation to $m_{t_0}^{\Omega}$ on the betting frame Ω. By relation (2), we get:

$$
\begin{aligned}
BetP_{t_0}^{\Omega}(\text{Peter}) &= .25 \\
BetP_{t_0}^{\Omega}(\text{Paul}) &= .25 \\
BetP_{t_0}^{\Omega}(\text{Mary}) &= .50
\end{aligned}
$$

Given the information available at time t_0, the bet on the killer’s sex (male versus female) is held at odds 1 to 1.

Peter’s alibi induces a revision of $m_{t_0}^{\Omega}$ into $m_{t_1}^{\Omega}$ by Dempster’s rule of conditioning:

$$
\begin{aligned}
m_{t_1}^{\Omega}(\text{Mary}) &= .50 \\
m_{t_1}^{\Omega}(\text{Paul}) &= .50
\end{aligned}
$$

The basic belief mass that was given to ‘Peter or Paul’ is transferred to Paul.

Let $BetP_{t_1}^{\Omega}$ be the pignistic probability obtained by applying the pignistic transformation to $m_{t_1}^{\Omega}$ on the betting frame $\Omega*$ whose elements are Paul and Marry.

$$
\begin{aligned}
BetP_{t_1}^{\Omega *}(Paul) &= .50 \\
BetP_{t_1}^{\Omega *}(Mary) &= .50
\end{aligned}
$$

Your odds for betting on male versus female would still be 1 to 1.

5.5.3 The probabilistic solution

The probabilistic solution is not obvious as one data is missing: the value α of the probability that Big Boss selects Peter if he must select a male killer. Any value could be accepted for α, but given the total ignorance in which we are about this value, let us assume that $\alpha = .5$, the most natural solution (any value could be used without changing the problem we raise). Then the odds on male versus female before learning about Peter’s alibi is 1 to 1, and after learning about Peter’s alibi, it becomes 1 to 2. The probabilities are then:

$$
\begin{aligned}
P_{t_1}^{\Omega^*}(Paul) &= 0.33 \\
P_{t_1}^{\Omega^*}(Mary) &= 0.66.
\end{aligned}
$$

The 1 to 1 odds of the transferable belief model solution can only be obtained in a probabilistic approach if $\alpha = 0$. Some critics would claim that the transferable belief model solution is valid as it fits with $\alpha = 0$. The only trouble with this answer is that if the alibi story had applied to Paul, than we would still bet on the frame Ω^{**} ={Peter, Marry} at 1 to 1 odds within the TBM approach. Instead the probabilistic solution with $\alpha = 0$ would lead to a 0 to 1 bet, as the probabilities are:

$$\begin{aligned} P_{t_1}^{\Omega^{**}}(Peter) &= 0.00 \\ P_{t_1}^{\Omega^{**}}(Mary) &= 1.00 \end{aligned}$$

So the classical probabilistic analysis does not lead to the transferable belief model solution.

5.5.4 Which solution is 'good'?

We are facing two solutions for the bet on male versus female after learning about Peter's alibi: the 1 to 1 or the 1 to 2 odds? Which solution is 'good' is not decidable, as it requires the definition of 'good'.

Computer simulations have been suggested for solving the dilemma, but they are useless. In any finite sequence of simulations, the proportion of cases when Peter is selected when Big Boss must choose a male killer is well defined. If this proportion is close from the α used by the Bayesian, the Bayesian solution will be the best, otherwise the TBM can be the best. But this implies that we introduce a probability α equal to the probability that Peter is selected when the killer is a male. In that case the problem is no more the one we had consider in the initial story. If such an α were known, then it would have been included in the TBM analysis, and in that case the TBM and the Bayesian solutions become identical, as it should be.

So in order to compare the TBM and the Bayesian solution of the initial saga, we are only left over with a subjective comparison of the two solutions... or an in depth comparison of the theoretical foundations that led to these solutions.

Other examples illustrating the difference between the Bayesian approach and the TBM approach can be found in Smets (1994b) (see the breakable sensors example)

5.6 An operational definition of *bel*.

The pignistic transformation can be used in order to provide both an operational definition of the degrees of belief, and a method to assess them. The approach is essentially identical to the one encountered in subjective probability theory except we use the possibility to construct several betting frames (see section 3.1).

Operational definition of degrees of belief. Suppose a finite space Ω, a family of games $G = \{G_1, G_2 \ldots\}$ built on the betting frames $F_i, i = 1, 2 \ldots$, respectively, where each frame is compatible with Ω. Suppose a player and a banker. Consider one game $G_i \in G$ and its betting frame F_i. Suppose A is discerned by F_i. We have $'BetP^{F_i}_{You,t}(A) = x'$ iff You consider at time t that the player must pay \$$x$ to the banker to enter the game G_i where the player wins \$1 from the banker if the actual world belongs to A and \$0 otherwise, and You are ready to be any of the player or the banker. Consider then all possible games G_i on G. Then $bel^{\Omega}_{You,t}$ is the belief function on Ω such that:

1. $BetP^{F_i}_{You,t} = \Gamma_{F_i}(bel^{\Omega}_{You,t}),\ \forall i = 1, 2....$
2. $bel^{\Omega}_{You,t}(A) = \min\{BetP^{F_i}_{You,t}(A) : i = 1, 2\ldots\}.$

where the two cases are equivalent.

It is important to realize that the pignistic probability functions obtained with different frames are not necessarily related between them by the laws of probability. So you could bet on A versus B where $B = \overline{A}$ with pignistic probabilities of 1/2 and 1/2, and on A versus B_1 versus B_2 where $B_1 \cup B_2 = B$ with pignistic probabilities of 1/3, 1/3 and 1/3 (this is encountered in case of total ignorance on $\Omega = A \cup B$, (see (Smets & Kennes, 1994)).

5.7 The assessment of bel.

In (Smets, 1998a), we explain in detail and illustrate how the bba's can be assessed. Here, we present only the general procedure.

The assessment of a belief function is essentially obtained through a schema based on preference between gambles (see section 3.2).

The method proposed in probability theory extends directly to belief functions. It is based on using several betting frames. Let a finite set Ω and a family of compatible betting frames F_1, F_2... For each F_i, we assess $BetP^{F_i}$ using the preference ordering between two bets as done in section 3.2. We then determine the set BF_i^{Ω} of belief function on Ω which pignistic transformation on F_i is $BetP^{F_i}$. We repeat the procedure with each F_i's. The bel^{Ω} belongs to the intersection of all the BF_i^{Ω}. If the intersection is empty, then it means the pignistic probability functions are inconsistent, what ideally should not occur, but it happens of course in practice, just as in probability theory where people assess probabilities that usually violate Kolmogorof axioms. It only translates the imprecision of the assessment tool. Thanks to the fact that a belief function is defined by a finite number of values and the possibility to build as many betting frames as one needs, the intersection can be such that it contains only one belief function.

The procedure is illustrated in the next example.

Example 5.4. Assessing *bel*. Suppose $\Omega = \{a, b\}$ where a denotes 'Circuit X is broken' and b denotes 'Circuit X is not broken'. Consider the betting frame F_1 with elements a and b. Suppose Your pignistic probabilities on that frame are:

$$\begin{aligned} BetP^{F_1}(a) &= 4/9 \\ BetP^{F_1}(b) &= 5/9. \end{aligned}$$

Suppose ψ and $\overline{\psi}$ are two complementary but otherwise unknown sets of worlds that denote that some circuit C whose properties are completely unknown to You is broken or not broken, respectively. The event $a \cap \psi$ will occur if circuits X and C are broken. The event $a \cap \overline{\psi}$ will occur if circuit X is broken and circuit C is not broken. Let us consider the betting frame F_2 with elements $\{a \cap \psi, a \cap \overline{\psi}, b\}$, and suppose Your pignistic probabilities on that new frame are:

$$\begin{aligned} BetP^{F_2}(a \cap \psi) &= 7/27 \\ BetP^{F_2}(a \cap \overline{\psi}) &= 7/27 \\ BetP^{F_2}(b) &= 13/27. \end{aligned}$$

Then it can be proven that the unique solution for m^Ω is:

$$\begin{aligned} m^\Omega(a) &= 2/9, \\ m^\Omega(b) &= 3/9 \\ m^\Omega(a,b) &= 4/9. \end{aligned}$$

To show it, let m' be the bba induces by m^Ω on F_2 by the uninformative refinement:

$$\begin{aligned} m'(a \cap \psi) &= 0 \\ m'(a \cap \overline{\psi}) &= 0, \\ m'(b) &= m^\Omega(b) \\ m'(a \cap \psi, a \cap \overline{\psi}) &= m^\Omega(a), \\ m'(a \cap \psi, b) &= 0 \\ m'(a \cap \overline{\psi}, b) &= 0 \\ m'(a \cap \psi, a \cap \overline{\psi}, b) &= m^\Omega(a,b). \end{aligned}$$

The solution for m^Ω must solve two linear equations derived from (2):

$$\begin{aligned} 4/9 &= m^\Omega(a) + \frac{m^\Omega(a,b)}{2} \\ 7/27 &= m'(a \cap \psi) + \frac{m'(a \cap \psi, a \cap \overline{\psi})}{2} + \frac{m'(a \cap \psi, b)}{2} + \frac{m'(a \cap \psi, a \cap \overline{\psi}, b)}{3} \\ &= \frac{m^\Omega(a)}{2} + \frac{m^\Omega(a,b)}{3}. \end{aligned}$$

The values of m^Ω are the only one that satisfy these two equations (given their sum is 1 as they are bbm). □

It might seem odd that b receives the pignistic probabilities $\frac{5}{9}$ or $\frac{13}{27}$ according to the betting context. It reflects the fact that a large amount ($\frac{4}{9}$) of Your initial belief was left unassigned, i.e., given to {a,b}. This example corresponds to a state in which You have very weak support for a and for b. You are not totally ignorant as in example 5.4.1, but still in a state of 'strong' ignorance. Part of $BetP^{F_1}(b) = \frac{5}{9}$ is due to justified beliefs $(\frac{3}{9})$ but the remainder results from a completely unassigned part of belief that You distribute equally by the pignistic transformation among the alternatives of Your betting frame.

5.8 Kinematic Dutch books.

A classical criticism against any non probabilistic model for quantified beliefs is based on Dutch books arguments, i.e. on the possibility to build a set of forced bets such that the player will lose for sure. Dutch books are avoided only if beliefs, when used for decision making, are quantified by probability measures. The transferable belief model with its pignistic transformation resists to such a criticism. Static (also called synchronic) Dutch books are of course avoided inasmuch as bets are based on pignistic probabilities. The real delicate point is to resist diachronic Dutch books, i.e., those built when changes in beliefs are considered and bets can be reconsidered after new information has been collected by the participants in the betting game (Teller, 1973; Jeffrey, 1988).

In (Smets, 1993c), we show how the transferable belief model can resist diachronic Dutch books criticism, and how to build pignistic probabilities when the player knows that intermediate experiments will be run whose outcomes could affect the involved bets.

The overall principle is based on the decision tree. There are two types of nodes, the decision nodes and the uncertainty nodes, i.e., those usually called the 'chance' nodes, but we prefer to avoid the term 'chance' as it might create confusion. At each 'uncertainty' node, one conditions by Dempster's rule of conditioning the belief function that concerns the variable concerned by that node on all the pieces of evidence that will have been collected by the time You reach that node. You build the pignisitic probabilities on the betting frame encountered at that node. When the whole tree is built and all pignistic probabilities are established, the tree is folded back in order to select the best decision. We insist that the adequate choice of the betting frames is a delicate matter, and should be kept consistent for the whole tree.

The originality of the models we obtain in that way is that the player will use different probabilities depending on whether he knows about the experiments to be run or not. The transferable belief model analysis is rich enough to quantify the impact of the knowledge that some relevant intermediate experiments will be run, which is usually not the case within probability theory. Full details about the construction of the pignistic probabilities in a context of diachronic Dutch book are presented in (Smets, 1993c). Nevertheless the solution proposed to resist diachronic Dutch books still needs further developments. It will not be further discussed here as we focus only on static decision making.

5.9 Decision based on the maximum plausibility.

Suppose a set H of n hypothesis $\{H_1, H_2, \ldots H_n\}$. In a context where You want to select the most 'probable/likely' hypothesis $H_i \in H$, authors like Appriou (1991) propose that one should select the most plausible hypothesis, a very natural proposal. Nevertheless we feel that this solution is poised by counterexamples like the next one.

Example 5.5. Decisions based on maximal plausibility. Suppose

$$\begin{aligned} m^H(H_1) &= .30, \\ m^H(H_2) &= .01, \\ m^H(H_2, H_3, \ldots, H_{70}) &= .69. \end{aligned}$$

Then

$$\begin{aligned} pl^H(H_1) &= .30, \\ pl^H(H_2) &= .70, \\ pl^H(H_3) = \ldots = pl^H(H_{70}) &= .69. \end{aligned}$$

So H_2 is the most plausible hypothesis, and according to that approach, the one to be selected.

With the pignistic transformation, we get:

$$\begin{aligned} BetP^H(H_1) &= .30, \\ BetP^H(H_2) &= .02, \\ BetP^H(H_3) = \ldots = BetP^H(H_{70}) &= .01. \end{aligned}$$

So now H_1 is the most 'probable' hypothesis. □

We feel the solution obtained with the pignistic transformation is better than the one based on the maximum plausibility. Of course, this is a matter of personal opinion, and we don't have a way to 'prove' which solution is definitively the best one.

5.10 An example: the Oil-Drilling Problem.

In order to illustrate the use of the pignistic transformation, we present the classical Oil-Drilling example. It was analyzed in (Strat, 1994) within an upper and lower probability context. We present its analyze into the TBM context.

Example 5.6 The Oil-Drilling Problem. A wildcatter must decide whether or not to drill for oil. He is uncertain whether the hole will be dry, have a trickle of oil, or be a gusher. Drilling a hole costs \$70.000. The payoffs for hitting a gusher, a trickle, or a dry hole are \$270.000, \$120.000 and \$0, respectively. At a cost of \$10.000 the wildcatter could take seismic sounding that would help determine the underlying geological structure. The sounding will determine whether the terrain has no structure, open structure, or closed structure.

Let H be the set of hole state: $H = \{$Dry (Dr), Trickle (Tr), Gusher (Gu)$\}$. Let S be the set of terrain structures: $S = \{$No struct, Open, Closed$\}$.

The experts have provided their opinions about which state prevails given the result of the seismic soundings test (see table 1). These are represented by the bba $m^H[s]$ for $s \in S$. When $s =$ No Struct, a Dry hole is essentially expected (at level .45), but there is nevertheless some support (.40) that the the hole might also be Trickle, and there is some ignorance (.15). When $s =$ Open, there is also some support for Gusher, and when $s =$ Closed, Gusher gets some support.

Besides the experts have some a priori belief about what the outcome of the seismic soundings test might be. It is represented by the bba m^S with $m^S(\{$No struct$\}) = .5$, $m^S(\{$No struct, Open$\}) = .3$ and $m^S(\{$Open, Closed$\}) = .2$. No struct (with Open) is supported, but there is also some support for Open or Closed. The way these bba can be assesses is discussed in section 5.7.

The bba $m^H[m^S]$ (presented in the 'Marginal' column of table 1) is computed by applying the disjunctive rule of combination (Smets, 1993a). One has:

$$b^H[m^S](h) = \sum_{X \subseteq S} m^S(X) b^H[X](h), \ \forall h \subseteq H,$$

with

$$b^H[X](h) = \prod_{s \in X} b^H[s](h),$$

where the b functions are the implicability functions with,

$$b(A) = bel(A) + m(\emptyset).$$

The first relation is just the explicit solution of the next relation:

$$m^H[m^S] = (m^{H \times S} \oplus \ m^{S \uparrow S \times H})^{\downarrow H},$$

where $m^{H \times S}$ is the bba on $H \times S$ that is built from the set of conditional bba $\{m^H[s], s \in S\}$ provided by the experts. The next one results from the disjunctive combination rule which is fully detailed in (Smets, 1993a). It corresponds in building $m[E_1$ or $E_2]$ from $m[E_1]$ and $m[E_2]$, just as the conjunctive combination rule consists in computing $m[E_1$ and $E_2]$ where E_1 and E_2 are two distinct pieces of evidence.

Figure 1 presents the decision tree for the present example. Decision nodes have their expected utilities in italics. Optimal decisions are indicated in capital letters. Suppose the test result is 'no structure' and drilling in done. The pignistic probabilities $BetP^H[No\ struct] = \Gamma_H(m^H[No\ struct])$ are indicated on the three possible branches. The expected utility is then $-\$36.500 = .05 \times 190.000 + .25 \times 40.000 + .70 \times (-80.000)$. The cost if one does not drill is -\$10.000 what is better, hence the optimal decision is not to drill. Identical computations

H State	No struct	Open	Closed	Marginal
Dry	.45	.10	0	.239
Trickle	0	0	0	0
Gusher	0	0	.25	0
Dr, Tr	.40	.30	0	.289
Dr, Gu	0	0	0	.005
Tr, Gu	0	.30	.30	.033
Dr, Tr, Gu	.15	.30	.45	.435
Dry	.70	.35	.15	.530
Trickle	.25	.40	.30	.306
Gusher	.05	.25	.55	.164
$BetP^S$	.65	.25	.10	

Table 1: The upper part presents the conditional bba $m^H[s]$ on the hole state H given the terrain structure $s \in S$. The 'Marginal' column presents the bba $m^H[m^S]$ on H given the prior belief m^S with $m^S(\{\text{No struct}\}) = .5$, $m^S(\{\text{No struct, Open}\}) = .3$ and $m^S(\{\text{Open, Closed}\}) = .2$. The middle part presents the pignistic probabilities $BetP^H[s]$ over H for each $s \in S$ and $BetP^H[m^S]$ for $m^H[m^S]$. The bottom part presents the pignistic probabilities $BetP^S$ over S computed from m^S.

are performed for the the other two results of the test. In these two cases, drilling is the optimal decision. The pignistic probabilities $BetP^S = \Gamma_S(m^S)$, presented at the bottom of table 1, are used to compute the expected utility when the test is performed. Its value is \$12.825.

We then proceed with the 'no test' option. One must first compute the bba $m^H[m^S]$ on H based on the conditional bba's $m^H[s], s \in S$, and the prior bba m^S on S. The column 'Marginal' in table 1 presents this bba. It is compute by the disjunctive combination rule detailed here above. Its pignistic transformation $BetP^H[m^S] = \Gamma_H(m^H[m^S])$ are given on the three related branches, and the expected utility is \$10.970. This result is not as good as the one obtained when performing the test, hence testing is the optimal decision.

In conclusions, the decision analysis shows that it is best to test, and that if the test result is open or closed structure, the optimal decision is to drill, whereas if the test result is no structure, the optimal decision is not to drill. This example is just given in order to illustrate the kind of computation to be performed.

Similar examples are studied by Xu (1992, 1993a), (Xu, Hsia, & Smets, 1993b). She explains the architecture of the software for decision making based on the valuation based system (Shenoy, 1991, 1992). In (Xu & Smets, 1996, 1995), the authors explain how to determine the sensitivity of the conclusions to variations in the bba.

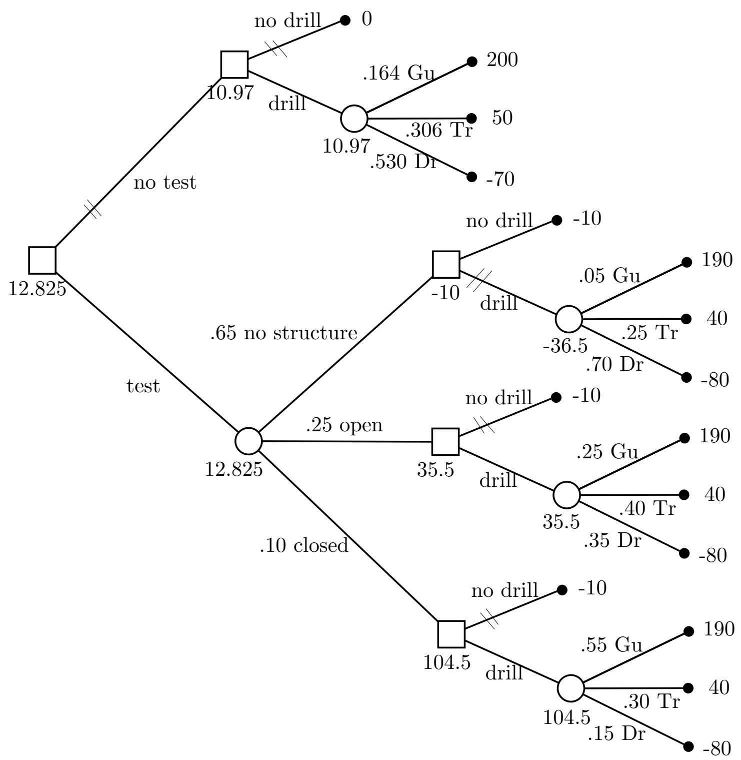

Figure 1: The decision tree for the Oil-Driling Problem, expectation in K$.

6 Decision Making in the upper and lower probabilities case.

6.1 Decisions based on upper and lower expectations.

We consider the case where there is a probability function P^Ω that describes Your belief about which world in Ω is the actual world. The closed world assumption is accepted. It happens that You don't know what is exactly P^Ω, and all You know about P^Ω is that it belongs to a family of probability functions defined on Ω. Let Π denote this family. Furthermore assume the lower envelope of Π is a belief function, what is not a necessity, but there are cases where it happens.

To make optimal decision, one must compute the expected utility of each act (and chose the best accordingly). Normally this is achieved by computing for each act a in a set A of possible acts the expected utility $\overline{u}(a)$ of that act, considering the utility $u(a, \omega)$ of each act $a \in A$ in each possible context $\omega \in \Omega$.

$$\overline{u}(a) = \sum_{\omega \in \Omega} u(a, \omega) P^{\Omega}(\omega) \ldots$$

Unfortunately $P^{\Omega}(\omega)$ is not available, as all we know is that $P^{\Omega} \in \Pi$. One approach consists in computing what would be $\overline{u}(a)$ if we consider all possible elements of Π. We can then compute the upper and the lower expectation of $\overline{u}(a)$, denoted $\overline{u}^*(a)$ and $\overline{u}_*(a)$, respectively. We have:

$$\begin{aligned} \overline{u}^*(a) &= \max\{\sum_{\omega \in \Omega} u(a, \omega) P^{\Omega}(\omega) : \ P^{\Omega} \in \Pi\}, \\ \overline{u}_*(a) &= \min\{\sum_{\omega \in \Omega} u(a, \omega) P^{\Omega}(\omega) : \ P^{\Omega} \in \Pi\}. \end{aligned}$$

So we know that $\overline{u}(a) \in [\overline{u}_*(a), \overline{u}^*(a)]$. Unfortunately, such an information does not lead to a total order, and ambiguities can appear when we have two acts a and b such that $\overline{u}_*(a) \leq \overline{u}^*(b)$ and $\overline{u}^*(a) \geq \overline{u}_*(b)$. In that case, Jaffray (1988, 1989a, 1989b, 1994) and Strat (Strat, 1990b, 1990a, 1994) have defended the suggestion to take a weighted average of the two expected utilities and select the best act using that value:

$$\overline{u}(a, \alpha) = \alpha \overline{u}_*(a) + (1 - \alpha) \overline{u}^*(a), \tag{3}$$

where α is an optimism/pessimism criteria. This solution is Hurwicz α-criteria (Luce & Raiffa, 1957).

Let m^{Ω} be the Möbius transform of P_*, the lower envelope of Π (just like the bba m^{Ω} is the Möbius transform of bel^{Ω}). It can be shown that:

$$\begin{aligned} \overline{u}^*(a) &= \sum_{X \subseteq \Omega} \max_{\omega \in X} u(a, \omega) \ m^{\Omega}(X) \\ \overline{u}_*(a) &= \sum_{X \subseteq \Omega} \min_{\omega \in X} u(a, \omega) \ m^{\Omega}(X) \end{aligned}$$

Using $\alpha = 1$ (0) means that You feel that the worst (best) possible outcome of X will occur when all You know is that $\omega \in X$. (Schubert, 1995) discusses about the choice of the value of α.(Nguyen & Walker, 1994) presents an overview of decision making with belief functions using Choquet integrals and random sets. Unfortunately, the nature of the uncertainty represented by the belief functions is not explained (see section 4).

Note: the expected utility one would obtain with the pignistic probabilities is:

$$\overline{u}(a) = \sum_{\omega \in \Omega} u(a, \omega) BetP^{\Omega}(\omega) = \sum_{X \subseteq \Omega} \sum_{\omega \in X} \frac{u(a, \omega)}{| X |} \ m^{\Omega}(X).$$

In that case, You consider that the utility when all You know is that $\omega \in X$ is the average of the utilities given to the elements of X.

Schmeidler (1989) had also proposed to chose the best act using the lower expectation (the case with $\alpha = 1$).

6.2 Justification of Jaffray-Strat's solution.

Jaffray (1988, 1989a, 1989b), (Jaffray & Wakker, 1994) presents a justification for using the weighted average of the upper and lower expected utilities as given by relation (3) based on von Neumann-Morgenstern linear utility theory.

It is assumed that the decision maker bases his/her judgment on the sole comparison of the lower probabilities on the outcome set associated with each decision. Thus ordering can be defined directly on the set $\boldsymbol{F}$ of lower probabilities on Ω. Let f_1 and f_2 be the lower envelopes of the two convex sets Π_1 and Π_2 of probability functions defined on Ω, respectively. Let $\succeq$ denote the decision maker preference ordering., with $\succ$ denoting its asymmetric part). The axioms for $\succeq$ are (adapted from (Jaffray, 1994). They are the transposition to lower probability of the standard axioms for probability (Jensen, 1967). Note that their justification in (Jaffray, 1994) requires these lower probabilities to be at least 2-monotone.

Proposition 6.1 ***Transitivity and Completeness.*** *$\succeq$ is a transitive and complete relation on $\boldsymbol{F}$.*

Proposition 6.2 ***Independence.*** *For all f_1 and f_2 in $\boldsymbol{F}$ and λ in (0,1),*

$$f_1 \succ f_2 \Rightarrow \lambda f_1 + (1-\lambda)g \succ \lambda f_2 + (1-\lambda)g.$$

Proposition 6.3 ***Continuity.*** *For all f, g, k in $\boldsymbol{F}$ such that $f \succ k \succ g$, there exist λ, μ in (0,1] such that:*

$$\lambda f + (1-\lambda)g \succ k \succ \mu f + (1-\mu)g.$$

These three propositions are necessary and sufficient conditions for the existence of a linear utility V on $\boldsymbol{F}$ representing $\succeq$, i.e., for the existence of $V : \boldsymbol{F} \rightarrow \mathbf{R}$ (where $\mathbf{R}$ is the set of reals) satisfying;

$$V(f) \geq V(g) \Leftrightarrow f \succ g.$$

and

$$V(f) = \sum_{B \subseteq \Omega} V(1_B) m(B)$$

where m is the Möbius transform of f, and 1_B is the lower probability function that satisfies $1_B(A) = 1$ if $B \subseteq A$, $= 0$ otherwise. In fact 1_B is a belief function focused on B; it represents 'all You know is that the actual world belongs to

B'. In the previous section, we had presented successively the following three solutions for V (where the act is not explicitated):

$$\begin{aligned} V^*(1_B) &= \max_{\omega \in B} u(\omega) \\ V_*(1_B) &= \min_{\omega \in B} u(\omega) \\ \overline{V}(1_B) &= \sum_{\omega \in B} u(\omega)/|B] \end{aligned}$$

If we add the next axiom, then V depends only on the upper and lower expected utilities.

Proposition 6.4 ***Dominance.*** *For all B', B" in* Ω*,*

$$V_*(1_{B'}) \geq V_*(1_{B"}) \text{ and } V^*(1_{B'}) \geq V^*(1_{B"}) \Rightarrow 1_{B'} \succeq 1_{B"}.$$

In (Cohen & Jaffray, 1985; Jaffray & Philippe, 1997), authors introduce some rationality axioms so that we get the solution based on Hurwicz α-criteria (see relation 3):

$$V(f) = \sum_{B \subseteq \Omega} (\alpha V_*(1_B) + (1-\alpha) V^*(1_B))\ m(B).$$

This derivation provides a justification for the Jaffray-Strat's solution. Note that the dominance axiom (6.4) is not satisfied in the TBM framework when using the pignistic transformation, but this should not be embarrassing as the TBM and the ULP model are not concerned with the same problem.

6.3 Ellsberg paradox revisited.

It can been defended that Jaffray-Strat's solution has the advantage over the pignistic probability solution that the decision does not depend on any possible refinement of Ω. (This property is not shared by the solution based on the pignistic transformation). That this property is really worth defending might be challenged by the next example.

Let us suppose Ellsberg's urn made out of 90 balls colored red (R), black (B) and white (W). We only know that exactly 30 balls are R. We are offer four bets listed in the table 2.

It has been argued that people prefer bet I to bet II, and bet IV to bet III. There is no probability distribution on $\Omega = \{R, B, W\}$ that is compatible with these two preferences.

With Jaffray-Strat's solution, this ordering is satisfied when $\alpha > 1/2$ (so the decision maker is somehow pessimist). With the pignistic probabilities computed from the lower probabilities, this ordering is not found, and we face the same paradox as the classical Bayesians. This might be seen as a weakness of the pignistic transformation. Nevertheless, once we refine B or W, Jaffray-Strat's solution could be criticized.

bet	R	B	W	Value ULP	Value BetP
I	1	0	0	1/3	1/3
II	0	1	0	2(1-α)/3	1/3
III	1	0	1	1/3 + 2(1-α)/3	2/3
IV	0	1	1	2/3	2/3

Table 2: Set S_1 of four bets. Columns 2 to 4 present the gains according to the color of the randomly selected ball. The 'Value ULP' column gives the value of the game using the Jaffray-Strat's solution. The 'Value BetP' column gives the value of the game using the solution based on the pignistic probabilities.

bet	R	B_1	...	B_{1000}	W	Value ULP	Value BetP
I	1	0	...	0	0	1/3	1/3
II	0	1	...	1	0	2(1-α)/3	2/3 - ϵ
III	1	0	...	0	1	1/3 + 2(1-α)/3	1/3 + ϵ
IV	0	1	...	1	1	2/3	2/3

Table 3: Set of bets S_2. Gains according to the color of the randomly selected ball.ϵ denotes 1/1001.

Suppose the set of bets S_2 where B is refined into 1000 B_i's (see table 3). When facing the set of bets S_2, we personally feel people might prefer bet II to bet I, and bet IV to bet III.

Suppose the set of bets S_3 where now W is refined into 1000 W_i's (see table 4). When facing the set of bets S_3, we personally feel people might prefer bet I to bet II, and bet III to bet IV.

It happens that in the three sets S_1, S_2 and S_3, the $\overline{u}(a)$ computed by Jaffray-Strat's formula give the same results, and therefore it does not produce the order that we have considered. Instead the pignistic probabilities are justifying the ordering for cases S_2 and S_3.

These examples only illustrate the difficulty to find 'rational' requirements that resist all criticisms. Cohen and Jaffray (1985) argue that it is 'rational' to require invariance to refinement, in which case their solution is appropriate. This sounds to be a good 'rational requirement', but its adequacy is nevertheless arguable as shown by our example.

bet	R	B	W_1	...	W_{1000}	Value ULP	Value BetP
I	1	0	0	...	0	1/3	1/3
II	0	1	0	...	0	2(1-α)/3	ϵ
III	1	0	1	...	1	1/3 + 2(1-α)/3	1 - ϵ
IV	0	1	1	...	1	2/3	2/3

Table 4: Set of bets S_3. Gains according to the color of the randomly selected ball.ϵ denotes 1/1001.

Note: In the TBM, the Ellsberg's urn experiment is to be handled with care. In Ellsberg's problem, we are facing a problem of imprecise probabilities. Each ball has a fixed color. The selection procedure of the ball obeys to a probability law (that we know only partially) and this knowledge about the selection procedure has to be taken in consideration when building Your beliefs. The belief induced by the knowledge of the structure of Ellsberg's urn is not equal to the lower probability function. In (Smets, 1994a), we explain what is the belief induced on the $\{R, B, W\}$ space by the knowledge that the selection procedure is random and the probability law that governs that selection belongs to a subset of probability functions.

6.4 The pignistic transformation within the ULP context.

Can we justify the applicability of the pignistic transformation within the ULP framework? This idea seems new. To check it, let us reconsider the example concerning the drink I have to buy for my friend (see section 5.1). Let $V = \{G, J\}$ and $D = \{c, w, b\}$ as before. Now we assume that there exists, for $v \in V$, probability functions $P^D[v]$, but their values are only known to belong to some given sets of probability functions, denoted $\Pi^D[v]$, respectively. Let us assume that these sets are uniquely characterized by their lower envelopes, denoted $P_*^D[v]$ for $v \in V$. We don't require that the lower envelopes $P_*^D[v]$ are belief functions as it is not important here.

The first approach consists in building, for each $v \in V$, the probability function $BetP^D[v]$ that will be used to make decisions. Let $BetP^D[v] = \Gamma_D(P_*^D[v])$, where Γ is still to be found. In fact all we need is to show that it satisfies the linearity constraint (see 5.2), the other assumptions being obviously satisfied.

By construction, the overall frame is $D \times V$, and $BetP^D[v]$ are the conditional probability functions over D given the visitor is v. Knowing the probability about who will be the visitor, then the probability function to be used to buy the drink is:

$$P^D = .5BetP^D[G] + .5BetP^D[J] = .5\Gamma_D(P_*^D[G]) + .5\Gamma_D(P_*^D[J]),$$

just as in the previous derivation.

Let us now consider the second approach. The probability distribution $P^{V\times D}$ on the product space $V \times D$ is known to belong to the set $\Pi^{V\times D}$ with:

$$\begin{aligned}\Pi^{V\times D} \quad = \quad & \{P^{V\times D} : P^{V\times D}(v, d) = .5P^D[v](d),\ \forall d \in D, \\ & \text{where } P^D[v] \in \Pi^D[v],\ v \in V\}.\end{aligned}$$

In order to make a decision about which drink to buy, we need the probability function P^D obtained by marginalizing on D the probabilities defined on $V \times D$. The family of probability functions so derived from $\Pi^{V\times D}$ is:

$$\Pi^D = \{P^D : P^D = .5P^D[G] + .5P^D[J], \text{ where } P^D[v] \in \Pi^D[v], v \in V = \{G, J\}\}.$$

This family Π^D is convex and its lower envelope is given by:

$$P_*^D = .5P_*^D[G] + .5P_*^D[J].$$

The needed pignistic probability function is then $BetP^D = \Gamma_D(P_*^D)$. Therefore, the equality of the two approach implies that:

$$.5\Gamma_D(P_*^D[G]) + .5\Gamma_D(P_*^D[J]) = \Gamma_D(.5P_*^D[G] + .5P_*^D[J]).$$

Generalizing to any value for the probability that G comes tonight, we get the linearity constraint, and thus the pignistic transformation seems to be justified.

7 Conclusions.

We have presented some of the procedures developed for the 'rational' decision making in the TBM context and in the ULP context. There are other proposals, but usually they are essentially *ad hoc* or proposed without justification.

We developed in full detail the pignistic transformation that provides, probably, the adequate way to build the probability function required for rational decisions. Its applicability seems not to be restricted to the TBM, and we feel it might also produce a useful tool in the upper and lower probabilities framework. Using any other probability function would violate the requirement behind the justification of the pignistic transformation. We feel such a violation must be avoided, and therefore the pignistic transformation seems to be inevitable.

These non-probabilistic theories are young. So it should not be surprising that in some future some of the proposed solutions turned out to be inadequate. Only time will tell us how robust these solutions are. We feel nevertheless confident the justifications given to the various methods are strong enough to give us hope that the proposed procedures will resist to the test of time.

References

Appriou, A. (1991). Probabilités et incertitude en fusion de données multi-senseurs. *Revue Scientifique et Technique de la Défense*, *11*, 27-40.

Ben Yaghlane, B., Smets, P., & Mellouli, K. (2000). Independence concepts for belief functions. In IPMU-2000 (Ed.), *Information processing and management of uncertainty* (Vol. 1, pp. 357–364).

Cohen, M., & Jaffray, J. Y. (1985). Decision making in a case of mixed uncertainty: a normative approach. *J. of Math. Psychology*, *29*, 428–442.

De Finetti, B. (1974). *Theory of probability.* Wiley, London.

DeGroot, M. H. (1970). *Optimal statistical decisions.* McGraw-Hill, New York.

Dempster, A. P. (1967). Upper and lower probabilities induced by a multiple valued mapping. *Ann. Math. Statistics*, *38*, 325–339.

Dempster, A. P. (1968). A generalization of Bayesian inference. *J. Roy. Statist. Soc. B.*, *30*, 205–247.

Dempster, A. P. (1972). A class of random convex polytopes. *Annals of Mathematical Statistics*, *43*, 250–272.

Dubois, D., Prade, H., & Smets, P. (1996). Representing partial ignorance. *IEEE System Machine and Cybernetic*, *26*, 361–377.

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

Dubois, D., Wellman, M. P., D'Ambrosio, B., & Smets, P. (Eds.). (1992). *Uncertainty in artificial intelligence 92.* Morgan Kaufman, San Mateo, Ca.

Fagin, R., & Halpern, J. (1991). A new approach to updating beliefs. In P. P. Bonissone, M. Henrion, L. N. Kanal, & J. F. Lemmer (Eds.), *Uncertainty in artificial intelligence 6* (pp. 347–374). North Holland, Amsteram.

Gardenfors, P., & Sahlin, N. E. (Eds.). (1988). *Decision, probability and utility.* Cambridge Univ. Press, Cambridge, G.B.

Gilboa, I., & Schmeidler, D. (1992). Updating ambiguous beliefs. In Y. Moses (Ed.), *Tark-92, theoretical aspects of reasoning about knowledge* (pp. 1xxx–Xxx). Morgan Kaufman, San Mateo, Ca.

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

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

Hacking, I. (1988). Slightly mfore realistic personal probability. In P. Gardenfors & N. E. Sahlin (Eds.), *Decision, probability and utility* (pp. 118–135). Cambridge Univ. Press, Cambridge, G.B.

Harper, X., & Hooker, X. (Eds.). (1976). *Foundations of probability theory, statistical inference, and statistical theories of science.* Reidel, Doordrecht-Holland.

Heckerman, D., & Mamdani, A. (Eds.). (1993). *Uncertainty in artificial intelligence 93.* Morgan Kaufmann, San Mateo, Ca, USA.

Henrion, M., Shachter, R. D., Kanal, L. N., & Lemmer, J. F. (Eds.). (1990). *Uncertainty in artificial intelligence 5.* North Holland, Amsteram.

IJCAI-87 (Ed.). (1987). *Intern. joint conf. on artificial intelligence.* Morgan Kaufman, San Mateo, Ca.

Jaffray, J. Y. (1988). Application of linear utility theory for belief functions. In B. Bouchon, L. Saitta, & R. R. Yager (Eds.), *Uncertainty and intelligent systems* (pp. 1–8). Springer-Verlag, Berlin.

Jaffray, J. Y. (1989a). Coherent bets under partially resolving uncertainty and belief functions. *Theory and Decision*, *26*, 99–105.

Jaffray, J. Y. (1989b). Linear utility theory for belief functions. *Op. Res. Letters*, *8*, 107–112.

Jaffray, J. Y. (1992). Bayesian updating and belief functions. *IEEE Trans. SMC*, *22*, 1144–1152.

Jaffray, J. Y. (1994). Dynamic decision making with belief functions. In R. R. Yager, J. Kacprzyk, & M. Fedrizzi (Eds.), *Advances in theDempster-Shafer theory of evidence* (pp. 331–352). Wiley, New York.

Jaffray, J. Y., & Philippe, F. (1997). On the existence of subjective upper and lower probabilities. *Mathematics of Operations Research*, *22*, 165–185.

Jaffray, J. Y., & Wakker, P. (1994). Decision making with belief functions: compatibility and incompatibility with the sure-thing principle. *Journal of Risk and Uncertainty*, *8*, 255–271.

Jeffrey, R. C. (1988). Conditioning, kinematics, and exchangeability. In B. Skyrms & W. L. Harper (Eds.), *Causation, chance, and credence* (pp. 21–255). Reidel, Dordrecht.

Jensen, N. E. (1967). An introduction to Bermoullian utility theory. I: Utility functions. *Swedish J. of Economics*, *69*, 163–183.

Kohlas, J., & Monney, P. A. (1994). Theory of evidence: A survey of its mathematical foundations, applications and computations. *ZOR-Mathematical Methods of Operatioanl Research*, *39*, 35–68.

Kohlas, J., & Monney, P. A. (1995). *A mathematical theory of hints: An approach to Dempster-Shafer theory of evidence.* Lecture Notes in Economics and Mathematical Systems No. 425. Springer-Verlag.

Kyburg, H. E., Jr. (1987a). Bayesian and non-Bayesian evidential updating. *Artificial Intelligence*, *31*, 271–294.

Kyburg, H. E., Jr. (1987b). Objectives probabilities. In IJCAI-87 (Ed.), *Intern. joint conf. on artificial intelligence* (pp. 02–904). Morgan Kaufman, San Mateo, Ca.

Kyburg, H. E., Jr. (1988). Bets and beliefs. In P. Gardenfors & N. E. Sahlin (Eds.), *Decision, probability and utility* (pp. 101–117). Cambridge Univ. Press, Cambridge, G.B.

Levi, I. (1980). *The enterprise of knowledge.* MIT Press, Cambridge, Mass.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions.* Wiley, New York.

Moral, S., & de Campos, L. M. (1990). Updating uncertain information. In B. Bouchon-Meunier, R. R. Yager, & L. A. Zadeh (Eds.), *Uncertainty in knowledge bases* (pp. 58–67). Springer Verlag, Berlin.

Nguyen, H. T., & Walker, E. A. (1994). On decision making using belief functions. In R. R. Yager, J. Kacprzyk, & M. Fedrizzi (Eds.), *Advances in theDempster-Shafer theory of evidence* (pp. 275–309). Wiley, New York.

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

Ramsey, F. P. (1964). Truth and probability. In H. E. Kyburg, Jr. & H. E. Smokler (Eds.), *Studies in subjective probability* (pp. 61–92). Wiley, New York.

Rivett, B. H. P. (1975). Behavioural problems of utility theory. In D. J. White & K. C. Bowen (Eds.), *The role and effectiveness of theories of decision in practice* (pp. 21–27). Hodder and Stoughton, London.

Roth, A. E. (Ed.). (1988). *The Shapley value: Essays in honor of l.s. shapley.* Cambridge University Press, Cambridge, GB.

Ruspini, E. H. (1986). *The logical foundations of evidential reasoning* (Tech. Rep.). Technical note 408, SRI International, Menlo Park, Ca, US.

Ruspini, E. H. (1987). Epistemic logics, probability and the calculus of evidence. In IJCAI-87 (Ed.), *Intern. joint conf. on artificial intelligence* (pp. 924–931). Morgan Kaufman, San Mateo, Ca.

Savage, L. J. (1954). *Foundations of statistics.* Wiley, New York.

Schaller, J.-P. (1991). *Multiple criteria decision aid under incomplete information.* Unpublished doctoral dissertation, University of Fribourg, Switzerland.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, *57*, 571–587.

Schubert, J. (1995). On rho in a decision-theoretic apparatus of Dempster-Shafer theory. *Int. J. Approx. Reas.*, *13*, 185–200.

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

Shafer, G. (1976b). A theory of statistical evidence. In X. Harper & X. Hooker (Eds.), *Foundations of probability theory, statistical inference, and statistical theories of science* (pp. xx–xx). Reidel, Doordrecht-Holland.

Shapley, L. S. (1953). A value for n-person games. In H. Kuhn & A. W. Tucker (Eds.), *Contributions to the theory of games* (Vol. 2, pp. 307–317). Princeton University Press, NJ, US.

Shenoy, P. P. (1991). A fusion algorithm for solving bayesian decision problems. In B. D'ambrosio, P. Smets, & B. P. P. and (Eds.), *Uncertainty in artificial intelligence 91* (pp. 323–331). Morgan Kaufmann, San Mateo, Ca, USA.

Shenoy, P. P. (1992). Valuation-based systems for Bayesian decision analysis. *Operations Research*, *40*, 463–484.

Smets, P. (1988). Belief functions. In P. Smets, E. H. Mamdani, D. Dubois, & H. Prade (Eds.), *Non standard logics for automated reasoning* (pp. 253–286). Academic Press, London.

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

Smets, P. (1992a). The concept of distinct evidence. In IPMU-92 (Ed.), *Information processing and management of uncertainty* (pp. 89–94).

Smets, P. (1992b). The nature of the unnormalized beliefs encountered in the transferable belief model. In D. Dubois, M. P. Wellman, B. D'Ambrosio, & P. Smets (Eds.), *Uncertainty in artificial intelligence 92* (pp. 292–297). Morgan Kaufman, San Mateo, Ca.

Smets, P. (1993a). Belief functions: the disjunctive rule of combination and the generalized bayesian theorem. *International Journal of Approximate Reasoning*, *9*, 1–35.

Smets, P. (1993b). Jeffreys rule of conditioning generalized to belief functions. In D. Heckerman & A. Mamdani (Eds.), *Uncertainty in artificial intelligence 93* (pp. 500–505). Morgan Kaufmann, San Mateo, Ca, USA.

Smets, P. (1993c). No Dutch book can be built against the TBM even though update is not obtained by bayes rule of conditioning. In R. Scozzafava (Ed.), *Workshop on probabilisitic expert systems* (pp. 181–204). Societa Italiana di Statistica , Roma.

Smets, P. (1993d). An axiomatic justifiaction for the use of belief function to quantify beliefs. In IJCAI-93 (Ed.), *Intern. joint conf. on artificial intelligence* (pp. 598–603). Morgan Kaufman, San Mateo, Ca.

Smets, P. (1994a). Belief induced by the knowledge of some probabilities. In D. Heckerman, D. Poole, & R. Lopez De Mantaras (Eds.), *Uncertainty in artificial intelligence 94* (pp. 523–530). Morgan Kaufmann, San Mateo, Ca.

Smets, P. (1994b). What is Dempster-Shafer's model? In R. R. Yager, J. Kacprzyk, & M. Fedrizzi (Eds.), *Advances in theDempster-Shafer theory of evidence* (pp. 5–34). Wiley, New York.

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

Smets, P. (1998a). The application of the transferable belief model to diagnostic problems. *Int. J. Intelligent Systems*, *13*, 127–157.

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

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

Smith, C. A. B. (1961). Consistency in statistical inference and decision. *J. Roy. Statist. Soc. B*, *23*, 1–37.

Smith, P., & Jones, O. R. (1986). *The philosophy of mind, an introduction.* Cambridge University Press, Cambridge.

Strat, T. M. (1990a). Decision analysis using belief functions. *Int. J. Approx. Reasoning*, *4*, 391–418.

Strat, T. M. (1990b). Making decisions with belief functions. In M. Henrion, R. D. Shachter, L. N. Kanal, & J. F. Lemmer (Eds.), *Uncertainty in artificial intelligence 5* (pp. 51–360). North Holland, Amsteram.

Strat, T. M. (1994). Decision analysis using belief functions. In R. R. Yager, J. Kacprzyk, & M. Fedrizzi (Eds.), *Advances in theDempster-Shafer theory of evidence* (pp. 275–309). Wiley, New York.

Suppes, P., & Zanotti, M. (1977). On using random relations to generate upper and lower probabilities. *Synthesis*, *36*, 427–440.

Teller, P. (1973). Conditionalization and observation. *Synthesis*, *26*, 218–258.

Teller, P. (1976). Conditionalization, observation and change of prefetrence. In X. Harper & X. Hooker (Eds.), *Foundations of probability theory, statistical inference, and statistical theories of science* (pp. 205–259). Reidel, Doordrecht-Holland.

Tsiporkova, E., Boeva, V., & De Baets, B. (1999). Dempster-Shafer's theory framed in modal logic. *Int. J. Approx. Reasoning*, *to appear.*

Tsiporkova, E., De Baets, B., & Boeva, V. (1999a). Dempster's rule of conditioning translated into modal logic. *Fyzzy Sets and Systems*, *102*, 371–383.

Tsiporkova, E., De Baets, B., & Boeva, V. (1999b). Evidence theory in multi-valued models of modal logic. *J. Appl. Non-Classical Logics, to appear.*

Voorbraak, F. (1993). *As far as i know: Epistemic logic and uncertainty.* Dissertation, Utrecht University.

Voorbracht, F. (1991). On the justification of Dempster's rule of combination. *Artificial Intelligence*, *48*, 171–197.

Walley, P. (1987). Belief functions representations of statistical of evidence. *The Annals of Statistics*, *15*, 1439–1465.

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

Williams, P. M. (1982). Discussion of Shafer's paper. *J. Roy. Statist. Soc.B*, *44*, 322-352.

Wilson, N. (1993). Decision making with belief functions and pignistic probabilities. In *European conference on symbolic and quantitative approaches to reasoning and uncertainty* (pp. 364–371). Granada: Springer Verlag.

Xu, H. (1992). A decision calculus for belief functions in valuation based systems. In D. Dubois, M. P. Wellman, B. D'Ambrosio, & P. Smets (Eds.), *Uncertainty in artificial intelligence 92* (pp. 352–359). Morgan Kaufman, San Mateo, Ca.

Xu, H., Hsia, Y. T., & Smets, P. (1993a). A belief function based decision support system. In D. Heckerman & A. Mamdani (Eds.), *Uncertainty in artificial intelligence 93* (pp. 535–542). Morgan Kaufmann, San Mateo, Ca, USA.

Xu, H., Hsia, Y. T., & Smets, P. (1993b). A belief function based decision support system. In D. Heckerman & A. Mamdani (Eds.), *Uncertainty in artificial intelligence 93* (pp. 535–542). Morgan Kaufmann, San Mateo, Ca, USA.

Xu, H., & Smets, P. (1995). Generating explanations for evidential reasoning. In P. Besnard & S. Hanks (Eds.), *Uncertainty in artificial intelligence 95* (pp. 574–581). Morgan Kaufman, San Francisco, Ca.

Xu, H., & Smets, P. (1996). Some strategies for explanation in evidential reasoning. *IEEE Trans. SMC:*, *26*, 599–607.

Yager, R. R., Kacprzyk, J., & Fedrizzi, M. (Eds.). (1994). *Advances in the Dempster-Shafer theory of evidence.* Wiley, New York.