Reasoning with imprecise belief structures[1]

Thierry Denœux

U.M.R. CNRS 6599 Heudiasyc
Université de Technologie de Compiègne
BP 20529 - F-60205 Compiègne cedex - France
email: Thierry.Denoeux@hds.utc.fr

[1]Technical Report Heudiasys 97/44.

Abstract

This paper extends the theory of belief functions by introducing new concepts and techniques, allowing to model the situation in which the beliefs held by a rational agent may only be expressed (or are only known) with some imprecision. Central to our approach is the concept of interval-valued belief structure, defined as a set of belief structures verifying certain constraints. Starting from this definition, many other concepts of Evidence Theory (including belief and plausibility functions, pignistic probabilities, combination rules and uncertainty measures) are generalized to cope with imprecision in the belief numbers attached to each hypothesis. An application of this new framework to the classification of patterns with partially known feature values is demonstrated.

Keywords: evidence theory, Dempster-Shafer theory, belief functions, Transferable Belief model, uncertainty modeling, imprecision, pattern classification.

1 Introduction

The representation of uncertainty is a major issue in many areas of Science and Engineering. Although Probability Theory has been regarded as a reference framework for about two centuries, it has become increasingly apparent that the concept of probability is not general enough to account for all kinds of uncertainty. More specifically, the requirement that precise numbers be assigned to every individual hypotheses is often regarded as too restrictive, particularly when the available information is poor.

The use of completely monotone capacities, also called *credibility* or *belief functions*, as a general tool for representing someone's degrees of belief was proposed by Shafer [17] in the seventies (belief functions had already been introduced by Dempster ten years earlier as lower probabilities generated by a multivalued mapping [1]). In the last two decades, the so-called Dempster-Shafer (D-S) theory of evidence has attracted considerable interest, but debates concerning the relevance of belief functions for uncertainty representation have sometimes been obscured by misunderstandings about their meaning [22, 27].

In the Dempster's model, the possibility space of interest Ω is assumed to be related through a many to one mapping M to an underlying space Ψ on which the existence of a precise probability measure P is assumed [1]. The lower probability $P_*(A)$ of $A \subseteq \Omega$ is then defined as the probability of the largest subset of Ψ whose image under M is included in A, while the upper probability $P^*(A)$ of A is the probability of the largest subset of Ψ such that the images under M of all its elements have a non empty intersection with A. Function P_* happens to be a belief function, but, as pointed out by Smets [28], it does not necessarily quantify an agent's belief. Rather, the interval $[P_*(A), P^*(A)]$ receives in this approach a natural interpretation as an imprecise specification of some unknown probability $P(A)$. Dempster's model is thus a particular form of imprecise probability model [31].

Smets' Transferable Belief Model (TBM) differs radically from Dempster's model (and from other non-standard probabilistic interpretations such as based, e.g., on random sets) in that it introduces belief functions independently from any probabilistic model [20, 30]. The main assumptions underlying the TBM are that (1) degrees of belief are quantified by numbers between 0 and 1; (2) there exists a two-level structure composed of a credal level where beliefs are entertained, and a pignistic level where decisions are made; (3) beliefs at the credal level are quantified by belief functions, while decisions at the pignistic level are based on probability functions; (4) when a decision has to be made, beliefs are transformed into probabilities using the so-called pignistic transformation. A complete axiomatic justification of that model is presented in [29].

The main reason why the multiple interpretations of the D-S theory must be carefully distinguished is that they have different implications with regard to the way beliefs should be updated when new evidence becomes available [28]. In fact, it is only within the TBM that Dempster's rules of conditioning and combination seem to be fully justified [20, 22, 24]. In the rest of this paper, we shall therefore adopt Smets' view of belief functions as an alternative to probability functions for *pointwise* representation of the beliefs held by a rational agent.

A consequence of the above point of view is that a belief function and the associated plausibility function cannot be regarded as defining probability intervals. Of course, such intervals mathematically exist, but they are meaningless since we do not assume the existence of any objective or subjective probability function: the TBM is not a model of poorly known probabilities. It may then be wondered whether the necessity to assign precise numbers to each subset of the possibility space is not too constraining. If beliefs – represented by belief masses and not by probabilities – are the quantities of interest, then uncertainty about their values can no longer be neglected as being of second order. In other words, one may wonder whether one of the main criticisms raised against Bayesian Probability Theory – its unreasonable requirement for precision – cannot also be raised against the TBM.

To help clarify this point, it is useful to examine some features of beliefs, evidence and the assessment process which can lead to imprecision in models of uncertainty. Walley [31] makes an interesting distinction between two main reasons for introducing imprecision in models uncertainty: *indeterminacy* and *model incompleteness*.

Indeterminacy may be defined as an absence of preference, due to limitations of the available information. The TBM effectively allows to cope with such type of imprecision, since a belief function may assign a positive mass of belief to a proposition A without supporting any strict subproposition. It is this feature of D-S theory that gives it a distinctive advantage over Probability Theory for representing ignorance.

The second main factor leading to imprecision is incompleteness of the model, which may be due to difficulties in analyzing evidence and assessing beliefs. Many of the sources of imprecision mentioned by Walley in the case of probability models are still relevant when dealing with belief functions, namely:

- lack of introspection or of assessment strategies (precise degrees of belief may exist, but it may be too difficult, too costly, or unnecessary to elicit them with great precision),
- instability (underlying beliefs may be unstable, or elicited beliefs may be influenced by the conditions of elicitation),
- ambiguity (beliefs may be elicited through ambiguous judgments such as “about 0.3”), etc.

For all these reasons, we need a generalization of the TBM allowing to work with imprecisely specified beliefs, which is the subject of this paper.

Whereas there is a rich literature on imprecise probabilities (see an excellent survey in [31]), attempts to define a rigorous, yet mathematically tractable generalization of D-S theory allowing to assign imprecise belief masses to propositions seem to have been until now limited[1]. The reason for that may be that

[1]This approach should not be confused with the concept of a belief function with fuzzy focal elements studied by several authors [38, 18, 9, 32, 15, 37, 36]. However, both approaches are compatible, i.e., imprecise masses may be assigned to fuzzy events.

credibility functions have very often been considered as defining probability intervals instead of pointwise beliefs, as explained above. Zadeh [38] mentioned the possibility of generalizing the concepts of expected possibility and necessity (which themselves generalize the notions of credibility and plausibility) to the case where belief masses are fuzzy or linguistic, and Smets [23] proposed a method for defining fuzzy degrees of belief and plausibility by conditioning with a fuzzy event. However, none of these authors seem to have really explored this research avenue. Lamata and Moral [14] have proposed "ad hoc" methods of calculus with interval-valued and linguistic probabilities and beliefs, but many of their procedures appear to be insufficiently justified.

The objective of this paper is to extend the main concepts of D-S theory, including those of credibility, plausibility, Dempster's rule of combination, normalization, and entropy-like uncertainty measures, to the case where degrees of belief in the various propositions are only known to lie within certain intervals. Further generalization to fuzzy belief numbers is possible, but will not be considered in this paper.

The organization of the paper is as follows. First, the necessary background concerning D-S theory is recalled in Section 2. Interval-valued belief structures (IBS), and the associated concepts of imprecise evidential functions and pignistic probabilities are then introduced in Section 3. Section 4 addresses the dynamic part of the model, i.e., the combination of IBSs induced by multiple sources, as well as normalization rules. It is then shown how uncertainty and information measures may be extended to IBSs, allowing to quantify the imprecision of an IBS, and to select a "maximally uncertain" belief structure compatible with an IBS (Section 5). Finally, an application of these concepts to pattern classification with partial knowledge of feature vectors is demonstrated in Section 6.

2 Background

2.1 Belief structures

Let Ω be a finite set called the possibility space, or the frame of discernment, and m a function from 2^{Ω} to $[0,1]$ verifying:

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

Such a function is called a *basic probability structure* by Shafer [17], a *basic belief assignment* by Smets [20], and a *belief structure* (BS) by Yager [32]. The latter terminology will be adopted in this paper. The quantity $m(A)$ (called a mass of belief or a belief number) may be interpreted as a "part of belief" that is committed to A given the available evidence, and that cannot be committed to any strict subset of A because of lack of sufficient information. Let

$$\mathcal{F}(m) = \{A \subseteq \Omega | m(A) > 0\}$$

The elements of $\mathcal{F}(m)$ are called the *focal elements* of m. Shafer [17] initially imposed a normality condition for belief structures ($\emptyset \notin \mathcal{F}(m)$). Smets [20] proposed to relax this condition, and to interpret $m(\emptyset)$ as the part of belief committed to the assumption that none of the hypotheses in Ω might be true (open-world assumption). If however the truth is known with absolute certainty to lie in Ω (closed-world assumption), then the normality condition is justified.

2.2 Evidential functions

Let m be a BS. For all $A \subseteq \Omega$, the belief (or credibility) and the plausibility of A are defined respectively as [17, 30]:

$$\mathrm{bel}_m(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \quad (1)$$

$$\mathrm{pl}_m(A) = \sum_{B \cap A \neq 0} m(B) \quad (2)$$

$$= \mathrm{bel}_m(\Omega) - \mathrm{bel}_m(\bar{A}) \quad (3)$$

where $\bar{A}$ denotes the complement of A. The quantity $\mathrm{bel}_m(A)$ may be interpreted as the total amount of justified support given to A, while $\mathrm{pl}_m(A)$ quantifies the maximum amount of specific support that *could* be given to A, if justified by additional information [30]. Note that $\emptyset$ is excluded from the sum in Equation 1, which under the open-world assumption is justified by the particular interpretation given to $m(\emptyset)$, and guarantees that $\mathrm{bel}_m(A) \leq \mathrm{pl}_m(A)$ for all A.

Shafer [17] also introduced the commonality function defined for all $A \subseteq \Omega$ as:

$$q_m(A) = \sum_{B \supseteq A} m(B). \quad (4)$$

Although the meaning of $q_m(A)$ in not so obvious as those of $\mathrm{bel}_m(A)$ and $\mathrm{pl}_m(A)$, function q plays an important role in relation to the conjunctive combination of BSs (see Section 2.4).

The credibility, plausibility and commonality functions (henceforth referred to as *evidential functions*) are in one-to-one correspondence with belief structures [17]. Therefore, they can be regarded as different expressions of the same information.

2.3 Pignistic transformation

The problem of decision making in solved in the TBM using the concept of pignistic probability function, which was shown by Smets [21, 30] to be the only solution compatible with simple rationality requirements. Given a normalized belief structure m quantifying one's beliefs at the credal level, a pignistic probability distribution BetP_m is defined as:

$$\mathrm{BetP}_m(A) = \sum_{B \subseteq \Omega} m(B) \frac{|A \cap B|}{|B|} \quad (5)$$

for all $A \subseteq \Omega$. In this transformation, the mass $m(B)$ is thus distributed equally among the elements of B. The same solution would also be obtained by applying the Insufficient Reason Principle at the level of each mass of belief, although this principle need not be postulated in the TBM [30].

2.4 Combination of belief structures

Let m_1 and m_2 be two BSs on the same frame of discernment Ω induced by distinct items of evidence. Dempster [1], followed by Shafer [17], suggested a procedure for combining m_1 and m_2, known as Dempster's rule of combination. According to this rule, the orthogonal sum of m_1 and m_2 is defined by:

$$(m_1 \oplus m_2)(A) = \frac{1}{K} \sum_{B \cap C = A} m_1(B) m_2(C) \tag{6}$$

with $K = \sum_{B \cap C \neq \emptyset} m_1(B) m_2(C)$ for $A \neq \emptyset$, and $m(\emptyset) = 0$. Under the open-world assumption, the normalizing factor may be dropped, leading to the simpler conjunctive sum operation [20]:

$$(m_1 \wedge m_2)(A) = \sum_{B \cap C = A} m_1(B) m_2(C). \tag{7}$$

The computation of the conjunctive sum is sometimes made simpler by using the following property of the commonality function:

$$q_{m_1 \wedge m_2}(A) = q_{m_1}(A) q_{m_2}(A) \quad \forall A \subseteq \Omega. \tag{8}$$

The conjunctive sum has a natural disjunctive counterpart, which is applicable when we only know that one of the two pieces of evidence actually holds [25]:

$$(m_1 \vee m_2)(A) = \sum_{B \cup C = A} m_1(B) m_2(C). \tag{9}$$

Yager[2] [34] suggested to further generalize these rules to any binary set operation $*$ as:

$$(m_1 \circledast m_2)(A) = \sum_{B * C = A} m_1(B) m_2(C). \tag{10}$$

where $\circledast$ is the operation on belief structures induced by $*$.

3 Interval-valued belief structures

3.1 Definition

In the rest of this paper, we now assume that belief masses are only known to lie within certain intervals. Uncertainty is then no longer described by a unique belief structure, but by a convex set of belief structures verifying certain constraints. More precisely, let us introduce the following definition.

[2] Yager actually assumed $*$ to be a *non-null forming* operation (i.e., $B * C \neq \emptyset$ for any B and C), because he only considered normalized BSs. This assumption is unnecessary here.

Definition 1 (Interval-valued belief structure)
Let $\mathcal{S}_\Omega$ denote the set of all belief structures on Ω. An interval-valued belief structure (IBS) is a non empty subset $\mathbf{m}$ of $\mathcal{S}_\Omega$ such that there exist n subsets $F_1, \ldots, F_n$ of Ω, and n intervals $[a_i, b_i], 1 \leq i \leq n$ (with $b_i > 0$) such that

$$\mathbf{m} = \{m \in \mathcal{S}_\Omega | a_i \leq m(F_i) \leq b_i, 1 \leq i \leq n, \text{ and } m(A) = 0, \forall A \notin \{F_1, \ldots, F_n\}\}$$

We note $\mathcal{F}(\mathbf{m}) = \{F_1, \ldots, F_n\}$.

Obviously, the condition that $\mathbf{m}$ be non-empty imposes certain constraints on the a_i and b_i; these constraints are expressed in the following proposition:

Proposition 1
A necessary and sufficient condition for $\mathbf{m}$ to be non-empty is that $\sum_{i=1}^n a_i \leq 1$ and $\sum_{i=1}^n b_i \geq 1$.

Proof. The condition is obviously necessary: if there exists some $m \in \mathbf{m}$, then we have

$$\sum_{i=1}^n a_i \leq \sum_{i=1}^n m(F_i) = 1$$

and

$$\sum_{i=1}^n b_i \geq \sum_{i=1}^n m(F_i) = 1$$

To prove that it is sufficient, let us note $s = \sum_{i=1}^n a_i$, $S = \sum_{i=1}^n b_i$, and

$$\lambda = \frac{S - 1}{S - s}$$

Let m be the belief structure defined as $m(F_i) = \lambda a_i + (1 - \lambda) b_i$ for $1 \leq i \leq n$, and $m(A) = 0$ for all $A \notin \{F_1, \ldots, F_n\}$. We have $a_i \leq m(F_i) \leq b_i$ for $1 \leq i \leq n$, and

$$\sum_{i=1}^n m(F_i) = \lambda s + (1 - \lambda) S = 1$$

Hence $m \in \mathbf{m}$, and $\mathbf{m} \neq \emptyset$. □

Remarks

1. If m is a belief structure, then the singleton $\{m\}$ is an IBS with $a_i = b_i = m(F_i)$ for all $F_i \in \mathcal{F}(m)$. Hence, the concept of IBS generalizes that of BS.

2. The set $\mathcal{S}(\Omega)$ is an IBS in which the interval associated to each subset of Ω is $[0, 1]$. It may be interpreted as reflecting "second-order" ignorance, i.e., ignorance of what the state of belief of an agent may be.

3.2 Bounds of an IBS

It is important to note that the intervals $[a_i, b_i]$ specifying an IBS $\mathbf{m}$ are not necessarily unique. Since

$$m(F_i) \leq \min \left[b_i, 1 - \sum_{j \neq i} a_i \right],$$

it is obvious that, whenever $b_i \geq 1 - \sum_{j \neq i} a_i$, b_i may be replaced by any $b'_i \geq b_i$. To obtain a unique characterization of $\mathbf{m}$, we must introduce the concepts of lower and upper bounds of $\mathbf{m}$, defined respectively as:

$$\begin{aligned} m^-(A) &= \min_{m \in \mathbf{m}} m(A) \\ m^+(A) &= \max_{m \in \mathbf{m}} m(A) \end{aligned}$$

for all $A \in [0, 1]^\Omega$. These functions may easily obtained from any set of intervals $[a_i, b_i]$ defining $\mathbf{m}$ by:

$$\begin{aligned} m^-(F_i) &= \max \left[a_i, 1 - \sum_{j \neq i} b_j \right] \\ m^+(F_i) &= \min \left[b_i, 1 - \sum_{j \neq i} a_j \right] \end{aligned}$$

for all $1 \leq i \leq n$, and $m^-(A) = m^+(A) = 0$, for all $A \notin \mathcal{F}(\mathbf{m})$. It is easy to check that:

$$\begin{aligned} m^-(F_i) &\geq 1 - \sum_{j \neq i} m^+(F_j) \\ m^+(F_i) &\leq 1 - \sum_{j \neq i} m^-(F_j) \end{aligned}$$

for all $1 \leq i \leq n$. It the rest of this paper, we note $\mathbf{m}(F_i) = [m^-(F_i), m^+(F_i)]$.

3.3 Example

Before introducing other definitions, let us describe as an example a typical situation in which the concept of IBS may be useful.

EXAMPLE 1 Let us consider a situation in which n balls have been drawn with replacement from an urn containing white and black balls. Knowing that exactly n_b black balls have been selected, what is your belief that the next randomly selected ball will be black ? Let $\Omega = \{b, w\}$ be the possibility space for that experiment. In [26], Smets has shown that, under the TBM, one's belief about which event in Ω will occur should be modeled by the following belief structure:

$$\begin{aligned} m(\{b\}) &= \frac{n_b}{n+1} \\ m(\{w\}) &= \frac{n-n_b}{n+1} \\ m(\{b,w\}) &= \frac{1}{n+1} \end{aligned}$$

Let us now assume that we only know that *between* 10 and 100 balls have been selected, among which *at least* 80 % were black. It is not clear in that case how beliefs could be represented by a single BS. One possibility could be to enumerate all possible cases, yielding a finite (but large) set of belief structures. However, it is much more convenient to compute lower and upper bounds for the mass of belief assigned to each hypothesis, which leads to the following IBS:

$$\begin{aligned} \mathbf{m}(\{b\}) &= [\frac{8}{11}, \frac{100}{101}] \\ \mathbf{m}(\{w\}) &= [0, \frac{20}{101}] \\ \mathbf{m}(\{w,b\}) &= [\frac{1}{101}, \frac{1}{11}] \end{aligned}$$

3.4 Imprecise evidential functions

The concepts of credibility, plausibility and commonality of a subset A of Ω induced by a BS may easily be generalized to the case of an IBS $\mathbf{m}$ by considering the range of $\mathrm{bel}_m(A)$, $\mathrm{pl}_m(A)$ and $q_m(A)$, respectively, for all $m \in \mathbf{m}$. Since these quantities are linear combinations of belief masses constrained to lie in closed intervals, their ranges are themselves closed intervals. We thus have the following definition:

Definition 2 (Imprecise evidential functions)
The credibility, plausibility and commonality of a subset A induced by an IBS $\mathbf{m}$ are the closed intervals defined respectively as:

$$\mathbf{bel_m}(A) = [\min_{m \in \mathbf{m}} bel_m(A), \max_{m \in \mathbf{m}} bel_m(A)] \quad (11)$$

$$\mathbf{pl_m}(A) = [\min_{m \in \mathbf{m}} pl_m(A), \max_{m \in \mathbf{m}} pl_m(A)] \quad (12)$$

$$\mathbf{q_m}(A) = [\min_{m \in \mathbf{m}} q_m(A), \max_{m \in \mathbf{m}} q_m(A)] \quad (13)$$

where bel_m, pl_m and q_m are the belief, plausibility and commonality functions defined by Equations 1, 2 and 4, respectively.

The calculation of each of these intervals requires to find the extrema of a function of n variables of the form:

$$f(x_1, \ldots, x_n) = \sum_{i \in I} x_i$$

with $I \subset \{1, \ldots, n\}$, under the constraints

$$a_i \leq x_i \leq b_i$$

for all $1 \leq i \leq n$ and

$$\sum_{i=1}^{n} x_i = 1$$

It is straightforward to show that, under these constraints, the extrema of f are given by:

$$\begin{aligned}
\min f &= \max \left(\sum_{i \in I} a_i, 1 - \sum_{i \notin I} b_i \right) \\
\max f &= \min \left(\sum_{i \in I} b_i, 1 - \sum_{i \notin I} a_i \right)
\end{aligned}$$

Using $-$ and $+$ superscripts to denote lower and upper bounds, respectively, we thus have, for example:

$$\mathrm{bel}^-_{\mathbf{m}}(A) = \max \left[\sum_{\emptyset \neq B \subseteq A} m^-(B), 1 - \sum_{B \not\subseteq A} m^+(B) - m^+(\emptyset) \right] \quad (14)$$

$$\mathrm{bel}^+_{\mathbf{m}}(A) = \min \left[\sum_{\emptyset \neq B \subseteq A} m^+(B), 1 - \sum_{B \not\subseteq A} m^-(B) - m^-(\emptyset) \right] \quad (15)$$

Similar expressions may be obtained for the bounds of $\mathbf{pl_m}$ and $\mathbf{q_m}$ without any difficulty. Note that, if $\mathbf{m}$ is normalized $(m^+(\emptyset) = 0)$, then the relation

$$\mathrm{pl}_m(A) = 1 - \mathrm{bel}_m(\bar{A})$$

for all $A \subseteq \Omega$ and $m \in \mathbf{m}$ has its counterpart in the following equalities:

$$\begin{aligned}
\mathrm{pl}^-_{\mathbf{m}}(A) &= 1 - \mathrm{bel}^+_{\mathbf{m}}(\bar{A}) \\
\mathrm{pl}^+_{\mathbf{m}}(A) &= 1 - \mathrm{bel}^-_{\mathbf{m}}(\bar{A}).
\end{aligned}$$

Example 2 An example of an IBS on $\Omega = \{a, b, c\}$ and the corresponding credibility, plausibility and commonality intervals is given in Table 1.

Remark: The one-to-one correspondence between belief structures and each of the evidential functions is not preserved when dealing with imprecise belief structures: an interval-valued evidential function does not uniquely specify an IBS. For instance, it is possible to find an IBS $\mathbf{m}$ with corresponding interval-valued belief function $\mathbf{bel_m}$, such that $\mathrm{bel}_m \in \mathbf{bel_m}$ for some $m \notin \mathbf{m}$. Stated differently, the set of all belief functions associated to a BS in $\mathbf{m}$ is strictly included in $\mathbf{bel_m}$ (similar statements hold for plausibilities and commonalities). To illustrate this point, let us consider $\mathbf{m}$ and $\mathbf{bel_m}$ as in Example 2, and the

BS m defined by Table 2. We have $\mathrm{bel}_m \in \mathbf{bel_m}$, whereas $m \notin \mathbf{m}$. This example proves that IBSs and interval-valued belief functions are not equivalent representations. Our choice of working with IBSs is essentially driven by practical considerations: in applications, belief structures are often constructed directly from the evidence; moreover, they have a simpler mathematical form than belief functions.

3.5 Interval-valued pignistic probability

An interval-valued pignistic probability function induced by an IBS may also be defined by considering the range of $\mathrm{BetP}_m(A)$ for all $m \in \mathbf{m}$.

Definition 3 (Interval-valued pignistic probabilities)
Let $\mathbf{m}$ be a normal IBS, i.e., an IBS such that $\emptyset \notin \mathcal{F}(\mathbf{m})$. The pignistic probability of a subset A induced by $\mathbf{m}$ is the closed interval defined as:

$$\mathbf{BetP_m}(A) = [\min_{m \in \mathbf{m}} BetP_m(A), \max_{m \in \mathbf{m}} BetP_m(A)]$$

where $BetP_m$ denotes the pignistic probability function induced by m.

The bounds of $\mathbf{BetP_m}(A)$ may be found as the solutions of a class of relatively simple linear programming (LP) problems which has been extensively studied by Dubois and Prade [5] (see also [6] and [8] p. 55). The fundamental result is expressed by the following theorem:

Theorem 1 (Dubois and Prade, 1981)
Let $x_1, \ldots, x_n$ be n variables linked by the following constraints:

$$\sum_{i=1}^{n} x_i = 1$$

$$a_i \leq x_i \leq b_i \quad 1 \leq i \leq n$$

and let f be a function defined by $f(x_1, \ldots, x_n) = \sum_{i=1}^{n} c_i x_i$ with

$$0 \leq c_1 \leq c_2 \leq \ldots \leq c_n.$$

Then

$$\begin{aligned}
\min f &= \max_{k=1,n} \left(\sum_{j=1}^{k-1} b_j c_j + \left(1 - \sum_{j=1}^{k-1} b_j - \sum_{j=k+1}^{n} a_j \right) c_k + \sum_{j=k+1}^{n} a_j c_j \right) \\
\max f &= \min_{k=1,n} \left(\sum_{j=1}^{k-1} a_j c_j + \left(1 - \sum_{j=1}^{k-1} a_j - \sum_{j=k+1}^{n} b_j \right) c_k + \sum_{j=k+1}^{n} b_j c_j \right)
\end{aligned}$$

This theorem may be directly applied by posing $a_i = m^-(F_i)$, $b_i = m^+(F_i)$ and

$$c_i = \frac{|A \cap F_i|}{|F_i|},$$

the focal elements F_i being arranged in such a way that

$$0 \leq c_1 \leq c_2 \leq \ldots \leq c_n.$$

EXAMPLE 3 An IBS and its associated pignistic probability function are shown in Table 3.

4 Combination of IBSs

In this section, we start by generalizing a class of combination operations, including the conjunctive and the disjunctive sums, to the case of IBSs. We then show how normalization procedures may be extended to IBSs.

4.1 Combination of two IBSs

4.1.1 Definition

Let $\circledast$ denote a binary operation on BSs induced by some set operation $*$. This operation may be extended to IBSs by considering the lower and upper bounds of $m_1 \circledast m_2$, for any $A \subseteq \Omega$.

DEFINITION 4 (COMBINATION OF TWO IBSS)
Let $\mathbf{m}_1$ and $\mathbf{m}_2$ be two IBSs on the same frame Ω, and let $\circledast$ be a binary operation on BSs induced by some set operation $*$. The combination of $\mathbf{m}_1$ and $\mathbf{m}_2$ by $\circledast$ is defined as the IBS $\mathbf{m} = \mathbf{m}_1 \circledast \mathbf{m}_2$ with bounds:

$$
\begin{aligned}
m^-(A) &= \min_{(m_1,m_2)\in\mathbf{m}_1\times\mathbf{m}_2} (m_1 \circledast m_2)(A) \\
m^+(A) &= \max_{(m_1,m_2)\in\mathbf{m}_1\times\mathbf{m}_2} (m_1 \circledast m_2)(A)
\end{aligned}
$$

for all $A \subseteq \Omega$.

Remark: An alternative approach to define the combination of two IBSs $\mathbf{m}_1$ and $\mathbf{m}_2$ could be to consider the set M of all the BSs obtained by combining one BS in $\mathbf{m}_1$ with one BS in $\mathbf{m}_2$:

$$M = \{m | \exists (m_1, m_2) \in \mathbf{m}_1 \times \mathbf{m}_2, m = m_1 \circledast m_2\} \tag{16}$$

We have obviously $M \subseteq \mathbf{m}_1 \circledast \mathbf{m}_2$, but this inclusion is strict in general, as shown by the following example.

EXAMPLE 4 Let us assume that $\mathcal{F}(\mathbf{m}_1) = \{A, \Omega\}$, $\mathcal{F}(\mathbf{m}_2) = \{B, \Omega\}$, with $C = A \cap B \notin \{A, B\}$, and:

$$
\begin{array}{ll}
\mathbf{m}_1(A) = [0, 0.5] & \mathbf{m}_1(\Omega) = [0.5, 1] \\
\mathbf{m}_2(B) = [0, 0.5] & \mathbf{m}_2(\Omega) = [0.5, 1]
\end{array}
$$

Let us compute the conjunctive sum $\mathbf{m}$ of $\mathbf{m}_1$ and $\mathbf{m}_2$. Let m be the conjunctive sum of $m_1 \in \mathbf{m}_1$ and $m_2 \in \mathbf{m}_2$. It has four focal elements: A, B, C and Ω. The belief masses are:

$$
\begin{aligned}
m(A) &= m_1(A)m_2(\Omega) \\
m(B) &= m_1(\Omega)m_2(B) \\
m(C) &= m_1(A)m_2(B) \\
m(\Omega) &= m_1(\Omega)m_2(\Omega)
\end{aligned}
$$

Hence, we have:

$$\begin{aligned}
\mathbf{m}(A) &= [m_1^-(A)m_2^-(\Omega), m_1^+(A)m_2^+(\Omega)] = [0, 0.5]\\
\mathbf{m}(B) &= [m_1^-(\Omega)m_2^-(B), m_1^+(\Omega)m_2^+(B)] = [0, 0.5]\\
\mathbf{m}(C) &= [m_1^-(A)m_2^-(B), m_1^+(A)m_2^+(B)] = [0, 0.25]\\
\mathbf{m}(\Omega) &= [m_1^-(\Omega)m_2^-(\Omega), m_1^+(\Omega)m_2^+(\Omega)] = [0.25, 1]
\end{aligned}$$

Let $m \in \mathbf{m}$ defined by $m(A) = 0.4$, $m(B) = 0.2$, $m(C) = 0.1$ and $m(\Omega) = 0.3$. Let us show that it is impossible to find $m_1 \in \mathbf{m}_1$ and $m_2 \in \mathbf{m}_2$ such that $m = m_1 \wedge m_2$. Let $x = m_1(A)$ and $y = m_2(B)$. These quantities must be solutions of a system of four equations:

$$x(1-y) = 0.4 \tag{17}$$

$$(1-x)y = 0.2 \tag{18}$$

$$xy = 0.1 \tag{19}$$

$$(1-x)(1-y) = 0.3 \tag{20}$$

It is easy to see that this system is incompatible: Equations 17 and 19 imply that $x = 0.5$ and $y = 0.2$, which makes it impossible to satisfy the other two equations.

This example shows that Equation 16 is not a good candidate for defining the combination of two IBS, since M is not, in general, an IBS.

4.1.2 Practical calculation

In practice, the computation of $m^-(A)$ and $m^+(A)$ requires to search for the minimum and the maximum of

$$\varphi_A(m_1, m_2) = \sum_{B*C=A} m_1(B)m_2(C) \tag{21}$$

under the constraints:

$$\begin{aligned}
\sum_{B\in\mathcal{F}(\mathbf{m}_1)} m_1(B) &= 1\\
\sum_{C\in\mathcal{F}(\mathbf{m}_2)} m_2(C) &= 1
\end{aligned}$$

$$\begin{array}{lllll}
m_1^-(B) & \leq m_1(B) & \leq m_1^+(B) & & \forall B \in \mathcal{F}(\mathbf{m}_1)\\
m_2^-(C) & \leq m_2(C) & \leq m_2^+(C) & & \forall C \in \mathcal{F}(\mathbf{m}_2)
\end{array}$$

The solution of this quadratic programming problem is trivial when the right-hand side of Equation 21 contains only one term (as in Example 4), since we then have a function of non interactive variables. In Appendix A, we give an analytic solution for a more general case of particular interest: the conjunctive sum of an arbitrary IBS with a simple IBS (i.e., an IBS with only one focal element in addition to the possibility space).

In the most general case (combination of two arbitrary IBSs), an explicit solution seems difficult to obtain, and we have to resort to some kind of iterative optimization procedure. The particular form of the function to be optimized suggests to employ the following alternate directions scheme, which proved experimentally to be very effective.

Consider for example the minimization of $\varphi_A(m_1, m_2)$. Let us fix m_1 and m_2 to some admissible values $m_1^{(0)}$ and $m_2^{(0)}$, respectively. Then $\varphi_A(m_1, m_2^{(0)})$ is a linear function of the $m_1(B)$, for $B \in \mathcal{F}(m_1)$:

$$\varphi_A(m_1, m_2^{(0)}) = \sum_{B \in \mathcal{F}(m_1)} m_1(B) \left(\sum_{B * C = A} m_2^{(0)}(C) \right)$$

The search for m_1 minimizing this expression is a linear programming problem that may be solved directly using Theorem 1. Let $m_1^{(1)}$ be a solution (if $m_1^{(0)}$ was already a solution, then we pose $m_1^{(1)} = m_1^{(0)}$). We then proceed by searching $m_2^{(1)}$ minimizing $\varphi_A(m_1^{(1)}, m_2)$. The procedure is iterated until a fixed point has been found, i.e., until we have reached k such that $m_1^{(k)} = m_1^{(k-1)}$ and $m_2^{(k)} = m_2^{(k-1)}$. More formally, the algorithm may be described as follows:

1. Initialize $m_1^{(0)}$ and $m_2^{(0)}$ to random admissible values.
2. $k \leftarrow 0$.
3. **repeat**
   (a) $k \leftarrow k + 1$.
   (b) Find m_1^* solution of
   $$\min_{m_1} \varphi_A(m_1, m_2^{(k-1)})$$
   (c) **If** $\varphi_A(m_1^*, m_2^{(k-1)}) < \varphi_A(m_1^{(k-1)}, m_2^{(k-1)})$ **then** $m_1^{(k)} \leftarrow m_1^*$ **else** $m_1^{(k)} \leftarrow m_1^{(k-1)}$ **endif.**
   (d) Find m_2^* solution of
   $$\min_{m_2} \varphi_A(m_1^{(k)}, m_2)$$
   (e) **If** $\varphi_A(m_1^{(k)}, m_2^*) < \varphi_A(m_1^{(k)}, m_2^{(k-1)})$ **then** $m_2^{(k)} \leftarrow m_2^*$ **else** $m_2^{(k)} \leftarrow m_2^{(k-1)}$ **endif.**

   until $m_1^{(k)} = m_1^{(k-1)}$ and $m_2^{(k)} = m_2^{(k-1)}$.

To prove that a fixed point is always reached, it is sufficient to notice that the minimization and maximization problems considered in Theorem 1 may only have a finite number of possible solutions when the values of the coefficients c_i, $i = 1, \ldots, n$ are varied. Hence, at each iteration of the algorithm, the pair $s = (m_1, m_2)$ may only take a finite number of possible values. Let us denote as $s_1, \ldots, s_r$ these values, and $\varphi_{A,1}, \ldots, \varphi_{A,r}$ the corresponding values of the objective function. At each iteration, s is changed from s_i to s_j only if $\varphi_{A,i} < \varphi_{A,j}$ (we still consider the minimization problem, but similar arguments obviously hold for the maximization one). Hence, the objective function strictly decreases at each iteration. Since it may only take a finite number of values, a stable point must be reached after a finite number of iterations.

EXAMPLE 5 Two IBSs $\mathbf{m}_1$ and $\mathbf{m}_2$ on $\Omega = \{a, b, c\}$ and their conjunctive sum computed using the above algorithm are shown in Table 4.

4.2 Combination of several IBSs

The extension of the $\circledast$ operation from BSs to IBSs as presented in the previous section does not generally preserve the associativity property. This can be shown using the following counterexample.

EXAMPLE 6 Let $\mathbf{m}_1$, $\mathbf{m}_2$ and $\mathbf{m}_3$ be three simple IBSs defined as shown in Table 5. Using the formula established in Appendix A, we find that

$$(\mathbf{m}_1 \wedge \mathbf{m}_2) \wedge \mathbf{m}_3 \neq \mathbf{m}_1 \wedge (\mathbf{m}_2 \wedge \mathbf{m}_3)$$

This lack of associativity is obviously a drawback, since it makes the result of the combination of several IBSs dependent on the order in which they are combined. An approach to solve this problem is to generalize Definition 4 to allow the combination of n IBSs in one step, as proposed in this section.

DEFINITION 5 (COMBINATION OF n IBSS)
Let $\mathbf{m}_1, \ldots, \mathbf{m}_n$ be n IBSs on the same frame Ω, and let $\circledast$ be a transitive operation on BSs induced by some set operation $*$. The combination of $\mathbf{m}_1, \ldots, \mathbf{m}_n$ by $\circledast$ is defined as the IBS $\mathbf{m} = \mathbf{m}_1 \circledast \ldots \circledast \mathbf{m}_n$ with bounds:

$$\begin{aligned} m^-(A) &= \min_{(m_1,\ldots,m_n)\in \mathbf{m}_1 \times \ldots \times \mathbf{m}_n} (m_1 \circledast \ldots \circledast m_n)(A) \\ m^+(A) &= \max_{(m_1,\ldots,m_n)\in \mathbf{m}_1 \times \ldots \times \mathbf{m}_n} (m_1 \circledast \ldots \circledast m_n)(A) \end{aligned}$$

for all $A \in [0,1]^\Omega$.

The alternate directions algorithm proposed in Section 4.1.2 may easily be generalized to the combination of n IBSs, the value of m_i at iteration k being determined by finding the solution of

$$\min_{m_i} \varphi(A)(m_1^{(k)}, \ldots, m_{i-1}^{(k)}, m_i, m_{i+1}^{(k-1)}, \ldots, m_n^{(k-1)}).$$

EXAMPLE 7 With the data of Example 6, we find

$$\mathbf{m}_1 \wedge \mathbf{m}_2 \wedge \mathbf{m}_3 = (\mathbf{m}_1 \wedge \mathbf{m}_2) \wedge \mathbf{m}_3.$$

Remark: If the IBSs are provided one at a time and the storage resources are limited, then it may more convenient to combine the IBSs in any order one by one according to Definition 4, and to regard the result as an approximation to the n-ary combination. This approach is justified by the following proposition.

PROPOSITION 2

Let $\mathbf{m}_1, \ldots, \mathbf{m}_n$ be n IBSs, and $\circledast$ an operation on IBSs induced by some transitive operation on BSs. Then

$$(\ldots((\mathbf{m}_1 \circledast \mathbf{m}_2) \circledast \mathbf{m}_3) \circledast \ldots) \circledast \mathbf{m}_n \supseteq \mathbf{m}_1 \circledast \ldots \circledast \mathbf{m}_n$$

Proof. We detail the proof for $n = 3$. Extension to arbitrary n is easily performed by recurrence on n.

Let $M_{1,2} = \{m | \exists m_1 \in \mathbf{m}_1, \exists m_2 \in \mathbf{m}_2, m = m_1 \circledast m_2\}$. We have already noticed that $M_{1,2} \subseteq \mathbf{m}_1 \circledast \mathbf{m}_2$. We thus have for any subset A of Ω:

$$\begin{aligned}
(\mathbf{m}_1 \circledast \mathbf{m}_2 \circledast \mathbf{m}_3)^-(A) &= \min_{(m_1,m_2,m_3) \in \mathbf{m}_1 \times \mathbf{m}_2 \times \mathbf{m}_3} (m_1 \circledast m_2) \circledast m_3 \\
&= \min_{(m,m_3) \in M_{1,2} \times \mathbf{m}_3} m \circledast m_3 \\
&\geq \min_{(m,m_3) \in (\mathbf{m}_1 \circledast \mathbf{m}_2) \times \mathbf{m}_3} m \circledast m_3 = \\
&\quad ((\mathbf{m}_1 \circledast \mathbf{m}_2) \circledast \mathbf{m}_3)^-(A)
\end{aligned}$$

Similarly,

$$(\mathbf{m}_1 \circledast \mathbf{m}_2 \circledast \mathbf{m}_3)^+(A) \leq ((\mathbf{m}_1 \circledast \mathbf{m}_2) \circledast \mathbf{m}_3)^+(A).$$

□

4.3 Normalization

If one is absolutely sure that the true value of the parameter of interest lies in the possibility space Ω, then the normality condition should be imposed on BSs. The normalization rule initially proposed by Dempster consists in dividing $m(A)$ by $1 - m(\emptyset)$ for all $A \in \mathcal{F}^*(m) = \mathcal{F}(m) \setminus \emptyset$ (Equation 6). To avoid some counterintuitive effects of this rule in the case of strongly conflicting items of evidence, Yager [35] proposed a different normalization principle (hereafter referred to as "Yager normalization") in which the mass assigned to the empty set is transferred to the possibility space. In this section, we show how these two procedures may be extended to IBSs.

4.3.1 Dempster normalization

Let $\mathbf{m}$ be an IBS such that $m^-(\emptyset) < 1$. We may define a *normalized* version of $\mathbf{m}$ as the IBS $\mathbf{m}_d^*$ defined by

$$\mathcal{F}(\mathbf{m}_d^*) = \mathcal{F}^*(\mathbf{m})$$

and the following bounds:

$$m_d^{*-}(A) = \min_{m \in \mathbf{m}} \frac{m(A)}{1 - m(\emptyset)} \tag{22}$$

$$m_d^{*+}(A) = \max_{m \in \mathbf{m}} \frac{m(A)}{1 - m(\emptyset)} \tag{23}$$

for all $A \in \mathcal{F}(\mathbf{m}_d^*)$. The following theorem gives the values of $m_d^{*-}(A)$ and $m_d^{*+}(A)$.

Theorem 2 (Dempster normalization of an IBS)
The normalized version $\mathbf{m}_d^*$ of $\mathbf{m}$ has bounds:

$$
\begin{aligned}
m_d^{*-}(A) &= \frac{m^-(A)}{1 - \max\left[m^-(\emptyset), 1 - \sum_{B \neq A, B \neq \emptyset} m^+(B) - m^-(A)\right]} \\
m_d^{*+}(A) &= \frac{m^+(A)}{1 - \min\left[m^+(\emptyset), 1 - \sum_{B \neq A, B \neq \emptyset} m^-(B) - m^+(A)\right]}
\end{aligned}
$$

for all $A \in \mathcal{F}(\mathbf{m}_d^*)$.

Proof. Let us note $x = m(A)$ and $y = m(\emptyset)$ for arbitrary $m \in \mathbf{m}$. We have to find the extrema of

$$f(x, y) = \frac{x}{1-y}$$

under the constraints:

$$\alpha_1 = m^-(A) \leq x \leq m^+(A) = \beta_1$$

$$\alpha_0 = m^-(\emptyset) \leq y \leq m^+(\emptyset) = \beta_0$$

$$\alpha_2 = 1 - \sum_{B \neq A, B \neq \emptyset} m^+(B) \leq x + y \leq 1 - \sum_{B \neq A, B \neq \emptyset} m^-(B) = \beta_2$$

The admissible region $\mathcal{R}$ may be represented graphically in the (x, y) plane (Figure 1). It is delimited by 6 lines corresponding to the 6 inequality constraints. Points A and C of coordinates (α_1, β_0) and (β_1, α_0), respectively, always belong to $\mathcal{R}$, since we have:

$$
\begin{aligned}
\alpha_1 \geq \alpha_2 - \beta_0 &\Rightarrow \alpha_1 + \beta_0 \geq \alpha_2 \\
\beta_0 \leq \beta_2 - \alpha_1 &\Rightarrow \alpha_1 + \beta_0 \leq \beta_2 \\
\alpha_0 \geq \alpha_2 - \beta_1 &\Rightarrow \alpha_0 + \beta_1 \geq \alpha_2 \\
\beta_1 \leq \beta_2 - \alpha_0 &\Rightarrow \alpha_0 + \beta_1 \leq \beta_2
\end{aligned}
$$

The points (x, y) such that $f(x, y) = \eta$ are situated along the line (L_η) with equation

$$x + \eta y = \eta$$

When η is gradually increased from 0 to 1, (L_η) first intersects $\mathcal{R}$ at the point $P = (\alpha_1, \max(\alpha_0, \alpha_2 - \alpha_1))$. Hence,

$$\min f = \frac{\alpha_1}{1 - \max(\alpha_0, \alpha_2 - \alpha_1)}.$$

Similarly, when η is decreased from 1 to 0, (L_η) meets $\mathcal{R}$ at the point $Q = (\beta_1, \min(\beta_0, \beta_2 - \beta_1))$, which entails that

$$\max f = \frac{\beta_1}{1 - \min(\beta_0, \beta_2 - \beta_1)}.$$

□

4.3.2 Yager normalization

In [35], Yager suggests an alternative normalization procedure in which the mass $m(\emptyset)$ is simply transferred to Ω. We thus have:

$$m_y^*(A) = \begin{cases} m(A) & \text{if } \; A \notin \{\emptyset, \Omega\} \\ 0 & \text{if } \; A = \emptyset \\ m(\Omega) + m(\emptyset) & \text{if } \; A = \Omega \end{cases}$$

This simple rule may be readily extended to IBSs: we have $m_y^{*-}(A) = m^-(A)$ and $m_y^{*+}(A) = m^+(A)$ for all $A \notin \{\emptyset, \Omega\}$,

$$m_y^{*-}(\emptyset) = m_y^{*+}(\emptyset) = 0,$$

and

$$\begin{aligned} m_y^{*-}(\Omega) &= \max\left(m^-(\Omega) + m^-(\emptyset), 1 - \sum_{B \notin \{\emptyset, \Omega\}} m^+(B)\right) \\ m_y^{*+}(\Omega) &= \min\left(m^+(\Omega) + m^+(\emptyset), 1 - \sum_{B \notin \{\emptyset, \Omega\}} m^-(B)\right). \end{aligned}$$

Example 8 Typical results using the two normalization procedures described above are shown in Table 6.

5 Uncertainty and information

5.1 Evidential uncertainty measures

Measuring the "amount of uncertainty" involved in a problem-solving situation, or the "amount of information" contained in an item of evidence, is an important problem which is usually tackled, within the framework of Probability Theory, using the concepts of entropy and information established by Claude Shannon [13]. More recently, many research efforts have been devoted to the development of a "Generalized Information Theory" whose purpose is to allow for uncertainty measurement in the wider context of non-probabilistic models such as that of belief functions (see e.g. [33, 19, 7, 16, 11, 13]). Although no single measure seems to have unquestionably emerged from this research work, some available results may already find interesting applications in such contexts as knowledge elicitation or statistical inference. The purpose of this section is to present a few evidential uncertainty measures which appear to us as particularly promising, and to study their extension to imprecise belief structures. An immediate application of these measures is to allow the definition of a "maximally uncertain" BS compatible with an IBS.

One of the first attempts to characterize the amount of information provided by a piece of evidence was performed by Smets [19], who defined a function I over the set of belief structure as:

$$I(m) = \sum_{A \subseteq \Omega} q_m(A) \tag{24}$$

The main justification of this measure of information lies in the following additivity property, which is a direct consequence of Equation 8:

$$I(m_1 \wedge m_2) = I(m_1) + I(m_2)$$

Hence, the amount of information of the conjunctive combination of two distinct pieces of evidence is the sum of the information of these two pieces of evidence. It should be noted that function I is infinite in the case of a dogmatic BS, i.e., a BS m such that $m(\Omega) = 0$.

Other authors took a different path and tried to directly generalize the Shannon entropy to belief functions. An interesting approach, proposed by Klir [11, 13], relies on a distinction between two types of uncertainty, both modeled by a belief function: nonspecifity, and discord or strife. Nonspecificity is properly measured by the following function:

$$N(m) = \sum_{A \in \mathcal{F}^*(m)} m(A) \log_2 |A| \tag{25}$$

which was shown to be unique under some well-defined requirements [16, 12]. On the other hand, the concept of strife refers to the total conflict of evidential claims within a body of evidence, and may be defined by the form:

$$S(m) = -\sum_{A \in \mathcal{F}(m)} m(A) \log_2 \left[\sum_{B \in \mathcal{F}(m)} m(B) \frac{|A \cap B|}{|A|} \right]. \tag{26}$$

It is then natural (although not theoretically justified) to define the total uncertainty by the equation:

$$NS(m) = N(m) + S(m). \tag{27}$$

5.2 Application to IBSs

The different uncertainty measures proposed for belief structures may be extended to imprecise belief structures in the following way. Let

$$U : \mathcal{S}(\Omega) \mapsto \mathbb{R}$$

denote an uncertainty measure for BSs. We may define $U(\mathbf{m})$ for some IBS $\mathbf{m}$ as an interval whose bounds are the minimum and maximum values of $U(m)$ for all $m \in \mathbf{m}$:

$$U(\mathbf{m}) = [\min_{m \in \mathbf{m}} U(m), \max_{m \in \mathbf{m}} U(m)].$$

When U denotes the nonspecificity measure, then the calculation of $U(\mathbf{m})$ is straightforward using Theorem 1, since $N(m)$ is a linear function of the belief masses. The computation of the other measures is more complex, as it involves the numerical resolution of non linear programming problems.

The concept of uncertainty measure of an IBS has interesting applications. For example, the width of the uncertainty interval $\Delta U(\mathbf{m}) = U^+(\mathbf{m}) - U^-(\mathbf{m})$ may be used to quantify the imprecision of an IBS. Let us assume that an agent's degrees of belief are properly described by some unknown BS m, and that two elicitation procedures have produced two IBSs $\mathbf{m}_1$ and $\mathbf{m}_2$, considered to be imprecise estimates of m. Then, $\mathbf{m}_1$ can be said to be more precise than $\mathbf{m}_2$ (according to uncertainty measure U), whenever $\Delta U(\mathbf{m}_1) \leq \Delta U(\mathbf{m}_2)$.

Another application of uncertainty measures concerns the definition of the "most uncertain" BS compatible with an IBS. Assume that your beliefs are described by an IBS $\mathbf{m}$, and that, for some purpose (e.g., decision making), you have to select a particular BS $m \in \mathbf{m}$. This is a typical instance of ampliative reasoning (i .e., reasoning in which conclusions are not entailed by the premises). In such a situation, it is wise to apply the Principle of Maximum Uncertainty [13], which requires that no unavailable information be used in the inference process. In this case, the application of this principle leads to the selection of the BS $m_{\max}$ reflecting maximum uncertainty (or minimal information content).

EXAMPLE 9 Let $\mathbf{m}$ denote the IBS on $\Omega = \{a, b, c\}$ defined by

$$\mathcal{F}(\mathbf{m}) = \{\{a\}, \{a, b\}, \{b, c\}, \{a, b, c\}\}$$

and

$$m(\{a\}) = [0.1, 0.7] \quad m(\{a, b\}) = [0, 0.4]$$
$$m(\{b, c\}) = [0, 0.6] \quad m(\{a, b, c\}) = [0.2, 0.5]$$

Table 7 shows the uncertainty measure intervals and the corresponding minimally and maximally uncertain BSs, for each of the uncertainty measures considered in this paper. It may be noticed that the nonspecificity criterion has many minima and maxima, and consequently does not allow in this case to select a single maximally uncertain BS. The S and NS criteria tend to favor maximally conflicting BSs (i.e., BSs in which the mass is distributed among disjoint subsets). It may also be noticed that the BS with minimal information content is also that with minimal strife, and is very different from that with maximal total uncertainty.

6 Application to the classification of imprecise data

6.1 Pattern classification and the TBM

Discriminant analysis, or pattern classification, is concerned with the assignment of entities, represented by feature vectors in $\mathbb{R}^d$, to predefined categories or classes. Typically, a decision rule is elaborated using a learning set of N vectors with known classification. In conventional statistical approaches to this problem, each vector is assumed to be drawn from a certain probability distribution, and decisions regarding previously unseen vectors are based on the estimated posterior probabilities of each class.

In [2], we showed that this problem could be addressed in a radically different manner using the TBM. In this new approach, a belief structure is computed for each new pattern to be classified, on the basis of its similarity to training vectors. Among other advantages, this method allows a decision rule to be established from training data with imprecise labeling. Such a situation typically arises, e.g., in medical diagnosis problems in which some records in a data base are related to patients for which only a partial or uncertain diagnosis is available. An incremental learning procedure for this method was described in [39, 40], and decision-theoretic issues were examined in some detail in [3]. In the following section, we show how the new concepts presented in this paper may be used to extend, in a natural way, this classification method to the more general case where each component of the feature vector is only known to lie within a certain interval.

Before describing this new approach, let us first summarize the method exposed in [2] and introduce some basic notation. Let $\Omega = \{\omega_1, \ldots, \omega_M\}$ denote the set of categories. The learning set is composed of N examples of the form $z^i = (\mathbf{x}^i, A^i)$, where $\mathbf{x}^i$ denotes a vector in $\mathbb{R}^d$, and A^i a subset of Ω containing the category of $\mathbf{x}^i$ (the class of the entity described by $\mathbf{x}^i$ is only known to belong to A^i)[3]. For each new vector $\mathbf{x}$, the consideration of example z^i induces a simple BS $m(\cdot|z^i)$ focused on A^i and Ω, the mass of belief assigned to A^i being defined as a decreasing function of the dissimilarity (according to some relevant measure δ) between $\mathbf{x}$ and $\mathbf{x}^i$:

$$m(A|z^i) = \begin{cases} \varphi(\delta(\mathbf{x}, \mathbf{x}^i)) & \text{if } A = A^i \\ 1 - \varphi(\delta(\mathbf{x}, \mathbf{x}^i)) & \text{if } A = \Omega \\ 0 & \text{otherwise} \end{cases} \tag{28}$$

where φ is a decreasing function verifying $\varphi(0) \leq 1$ and $\lim_{d \to \infty} \varphi(d) = 0$. When δ denotes the Euclidean distance, a rational choice for φ was shown in [4] to be:

$$\varphi(d) = \alpha e^{-\gamma d^2} \tag{29}$$

where α and γ are parameters that may be learnt form training data [39]. The BSs induced by each learning sample are then combined using Dempster's rule:

$$m = m(\cdot|z^1) \oplus \ldots \oplus m(\cdot|z^N) \tag{30}$$

[3] In [2], we actually considered an even more general situation in which each training example is of the form $(\mathbf{x}^i, m^i)$, where m^i denotes a BS describing one's belief in the class of $\mathbf{x}^i$. Our analysis can easily be generalized to this general case (with slightly more complex notations however).

6.2 Extension to interval-valued features

Let us now consider the more general case in which each component x_j^i of a training pattern $\mathbf{x}^i$ is only known to belong to a certain interval: $x_j^i \in [x_j^{i-}, x_j^{i+}]$. Let $\mathbf{x} = (x_1, \ldots, x_d)^t$ denote an arbitrary interval-valued feature vector such that $x_i \in [x_i^-, x_i^+]$ for all $1 \leq i \leq d$. Using the rules of interval arithmetics [10], it is possible to compute lower and upper bounds for the dissimilarity between $\mathbf{x}$ and $\mathbf{x}^i$:

$$\delta(\mathbf{x}, \mathbf{x}^i) \in [\delta^{i-}, \delta^{i+}] \tag{31}$$

Assuming the choice of the exponential function for φ such as described by Equation 29, we can then compute lower and upper bounds for the BS $m(\cdot|z^i)$ defined by Equation 28, which yields an IBS $\mathbf{m}(\cdot|z^i)$ given by:

$$\mathbf{m}(A|z^i) = \begin{cases} \left[\alpha e^{-\gamma(\delta^{i+})^2}, \alpha e^{-\gamma(\delta^{i-})^2}\right] & \text{if } A = A^i \\ \left[1 - \alpha e^{-\gamma(\delta^{i-})^2}, 1 - \alpha e^{-\gamma(\delta^{i+})^2}\right] & \text{if } A = \Omega \\ 0 & \text{otherwise} \end{cases} \tag{32}$$

The IBSs induced by each of the N training samples may then be combined using the n-ary conjunctive sum operation defined in Section 4.

Remark: The learning procedure proposed in [39] for optimizing the parameters α and γ has to be generalized in that case. This may be achieved by defining a suitable performance criterion. This point would require substantially more developments and will not be studied here. For simplicity, we shall assume parameters α and γ to be fixed by the user.

6.3 Example

A simple three-class data set composed of 6 patterns in two dimensions is shown in Table 8 and represented graphically in Figure 2. Each learning example consists in two intervals corresponding to each of the two features, and a class label. Three samples (z^2, z^4 and z^6) have incomplete class specification. An additional interval-valued feature vector $\mathbf{x}$ has unknown class membership. The problem is to quantify our belief concerning the class of $\mathbf{x}$, based on the evidence of the training set. Parameters α and γ were given arbitrary values of 0.9 and 0.5, respectively.

The unnormalized IBS $\mathbf{m}$ characterizing the uncertainty on the class of $\mathbf{x}$ was computed using the iterative algorithm described in Section 4.2 as:

$$\mathbf{m} = \mathbf{m}(\cdot|z^1) \wedge \mathbf{m}(\cdot|z^2) \wedge \mathbf{m}(\cdot|z^3) \wedge \mathbf{m}(\cdot|z^4) \wedge \mathbf{m}(\cdot|z^5) \wedge \mathbf{m}(\cdot|z^6)$$

Table 9 shows the result of this calculation, as well as the IBS $\widehat{\mathbf{m}}$ computed by combining the IBS $\mathbf{m}(\cdot|z^i)$ two by two using the formula presented in Appendix A:

$$\widehat{\mathbf{m}} = ((((\mathbf{m}(\cdot|z^1) \wedge \mathbf{m}(\cdot|z^2)) \wedge \mathbf{m}(\cdot|z^3)) \wedge \mathbf{m}(\cdot|z^4)) \wedge \mathbf{m}(\cdot|z^5)) \wedge \mathbf{m}(\cdot|z^6)$$

As can be seen from Table 9, $\widehat{\mathbf{m}}$ is a reasonably good approximation of $\mathbf{m}$, while requiring far less computations.

The interval-valued belief and plausibility functions induced by $\mathbf{m}$ are also shown in Table 9, as well as the BS m_0 obtained by replacing each interval $[x_j^{i-}, x_j^{i+}]$ by a single value $x_j^i = (x_j^{i-} + x_j^{i+})/2$ (which amounts to ignoring the imprecision on the feature values).

Let us now assume that some decision has to made, regarding the assignment of $\mathbf{x}$ to one of the three classes. Several strategies may be applied, for instance:

1. Compute the interval-valued pignistic probabilities $\mathbf{BetP}(\{\omega_i\})$, $i = 1, 2, 3$; if, for some i, $\mathrm{BetP}^-(\{\omega_i\}) > \mathrm{BetP}^+(\{\omega_j\})$ for all $j \neq i$, then assign $\mathbf{x}$ to class ω_i.

2. Applying the Principle of Maximum Uncertainty, compute a "maximally uncertain" or "least informative" BS $\widetilde{m}$ compatible with $\mathbf{m}$ (according to some uncertainty measure), and make a decision according to the pignistic probability function induced by $\widetilde{m}$.

These two strategies are illustrated in Table 10. The IBS $\mathbf{m}$ computed as explained above was normalized using the formula of Theorem 2, yielding a normalized IBS $\mathbf{m}_d^*$. We then computed the least informative BS $\widetilde{m}$ compatible with $\mathbf{m}_d^*$, defined by:

$$I(\widetilde{m}) = \min_{m \in \mathbf{m}_d^*} I(m)$$

where I denotes the information measure defined by Equation 24. The values of $\mathbf{BetP}_{\mathbf{m}_d^*}$ and $\mathrm{BetP}_{\widetilde{m}}$ are shown in Table 10. The pignistic probability function induced by m_0 (after normalization) is also given for comparison. As can be seen, the strategy based on imprecise pignistic probabilities leads to indecision in that case, while the use of $\widetilde{m}$ leads to assigning the pattern to class 1. The same decision is prescribed if imprecision on feature values is neglected; note however that a higher confidence is attached to the decision in that case, as a result of ignoring an important source of uncertainty.

7 Conclusions

In this paper, the Transferable Belief Model has been enriched with new concepts and techniques allowing to model the situation in which the beliefs held by a rational agent may only be expressed (or are only known) with some imprecision. Central to our approach is the concept of interval-valued belief structure, which is defined as a set of belief structures verifying certain constraints. Starting from this definition, many other concepts of Evidence Theory (including belief and plausibility functions, pignistic probabilities, combination rules and uncertainty measures) have been extended to cope with imprecision in the belief numbers attached to each hypothesis. An application of this new framework to the classification of patterns with partially known feature values has been demonstrated.

Although intervals are probably the simplest formalism for representing imprecise numerical values, there are certain contexts in other models of uncertainty should be prefered. In particular, the theory of fuzzy sets has proved very efficient for representing vague quantities expressed though verbal statements. Our current work aims at further extending the concepts presented in this paper to the case where beliefs are expressed as fuzzy numbers.

A Conjunctive sum with a simple IBS

Let us consider the simple case where an arbitrary IBS $\mathbf{m}_1$ is combined with a *simple* IBS $\mathbf{m}_2$ with $\mathcal{F}(\mathbf{m}_2) = \{A, \Omega\}$ $(A \subset \Omega)$. Let us denote $\mathbf{m} = \mathbf{m}_1 \wedge \mathbf{m}_2$. For any $B \subseteq \Omega$, $m_1 \in \mathbf{m}_1$ and $m_2 \in \mathbf{m}_2$, we then have:

$$
\begin{aligned}
m(B) &= \sum_{C \cap D = B} m_1(C) m_2(D) && (33)\\
&= m_2(A) \sum_{A \cap C = B} m_1(C) + m_2(\Omega) m_1(B) && (34)
\end{aligned}
$$

To find the minimum and maximum of $m(B)$, let us consider two cases.

Case 1: $B \not\subseteq A$. In that case, the first term in Equation 34 vanishes, and we have:

$$m(B) = m_2(\Omega) m_1(B)$$

It is then obvious that

$$
\begin{aligned}
m^-(B) &= m_2^-(\Omega) m_1^-(B) && (35)\\
m^+(B) &= m_2^+(\Omega) m_1^+(B) && (36)
\end{aligned}
$$

Case 2: $B \subseteq A$. By noticing that $m_2(\Omega) = 1 - m_2(A)$ and rearranging the terms in Equation 34, we find:

$$m(B) = m_1(B) + m_2(A) \sum_{\substack{C : A \cap C = B \\ C \neq B}} m_1(C)$$

Let us denote

$$x = \sum_{A \cap C \neq B} m_1(C) \qquad y = \sum_{\substack{C : A \cap C = B \\ C \neq B}} m_1(C) \qquad z = m_1(B)$$

We have to find the minimum and maximum of:

$$m(B) = 0 \cdot x + m_2(A) y + z$$

under the constraints:

$$x + y + z = 1$$

$$\alpha_1 = \sum_{A \cap C \neq B} m_1^-(C) \leq x \leq \sum_{A \cap C \neq B} m_1^+(C) = \beta_1$$

$$\alpha_2 = \sum_{\substack{C : A \cap C = B \\ C \neq B}} m_1^-(C) \leq y \leq \sum_{\substack{C : A \cap C = B \\ C \neq B}} m_1^+(C) = \beta_2$$

$$\alpha_3 = m_1^-(B) \leq z \leq m_1^+(B) = \beta_3$$

By applying Theorem 1, we obtain:

$$\min_{m_1 \in \mathbf{m}_1} m(B) =$$
$$\max[m_2(A)\alpha_2 + \alpha_3, m_2(A)(1-\beta_1-\alpha_3)+\alpha_3, m_2(A)\beta_2 + 1 - \beta_1 - \beta_2] \quad (37)$$

$$\max_{m_1 \in \mathbf{m}_1} m(B) =$$
$$\min[m_2(A)\beta_2 + \beta_3, m_2(A)(1-\alpha_1-\beta_3)+\beta_3, m_2(A)\alpha_2 + 1 - \alpha_1 - \alpha_2] \quad (38)$$

Since for given m_1 the minimum (resp., the maximum) of $m(B)$ is realized for $m_2(A) = m_2^-(A)$ (resp., $m_2(A) = m_2^+(A)$), we finally have proved the following proposition:

Proposition 3
In the case where $B \subseteq A$, the bounds of m in Equation 34 are given by:

$$\begin{aligned} m^-(B) &= \max(X_1, X_2, X_3) \\ m^+(B) &= \min(Y_1, Y_2, Y_3) \end{aligned}$$

with

$$\begin{aligned} X_1 &= m_1^-(B) + m_2^-(A) \sum_{\substack{C: A\cap C = B \\ C \neq B}} m_1^-(C) \\ X_2 &= m_1^-(B) + m_2^-(A)\left(1 - \sum_{A\cap C \neq B} m_1^+(C) - m_1^-(B)\right) \\ X_3 &= 1 - \sum_{A\cap C \neq B} m_1^+(C) + (m_2^-(A) - 1) \sum_{\substack{C: A\cap C = B \\ C \neq B}} m_1^+(C) \end{aligned}$$

and

$$\begin{aligned} Y_1 &= m_1^+(B) + m_2^+(A) \sum_{\substack{C: A\cap C = B \\ C \neq B}} m_1^+(C) \\ Y_2 &= m_1^+(B) + m_2^+(A)\left(1 - \sum_{A\cap C \neq B} m_1^-(C) - m_1^+(B)\right) \\ Y_3 &= 1 - \sum_{A\cap C \neq B} m_1^-(C) + (m_2^+(A) - 1) \sum_{\substack{C: A\cap C = B \\ C \neq B}} m_1^-(C) \end{aligned}$$

References

[1] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, AMS-38:325–339, 1967.

[2] T. Denœux. A k-nearest neighbor classification rule based on Dempster-Shafer theory. *IEEE Transactions on Systems, Man and Cybernetics*, 25(05):804–813, 1995.

[3] T. Denœux. Analysis of evidence-theoretic decision rules for pattern classification. *Pattern Recognition*, 30(7):1095–1107, 1997.

[4] T. Denœux. Application du modèle des croyances transférables en reconnaissance de formes. *Traitement du Signal (In press)*, 1997.

[5] D. Dubois and H. Prade. Addition of interactive fuzzy numbers. *IEEE Transactions on Automatic Control*, 26(4):926–936, 1981.

[6] D. Dubois and H. Prade. The use of fuzzy numbers in decision analysis. In M. M. Gupta and E. Sanchez, editors, *Fuzzy Information and Decision Processes*, pages 309–321. North-holland, New-York, 1982.

[7] D. Dubois and H. Prade. Properties of measures of information in evidence and possibility theories. *Fuzzy sets and systems*, 124:161–182, 1987.

[8] D. Dubois and H. Prade. *Possibility Theory: An approach to computerized processing of uncertainty.* Plenum Press, New-York, 1988.

[9] M. Ishizuka and K. S. Fu. Inference procedures under uncertainty for the problem-reduction method. *Information Sciences*, 28:179–206, 1982.

[10] A. Kaufmann and M. M. Gupta. *Introduction to fuzzy arithmetic. Theory and applications.* International Thomson Computer Press, London, 1991.

[11] G. J. Klir. Measures of uncertainty in the Dempster-Shafer theory of evidence. In R. R. Yager, M. Fedrizzi, and J. Kacprzyk, editors, *Advances in the Dempster-Shafer theory of evidence*, pages 35–49. John Wiley and Sons, New-York, 1994.

[12] G. J. Klir and M. Mariano. On the uniqueness of possibilistic measure of uncertainty and information. *Fuzzy sets and systems*, 24:197–219, 1987.

[13] G. J. Klir and B. Yuan. *Fuzzy sets and fuzzy logic. Theory and applications.* Prentice-Hall, 1995.

[14] M. Lamata and S. Moral. Calculus with linguistic probabilities and beliefs. In R. R. Yager, M. Fedrizzi, and J. Kacprzyk, editors, *Advances in the Dempster-Shafer theory of evidence*, pages 133–152. Wiley, 1994.

[15] H. Ogawa, K. S. Fu, and J. T. P. Yao. An inexact inference for damage assessment of existing structures. *Int. J. Man-Machine Studies*, 22:295–306, 1985.

[16] A. Ramer. Uniqueness of information measure in the theory of evidence. *Fuzzy sets and systems*, 24:183–196, 1987.

[17] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, Princeton, N.J., 1976.

[18] P. Smets. The degree of belief in a fuzzy event. *Information Sciences*, 25:1–19, 1981.

[19] P. Smets. Information content of an evidence. *International Journal of Man-Machine Studies*, 19:33–43, 1983.

[20] P. Smets. The combination of evidence in the Transferable Belief Model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5):447–458, 1990.

[21] P. Smets. Constructing the pignistic probability function in a context of uncertainty. In M. Henrion, R. D. Schachter, L. N. Kanal, and J. F. Lemmer, editors, *Uncertainty in Artificial Intelligence 5*, pages 29–40. North-Holland, Amsterdam, 1990.

[22] P. Smets. Resolving misunderstandings about belief functions. *International Journal of Approximate Reasoning*, 6:321–344, 1990.

[23] P. Smets. The Transferable Belief Model and Possibility Theory. Technical Report TR/IRIDIA/90-2, IRIDIA, Bruxelles, 1990.

[24] P. Smets. The Transferable Belief Model and random sets. *International Journal of Intelligent Systems*, 7:37–46, 1990.

[25] P. Smets. Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem. *International Journal of Approximate Reasoning*, 9:1–35, 1993.

[26] P. Smets. Belief induced by the partial knowledge of the probabilities. In D. Heckerman, D. Poole, and R. Lopez de mantaras, editors, *Uncertainty in AI'94*, pages 523–530. Morgan Kaufmann, San Mateao, 1994.

[27] P. Smets. What is Dempster-Shafer's model ? In R. R. Yager, M. Fedrizzi, and J. Kacprzyk, editors, *Advances in the Dempster-Shafer theory of evidence*, pages 5–34. Wiley, 1994.

[28] P. Smets. Non standard probabilistic and non probabilistic representations of uncertainty. Technical Report TR/IRIDIA/95-2, IRIDIA, Bruxelles, 1995.

[29] P. Smets. The normative representation of quantified beliefs by belief functions. *Artificial Intelligence*, 92(1–2):229–242, 1997.

[30] P. Smets and R. Kennes. The Transferable Belief Model. *Artificial Intelligence*, 66:191–243, 1994.

[31] P. Walley. *Statistical Reasoning with Imprecise Probabilities.* Chapman and Hall, London, 1991.

[32] R. R. Yager. Generalized probabilities of fuzzy events from fuzzy belief structures. *Information Sciences*, 28:45–62, 1982.

[33] R. R. Yager. Entropy and specificity in a mathematical theory of evidence. *International Journal og general Systems*, 9:249–260, 1983.

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

[35] R. R. Yager. On the normalization of fuzzy belief structure. *International Journal of Approximate Reasoning*, 14:127–153, 1996.

[36] R. R. Yager and D. P. Filev. Including probabilistic uncertainty in fuzzy logic controller modeling using Dempster-Shafer theory. *IEEE Transactions on Systems, Man and Cybernetics*, 25(8):1221–1230, 1995.

[37] J. Yen. Generalizing the Dempster-Shafer theory to fuzzy sets. *IEEE Transactions on Systems, Man and Cybernetics*, 20(3):559–569, 1990.

[38] L. A. Zadeh. Fuzzy sets and information granularity. In R. K. Ragade M. M. Gupta and R. R. Yager, editors, *Advances in Fuzzy Sets Theory and Applications*, pages 3–18. North-Holland Publishing Co., 1979.

[39] L. M. Zouhal and T. Denœux. An adaptive k-NN rule based on Dempster-Shafer theory. In *Proc. of the 6th Int. Conf. on Computer Analysis of Images and Patterns (CAIP'95)*, pages 310–317, Prague, September 1995. Springer Verlag.

[40] L. M. Zouhal and T. Denœux. A comparison between fuzzy and evidence-theoretic k-NN rules for pattern recognition. In *Proceedings of EUFIT'95*, volume 3, pages 1319–1325, Aachen, August 1995.

List of Tables

Table 1: Example of credibility, plausibility and commonality intervals induced by an IBS (Example 2).

A	$\mathbf{m}(A)$	$\mathbf{bel}_{\mathbf{m}}(A)$	$\mathbf{pl}_{\mathbf{m}}(A)$	$\mathbf{q}_m(A)$
$\emptyset$	[0,0.2]	0	0	1
$\{a\}$	0	0	$[0.6, 0.8]$	$[0.6, 0.8]$
$\{b\}$	$[0.2, 0.4]$	$[0.2, 0.4]$	$[0.4, 0.6]$	$[0.5, 0.6]$
$\{c\}$	0	0	$[0.5, 0.7]$	$[0.5, 0.7]$
$\{a, b\}$	$[0.1, 0.2]$	$[0.3, 0.5]$	[0.8,1]	$[0.2, 0.4]$
$\{a, c\}$	$[0.4, 0.5]$	$[0.4, 0.5]$	$[0.6, 0.8]$	$[0.5, 0.7]$
$\{b, c\}$	0	[0.2 0.4]	[0.8,1]	$[0.1, 0.3]$
$\{a, b, c\}$	$[0.1, 0.3]$	[0.8,1]	[0.8,1]	$[0.1, 0.3]$

Table 2: A belief structure and its associated belief function.

	$\emptyset$	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	$\{a, b, c\}$
m	0	0	0.3	0	0	0.5	0.1	0.1
bel_m	0	0	0.3	0	0.3	0.5	0.4	1

Table 3: An IBS and its associated pignistic probability function (Example 3).

A	$\mathbf{m}(A)$	$\mathbf{BetP}_{\mathbf{m}}(A)$
$\{a\}$	0	$[0.283, 0.383]$
$\{b\}$	$[0.2, 0.4]$	$[0.317, 0.483]$
$\{c\}$	0	$[0.233, 0.317]$
$\{a, b\}$	$[0.1, 0.2]$	[0.683,0.767]
$\{a, c\}$	$[0.4, 0.5]$	$[0.517, 0.683]$
$\{b, c\}$	0	[0.617,0.717]
$\{a, b, c\}$	$[0.1, 0.3]$	1

Table 4: Conjunctive sum of two IBSs (Example 5).

	$\{a\}$	$\{a, b\}$	$\{a, c\}$	$\{a, b, c\}$
$\mathbf{m}_1$	[0.2,0.5]	[0.3,0.7]	[0,0.4]	[0.1,0.5]
$\mathbf{m}_2$	[0.2,0.5]	[0.1,0.2]	[0.3,0.7]	[0,0.4]
$\mathbf{m}_1 \wedge \mathbf{m}_2$	[0.45,0.91]	[0.04,0.37]	[0.03,0.35]	[0,0.2]

Table 5: Non associativity of the conjunctive sum in the case of IBSs (Example 6).

	$\{a\}$	$\{a,b\}$	$\{a,c\}$	$\{a,b,c\}$
$\mathbf{m}_1$		[0.6,0.8]		[0.2,0.4]
$\mathbf{m}_2$		[0.1,0.6]		[0.4,0.9]
$\mathbf{m}_3$			[0.2,0.6]	[0.4,0.8]
$\mathbf{m}_1 \wedge \mathbf{m}_2$		[0.64,0.92]		[0.08,0.36]
$\mathbf{m}_2 \wedge \mathbf{m}_3$	[0.02,0.36]	[0.04,0.48]	[0.08,0.54]	[0.16,0.72]
$(\mathbf{m}_1 \wedge \mathbf{m}_2) \wedge \mathbf{m}_3$	[0.128,0.552]	[0.256,0.736]	[0.016,0.216]	[0.032,0.288]
$\mathbf{m}_1 \wedge (\mathbf{m}_2 \wedge \mathbf{m}_3)$	[0.068,0.712]	[0.136,0.816]	[0.016,0.216]	[0.032,0.288]

Table 6: Example of Dempster and Yager normalization (Example 8).

A	$\mathbf{m}(A)$	$\mathbf{m}_d^*$	$\mathbf{m}_y^*$
$\emptyset$	[0.1,0.3]	0	0
$\{a\}$	[0.2 0.5]	[0.222, 0.714]	[0.2,0.5]
$\{b\}$	[0, 0.4]	[0, 0.5]	[0.0.4]
$\{a,b\}$	[0.2, 0.7]	[0.222,0.778]	[0.3,0.8]

Table 7: Uncertainty measure intervals and corresponding minimally and maximally uncertain BSs for 4 uncertainty criteria. The belief masses for $m_{\min}$ and $m_{\max}$ are given in the following order: $(m(\{a\}), m(\{a,b\}), m(\{b,c\}), m(\{a,b,c\}))$ (Example 9).

U	$U(\mathbf{m})$	$m_{\min}$	$m_{\max}$
N	[0.417,1.193]	(0.7, *, * , 0.2)	(0.1, *, * , 0.5)
S	[0.191,0.739]	(0.1, 0.4, 0 , 0.5)	(0.425, 0, 0.375 , 0.2)
NS	[0.679,1.509]	(0.7, 0.1, 0 , 0.2)	(0.271, 0, 0.530 , 0.2)
I	[1.492,4.529]	(0.1, 0.4, 0 , 0.5)	(0.7, 0, 0.1 , 0.2)

Table 8: Data set

name	x_1^i	x_2^i	A^i
z^1	[1.9,2.1]	[1.0,1.4]	$\{\omega_1\}$
z^2	[1.4,1.7]	[1.8,2.0]	$\{\omega_1,\omega_2\}$
z^3	[0.8,1.0]	[2.7,3.0]	$\{\omega_2\}$
z^4	[1.9,2.0]	[2.9,3.4]	$\{\omega_2,\omega_3\}$
z^5	[2.8,3.2]	[2.9,3.0]	$\{\omega_3\}$
z^6	[2.4,2.6]	[2.0,2.2]	$\{\omega_1,\omega_3\}$
$\mathbf{x}$	[2.0,2.3]	[1.8 1.9]	?

Table 9: Results of the pattern classification experiment (see notations in text)

focals	$\widehat{\mathbf{m}}$	$\mathbf{m}$	$\mathbf{bel_m}$	$\mathbf{pl_m}$	m_0
$\emptyset$	[0.32,0.88]	[0.41,0.83]	0	0	0.65
$\{\omega_1\}$	[0.09,0.58]	[0.11,0.48]	[0.11,0.48]	[0.11,0.59]	0.27
$\{\omega_2\}$	[0.00,0.08]	[0.00,0.07]	[0.00,0.07]	[0.01,0.18]	0.02
$\{\omega_3\}$	[0.00,0.16]	[0.01,0.12]	[0.01,0.12]	[0.01,0.25]	0.03
$\{\omega_1, \omega_2\}$	[0.00 0.06]	[0.00,0.06]	[0.11,0.58]	[0.12,0.59]	0.01
$\{\omega_1, \omega_3\}$	[0.00,0.08]	[0.00,0.08]	[0.12,0.59]	[0.12 0.59]	0.02
$\{\omega_2, \omega_3\}$	[0.00,0.02]	[0.00,0.02]	[0.01,0.21]	[0.01,0.38]	0.00
$\{\omega_1, \omega_2, \omega_3\}$	[0.00,0.03]	[0.00,0.03]	[0.17,0.59]	[0.17,0.59]	0.00

Table 10: Results of the pattern classification experiment (continued)

focals	$\mathbf{m}_d^*$	$\mathbf{BetP}_{\mathbf{m}_d^*}$	$\widetilde{m}$	$\mathrm{BetP}_{\widetilde{m}}$	m_{0d}^*	$\mathrm{BetP}_{m_{0d}^*}$
$\{\omega_1\}$	[0.22,0.97]	[0.23,0.98]	0.22	0.52	0.76	0.80
$\{\omega_2\}$	[0.00,0.37]	[0.01,0.57]	0.00	0.24	0.04	0.07
$\{\omega_3\}$	[0.01,0.50]	[0.01,0.64]	0.01	0.24	0.09	0.13
$\{\omega_1, \omega_2\}$	[0.00,0.32]	[0.36,0.99]	0.24	0.76	0.03	0.87
$\{\omega_1, \omega_3\}$	[0.01,0.41]	[0.43,0.99]	0.24	0.76	0.06	0.93
$\{\omega_2, \omega_3\}$	[0.00,0.12]	[0.02,0.77]	0.12	0.48	0.01	0.20
$\{\omega_1, \omega_2, \omega_3\}$	[0.00,0.16]	1	0.16	1	0.01	1

List of Figures

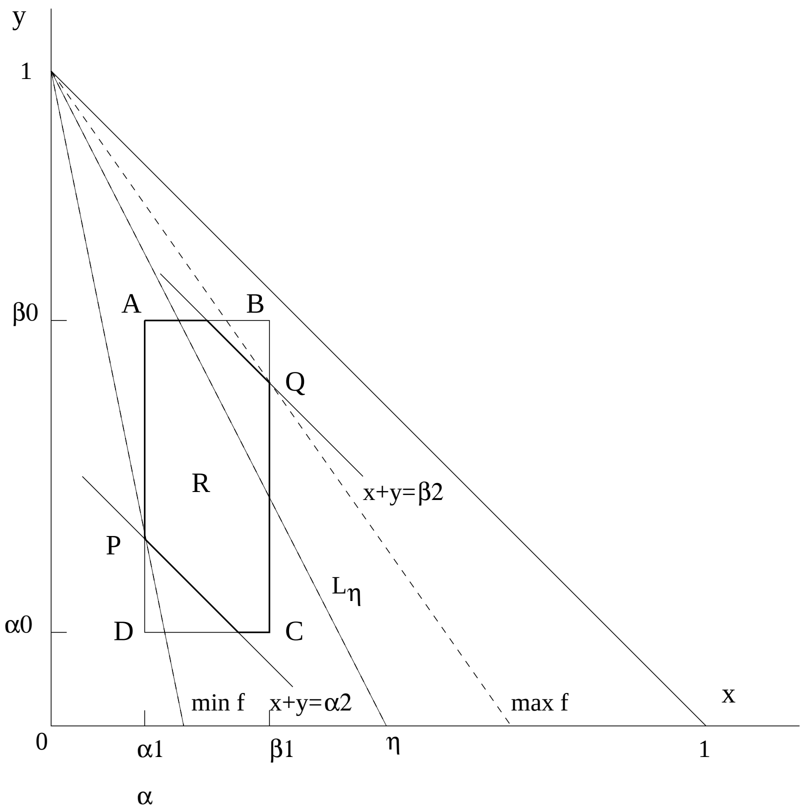

Figure 1: Normalization of an IBS.

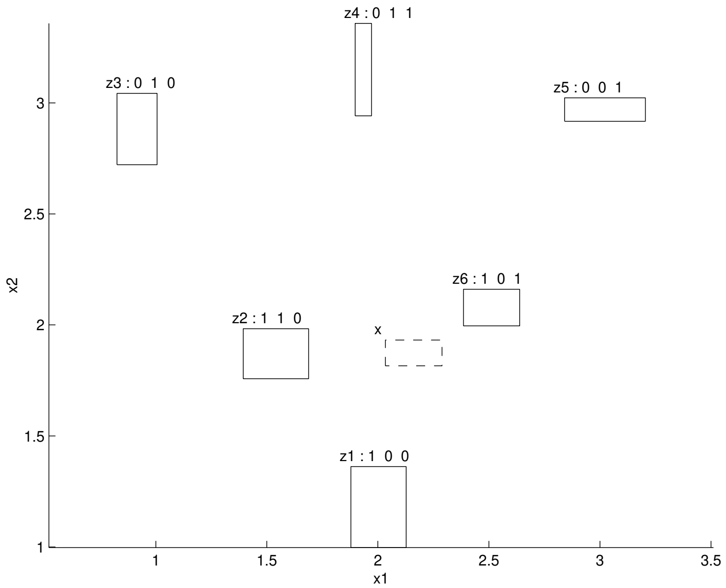

Figure 2: Data set.