Conjunctive and Disjunctive Combination of Belief Functions Induced by Non Distinct Bodies of Evidence [1]

Thierry Denœux
UMR CNRS 6599 Heudiasyc
Université de Technologie de Compiègne
BP 20529 - F-60205 Compiègne cedex - France
Thierry.Denoeux@hds.utc.fr

October 23, 2007

[1]This paper is an extended version of [6].

Abstract

Dempster's rule plays a central role in the theory of belief functions. However, it assumes the combined bodies of evidence to be distinct, an assumption which is not always verified in practice. In this paper, a new operator, the cautious rule of combination, is introduced. This operator is commutative, associative and idempotent. This latter property makes it suitable to combine belief functions induced by reliable, but possibly overlapping bodies of evidence. A dual operator, the bold disjunctive rule, is also introduced. This operator is also commutative, associative and idempotent, and can be used to combine belief functions issues from possibly overlapping *and* unreliable sources. Finally, the cautious and bold rules are shown to be particular members of infinite families of conjunctive and disjunctive combination rules based on triangular norms and conorms.

Keywords: Evidence theory, Dempster-Shafer theory, Transferable Belief Model, Distinct Evidence, Idempotence, Information fusion.

1 Introduction

Dempster's rule of combination [3, 29] is known to play a pivotal role in the theory of belief functions, together with its unnormalized version introduced by Smets in the Transferable Belief Model (TBM) [31], hereafter referred to as the TBM conjunctive rule. Justifications for the origins and uniqueness of these rules have been provided by several authors [9, 31, 23, 22]. However, although they appear well founded theoretically, the need for greater flexibility through a larger choice of combination rules has been recognized by many researchers involved in real-world applications. Two limitations of Dempster's rule and its unnormalized version seem to be their lack of robustness with respect to conflicting evidence (a criticism which mainly applies to Dempster's rule), and the requirement that the items of evidence combined be distinct.

The issue of conflict management has been addressed by several authors, who proposed alternative rules which, unfortunately, are generally not associative (see, e.g., [41, 12, 26], and reviews in [28] and [38]). The disjunctive rule of combination [10, 32] (hereafter referred to as the TBM disjunctive rule) is both associative and more robust than Dempster's rule in the presence of conflicting evidence, and its use is appropriate when the conflict is due to poor reliability of some of the sources. It may also be argued that problems with Dempster's rule (and, to a lesser extent, with the TBM conjunctive rule) are often due to incorrect or incomplete modelisation of the problem at hand, and that these rules often yield reasonable results when they are properly applied [18]. In [38], an expert system approach is advocated in case of large conflict, to determine its origin and revise the underlying hypotheses accordingly.

The other, and perhaps more fundamental, limitation of Dempster's rule lies in the assumption that the items of evidence combined be distinct or, in other words, that the information sources be independent. As remarked by Dempster [3], the real-world meaning of this notion is difficult to describe. The general idea is that, in the combination process, no elementary item of evidence should be counted twice. Thus, non overlapping random samples from a population are clearly distinct items of evidence, whereas "opinions of different people based on overlapping experiences could not be regarded as independent sources" [3]. When the nature of the interaction between items of evidence can be described mathematically, then it is possible to extend Dempster's rule or the TBM conjunctive rule so as to include this knowledge (see, e.g., [9, 30]). However, it is often the case that, although two items of evidence (such as, e.g., opinions expressed by two experts sharing some experiences, or observations of correlated random quantities) can clearly not be regarded as distinct, the interaction between them is ill known and, in many cases, almost impossible to describe.

In such a common situation, it would be very helpful to have a combination rule that would not rely on the distinctness assumption. An early attempt to provide such a rule is reported in [27], but it was limited to the combination of simple belief functions (i.e., belief functions having at most two focal sets, including the frame of discernment). This method was extended to separable belief functions (i.e., belief functions that can be decomposed as the conjunctive sum of simple belief functions) in [16]. However, not all belief functions are separable, and the justification for this approach was unclear.

A natural requirement for a rule allowing the combination of overlapping bodies of evidence is idempotence. The arithmetic mean does possess this property, but it is not associative, another requirement often regarded as essential. Following an approach initiated by Dubois and Prade in [9], Cattaneo [1] studied a family of rules generalizing the TBM conjunctive rule, based on the definition of a joint belief function on a product space, whose marginals are the belief functions to be combined. Inside this family, he proposed a rule minimizing the conflict, which happens to be idempotent. However, he showed that, within this particular family of rules, associativity is incompatible both with idempotency, and with conflict minimization.

In contrast, associative and idempotent operators exist in possibility theory, based on the minimum triangular norm and its dual, the maximum triangular conorm. Dubois and Yager [15] showed that agregation operators for possibility distributions (or, equivalently, fuzzy set connectives) can be deduced from assumptions on multi-valued mappings underlying the possibility distributions viewed as consonant belief functions. This approach, however, has not made it possible to extend possibilistic agregation operators to arbitrary belief functions while maintaining such properties as associativity and idempotency. New operators satisfying these properties are proposed in this paper, following a completely different approach based on some ideas suggested to the author by the late Philippe Smets [36].

The rest of this paper is organized as follows. The underlying fundamental concepts, including the canonical decomposition and the relative information content of belief functions, are first recalled in Section 2. The cautious conjunctive rule and its dual, the bold disjunctive rule are then introduced in Sections 3 and 4, respectively. The cautious and bold rules are shown in Section 5 to be particular members of infinite families of conjunctive and disjunctive combination rules based on triangular norms and conorms. Finally, the efficiency of the cautious rule to combine information from dependent features in a classifier fusion problem is demonstrated experimentally in Section 6, and Section 7 concludes the paper.

2 Fundamental Concepts

In this section, the main building blocks of new combination rules defined later are introduced. The basic concepts and terminology related to belief functions are first summarized in Section 2.1. Section 2.2 then focuses on the canonical conjunctive decomposition of non dogmatic belief functions, which allows their representation in the form of conjunctive weight functions taking values in $(0, +\infty)$. This section is essential, as the cautious conjunctive rule introduced in this paper will be expressed as a function of conjunctive weights. Finally, Section 2.3 recalls known definitions and results related to the ordering of belief functions according to their information content; a new partial ordering relation based on conjunctive weights is also introduced. This ordering relation will play an important role in the derivation of the new combination rules.

2.1 Basic Definitions and Notations

In this paper, the TBM [39, 34] is accepted as a model of uncertainty. An agent's state of belief expressed on a finite frame of discernment $\Omega = \{\omega_1, \ldots, \omega_K\}$ is represented by a basic belief assignment (BBA) m, defined as a mapping from 2^Ω to $[0, 1]$ verifying

$\sum_{A \subseteq \Omega} m(A) = 1$. Subsets A of Ω such that $m(A) > 0$ are called focal sets of m. A BBA m is said to be

- normal if $\emptyset$ is not a focal set (this condition is not imposed in the TBM);
- subnormal is $\emptyset$ is a focal set;
- dogmatic if Ω is not a focal set;
- vacuous if Ω is the only focal set;
- simple if it has at most two focal sets and, if it has two, Ω is one of those;
- categorical if it has only one focal set;
- Bayesian if its focal sets are singletons.

A subnormal BBA m can be transformed into a normal BBA m^* by the normalization operation defined as follows:

$$m^*(A) = \begin{cases} k \cdot m(A) & \text{if } A \neq \emptyset, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

for all $A \subseteq \Omega$, with $k = (1 - m(\emptyset))^{-1}$.

A simple BBA (SBBA) m such that $m(A) = 1 - w$ for some $A \neq \Omega$ and $m(\Omega) = w$ can be noted A^w (the advantage of this notation will become apparent later). The vacuous BBA can thus be noted A^1 for any $A \subset \Omega$, and a categorical BBA can be noted A^0 for some $A \neq \Omega$. A BBA m can equivalently be represented by its associated belief, implicability, plausibility and commonality functions defined, respectively, as:

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B), \tag{2}$$

$$b(A) = \sum_{B \subseteq A} m(B) = bel(A) + m(\emptyset), \tag{3}$$

$$pl(A) = \sum_{B \cap A \neq \emptyset} m(B), \tag{4}$$

and

$$q(A) = \sum_{B \supseteq A} m(B), \tag{5}$$

for all $A \subseteq \Omega$. BBA m can be recovered from any of these functions. For instance:

$$m(A) = \sum_{B \supseteq A} (-1)^{|B|-|A|} q(B), \quad \forall A \subseteq \Omega, \tag{6}$$

and

$$m(A) = \sum_{B \subseteq A} (-1)^{|A|-|B|} b(B), \quad \forall A \subseteq \Omega, \tag{7}$$

where $|A|$ denotes the cardinality of A.

The negation (or complement) $\overline{m}$ of a BBA m is defined as the BBA verifying $\overline{m}(A) = m(\overline{A})$ for all $A \subseteq \Omega$, where $\overline{A}$ denotes the complement of A [10]. It may easily be shown that the implicability function $\overline{b}$ associated to $\overline{m}$ and the commonality function q associated to m are linked by the following relation:

$$\overline{b}(A) = q(\overline{A}), \quad \forall A \subseteq \Omega. \tag{8}$$

A BBA m is said to be consonant if its focal sets are nested. This is known to be equivalent to the following condition [29]:

$$pl(A \cup B) = pl(A) \vee pl(B), \quad \forall A, B \subseteq \Omega,$$

where $\vee$ denote the maximum operator. The above equation defines a possibility measure [42]. Consequently, a consonant BBA uniquely defines a possibility measure. The corresponding possibility distribution is then given by

$$\pi(\omega) = pl(\{\omega\}) = q(\{\omega\}), \quad \forall \omega \in \Omega.$$

Given a BBA m and a coefficient $\alpha \in [0, 1]$, the discounting of m with discount rate α yields the new BBA ${}^{\alpha}m$ defined by:

$${}^{\alpha}m = (1 - \alpha)m + \alpha\, m_{\Omega},$$

where m_{Ω} denotes the vacuous BBA [29, page 252]. The discounting operation is used to model a situation where a source S provides a BBA m, and the reliability of S is measured by $1 - \alpha$. If S is fully reliable ($1 - \alpha = 1$), then m is left unchanged. If S is not reliable at all, m is transformed into the vacuous BBA. In intermediate situations, m is replaced by a convex combination of m and the vacuous BBA.

The TBM conjunctive rule and Dempster's rule are noted $\bigcirc\!\!\!\!\cap$ and $\oplus$, respectively. They are defined as follows. Let m_1 and m_2 be two BBAs, and let $m_{1\bigcirc\!\!\!\!\cap 2}$ and $m_{1\oplus 2}$ be the result of their combination by $\bigcirc\!\!\!\!\cap$ and $\oplus$. We have:

$$m_{1\bigcirc\!\!\!\!\cap 2}(A) = \sum_{B \cap C = A} m_1(B) m_2(C), \quad \forall A \subseteq \Omega, \tag{9}$$

and, assuming that $m_{1\bigcirc\!\!\!\!\cap 2}(\emptyset) \neq 1$:

$$m_{1\oplus 2}(A) = \begin{cases} 0 & \text{if } A = \emptyset, \\ \frac{m_{1\bigcirc\!\!\!\!\cap 2}(A)}{1 - m_{1\bigcirc\!\!\!\!\cap 2}(\emptyset)} & \text{otherwise.} \end{cases} \tag{10}$$

Dempster's rule is just equivalent to the TBM conjunctive rule followed by normalization using (1). Both rules are commutative, associative, and admit a unique neutral element: the vacuous BBA. They both assume the combined items of evidence to be distinct. Let A^{w_1} and A^{w_2} be two SBBAs with the same focal element $A \neq \Omega$. The result of their $\bigcirc\!\!\!\!\cap$-combination is the SBBA $A^{w_1 w_2}$. The $\oplus$ operator yields the same result as long as $A \neq \emptyset$. The TBM conjunctive rule has a simple expression in terms of commonality functions: with obvious notations, we have:

$$q_{1\bigcirc\!\!\!\!\cap 2} = q_1 \cdot q_2. \tag{11}$$

In the TBM, conditioning by $B \subseteq \Omega$ is equivalent to conjunctive combination with a categorical BBA m_B focused on B. The result is noted $m[B]$, with $m[B] = m \cap\!\!\!\!\bigcirc m_B$. This conditional BBA quantifies our belief on Ω, in a context where B holds.

Let us now assume that $m_{1 \cap\!\!\!\!\bigcirc 2}$ has been obtained by combining two BBAs m_1 and m_2, and then we learn that m_2 is in fact not supported by evidence and should be "removed" from $m_{1 \cap\!\!\!\!\bigcirc 2}$. This "decombination" operation was introduced in [33]. It is well defined if m_2 is non dogmatic. Denoting $\oslash$ this operator, we can write:

$$m_{1 \cap\!\!\!\!\bigcirc 2} \oslash m_2 = m_1.$$

Decombination can easily be computed for any two BBAs m_1 and m_2 using the corresponding commonality functions as:

$$q_{1 \oslash 2}(A) = \frac{q_1(A)}{q_2(A)}, \quad \forall A \subseteq \Omega. \tag{12}$$

Note that $q_2(A) > 0$ for all A as long as m_2 is non dogmatic. One should also be aware that the quotient of two commonality functions is not always a commonality function. Consequently, $m_1 \oslash m_2$ is not necessarily a BBA.

A disjunctive rule of combination $\cup\!\!\!\!\bigcirc$ also exists [10, 32]: it is defined as

$$m_{1 \cup\!\!\!\!\bigcirc 2}(A) = \sum_{B \cup C = A} m_1(B) m_2(C), \quad \forall A \subseteq \Omega. \tag{13}$$

This rule, called the TBM disjunctive rule, is also commutative and associative. It has a simple expression in terms of implicability functions, which is the counterpart of (11):

$$b_{1 \cup\!\!\!\!\bigcirc 2} = b_1 \cdot b_2. \tag{14}$$

As for the TBM conjunctive rule, an inverse operation may also be defined for the TBM disjunctive rule:

$$b_{1 \obslash 2}(A) = \frac{b_1(A)}{b_2(A)}, \quad \forall A \subseteq \Omega. \tag{15}$$

This operation is well-defined as long as m_2 is subnormal (in which case we have $b_2(A) > 0$ for all A). However, it does not necessarily produce a belief function. Its interpretation is similar to that of $\oslash$: it removes, or "decombines"', evidence which has been combined disjunctively with prior knowledge.

The dual nature of $\cap\!\!\!\!\bigcirc$ and $\cup\!\!\!\!\bigcirc$ becomes apparent when one notices that these two operators are linked by De Morgan's laws [10]:

$$\begin{aligned} \overline{m_1 \cup\!\!\!\!\bigcirc m_2} &= \overline{m_1} \cap\!\!\!\!\bigcirc \overline{m_2}, \quad (16) \\ \overline{m_1 \cap\!\!\!\!\bigcirc m_2} &= \overline{m_1} \cup\!\!\!\!\bigcirc \overline{m_2}, \quad (17) \end{aligned}$$

for all m_1 and m_2.

As remarked by Smets [32], the TBM conjunctive rule is based on the assumption that the belief functions to be combined are induced by reliable sources of information, whereas the TBM disjunctive rule only assume that at least one source of information is reliable, but we do not know which one. Both rules assume the sources of information to be independent (i.e., they are assumed to provide distinct, non overlapping pieces of evidence).

In the TBM, combination rules belong to the *credal level* where evidence aggregation takes place, whereas decisions are made at the *pignistic level* [39], where each BBA m is mapped to a pignistic probability function $Betp_m$ defined by

$$Betp_m(\omega) = \sum_{\{A:\omega\in A\}} \frac{m^*(A)}{|A|}, \quad \forall \omega \in \Omega, \tag{18}$$

where m^* denotes the normalized version of m.

2.2 Canonical Conjunctive Decomposition of a Belief Function

Shafer [29, Chapter 4] defined a separable BBA as the result of the $\oplus$ combination of SBBAs. For every separable BBA in the sense of Shafer, one has:

$$m = \bigoplus_{\emptyset \neq A \subset \Omega} A^{w(A)}, \tag{19}$$

with $w(A) \in [0,1]$ for all $A \subset \Omega$, $A \neq \emptyset$. This representation is unique if m is non dogmatic. Shafer named this representation the *canonical decomposition* of m.

The concept of separability can be extended to subnormal BBAs in two ways:

- We will say that a BBA m is u-separable (where "u" stands for "unnormalized") if we have
$$m = \bigcirc\!\!\!\!\cap_{A\subset\Omega} A^{w(A)}, \tag{20}$$
with $w(A) \in [0,1]$ for all $A \subset \Omega$;
- We will say that a BBA m is n-separable (where "n" stands for "normalized") if we have
$$m^* = \bigoplus_{\emptyset \neq A \subset \Omega} A^{w(A)}, \tag{21}$$
where $w(A) \in [0,1]$ for all $A \subset \Omega$, $A \neq \emptyset$, and m^* is the normalized form of m.

Again, the decompositions (20) and (21) are unique as long as m is non dogmatic. Clearly, (20) implies (21), but the converse is not true, as will be shown below. Consequently, u-separability is a stronger notion than n-separability.

2.2.1 Extension to non dogmatic BBAs

The canonical decomposition of a separable BBA was extended to any non dogmatic BBA by Smets [33]. The key to such a generalization is the notion of generalized simple BBA (GSBBA), defined as a function μ from 2^Ω to $\mathbb{R}$ verifying

$$\begin{aligned} \mu(A) &= 1 - w, && (22)\\ \mu(\Omega) &= w, && (23)\\ \mu(B) &= 0 \quad \forall B \in 2^\Omega \setminus \{A, \Omega\}, && (24) \end{aligned}$$

for some $A \neq \Omega$ and some $w \in [0, +\infty)$. Any GSBBA μ can thus be noted A^w for some $A \neq \Omega$ and $w \in [0, +\infty)$. When $w \leq 1$, μ is a SBBA. When $w > 1$, μ is not a BBA, since it is no longer a mapping from 2^Ω to $[0,1]$. Such a function can be referred to as an inverse simple BBA (ISBBA), using a terminology similar to that used in [33]. The TBM conjunctive rule can be trivially extended to combine SBBAs and ISBBAs alike. In particular, the relationship $A^{w_1} \textcircled{\cap} A^{w_2} = A^{w_1 w_2}$ still holds for $w_1, w_2 \in [0, +\infty)$.

In [33], Smets proposed an interpretation of an ISBBA as representing a state of belief in which one has some reasons *not* to believe in A. By combining A^w for some $w > 1$ with the SBBA $A^{1/w}$ using the TBM conjunctive rule, one obtains the vacuous bba A^1. Hence, the ISBBA A^w corresponds to a situation where the agent has a "debt of belief" in A, and some evidence has to be accumulated before it starts to believe in A.

Using the concept of GSBBA, and extending Shafer's approach, Smets showed that any non dogmatic BBA can be uniquely represented as the conjunctive combination of GSBBAs:

$$m = \textcircled{\cap}_{A \subset \Omega} A^{w(A)}, \tag{25}$$

with $w(A) \in (0, +\infty)$ for all $A \subset \Omega$. Equation (25) is clearly an extension of (19). It defines the canonical conjunctive decomposition of m (we will see in Section 4.1 that a canonical *disjunctive* decomposition also exists). The weights $w(A)$ for every $A \subset \Omega$ can be obtained from the commonalities using the following formula:

$$w(A) = \prod_{B \supseteq A} q(B)^{(-1)^{|B|-|A|+1}}, \tag{26}$$

$$= \begin{cases} \dfrac{\prod_{B \supseteq A, |B| \notin 2\mathbb{N}} q(B)}{\prod_{B \supseteq A, |B| \in 2\mathbb{N}} q(B)} & \text{if } |A| \in 2\mathbb{N} \\[2ex] \dfrac{\prod_{B \supseteq A, |B| \in 2\mathbb{N}} q(B)}{\prod_{B \supseteq A, |B| \notin 2\mathbb{N}} q(B)} & \text{otherwise,} \end{cases} \tag{27}$$

where $2\mathbb{N}$ denotes the set of even natural numbers. Eq. (26) can be equivalently written

$$\ln w(A) = -\sum_{B \supseteq A} (-1)^{|B|-|A|} \ln q(B), \quad \forall A \subset \Omega. \tag{28}$$

One notices the similarity with (6). Hence, any procedure for transforming q to m (such as the Fast Möbius Transform [21] or matrix multiplication [35]) can be used to compute $\ln w$ from $-\ln q$.

The function $w : 2^\Omega \setminus \{\Omega\} \to (0, +\infty)$ (hereafter referred to as the conjunctive weight function) is thus yet another equivalent representation of any non dogmatic BBA (together with *bel*, *pl*, *q*, etc.). This concept of conjunctive weight function can be extended to a dogmatic BBA m by discounting it with some discount rate ϵ and letting ϵ tend towards 0 [33]. However, this extension requires some mathematical subtleties. Furthermore, it may be argued that most (if not all) states of belief, being based on imperfect and not entirely conclusive evidence, should be represented by non dogmatic belief functions, even if the mass $m(\Omega)$ is very small. For instance, consider a coin tossing experiment. It is natural to define a BBA on $\Omega = \{Heads, Tails\}$ as $m(\{Heads\}) = 0.5$ and $m(\{Tails\}) = 0.5$. However, this assumes the coin to be perfectly balanced, a condition never exactly verified in practice. So, a more appropriate BBA may be $m(\{Heads\}) = 0.5(1-\epsilon)$, $m(\{Tails\}) = 0.5(1-\epsilon)$ and $m(\Omega) = \epsilon$ for some small $\epsilon > 0$.

EXAMPLE 1 Let $\Omega = \{a, b, c\}$ be a frame of discernment, and m the BBA shown in Table 1. The weights can be computed from the commonalities using (26) as follows:

$$
\begin{aligned}
w(\emptyset) &= \frac{q(\{a\})q(\{b\})q(\{c\})q(\{a,b,c\})}{q(\emptyset)q(\{a,b\})q(\{a,c\})q(\{b,c\})} = \frac{0.5 \times 1 \times 0.7 \times 0.2}{1 \times 0.5 \times 0.2 \times 0.7} = 1 \\
w(\{a\}) &= \frac{q(\{a,b\})q(\{a,c\})}{q(\{a\})q(\{a,b,c\})} = \frac{0.5 \times 0.2}{0.5 \times 0.2} = 1 \\
w(\{b\}) &= \frac{q(\{a,b\})q(\{b,c\})}{q(\{b\})q(\{a,b,c\})} = \frac{0.5 \times 0.7}{1 \times 0.2} = 7/4 \\
w(\{a,b\}) &= \frac{q(\{a,b,c\})}{q(\{a,b\})} = \frac{0.2}{0.5} = 2/5 \\
w(\{c\}) &= \frac{q(\{a,c\})q(\{b,c\})}{q(\{c\})q(\{a,b,c\})} = \frac{0.2 \times 0.2}{0.7 \times 0.2} = 1 \\
w(\{a,c\}) &= \frac{q(\{a,b,c\})}{q(\{a,c\})} = \frac{0.2}{0.1} = 1 \\
w(\{b,c\}) &= \frac{q(\{a,b,c\})}{q(\{b,c\})} = \frac{0.2}{0.7} = 2/7.
\end{aligned}
$$

We can see that m can be represented as the conjunctive combination of two SBBAs $\{a, b\}^{2/5}$ and $\{b, c\}^{2/7}$, and an ISBBA $\{b\}^{7/4}$.

Table 1: A BBA with its commonality and weight functions.

A	$m(A)$	$q(A)$	$w(A)$
$\emptyset$	0	1	1
$\{a\}$	0	0.5	1
$\{b\}$	0	1	7/4
$\{a,b\}$	0.3	0.5	2/5
$\{c\}$	0	0.7	1
$\{a,c\}$	0	0.2	1
$\{b,c\}$	0.5	0.7	2/7
Ω	0.2	0.2	

2.2.2 Special cases

In the following two propositions, we provide analytical formulas for the conjunctive weight functions associated to two important classes of BBAs.

PROPOSITION 1 *Let $A_1, \ldots, A_n$ be n subsets of Ω such that $A_i \cap A_j = \emptyset$ for all $i, j \in \{1, \ldots, n\}$, and let m be a BBA on Ω with focal sets $A_1, \ldots, A_n$, and Ω. We assume that $m(\Omega) + \sum_{k=1}^n m(A_k) \leq 1$, so that $\emptyset$ may also be a focal set. The conjunctive weight function associated to m is:*

$$
w(A) = \begin{cases} \frac{m(\Omega)}{m(A_k)+m(\Omega)}, & A = A_k, \\ m(\Omega)\prod_{k=1}^n \left(1 + \frac{m(A_k)}{m(\Omega)}\right), & A = \emptyset, \\ 1, & \textit{otherwise}. \end{cases}
$$

Proof: We have:

$$q(A) = \begin{cases} m(A_k) + m(\Omega), & A \subseteq A_k, \\ 1, & A = \emptyset, \\ m(\Omega), & \text{otherwise.} \end{cases}$$

Consequently, m may be expressed as a function of q as follows:

$$\begin{aligned} m(A_k) &= q(A_k) - q(\Omega), \quad k = 1, \ldots, n, && (29)\\ m(\Omega) &= q(\Omega) && (30)\\ m(\emptyset) &= q(\emptyset) - q(\Omega) - \sum_{k=1}^{n} (q(A_k) - q(\Omega)) && (31)\\ m(A) &= 0, \quad \forall A \notin \{A_1, \ldots, A_n, \Omega, \emptyset\}. && (32) \end{aligned}$$

As explained above, $\ln w$ may be obtained from $-\ln q$ using any procedure that transforms q to m. Consequently, we may, in the above equations, replace m by $\ln w$ and q by $-\ln q$ (except in (30), because $w(\Omega)$ is not defined). We obtain from (29):

$$\ln w(A_k) = -\ln q(A_k) + \ln q(\Omega) = \ln \frac{q(\Omega)}{q(A_k)},$$

which implies

$$w(A_k) = \frac{m(\Omega)}{m(A_k) + m(\Omega)}, \quad k = 1, \ldots, n.$$

Now, from (31) we get

$$\begin{aligned} \ln w(\emptyset) &= -\ln q(\emptyset) + \ln q(\Omega) + \sum_{k=1}^{n} (\ln q(A_k) - \ln q(\Omega)), \\ &= \ln \left(q(\Omega)^{1-n} \prod_{k=1}^{n} q(A_k) \right), \\ &= \ln \left(m(\Omega)^{1-n} \prod_{k=1}^{n} (m(\Omega) + m(A_k)) \right), \end{aligned}$$

which implies

$$w(\emptyset) = m(\Omega) \prod_{k=1}^{n} \left(1 + \frac{m(A_k)}{m(\Omega)} \right).$$

Finally, (32) implies that $w(A) = 1$, for all $A \notin \{A_1, \ldots, A_n, \Omega, \emptyset\}$. □

The BBAs studied in Proposition 1 may be termed "quasi-Bayesian", as they can be obtained by discounting Bayesian BBAs defined on a coarsening of Ω. This class of BBAs is closed under the TBM conjunctive rule. Quasi-Bayesian BBAs are defined by a small number of masses, and are frequently encountered in applications.

Another important case concerns consonant BBAs, whose focal sets are nested. The following proposition provides formulas to compute the weight function associated to a consonant BBA.

PROPOSITION 2 *Let m be a consonant BBA, with associated possibility distribution $\pi(\omega_k) = q(\{\omega_k\})$, $k = 1, \ldots, K$. We note $\pi_k = \pi(\omega_k)$ and we assume that the elements of Ω have been arranged in decreasing order of plausibility, i.e., we have*

$$1 \geq \pi_1 \geq \pi_2 \geq \ldots \geq \pi_K > 0.$$

Let $A_k = \{\omega_1, \ldots, \omega_k\}$, $k = 1, \ldots, K$. The focal sets of m are in $\{A_1, \ldots, A_K, \emptyset\}$ (m is subnormal if $\pi_1 < 1$, and it is non dogmatic since we have assumed $\pi_K > 0$). The conjunctive weight function associated to m is:

$$w(A) = \begin{cases} \pi_1 & A = \emptyset, \\ \frac{\pi_{k+1}}{\pi_k}, & A = A_k,\ 1 \leq k < K, \\ 1, & \textit{otherwise.} \end{cases}$$

Proof. As shown in [8], m can be computed from $\pi_1, \ldots, \pi_K$ as:

$$m(A) = \begin{cases} 1 - \pi_1, & A = \emptyset, \\ \pi_k - \pi_{k+1}, & A = A_k,\ 1 \leq k < K, \\ \pi_K, & A = \Omega, \\ 0 & \text{otherwise.} \end{cases}$$

Since $\pi_k = q(\{\omega_k\})$, we may deduce that

$$\ln w(A) = \begin{cases} \ln \pi_1, & A = \emptyset, \\ -\ln \pi_k + \ln \pi_{k+1}, & A = A_k,\ 1 \leq k < K, \\ 0 & \text{otherwise,} \end{cases}$$

from which the desired expression of w can be easily derived. □

2.2.3 Normalization and combination

It may be remarked that normalizing a subnormal BBA m using (1) amounts to combining it with the ISBBA $\emptyset^k$:

$$m^* = m \circledcap \emptyset^k.$$

Consequently, the weight function w^* associated to m^* is identical to w, except for the weight assigned to $\emptyset$. If $m = \circledcap_{A \subset \Omega} A^{w(A)}$, we have

$$\begin{aligned} m^* &= \emptyset^k \circledcap \emptyset^{w(\emptyset)} \circledcap \left(\circledcap_{\emptyset \neq A \subset \Omega} A^{w(A)} \right) \\ &= \emptyset^{k \cdot w(\emptyset)} \circledcap \left(\circledcap_{\emptyset \neq A \subset \Omega} A^{w(A)} \right), \\ &= \circledcap_{A \subset \Omega} A^{w^*(A)} \end{aligned}$$

with $w^*(\emptyset) = k \cdot w(\emptyset)$ and $w^*(A) = w(A)$ for all $A \in 2^\Omega \setminus \{\emptyset, \Omega\}$. We can write, equivalently:

$$m^* = \bigoplus_{\emptyset \neq A \subset \Omega} A^{w(A)}. \tag{33}$$

As a direct consequence of the above remark, it is easy to formulate criteria for u-separability and n-separability:

- A BBA m is u-separable iff $w(A) \leq 1$, for all $A \subset \Omega$;
- A BBA m is n-separable iff $w(A) \leq 1$, for all $A \subset \Omega$, $A \neq \emptyset$.

For instance, quasi-Bayesian BBAs studied in Proposition 1 are n-separable, but they are not u-separable in general (we may have $w(\emptyset) > 1$). In contrast, consonant BBAs are u-separable, since they satisfy the condition $w(A) \leq 1$ for all $A \subset \Omega$.

The w representation appears particularly interesting when it comes to combining BBAs using the TBM conjunctive rule or Dempster's rule. Indeed, let m_1 and m_2 be two BBAs with weight functions w_1 and w_2. We have:

$$\begin{aligned} m_1 \bigcirc\!\!\!\!\cap\, m_2 &= \left(\bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w_1(A)}\right) \bigcirc\!\!\!\!\cap \left(\bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w_2(A)}\right) && (34)\\ &= \bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w_1(A) w_2(A)}. && (35) \end{aligned}$$

We can thus write, with obvious notations:

$$w_{1 \bigcirc\!\!\!\!\cap 2} = w_1 \cdot w_2,$$

which is reminiscent of (11). The inverse TBM conjunctive rule $\oslash$ also has a simple expression in the w-space, similar to (33): we have $w_{1 \oslash 2} = w_1 / w_2$. Hence,

$$\begin{aligned} m_1 \oslash m_2 &= \left(\bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w_1(A)}\right) \oslash \left(\bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w_2(A)}\right) && (36)\\ &= \bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w_1(A) / w_2(A)}. && (37) \end{aligned}$$

Finally, using (33), it is easy to see that

$$m_1 \oplus m_2 = \bigoplus_{\emptyset \neq A \subset \Omega} A^{w_1(A) w_2(A)}. \tag{38}$$

2.2.4 Latent belief structure

Let m be a non dogmatic belief function, and w its associated conjunctive weight function. For each weight $w(A)$ let us define the following two quantities:

$$w^c(A) = 1 \wedge w(A), \tag{39}$$

and

$$w^d(A) = 1 \wedge \frac{1}{w(A)}, \tag{40}$$

where $\wedge$ denotes the minimum operator. It is clear that we have

$$w(A) = \frac{w^c(A)}{w^d(A)}. \tag{41}$$

Consequently, we can write

$$\begin{aligned} m &= \bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w^c(A) / w^d(A)} && (42)\\ &= \left(\bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w^c(A)}\right) \oslash \left(\bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w^d(A)}\right) && (43)\\ &= m^c \oslash m^d. && (44) \end{aligned}$$

Any non dogmatic BBA m can thus be decomposed into two u-separable BBAs m^c and m^d called, respectively, its confidence and diffidence components. The pair (m^c, m^d) forms what Smets called a latent belief structure (LBS) [33]. He proposed to interpret m^c as representing positive evidence, i.e., good reasons to believe in various propositions $A \subseteq \Omega$, and m^d as representing negative evidence, i.e., good reasons *not* to believe in the same propositions. The BBA m is obtained by removing the negative evidence m^d from the positive evidence m^c. Note that we have the following property with respect to the TBM conjunctive rule: if (m_1^c, m_1^d) and (m_2^c, m_2^d) are two LBSs associated to non dogmatic BBAs m_1 and m_2, respectively, then $(m_1^c \textcircled{\scriptsize$\cap$} m_2^c, m_1^d \textcircled{\scriptsize$\cap$} m_2^d)$ is a LBS associated to $m_1 \textcircled{\scriptsize$\cap$} m_2$.

2.3 Informational Comparison of Belief Functions

In the TBM, the *Least commitment Principle* (LCP) plays a role similar to the principle of maximum entropy in Bayesian Probability Theory. As explained in [32], the LCP indicates that, given two belief functions compatible with a set of constraints, the most appropriate is the least informative. To make this principle operational, it is necessary to define ways of comparing belief functions according to their information content. Three such partial orderings, generalizing set inclusion, were proposed in the 1980's by Yager [40] and Dubois and Prade [10]; they are defined as follows:

- *pl*-ordering: $m_1 \sqsubseteq_{pl} m_2$ iff $pl_1(A) \leq pl_2(A)$, for all $A \subseteq \Omega$;
- *q*-ordering: $m_1 \sqsubseteq_q m_2$ iff $q_1(A) \leq q_2(A)$, for all $A \subseteq \Omega$;
- *s*-ordering: $m_1 \sqsubseteq_s m_2$ iff there exists a square matrix S with general term $S(A, B)$, $A, B \in 2^\Omega$ verifying

  $$\sum_{B \subseteq \Omega} S(A, B) = 1, \quad \forall A \subseteq \Omega,$$

  $$S(A, B) > 0 \Rightarrow A \subseteq B, \quad A, B \subseteq \Omega,$$

  such that

  $$m_1(A) = \sum_{B \subseteq \Omega} S(A, B) m_2(B), \quad \forall A \subseteq \Omega. \tag{45}$$

  The term $S(A, B)$ may be seen as the proportion of the mass $m_2(B)$ which is transferred ("flows down") to A. Matrix S is named a specialization matrix [23, 35], and m_1 is said to be a specialization of m_2.

As shown in [10], these three definitions are not equivalent: $m_1 \sqsubseteq_s m_2$ implies $m_1 \sqsubseteq_{pl} m_2$ and $m_1 \sqsubseteq_q m_2$, but the converse is not true. Additionally, *pl*-ordering and *q*-ordering are not comparable. However, in the set of consonant BBAs, these three partial orders are equivalent. The interpretation of these ordering relations is discussed in [10] from a set-theoretical perspective, and in [13] from the point of view of the TBM. Whenever we have $m_1 \sqsubseteq_x m_2$, with $x \in \{pl, q, s\}$, we will say that m_1 is x-more committed than m_2.

Another concept which leads to an alternative definition of informational ordering is that of Dempsterian specialization [23]. m_1 is said to be a Dempsterian specialization of m_2, which we note $m_1 \sqsubseteq_d m_2$, iff there exists a BBA m such that

$m_1 = m \textcircled{\cap} m_2$. As shown in [23], this is a stronger condition than specialization, i.e., we have $m_1 \sqsubseteq_d m_2 \Rightarrow m_1 \sqsubseteq_s m_2$, but the converse is false. If $m_1 = m \textcircled{\cap} m_2$, then there is a specialization matrix S_m defined as a function of m, called a Dempsterian specialization matrix, allowing to compute m_1 from m_2 using relation (45).

Finally, we can think of one more definition of informational ordering based on the weight function recalled in Section 2.2: given two non dogmatic BBAs m_1 and m_2, we can say that m_1 is w-more committed than m_2, which we note $m_1 \sqsubseteq_w m_2$, iff $w_1(A) \leq w_2(A)$, for all $A \subset \Omega$. Because of (35), this is equivalent to the existence of a *u-separable* BBA m, with weight function $w = w_1/w_2$, such that $m_1 = m \textcircled{\cap} m_2$. Consequently, w-ordering is strictly stronger than d-ordering. The meaning of $\sqsubseteq_d$ and $\sqsubseteq_w$ is clear: if $m_1 \sqsubseteq_d m_2$ or $m_1 \sqsubseteq_w m_2$, it means that m_1 is the result of the combination of m_2 with some BBA m; consequently, m_1 has a higher information content than m_2. In the case of $\sqsubseteq_w$, the requirement that m be u-separable may seem artificial. However, it may be argued that most belief functions encountered in practice result from the pooling of simple evidence, and are therefore u-separable. As shown in Section 2.2.2, this is also the case for consonant belief funtions, a class of belief functions often encountered in applications because of its simplicity. Furthermore, we will see that w-ordering happens to be a simpler and more convenient notion, for some purposes, than other orderings. A slightly weaker notion based on n-separability will be defined later in Section 3.3. We defer the introduction of this additional notion for clarity of presentation.

In summary, we thus have, for any two non dogmatic BBAs m_1 and m_2:

$$m_1 \sqsubseteq_w m_2 \Rightarrow m_1 \sqsubseteq_d m_2 \Rightarrow m_1 \sqsubseteq_s m_2 \Rightarrow \begin{cases} m_1 \sqsubseteq_{pl} m_2 \\ m_1 \sqsubseteq_q m_2, \end{cases} \tag{46}$$

where all implications are strict.

The vacuous BBA m_Ω (with weight function $w_\Omega(A) = 1$, for all $A \subset \Omega$) is the unique greatest element for partial orderings $\sqsubseteq_x$ with $x \in \{pl, q, s, d\}$, i.e., we have

$$m \sqsubseteq_x m_\Omega, \quad \forall m, \forall x \in \{pl, q, s, d\}.$$

In contrast, m_Ω is only a maximal element for $\sqsubseteq_w$, i.e., we have the following weaker property

$$m_\Omega \sqsubseteq_w m \Rightarrow m = m_\Omega.$$

The BBAs that are w-less specific than m_Ω are the u-separable ones. Non u-separable BBAs are not comparable with m_Ω according to relation $\sqsubseteq_w$.

As emphasized by Dubois and Prade in [10], relations $\sqsubseteq_x$ with $x \in \{pl, q, s\}$ generalize set inclusion: if m_A and m_B are two categorical BBAs such that $m_A(A) = 1$ and $m_B(B) = 1$, then $m_A \sqsubseteq_x m_B$, with $x \in \{pl, q, s\}$, if and only if $A \subseteq B$. The same is true for relation $\sqsubseteq_d$. For relation $\sqsubseteq_w$, this property does not hold, since categorical BBAs, being dogmatic, cannot be compared according to $\sqsubseteq_w$. However, we can still have a similar property if we consider a categorical BBA as the limit of a sequence of non dogmatic BBAs. More precisely, let (ϵ_n), $n = 1, 2, \ldots, \infty$ be a real sequence such that $\epsilon_n \in [0, 1]$ for all n, and $\lim_{n \to \infty} \epsilon_n = 0$. For any $A \subseteq \Omega$, let m_A^n the BBA with following weight function:

$$w_A^n(C) = \begin{cases} \epsilon_n & \text{if } C \supseteq A \\ 1 & \text{otherwise,} \end{cases}$$

for all $C \subset \Omega$. It is clear that $m_A^n(A) \geq 1-\epsilon_n$. Consequently, we have $\lim_{n\to\infty} m_A^n(A) = 1$ and $\lim_{n\to\infty} m_A^n(C) = 0$, for all $C \neq A$. This means that the categorical BBA m_A may seen as the limit of the sequence (m_A^n), this sequence being uniquely defined, given (ϵ_n). Using this representation, we can state the following proposition.

PROPOSITION 3 *Let A and B be two subsets of Ω, m_A and m_B the categorical BBAs focused on A and B, and (m_A^n) and (m_B^n) their representations as sequences of BBAs as defined above. Then, we have*

$$A \subseteq B \Leftrightarrow m_A^n \sqsubseteq_w m_B^n, \quad \forall n \geq 1.$$

Proof: Assume that $A \subseteq B$. Let C be an arbitrary subset of Ω:

- if $C \supseteq B$, then $C \supseteq A$, and we have $w_A^n(C) = w_B^n(C) = \epsilon_n$;
- if $C \not\supseteq B$, then $w_B^n(C) = 1 \geq w_A^n(C)$.

Conversely, assume that $w_A^n(C) \leq w_B^n(C)$ for all $C \subset \Omega$. Then $w_A^n(B) \leq w_B^n(B) = \epsilon_n$. Consequently, $w_A^n(B) = \epsilon_n$, and $A \subseteq B$. □

Equipped with these definitions of the relative information content of belief functions, it is possible to give an operational meaning of the LCP. Let $\mathcal{M}$ be a set of BBA compatible with a set of constraints. The LCP dictates to choose a greatest element in $\mathcal{M}$, if one such element exists, according to one of the partial ordering $\sqsubseteq_x$, for some $x \in \{pl, q, s, d, w\}$. These partial orderings seem equally well justified and reasonable and, in the absence of any decisive argument to discard any of them, considerations of simplicity, existence of a solution and tractability of calculations can be invoked to choose a particular ordering for a given problem. For instance, q-ordering is adopted in [13] to derive the expression of the q-least committed BBA with given pignistic probability function. In the following section, the same principle is used to derive a rule of combination, using partial ordering $\sqsubseteq_w$.

3 The Cautious Conjunctive Rule

3.1 Derivation from the LCP

Just as relations $\sqsubseteq_x$ may be seen as generalizing set inclusion, it is possible to conceive conjunctive combination rules generalizing set intersection, by reasoning as follows. Assume that we have two sources of information, one of which indicates that the true value of the variable of interest ω lies in $A \subseteq \Omega$, while the other one indicates that it lies in $B \subseteq \Omega$, with $A \neq B$. If we consider both sources as reliable, then we can deduce that ω lies in some set C such that $C \subseteq A$, and $C \subseteq B$. The largest of these subsets is the intersection $A \cap B$ of A and B.

Let us now assume that the two sources provide BBAs m_1 and m_2, and the sources are both considered to be reliable. The agent's state of belief, after receiving these two pieces of information, should then be represented by a BBA m_{12} more informative than m_1, and more informative than m_2. Let us denote by $\mathcal{S}_x(m)$ the set of BBAs m' such that $m' \sqsubseteq_x m$, for some $x \in \{pl, q, s, d, w\}$. We thus have $m_{12} \in \mathcal{S}_x(m_1)$ and $m_{12} \in \mathcal{S}_x(m_2)$ or, equivalently, $m_{12} \in \mathcal{S}_x(m_1) \cap \mathcal{S}_x(m_2)$. According to the LCP, one should select the x-least committed element in $\mathcal{S}_x(m_1) \cap \mathcal{S}_x(m_2)$. This defines a conjunctive combination rule, provided that an x-least committed element (i.e., a greatest element with respect with partial order $\sqsubseteq_x$) exists and is unique.

In [13], this approach was used to justify the minimum rule for combining possibility distributions, from the point of view of the TBM. Let m_1 and m_2 be two consonant BBAs, and let q_1 and q_2 be their respective commonality functions. Then, the consonant BBA m_{12} with commonality function $q_{12}(A) = q_1(A) \wedge q_2(A)$ for all $A \subseteq \Omega$ is claimed in [13] to be the s-least committed element in the set $\mathcal{S}_s(m_1) \cap \mathcal{S}_s(m_2)$. This approach, however, cannot be blindly transposed to non consonant BBAs, since the minimum of two commonality functions is not, in general, a commonality function.

However, applying this approach with the $\sqsubseteq_w$ ordering does yield an interesting solution, as shown by the following lemma and proposition.

LEMMA 1 *Let m by a non dogmatic BBA with conjunctive weight function w, and let w' be a mapping from $2^\Omega \setminus \Omega$ to $(0, +\infty)$ such that $w'(A) \leq w(A)$ for all $A \subset \Omega$. Then w' is the conjunctive weight function of some BBA m'.*

Proof: We have

$$w'(A) = w(A) \cdot \frac{w'(A)}{w(A)}, \quad \forall A \subset \Omega.$$

Since $w'(A)/w(A) \leq 1$ for all $A \subset \Omega$, w'/w is the weigth function of a u-separable BBA. Consequently, w' is the weigth function of a BBA m' obtained by combining m with a u-separable BBA using the TBM conjunctive rule. □

PROPOSITION 4 *Let m_1 and m_2 be two non dogmatic BBAs. The w-least committed element in $\mathcal{S}_w(m_1) \cap \mathcal{S}_w(m_2)$ exists and is unique. It is defined by the following weight function:*

$$w_{1 \owedge 2}(A) = w_1(A) \wedge w_2(A), \quad \forall A \subset \Omega. \tag{47}$$

Proof: For $i = 1$ and $i = 2$, we have $m \in \mathcal{S}_w(m_i)$ iff $w(A) \leq w_i(A)$ for all $A \subset \Omega$. Hence, $m \in \mathcal{S}_w(m_1) \cap \mathcal{S}_w(m_2)$ iff $w(A) \leq w_1(A) \wedge w_2(A)$ for all $A \subset \Omega$. Let us consider function $w_{1 \owedge 2}$ defined by $w_{1 \owedge 2}(A) = w_1(A) \wedge w_2(A)$, for all $A \subset \Omega$. This is the conjunctive weight function of a valid BBA, as a consequence of Lemma 1. Consequently, it corresponds to the unique w-least committed element in $\mathcal{S}_w(m_1) \cap \mathcal{S}_w(m_2)$. □

Equation (47) defines a new rule which can be formally defined as follows.

DEFINITION 1 (CAUTIOUS CONJUNCTIVE RULE) *Let m_1 and m_2 be two non dogmatic BBAs. Their combination using the cautious conjunctive rule (or cautious rule for short) is noted $m_{1 \owedge 2} = m_1 \owedge m_2$. It is defined as the BBA with the following weight function:*

$$w_{1 \owedge 2}(A) = w_1(A) \wedge w_2(A), \quad \forall A \subset \Omega.$$

We thus have

$$m_1 \owedge m_2 = \bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w_1(A) \wedge w_2(A)}. \tag{48}$$

Note that this rule happens to generalize a method proposed by Kennes [20] for combining *u-separable* BBAs induced by non distinct items of evidence, based on an application of category theory to evidential reasoning. Using the canonical decomposition of non dogmatic belief functions and the concept of w-ordering, the new rule described in this paper proves to be well justified for combining the wider class of non dogmatic belief functions.

As another remark, it must also be emphasized that the cautious rule provides a BBA $m_1 \owedge m_2$ which is the w-least committed in the set $\mathcal{S}_w(m_1) \cap \mathcal{S}_w(m_2)$ of BBAs that are w-more committed than both m_1 and m_2. When either m_1 or m_2 is not u-separable, then $m_1 \ocap m_2$ does not belong to that set (because, e.g., $w_1(A)w_2(A) > w_1(A)$ whenever $w_2(A) > 1$). Consequently, we do not have $m_1 \ocap m_2 \sqsubseteq_w m_1 \owedge m_2$ in general, except when both m_1 and m_2 are separable.

In practice, the cautious combination of two non dogmatic BBAs m_1 and m_2 can thus be computed as follows:

- Compute the commonality functions q_1 and q_2 using (5);
- Compute the weight functions w_1 and w_2 using (26);
- Compute $m_{1 \owedge 2} = m_1 \owedge m_2$ as the $\ocap$ combination of GSBBAs $A^{w_1(A) \wedge w_2(A)}$, for all $A \subset \Omega$ such that $w_1 \wedge w_2(A) \neq 1$.

Example 2 Table 2 shows the weight functions of two BBAs m_1 and m_2 on $\Omega = \{a, b, c\}$, a well as the combined weight function $w_{1 \owedge 2}$ and BBA $m_1 \owedge m_2$. In this case, $m_{1 \owedge 2}$ is obtained as the TBM conjunctive combination of three SBBAs: $\{b\}^{0.7}$, $\{a, b\}^{2/5}$ and $\{b, c\}^{2/7}$. By combining the first two we get a mass 0.3 on $\{b\}$, $3/5 \times 0.7 = 0.42$ on $\{a, b\}$ and $2/5 \times 0.7 = 0.28$ on Ω. Combination with $\{b, c\}^{2/7}$ then yields

$$\begin{aligned}
m_{1 \owedge 2}(\{b\}) &= 0.3 \times (2/7 + 5/7) + 0.42 \times 5/7 = 0.6, \\
m_{1 \owedge 2}(\{a, b\}) &= 0.42 \times 2/7 = 0.12, \\
m_{1 \owedge 2}(\{b, c\}) &= 0.28 \times 5/7 = 0.2, \\
m_{1 \owedge 2}(\Omega) &= 0.28 \times 2/7 = 0.08.
\end{aligned}$$

The result of the combination of m_1 and m_2 using the TBM conjunctive rule directly from (9) is shown in the last column of Table 2 for comparison.

Table 2: Combination of two BBAs using the cautious rule.

A	$m_1(A)$	$w_1(A)$	$m_2(A)$	$w_2(A)$	$w_{1 \owedge 2}(A)$	$m_{1 \owedge 2}(A)$	$m_{1 \ocap 2}(A)$
$\emptyset$	0	1	0	1	1	0	0
$\{a\}$	0	1	0	1	1	0	0
$\{b\}$	0	7/4	0.3	0.7	0.7	0.6	0.42
$\{a, b\}$	0.3	2/5	0	1	2/5	0.12	0.09
$\{c\}$	0	1	0	1	1	0	0
$\{a, c\}$	0	1	0	1	1	0	0
$\{b, c\}$	0.5	2/7	0.4	3/7	2/7	0.2	0.43
Ω	0.2		0.3			0.08	0.06

3.2 Properties

Proposition 5 *The cautious conjunctive rule has the following properties:*

Commutativity: *for all* m_1 *and* m_2, $m_1 \owedge m_2 = m_2 \owedge m_1$*;*

Associativity: *for all* m_1, m_2 *and* m_3, $m_1 \owedge (m_2 \owedge m_3) = (m_1 \owedge m_2) \owedge m_3$*;*

Idempotence: *for all* m, $m \owedge m = m$*;*

Distributivity of $\textcircled{\cap}$ with respect to $\owedge$ *: for all* m_1, m_2 *and* m_3,

$$m_1 \textcircled{\cap} (m_2 \owedge m_3) = (m_1 \textcircled{\cap} m_2) \owedge (m_1 \textcircled{\cap} m_3).$$

Proof: Commutativity, associativity and idempotence result directly from corresponding properties of the minimum operator. Distributivity of $\textcircled{\cap}$ with respect to $\owedge$ is a consequence of distribution of the product with respect to the minimum:

$$w_1(w_2 \wedge w_3) = (w_1 w_2) \wedge (w_1 w_3), \quad \forall w_1, w_2, w_3.$$

$\square$

The last property (distributivity) is actually quite important, as it explains why the cautious rule can be considered to be more relevant than the TBM conjunctive rule $\textcircled{\cap}$ when combining non distinct items of evidence: if two sources provide BBAs $m_1 \textcircled{\cap} m_2$ and $m_1 \textcircled{\cap} m_3$ having some evidence m_1 in common, the shared evidence is not counted twice.

The following proposition is linked to the notion of LBS introduced in Section 2.2.4. It will be useful to explain some additional properties of the cautious rule.

PROPOSITION 6 *Let* m_1 *and* m_2 *be two non dogmatic BBAs with conjunctive weight functions* w_1 *and* w_2*. Let* (m_1^c, m_1^d) *and* (m_2^c, m_2^d) *denote the LBSs associated to* m_1 *and* m_2*, respectively, and let* (w_1^c, w_1^d) *and* (w_2^c, w_2^d) *denote the corresponding weights. Then the LBS* $(m_{1\owedge 2}^c, m_{1\owedge 2}^d)$ *associated to* $m_1 \owedge m_2$ *is defined by*

$$m_{1\owedge 2}^c = \textcircled{\cap}_{A \subseteq \Omega} A^{w_1^c \wedge w_2^c},$$

$$m_{1\owedge 2}^d = \textcircled{\cap}_{A \subseteq \Omega} A^{w_1^d \vee w_2^d},$$

where $\vee$ *denotes the maximum operation.*

Proof: For any $A \subset \Omega$, we have

$$\begin{aligned} w_{1\owedge 2}^c(A) &= 1 \wedge w_{1\owedge 2}(A) \\ &= 1 \wedge w_1(A) \wedge w_2(A) \\ &= (1 \wedge w_1(A)) \wedge (1 \wedge w_2(A)) \\ &= w_1^c(A) \wedge w_2^c(A), \end{aligned}$$

and

$$\begin{aligned} w^d_{1\owedge 2}(A) &= 1 \wedge \frac{1}{w_{1\owedge 2}(A)} \\ &= 1 \wedge \frac{1}{w_1(A) \wedge w_2(A)} \\ &= 1 \wedge \left(\frac{1}{w_1(A)} \vee \frac{1}{w_2(A)} \right) \\ &= \left(1 \wedge \frac{1}{w_1(A)} \right) \vee \left(1 \wedge \frac{1}{w_2(A)} \right) \\ &= w^d_1(A) \vee w^d_2(A). \end{aligned}$$

□

We thus see that, using the cautious rule, the confidence parts are combined conjunctively, whereas the diffidence parts are combined disjunctively by taking the maximum of the two weight functions w^d_1 and w^d_2. Note that such a disjunctive combination is well defined only for u-separable BBAs (see Section 4 for further discussion on this issue and the definition of a disjunctive counterpart of the cautious rule). Combining the diffidence components disjunctively does seem to make sense, as shown by the following informal argument. According to Smets [33], $m_d(A)$ should be interpreted as the strength of evidence that one should *not* believe A. If I receive two pieces of evidence, one of which tells me *not* to believe A while the other tells me *not* to believe B, then I am inclined not to believe $A \cup B$, hence the disjunctive nature of the combination. Consequently, there seems to be some form of duality between the confidence and diffidence components of a LBS, which translates into different mechanisms for combining each of the two components.

As is well known, the vacuous BBA is the neutral element for the TBM conjunctive rule, whereas it is an absorbing element for the TBM disjunctive rule. As a consequence of the dual conjunctive/disjunctive nature of the cautious rule, cautious combination of a BBA m with the vacuous BBA has the effect of absorbing the diffidence component, while leaving the confidence component unchanged. This is formalized in the following proposition.

Proposition 7 *For any non dogmatic BBA m with corresponding LBS (m^c, m^d):*

$$m \owedge m_\Omega = m^c.$$

Proof. This is a direct consequence of Proposition 6. Let w^c and w^d denote, respectively, the weight functions of m^c and m^d. The weights associated to the confidence component of $m \owedge m_\Omega$ are $w^c(A) \wedge 1 = w^c(A)$ for all $A \subset \Omega$, whereas those associated to the diffidence component are $w^d(A) \vee 1 = 1$ for all $A \subset \Omega$. □

The following proposition follows directly from the previous one.

Proposition 8 *For any non dogmatic BBA m, $m_\Omega \owedge m = m$ iff m is u-separable.*

Proof. m is u-seperable iff it is equal to its confidence component, i.e. $m = m^c$, hence the result. □

Proposition 8 implies that, when combining a BBA m with the vacuous BBA using the cautious rule, one does not in general recover m but only a *u-separable approximation* in the form of its confidence component. This is a consequence of Proposition 6, which shows that, for non u-separable BBAs, the cautious rule is not purely conjunctive as it combines the diffidence components in a disjunctive manner. Furthermore, it is easy to see that the cautious conjunctive rule has no neutral element, since the only BBA m_0 such that $m \owedge m_0 = m$ for any u-separable BBA m is the vacuous BBA, and this property is not satisfied for non u-separable BBAs.

Note that, in practice, the cautious rule will often behave in a purely conjunctive fashion because most belief functions encountered in applications are u-separable. This is the case, in particular, for belief functions elicited from experts, which are often obtained by discounting logical propositions. This is also the case for consonant BBAs, as shown in in Section 2.2.2, and for belief functions inferred using the General Bayesian Theorem [32, 2] (see Section 6) or the evidential case-based reasoning approach [4, 5], two widely used mechanisms for reasoning with belief functions in diagnosis and classification applications [7].

To conclude this description of the main properties of the cautious rule, it is important to mention some implications of selecting the $\sqsubseteq_w$ ordering in its definition. As a consequence of (46), we have, for any bbas m_1 and m_2;

$$\mathcal{S}_w(m_1) \cap \mathcal{S}_w(m_2) \subseteq \mathcal{S}_d(m_1) \cap \mathcal{S}_d(m_2) \subseteq \mathcal{S}_s(m_1) \cap \mathcal{S}_s(m_2) \subseteq \mathcal{S}_{pl}(m_1) \cap \mathcal{S}_{pl}(m_2)$$

and

$$\mathcal{S}_w(m_1) \cap \mathcal{S}_w(m_2) \subseteq \mathcal{S}_d(m_1) \cap \mathcal{S}_d(m_2) \subseteq \mathcal{S}_s(m_1) \cap \mathcal{S}_s(m_2) \subseteq \mathcal{S}_q(m_1) \cap \mathcal{S}_q(m_2),$$

with the subset relations being usually strict. Choosing the combined bba in $\mathcal{S}_w(m_1) \cap \mathcal{S}_w(m_2)$, as done by the cautious rule, then comes down to choosing the *smallest set* of possible combined bbas in the above relations. In that set, the cautious rule selects the w-least committed element, which exists and is unique, as stated by Proposition 4. In that sense, it may be termed "cautious" as it is derived from the LCP. However, it must be kept in mind that the choice of $\mathcal{S}_w(m_1) \cap \mathcal{S}_w(m_2)$ imposes severe restrictions on the combination. As a consequence, bbas x-less committed than $m_1 \owedge m_2$ (with $x \in \{w, d, s, pl, q\}$) may exist outside $\mathcal{S}_w(m_1) \cap \mathcal{S}_w(m_2)$. In particular, when m_1 or m_2 is not u-separable, $m_1 \bigcirc\!\!\!\!\!\cap\, m_2$ does not belong to $\mathcal{S}_w(m_1) \cap \mathcal{S}_w(m_2)$, and it is possible to have $m_1 \owedge m_2 \sqsubseteq_w m_1 \bigcirc\!\!\!\!\!\cap\, m_2$, as shown by the following example[1].

Example 3 Let us consider the following bbas on $\Omega = \{a, b, c, d, e\}$:

$$m_1(A) = \begin{cases} 0.4 & \text{if } A = \{a, b\} \text{ or } A = \{b, c\} \\ 0.2 & \text{if } A = \Omega, \\ 0 & \text{otherwise,} \end{cases}$$

$$m_2(A) = \begin{cases} 0.4 & \text{if } A = \{b, d\} \text{ or } A = \{b, e\} \\ 0.2 & \text{if } A = \Omega, \\ 0 & \text{otherwise.} \end{cases}$$

[1] This example was suggested to the author by Frédéric Pichon.

The corresponding weight functions are

$$w_1(A) = \begin{cases} 1/3 & \text{if } A = \{a, b\} \text{ or } A = \{b, c\} \\ 1.8 & \text{if } A = \{b\}, \\ 1 & \text{otherwise}, \end{cases}$$

$$w_2(A) = \begin{cases} 1/3 & \text{if } A = \{b, d\} \text{ or } A = \{b, e\} \\ 1.8 & \text{if } A = \{b\}, \\ 1 & \text{otherwise}. \end{cases}$$

It can easily be checked that $w_1(A) \wedge w_2(A) \leq w_1(A)w_2(A)$ for all $A \subset \Omega$ and, consequently, $m_1 \owedge m_2 \sqsubseteq_w m_1 \ocap m_2$.

As illustrated by the previous example, the cautious rule is not more "cautious" than the TBM conjunctive rule when applied to non u-separable bbas, even in the sense of the $\sqsubseteq_w$ ordering. As will be shown in Section 5, these two rules actually belong to two different families of rules with distinct algebraic properties, and as such they cannot easily be compared.

The strong constraint imposed to the cautious rule seems to be the price to pay for its ease of calculation and good properties (associativity, idempotence, distributivity of $\ocap$ with respect to $\owedge$)) as outlined above. It would, of course, be possible to consider a larger set $\mathcal{S}_x(m_1) \cap \mathcal{S}_x(m_2)$, with $x \in \{d, s, pl, q\}$ as the search space for the combination of two bbas m_1 and m_2. However, the existence and unicity of a x-least committed element would then no longer be verified in general, making it impossible to apply the LCP. One could also consider selecting a combined bba maximizing an uncertainty measure (see, e.g., [25]). Finding both a search space and an uncertainty measure ensuring interesting properties of the combination seems to be a very difficult problem which is left for further research.

3.3 The Normalized Cautious Rule

A normalized version of the cautious rule introduced in the previous section may be defined by replacing the conjunctive rule $\ocap$ by Dempster's rule $\oplus$ in (48). Denoting this rule by $\owedge^*$, we have:

$$m_{1 \owedge^* 2} = m_1 \owedge^* m_2 = \bigoplus_{\emptyset \neq A \subset \Omega} A^{w_1(A) \wedge w_2(A)}. \tag{49}$$

We thus have

$$m_{1 \owedge^* 2}(A) = k \cdot m_{1 \owedge 2}(A), \quad \forall A \subseteq \Omega, A \neq \emptyset,$$

with $k = (1 - m_{1 \owedge 2}(\emptyset))^{-1}$, and $m_{1 \owedge^* 2}(\emptyset) = 0$. Note that we can never have $m_{1 \owedge 2}(\emptyset) = 1$, because the cautious combination of two non dogmatic BBAs can never be dogmatic. As shown in Section 2.2, the weight functions of $m_1 \owedge^* m_2$ and $m_1 \owedge m_2$ differ only by the weight assigned to the empty set: with obvious notations, we have

$$w_{1 \owedge^* 2}(A) = w_{1 \owedge 2}(A) = w_1(A) \wedge w_2(A), \quad \forall A \in 2^\Omega \setminus \{\emptyset, \Omega\},$$

and $w_{1 \owedge^* 2}(\emptyset) = k \cdot w_{1 \owedge 2}(\emptyset)$. Clearly, this rule has the same properties as its unnormalized counterpart: it is commutative, associative and idempotent. When combining several BBAs $m_1, \ldots, m_n$ using the normalized cautious rule, we may either compute their unnormalized cautious combination and normalize the end result, or normalize intermediate results in the computation. The same property is known to hold for Dempster's rule [29].

The normalized cautious rule $\owedge^*$ can be justified using the LCP with a suitable ordering relation. Let $\sqsubseteq_w^*$ denote the following partial order between non dogmatic BBAS: $m_1 \sqsubseteq_w^* m_2$ iff $w_1(A) \leq w_2(A)$, for all $A \subset \Omega$, $A \neq \emptyset$. Obviously, we have

$$m_1 \sqsubseteq_w m_2 \Rightarrow m_1 \sqsubseteq_w^* m_2,$$

for all m_1 and m_2, and the implication is strict. When the condition $m_1 \sqsubseteq_w^* m_2$ holds, we will say that m_1 is w^*-more committed than m_2. Using the same line of reasoning as in Section 3.1, it is easy to show that $m_1 \owedge^* m_2$ is the w^*-least committed BBA among all *normal* BBAs which are w^*-more committed than m_1 and m_2.

The following proposition is a counterpart to Proposition 8:

PROPOSITION 9 *For any non dogmatic normal BBA m, $m_\Omega \owedge^* m = m$ iff m is n-separable.*

Proof. Let m be a non dogmatic normal BBA and w is weight function. Then $w(A) \wedge 1 = w(A)$ for $A \subset \Omega$, $A \neq \emptyset$ implies that $w(A) \leq 1$ for $A \subset \Omega$, $A \neq \emptyset$, i.e., m is n-separable. The converse is obvious. □

EXAMPLE 4 Table 3 shows intermediate steps for the computation of the normalized cautious combination of two BBAs m_1 and m_2. Their weight functions are first computed and combined using the minimum operator. The corresponding BBA is then computed, and normalized. It can be checked that the weight function $w_{1\owedge^*2}$ is equal to $w_{1\owedge 2}$, except for $w_{1\owedge^*2}(\emptyset) = 1/(1 - m_{1\owedge 2}(\emptyset)) = 2.57$.

Table 3: Combination of two BBAs using the normalized cautious rule.

A	$m_1(A)$	$w_1(A)$	$m_2(A)$	$w_2(A)$	$w_{1\owedge 2}(A)$	$m_{1\owedge 2}(A)$	$m_{1\owedge^* 2}(A)$
$\emptyset$	0	1	0	7/5	1	0.61	0
$\{a\}$	0	1	0.3	0.5	0.5	0.061	0.16
$\{b\}$	0	7/4	0	1	1	0.092	0.24
$\{a, b\}$	0.3	2/5	0	1	2/5	0.037	0.094
$\{c\}$	0	1	0.4	3/7	3/7	0.11	0.29
$\{a, c\}$	0	1	0	1	1	0	0
$\{b, c\}$	0.5	2/7	0	1	2/7	0.061	0.16
Ω	0.2		0.3			0.025	0.063

Note that we have, in Example 4, $m_{1\owedge 2}(\emptyset) = 0.61$, whereas it can be checked that $m_{1\ocap 2}(\emptyset) = 0.27$. This illustrates the fact that the cautious rule is not related to conflict minimization, contrary to, e.g., the rule proposed in [1].

4 The Bold Disjunctive Rule

Just as the cautious rule extends set intersection, as shown in Section 3.1, one may wonder whether the same principle could be used to derive a disjunctive operator generalizing set union. Just as the union of two sets A and B is the smallest set containing both A and B, we could attempt to define the disjunction of two BBAs m_1 and m_2 issued from two sources as the *most* x-committed BBA, among the set of BBAs less x-committed than m_1 and m_2. This approach seems appropriate when it is only known that at least one of the two sources is reliable, but we do not know which one. The combined BBA should then be less informative than each of the BBAs provided by the individual sources.

Formally, let us denote as $\mathcal{G}_x(m)$ the set of BBAs m' such that $m \sqsubseteq_x m'$. If the set $\mathcal{G}_x(m_1) \cap \mathcal{G}_x(m_2)$ possesses a most x-committed element, then this element could, by definition, be equated to the disjunction of m_1 and m_2. Since this BBA would be the most committed one, among those which are less informative than m_1 and m_2, such a rule could be named a *bold* disjunctive rule.

If we adopt w-ordering as the definition of inclusion, then $\mathcal{G}_w(m_1)$ is the set of BBAs m such that $w(A) \geq w_1(A)$ for all $A \subset \Omega$. Similarly, $\mathcal{G}_w(m_2)$ contains the BBAs m such that $w(A) \geq w_2(A)$ for all $A \subset \Omega$. The intersection $\mathcal{G}_w(m_1) \cap \mathcal{G}_w(m_2)$ thus contains the set of BBAs for which $w(A) \geq w_1(A) \vee w_2(A)$ for all $A \subset \Omega$, where $\vee$ denotes the maximum operator. The most w-committed element in that set, if it exists, has the weight function $w_1 \vee w_2$. This approach is valid in the case where m_1 and m_2 are both separable BBA: in that case, we still have $w_1(A) \vee w_2(A) \leq 1$ for all $A \subset \Omega$, and $w_1 \vee w_2$ defines a separable belief function [20]. However, this rule cannot be used to combine arbitrary non dogmatic BBAs, because $w_1 \vee w_2$ does not always correspond to a belief function, as shown by the following counterexample.

Example 5 Consider the two BBAs m_1 and m_2 of Example 2, and let $w = w_1 \vee w_2$ be the weight function obtained by taking the maximum of the weight functions of m_1 and m_2. We have $w(\{b\}) = 7/4 \vee 0.7 = 7/4$, $w(\{b,c\}) = 2/7 \vee 3/7 = 3/7$, and $w(A) = 1$ for all other $A \subset \Omega$. The corresponding mass function is thus the TBM combination of ISBBA $\{b\}^{7/4}$ and SBBA $\{b,c\}^{3/7}$. We get $m(\{b\}) = -3/4$, $m(\{b,c\}) = 7/4 \times 4/7 = 1$, and $m(\Omega) = 7/4 \times 3/7 = 3/4$, which does not correspond to a belief function.

In the rest of this section, we will show that the above approach does allow to define a disjunctive counterpart of the cautious rule, provided it is based on a proper informational ordering of belief functions. To define such an ordering, we will need to introduce a canonical *disjunctive* representation of belief functions, dual to the "conjunctive" one introduced in [33] and recalled in Section 2.2. This representation is presented in the following section.

4.1 Canonical Disjunctive Decomposition of a Subnormal BBA

Let m be a subnormal BBA. Its complement $\overline{m}$ is non dogmatic and can be decomposed as

$$\overline{m} = \bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{\overline{w}(A)}.$$

Consequently, m can be written

$$\begin{aligned} m &= \overline{\bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{\overline{w}(A)}} \\ &= \bigcirc\!\!\!\!\cup_{A \subset \Omega} \overline{A^{\overline{w}(A)}}. \end{aligned}$$

We recall that $A^{\overline{w}(A)}$ denotes the GSBBA assigning a mass $\overline{w}(A) > 0$ to Ω and a mass $1 - \overline{w}(A)$ to A. Consequently, its complement $\overline{A^{\overline{w}(A)}}$ is a generalized BBA with two focal sets: $\overline{A}$ with a mass $1 - \overline{w}(A)$, and $\emptyset$ with a mass $\overline{w}(A)$. Such a mapping can be called a negative GSBBA, as it is the negation of a GSBBA, and noted $\overline{A}_{v(\overline{A})}$, with $v(\overline{A}) = \overline{w}(A)$. We can thus write

$$
\begin{aligned}
m &= \bigcirc\!\!\!\!\cup_{A \subset \Omega} \overline{A}_{v(\overline{A})} && (50)\\
&= \bigcirc\!\!\!\!\cup_{A \neq \emptyset} A_{v(A)}. && (51)
\end{aligned}
$$

We have proved the following proposition:

Proposition 10 (Canonical disjunctive decomposition) *Any subnormal BBA m can be uniquely decomposed as the $\bigcirc\!\!\!\!\cup$-combination of negative generalized BBAs $A_{v(A)}$ assigning a mass $v(A) > 0$ to $\emptyset$, and a mass $1 - v(A)$ to A, for all $A \subseteq \Omega$, $A \neq \emptyset$:*

$$m = \bigcirc\!\!\!\!\cup_{A \neq \emptyset} A_{v(A)}. \quad (52)$$

Function $v : 2^\Omega \setminus \{\emptyset\} \to (0, +\infty)$ will be referred to as the disjunctive weight function. It is related to the conjunctive weight function $\overline{w}$ associated to the negation $\overline{m}$ of m by the equation

$$v(A) = \overline{w}(\overline{A}), \quad \forall A \neq \emptyset. \quad (53)$$

A comparison between equations (8) and (53) shows that the relation between v and w parallels that between b and q. As a consequence, v can be obtained from b using a formula similar to (28), as expressed in the following proposition.

Proposition 11 *Let v and b the disjunctive weight and implicability functions associated to a subnormal BBA m. They are related by the following equation:*

$$\ln v(A) = -\sum_{B \subseteq A} (-1)^{|A|-|B|} \ln b(B). \quad (54)$$

Proof: Using (53) and (28), we have

$$
\begin{aligned}
\ln v(A) &= \ln \overline{w}(\overline{A})\\
&= -\sum_{\overline{B} \supseteq \overline{A}} (-1)^{|\overline{B}|-|\overline{A}|} \ln \overline{q}(\overline{B}).
\end{aligned}
$$

Now, (8) implies that $\overline{q}(\overline{B}) = b(B)$. Observing that $\overline{B} \supseteq \overline{A} \Leftrightarrow B \subseteq A$, and $|\overline{B}| - |\overline{A}| = |A| - |B|$, we get the desired result. □

Note that relation (54) between $\ln v$ and $-\ln b$ has the same form as relation (7) between m and b. Consequently, any procedure for transforming b to m can be used to compute $\ln v$ from $-\ln b$.

Example 6 Table 4 illustrates the computation of disjunctive weights using (53).

Just as the TBM conjunctive rule can be easily computed using conjunctive weights using (35), the TBM disjunctive rule has a simple expression in terms of the disjunctive weights, as shown by the following proposition.

Table 4: Computation of disjunctive weights.

A	$m(A)$	$\overline{m}(A)$	$\overline{w}(A)$	$v(A)$
$\emptyset$	0.1	0	2.8	
$\{a\}$	0	0.6	0.1429	1
$\{b\}$	0	0	1	1
$\{a,b\}$	0.3	0	1	0.25
$\{c\}$	0	0.3	0.25	1
$\{a,c\}$	0	0	1	1
$\{b,c\}$	0.6	0	1	0.1429
Ω	0	0.1		2.8

Proposition 12 *Let m_1 and m_2 be two subnormal BBAs with disjunctive weight functions v_1 and v_2. The disjunctive weight function $v_{1 \textcircled{\cup} 2}$ associated to $m_1 \textcircled{\cup} m_2$ is given by $v_{1 \textcircled{\cup} 2} = v_1 v_2$.*

Proof: It is easy to check that we have $A_v \textcircled{\cup} A_{v'} = A_{vv'}$. Consequently, we have:

$$\begin{aligned} m_1 \textcircled{\cup} m_2 &= \left(\textcircled{\cup}_{A \neq \emptyset} A_{v_1(A)} \right) \textcircled{\cup} \left(\textcircled{\cup}_{A \neq \emptyset} A_{v_2(A)} \right) && (55) \\ &= \textcircled{\cup}_{A \neq \emptyset} A_{v_1(A) v_2(A)}. && (56) \end{aligned}$$

□

It follows directly that the inverse $\oslash$ of the TBM disjunctive rule also has a simple expression in the v-space, as $v_{1 \oslash 2} = v_1 / v_2$.

Finally, the concept of latent belief structure recalled in Section 2.2.4 also has a disjunctive counterpart. For each disjunctive weight $v(A)$ let us define the following two quantities:

$$v^c(A) = 1 \wedge v(A), \tag{57}$$

and

$$v^d(A) = 1 \wedge \frac{1}{v(A)}. \tag{58}$$

It is clear that we have

$$v(A) = \frac{v^c(A)}{v^d(A)}. \tag{59}$$

Consequently, we can write

$$\begin{aligned} m &= \textcircled{\cup}_{A \subset \Omega} A_{v^c(A)/v^d(A)} && (60) \\ &= \left(\textcircled{\cup}_{A \subset \Omega} A_{v^c(A)} \right) \oslash \left(\textcircled{\cup}_{A \subset \Omega} A_{v^d(A)} \right) && (61) \\ &= m^c_{dis} \oslash m^d_{dis}. && (62) \end{aligned}$$

The pair (m^c_{dis}, m^d_{dis}) is the disjunctive counterpart of the LBS introduced in 2.2.4, and can be named the disjunctive LBS associated to m.

4.2 Informational Ordering Based on Disjunctive Weights

The concept of disjunctive weight function defined above makes it possible to define a new partial ordering relation between belief functions, which is the counterpart of $\sqsubseteq_w$ introduced in Section 2.3.

Let m_1 and m_2 be two subnormal BBAs with disjunctive weight functions v_1 and v_2. Assume that $v_1(A) \geq v_2(A)$, for all $A \neq \emptyset$. Let $v = v_2/v_1$, and m the corresponding BBA (it corresponds to a belief function, since $v(A) \leq 1$ for all $A \neq \emptyset$). We thus have $m_2 = m_1 \ocircle\!\!\!\!\cup\, m$, which implies that m_1 is a specialization of m_2. In that sense, m_1 is more informative than m_2. Consequently, the following new informational ordering can be introduced:

$$m_1 \sqsubseteq_v m_2 \Leftrightarrow v_1(A) \geq v_2(A), \quad \forall A \neq \emptyset.$$

If $m_1 \sqsubseteq_v m_2$, we will say that m_1 is v-more committed than m_2.

Just as $\sqsubseteq_v$ is a counterpart of $\sqsubseteq_w$ as a result of the duality between the conjunctive and disjunctive decompositions, we may observe that it is also possible to define a dual to the d-ordering (Dempsterian specialization) recalled in Section 2.3. Assume that $m_2 = m_1 \ocircle\!\!\!\!\cup\, m$ for some *arbitrary* bba m. Then m_1 is a particular kind of specialization of m_1, which we can write as: $m_1 \sqsubseteq_{dd} m_2$. This new ordering is obviously stronger than $\sqsubseteq_s$, but weaker than $\sqsubseteq_v$. The two new ordering relations $\sqsubseteq_v$ and $\sqsubseteq_{dd}$ allow us to complete the picture drawn in Section 2.3 as follows:

$$\left.\begin{array}{l} m_1 \sqsubseteq_w m_2 \Rightarrow m_1 \sqsubseteq_d m_2 \\ m_1 \sqsubseteq_v m_2 \Rightarrow m_1 \sqsubseteq_{dd} m_2 \end{array}\right\} \Rightarrow m_1 \sqsubseteq_s m_2 \Rightarrow \left\{\begin{array}{l} m_1 \sqsubseteq_{pl} m_2 \\ m_1 \sqsubseteq_q m_2, \end{array}\right.$$

where again all implications are strict.

4.3 Derivation of the Bold Disjunctive Rule

Using the general approach outlined at the beginning of this section, we can define a disjunctive rule based on the $\sqsubseteq_v$ ordering, as shown by the following proposition.

Proposition 13 *Let m_1 and m_2 be two subnormal BBAs. The v-most committed element in $\mathcal{G}_v(m_1) \cap \mathcal{G}_v(m_2)$ exists and is unique. It is defined by the following disjunctive weight function:*

$$v_{1 \varovee 2}(A) = v_1(A) \wedge v_2(A), \quad \forall A \in 2^\Omega \setminus \emptyset. \tag{63}$$

Proof: The proof is similar to that of Proposition 4. For any $m \in \mathcal{G}_v(m_1) \cap \mathcal{G}_v(m_2)$, we have $v(A) \leq v_1(A)$ and $v(A) \leq v_2(A)$, hence $v(A) \leq v_1(A) \wedge v_2(A)$ for all non empty subset A of Ω. The v-most committed element in $\mathcal{G}_v(m_1) \cap \mathcal{G}_v(m_2)$ is obtained by taking the minimum of $v_1(A)$ and $v_2(A)$ for all A. One can verify that it corresponds to a belief function, as a consequence of a counterpart of Lemma 1 for disjunctive weights. □

Equation (63) introduces a new rule which can be formally defined as follows.

Definition 2 (Bold disjunctive rule) *Let m_1 and m_2 be two subnormal BBAs. Their combination using the bold disjunctive rule is noted $m_{1 \varovee 2} = m_1 \varovee m_2$. It is defined as the BBA with the following disjunctive weight function:*

$$v_{1 \varovee 2}(A) = v_1(A) \wedge v_2(A), \quad A \in 2^\Omega \setminus \emptyset.$$

We thus have

$$m_1 \circledvee m_2 = \bigcirc\!\!\!\!\cup_{A \neq \emptyset} A_{v_1(A) \wedge v_2(A)}. \tag{64}$$

EXAMPLE 7 Table 5 shows two subnormal BBAs, together with their v-representation and their bold disjunction. The resulting BBA $m_{1\circledvee 2}$ may be computed from $v_{1\circledvee 2}$ by combining the three BBAs $\{b\}_{1/6}$, $\{a, b\}_{1/4}$ and $\{b, c\}_{1/7}$ using the disjunctive rule $\bigcirc\!\!\!\!\cup$. The result of the combination using the TBM disjunctive rule is shown in the last column for comparison.

Table 5: Combination of two BBAs using the bold disjunctive rule.

A	$m_1(A)$	$v_1(A)$	$m_2(A)$	$v_2(A)$	$v_{1\circledvee 2}(A)$	$m_{1\circledvee 2}(A)$	$m_{1\bigcirc\!\!\!\!\cup 2}(A)$
$\emptyset$	0.1	1	0.1	1	1	0.0060	0.01
$\{a\}$	0	1	0	1	1	0	0
$\{b\}$	0	1	0.5	1/6	1/6	0.0298	0.05
$\{a, b\}$	0.3	1/4	0	1	1/4	0.1071	0.18
$\{c\}$	0	1	0	1	1	0	0
$\{a, c\}$	0	1	0	1	1	0	0
$\{b, c\}$	0.6	1/7	0.4	0.6	1/7	0.2143	0.64
Ω	0	14/5	0	1	1	0.6429	0.12

The fact that the bold disjunctive rule is only applicable to subnormal BBAs may appear as a severe restriction, as most belief functions encountered in practice are normal (and subnormality often arises when combining conflicting items of evidence using the TBM conjunctive rule). However, practically, it is always possible to "de-normalize" a BBA by transferring a very small proportion of the unit mass of belief to the empty set. In the TBM, the mass $m(\emptyset)$ may be interpreted as being committed to the hypothesis than none of the elementary hypotheses in the frame of discernment is true. Assigning even an infinitesimal mass to this hypothesis may be justified in most cases, as a way to acknowledge the fact that the adopted model might not be complete, unlikely as it may be.

4.4 Properties of the Bold Disjunctive Rule

The bold disjunctive rule has properties which parallel those of the cautious conjunctive rule, due to the dual nature of these two rules. These properties are listed below.

PROPOSITION 14 *The bold disjunctive rule has the following properties:*

Commutativity: *for all m_1 and m_2, $m_1 \circledvee m_2 = m_2 \circledvee m_1$;*

Associativity: *for all m_1, m_2 and m_3, $m_1 \circledvee (m_2 \circledvee m_3) = (m_1 \circledvee m_2) \circledvee m_3$;*

Idempotence: *for all m, $m \circledvee m = m$;*

Distributivity of $\bigcirc\!\!\!\!\cup$ with respect to $\circledvee$ *: for all m_1, m_2 and m_3,*

$$m_1 \bigcirc\!\!\!\!\cup (m_2 \circledvee m_3) = (m_1 \bigcirc\!\!\!\!\cup m_2) \circledvee (m_1 \bigcirc\!\!\!\!\cup m_3).$$

Proof. The proof is similar to that of Proposition 18. □

The following proposition is the counterpart of Proposition 6, and can be derived in the same manner. It shows that the bold disjunctive rule treats differently the two components of disjunctive LBSs.

PROPOSITION 15 *Let m_1 and m_2 be two subnorm BBAs with disjunctive weight functions v_1 and v_2. Let $(m^c_{dis,1}, m^d_{dis,1})$ and $(m^c_{dis,2}, m^d_{dis,2})$ denote the disjunctive LBSs associated to m_1 and m_2, respectively, and let (v^c_1, v^d_1) and (v^c_2, v^d_2) denote the corresponding disjunctive weights. Then the disjunctive LBS $(m^c_{dis,1\circledvee 2}, m^d_{dis,1\circledvee 2})$ associated to $m_1 \circledvee m_2$ is defined by*

$$m^c_{dis,1\circledvee 2} = \textcircled{\cup}_{A\subseteq\Omega} A_{v^c_1 \wedge v^c_2},$$

$$m^d_{dis,1\circledvee 2} = \textcircled{\cup}_{A\subseteq\Omega} A_{v^d_1 \vee v^d_2}.$$

Finally, the following proposition shows that the $\circledvee$ and $\circledwedge$ operations are dual to each other with respect to complementation, i.e., they are linked by De Morgan laws analogous to (16) and (17).

PROPOSITION 16 (DE MORGAN'S LAWS) *Let m_1 and m_2 be two subnormal BBAs. We have:*

$$\overline{m_1 \circledvee m_2} = \overline{m}_1 \circledwedge \overline{m}_2, \tag{65}$$

for all subnormal BBAs m_1 and m_2, and

$$\overline{m_1 \circledwedge m_2} = \overline{m}_1 \circledvee \overline{m}_2 \tag{66}$$

for all non dogmatic BBAs m_1 and m_2.

Proof: Let m_1 and m_2 be two subnormal BBAs. We have

$$\begin{aligned}
\overline{m_1 \circledvee m_2} &= \overline{\bigcup_{A\neq\emptyset} A_{v_1(A)\wedge v_2(A)}} \\
&= \bigcap_{A\neq\emptyset} \overline{A_{v_1(A)\wedge v_2(A)}} \\
&= \bigcap_{A\neq\emptyset} \overline{A}^{\overline{w}_1(A)\wedge\overline{w}_2(\overline{A})} \\
&= \bigcap_{A\subset\Omega} A^{\overline{w}_1(A)\wedge\overline{w}_2(A)} \\
&= \overline{m}_1 \circledwedge \overline{m}_2.
\end{aligned}$$

The proof of (66) is similar. □

5 General Combination Rules Based on Triangular Norms and Conorms

As we have seen, the cautious and TBM conjunctive rules are based on pointwise combination of conjunctive weights (using, respectively, the minimum and the product), whereas the bold and TBM disjunctive rule are based on similar combination of disjunctive weights. One may wonder whether such operations on weights could be generalized to define other combination rules with interesting properties. To simplify the discussion, only operations on conjunctive weights will be considered here; extension to disjunctive weights is obvious.

Let m_1 and m_2 be two non dogmatic BBAs with conjunctive weight functions w_1 and w_2. We have seen that each weight $w(A)$ may be decomposed into two components in $[0,1]$: a confidence component $w^c(A) = 1 \wedge w(A)$ and a diffidence component $w^d(A) = 1 \wedge (w(A))^{-1}$, with $w(A) = w^c(A)/w^d(A)$.

Both the TBM conjunctive rule and the cautious rule can be described in terms of operations on conjunctive and disjunctive weights:

- The TBM conjunctive rule combines the confidence and diffidence components of the weights using the product:

$$\begin{aligned} w^c_{1\bigcirc\!\!\!\!\cap\, 2}(A) &= w^c_1(A) \cdot w^c_2(A), & (67)\\ w^d_{1\bigcirc\!\!\!\!\cap\, 2}(A) &= w^d_1(A) \cdot w^d_2(A); & (68) \end{aligned}$$

- The cautious rule combines the confidence components using the minimum, and the diffidence components using the maximum:

$$\begin{aligned} w^c_{1\bigcirc\!\!\!\!\wedge\, 2}(A) &= w^c_1(A) \wedge w^c_2(A), & (69)\\ w^d_{1\bigcirc\!\!\!\!\wedge\, 2}(A) &= w^d_1(A) \vee w^d_2(A). & (70) \end{aligned}$$

Can these operations be generalized? To answer this question, we may observe that, in the interval $[0,1]$, the product and the minimum are triangular norms (t-norms for short), whereas the maximum is a triangular conorm (or t-conorm) [24]. We recall that a t-norm is a commutative and associative binary operator $\top$ on the unit interval satisfying the monotonicity property

$$y \leq z \Rightarrow x \top y \leq x \top z, \quad \forall x, y, z \in [0,1],$$

and the boundary condition $x \top 1 = x$, $\forall x \in [0,1]$. A t-conorm $\perp$ has the same three basic properties (commutativity, associativity, monotonicity) and differs only by the boundary condition $x \perp 0 = x$. Because of their different boundary conditions, t-norms and t-conorms are usually interpreted, respectively, as generalized conjunction and disjunction operators in fuzzy logic.

A first consequence of this observation is that the TBM conjunctive rule and the cautious rule combine the diffidence weights using operations with different algebraic properties and, in that respect, they should be regarded as belonging to different families of combination rules. However, it does seem possible to define new combination rules with interesting properties by generalizing the cautious rule and the TBM conjunctive rule separately. This is done in the rest of this section, with emphasis on the generalization of the cautious rule, which is the main topic of this paper.

5.1 Generalized cautious rules

The following proposition shows that new rules for combining non dogmatic belief functions can be defined by replacing the minimum and the maximum in (69) and (70) by, respectively, a positive t-norm and a t-conorm. (A t-norm $\top$ is said to be positive iff $x > 0$ and $y > 0$ implies $x \top y > 0$).

Proposition 17 *Let m_1 and m_2 be two non dogmatic BBAs, w_1 and w_2 their weight functions, and (w_1^c, w_1^d) and (w_2^c, w_2^d) their decompositions into confidence and diffidence components. Let w_{1*2} be the mapping from $2^\Omega \setminus \Omega$ to $(0, +\infty)$ defined as:*

$$w_{1*2}(A) = \frac{w_1^c(A) \top w_2^c(A)}{w_1^d(A) \perp w_2^d(A)}, \quad \forall A \subset \Omega, \tag{71}$$

where $\top$ is a positive t-norm, and $\perp$ a t-cornorm. Then:

- *Function w_{1*2} is the conjunctive weight function of a non dogmatic BBA m_{1*2};*
- *We have $m_{1*2} \sqsubseteq_w m_1 \owedge m_2$.*

Proof. It is known [24] that the minimum is the largest t-norm, while the maximum is the weakest t-conorm. Consequently, we have $w_1^c(A) \top w_2^c(A) \leq w_1^c(A) \wedge w_2^c(A)$ and $w_1^d(A) \perp w_2^d(A) \geq w_1^d(A) \vee w_2^d(A)$, hence $w_{1*2}(A) \leq w_1(A) \wedge w_2(A)$ for all A. Using Lemma 1, this proves that w_{1*2} corresponds to a belief function. It is obviously w-more committed than $m_1 \owedge m_2$. Additionally, positivity of $\top$ ensures $w_{1*2}(A) > 0$ and, consequently, that m_{1*2} is non dogmatic. $\square$

Note that each combined weight $w_{1*2}(A)$ can be expressed directly as a function of $w_1(A)$ and $w_2(A)$ as $w_{1*2}(A) = w_1(A) *_{\top,\perp} w_2(A)$, where $*_{\top,\perp}$ is the following operator in $(0, +\infty)$:

$$x *_{\top,\perp} y = \begin{cases} x \top y & \text{if } x \vee y \leq 1, \\ x \wedge y & \text{if } x \vee y > 1 \text{ and } x \wedge y \leq 1, \\ \left(\frac{1}{x} \perp \frac{1}{y} \right)^{-1} & \text{otherwise,} \end{cases} \tag{72}$$

for all $x, y > 0$.

Given a positive t-norm $\top$ and a t-cornorm $\perp$, Proposition 17 allows us to define a belief function combination operator $\circledast_{\top,\perp}$ as

$$m_1 \circledast_{\top,\perp} m_2 = \bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w_1(A) *_{\top,\perp} w_2(A)},$$

where $*_{\top,\perp}$ is defined by (72). Note that the cautious rule corresponds to $\circledast_{\wedge,\vee}$.

The $\circledast_{\top,\perp}$ operator has some interesting properties. We start with the following lemma.

Lemma 2 *For any positive t-norm $\top$ and any t-conorm $\perp$, the operator $*_{\top,\perp}$ defined by (72) is commutative, associative, and satisfies the monotonicity property*

$$y \leq z \Rightarrow x * y \leq x * z, \quad \forall x, y, z > 0.$$

Proof: Commutativity results directly from the commutativity of $\top$, $\wedge$ and $\perp$. For associativity, we may consider different cases:

- If $x \leq 1$, $y \leq 1$ and $z \leq 1$, then

$$(x *_{\top,\perp} y) *_{\top,\perp} z = (x \top y) \top z = x \top (y \top z) = x *_{\top,\perp} (y *_{\top,\perp} z);$$

- If $x > 1$, $y > 1$ and $z > 1$, then

$$\begin{aligned}
(x *_{\top,\perp} y) *_{\top,\perp} z &= \left(\left[\left(\frac{1}{x} \perp \frac{1}{y} \right)^{-1} \right]^{-1} \perp \frac{1}{z} \right)^{-1} \\
&= \left(\left[\frac{1}{x} \perp \frac{1}{y} \right] \perp \frac{1}{z} \right)^{-1} \\
&= \left(\frac{1}{x} \perp \left[\frac{1}{y} \perp \frac{1}{z} \right] \right)^{-1} \\
&= \left(\frac{1}{x} \perp \left[\left(\frac{1}{y} \perp \frac{1}{z} \right)^{-1} \right]^{-1} \right)^{-1} \\
&= x *_{\top,\perp} (y *_{\top,\perp} z);
\end{aligned}$$

- If $x \leq 1$, $y \leq 1$ and $z > 1$, then

$$(x *_{\top,\perp} y) *_{\top,\perp} z = (x \top y) \wedge z = x \top y,$$

  and

$$x *_{\top,\perp} (y *_{\top,\perp} z) = x \top (y \wedge z) = x \top y;$$

- If $x > 1$, $y > 1$ and $z \leq 1$, then

$$(x *_{\top,\perp} y) *_{\top,\perp} z = (x *_{\top,\perp} y) \wedge z = z,$$

  and

$$x *_{\top,\perp} (y *_{\top,\perp} z) = x \wedge (y \wedge z) = z.$$

The other cases can be deduced from the above last two cases using the commutativity property. Finally, monotonicity can be proved in a similar manner, by considering the different cases:

Case 1: $y \leq z \leq 1$. Then:

- if $x \leq 1$, then $x *_{\top,\perp} y = x \top y$ and $x *_{\top,\perp} z = x \top z$, and $x \top y \leq x \top z$ by the monotonicity of $\top$;
- if $x > 1$, then $x *_{\top,\perp} y = y$ and $x *_{\top,\perp} z = z$, and the result follows directly.

Case 2: $y \leq 1 < z$. Then:

- if $x \leq 1$, then $x *_{\top,\perp} y = x \top y$ and $x *_{\top,\perp} z = x \wedge z$. Now, $x \top y \leq x \wedge y \leq x \wedge z$, since $\top$ is dominated by $\wedge$;
- if $x > 1$, then $x *_{\top,\perp} y = y$ and $x *_{\top,\perp} z = ((1/x) \perp (1/z))^{-1}$. Now, we have $1/x < 1/y$ and $1/z < 1/y$, hence $(1/x) \perp (1/z) < 1/y$ and $((1/x) \perp (1/z))^{-1} > y$.

Case 3: $1 < y \leq z$. Then:

- if $x \leq 1$, then $x *_{\top,\perp} y = x$, $x *_{\top,\perp} z = x$ and the result follows directly;

- if $x > 1$, then $x *_{\top,\perp} y = ((1/x)\perp(1/y))^{-1}$ and $x *_{\top,\perp} z = ((1/x)\perp(1/z))^{-1}$. Now, since $1/y \geq 1/z$, we have $(1/x)\perp(1/y) \geq (1/x)\perp(1/z)$ by the monotonicity of $\perp$, hence $((1/x)\perp(1/y))^{-1} \leq ((1/x)\perp(1/z))^{-1}$. $\square$

PROPOSITION 18 *The $\circledast_{\top,\perp}$ rule has the following properties:*

Commutativity: *for all m_1 and m_2, $m_1 \circledast_{\top,\perp} m_2 = m_2 \circledast_{\top,\perp} m_1$;*

Associativity: *for all m_1, m_2 and m_3,*

$$m_1 \circledast_{\top,\perp} (m_2 \circledast_{\top,\perp} m_3) = (m_1 \circledast_{\top,\perp} m_2) \circledast_{\top,\perp} m_3;$$

Monotonicity with respect to $\sqsubseteq_w$ *: for all m_1, m_2 and m_3,*

$$m_1 \sqsubseteq_w m_2 \Rightarrow m_1 \circledast_{\top,\perp} m_3 \sqsubseteq_w m_2 \circledast_{\top,\perp} m_3.$$

Proof: These properties follow directly from corresponding properties of $*_{\top,\perp}$ expressed in Lemma 2.

5.2 Discussion

We thus have defined a family of commutative and associative combination operators, which also have the property of monotonicity with respect to $\sqsubseteq_w$. The latter property means that, if a BBA m_1 is less informative than a BBA m_2 according to the $\sqsubseteq_w$ ordering, then this order is unchanged after combination with a third BBA. The cautious rule is the only idempotent rule in this family, since the minimum and the maximum are, respectively, the only idempotent t-norm and co-tnorm.

We may remark that, normalized combination rule can be defined in the same way, by combining the weights of non empty subsets A of Ω, and normalizing the result. These normalized combination rules are also commutative, associative, and monotonic with respect to $\sqsubseteq_w^*$.

The same approcah can also be used to generalize the TBM conjunctive rule, by replacing the product in (67) and (68) by two t-norms $\top_1$ and $\top_2$, respectively. However, some cautious should be exercised here, because the resulting combined weight function is not guaranteed to correspond to a belief function for any choice of $\top_1$ and $\top_2$. For instance, choosing $\top_1 = \top_2 = \wedge$ does not yield a belief function in general. A sufficient condition for obtaining a belief function is to choose $\top_1$ and $\top_2$ such that $\top_1 \leq \Pi \leq \top_2$, where Π denotes the product t-norm. Deeper investigation of this topic is beyond the scope of this paper.

One may object that these new combination rules, in spite of their interesting properties outlined in Proposition 18, are only weakly justified. However, we may remark that the same situation prevails in Possibility theory [11], where there are as many conjunctive and disjunctive operators as t-norms and t-conorms. Although this multiplicity of operators may be seen as a weakness of the axiomatic foundations of Possibility theory, it also proves beneficial from a practical point of view as it provides considerable flexibility to adjust the behavior of a system to user-defined requirements [14] or to learning examples. In contrast, Dempster-Shafer theory has sometimes been criticized for its lack of flexibility in the choice of combination operator [9], a criticism which, in light of the new results presented in this paper, appears to be unjustified.

Figure 1: Distance between the target BBA and the combined BBA, as a function of parameter s of the Frank family of t-norms.

5.3 Combination rule optimization

As shown in the previous section, a commutative and associative operator based on conjunctive (or disjunctive) weights can be associated to each pair of a t-norm and a t-conorm. By choosing parameterized families of t-norms and t-conorms [24], it is thus possible to defined parameterized families of belief function combination rules. This introduces the possibility to learn a combination rule from examples, as shown in the following simple illustrative example.

EXAMPLE 8 Assume that the BBAs m_1 and m_2 shown in Table 6 have been provided by two sensors, and expert knowledge regarding the true value of the variable of interest is represented by BBA m_e also shown in Table 6. For this simple illustrative example, m_e was artificially constructed by combining m_1 and m_2 using the $\circledast_{\top,\perp}$ operator based on the Frank t-norm and dual t-conorm with parameter $s = 0.5$, and adding a small amount of random noise. We recall that the Frank family of t-norms [24] is defined by

$$x \top_s y = \begin{cases} x \wedge y & \text{if } s = 0, \\ xy & \text{if } s = 1, \\ \log_s \left(1 + \frac{(s^x - 1)(s^y - 1)}{s - 1}\right) & \text{otherwise,} \end{cases}$$

for all $x, y \in [0, 1]$, where s is a positive parameter. The dual t-conorm $\perp_s$ is defined by $x \perp_s y = 1 - (1 - x) \top_s (1 - y)$.

We wish to find a combination rule of the form $\circledast_{\top_s,\perp_s}$, where $\top_s$ and $\perp_s$ are, respectively, the Frank t-norm and t-conorm with parameter s, such that the combination of m_1 and m_2 yields a BBA as close as possible to m_e. Note that, in real applications such as classifier fusion problems, a large number of such learning instances would typically be available. Parameter s was varied between 0 and 1, and for each s value the discrepancy between $m_{12} = m_1 \circledast_{\top_s,\perp_s} m_2$ and m_e was measured by Jousselme's distance [19] defined :

$$d(m_{12}, m_e) = \sqrt{\frac{1}{2}(\mathbf{m}_{12} - \mathbf{m}_e)^t D(\mathbf{m}_{12} - \mathbf{m}_e)},$$

where $\mathbf{m}_{12}$ and $\mathbf{m}_e$ are $2^{|\Omega|}$-dimensional vectors of basic belief masses corresponding to m_{12} and m_e, and D is the square matrix of size $2^{|\Omega|}$ defined by

$$D(A, B) = \begin{cases} 1 & \text{if } A = B = \emptyset, \\ \frac{|A \cap B|}{|A \cup B|} & \text{otherwise.} \end{cases}$$

Distance d is plotted as a function of s in Figure 1. The best fit between the combined BBA m_{12} and the target BBA m_e is obtained for $s \approx 0.33$, which is an estimate of the true value $s = 0.5$.

Table 6: BBAs of Example 8.

	$\emptyset$	$\{a\}$	$\{b\}$	$\{a,b\}$	$\{c\}$	$\{a,c\}$	$\{b,c\}$	$\{a,b,c\}$
m_1	0.11	0.10	0.13	0.01	0.10	0.07	0.19	0.29
m_2	0.08	0.19	0.16	0.05	0	0.10	0.11	0.31
m_e	0.3965	0.1210	0.1446	0.0236	0.0871	0.0498	0.0934	0.0841

6 Application to classifier fusion

Although the cautious rule and its relatives have nice mathematical properties, their usefulness in applications of belief functions might be questioned. In this section, we present numerical experiments showing the efficiency of the cautious rule and a t-norm based generalization to combine classifiers built from dependent features.

6.1 Problem statement and formalization

Let us consider a classification problem with K classes and d continuous features $X_1, \ldots, X_d$. Assume that each feature X_i has a known conditional probability distribution $f_k(x_i)$ in each class ω_k ($k = 1, \ldots, K$). The class prior probabilities are unknown. Additionally, nothing is known concerning the correlations between features.

This problem can be tackled in the TBM using the General Bayesian Theorem (GBT) [32, 2, 7]. Assume that the known probability density $f_k(x_i)$ of feature X_i in ω_k is interpreted as the pignistic probability of some unknown conditional belief function on $\mathbb{R}$. For simplicity, $f_k(x_i)$ will be assumed to be unimodal and symmetric. As shown in [37], the q-least committed belief function on $\mathbb{R}$ associated with a unimodal symmetric pignistic probability density f_k with mode ν_k is consonant (and, consequently, equivalent to a plausibility measure). The corresponding possibility distribution (called "contour function" by Shafer [29, page 221]) is:

$$pl_k(x_i) = \begin{cases} 2(x_i - \nu_k) f_k(x_i) + 2\int_{x_i}^{+\infty} f_k(t_i)dt_i & \text{if } x_i \geq \nu_k \\ 2(\nu_k - x_i) f_k(x_i) + 2\int_{-\infty}^{x_i} f_k(t_i)dt_i & \text{otherwise.} \end{cases} \quad (73)$$

The quantity $pl_k(x_i)$ is the plausibility that feature X_i takes value x_i, given that the object belongs to class ω_k.

Assume that the value x_i of feature X_i has been observed for a certain object. What is our belief state concerning the class of this object ? In the TBM, the answer is provided by the GBT. The induced BBA on the set of classes $\Omega = \{\omega_1, \ldots, \omega_K\}$, conditional on $X_i = x_i$, is u-separable [7]. It is given by:

$$m^\Omega[x_i] = \textcircled{\cap}_{k=1}^{K} \overline{\{\omega_k\}}^{pl_k(x_i)}. \quad (74)$$

If the features X_i are assumed to be conditionally independent given the class, then the evidence of the d feature values can be considered as distinct and, as such, can be combined by the TBM conjunctive rule:

$$\begin{aligned} m^\Omega[x_1, \ldots, x_d] &= \textcircled{\cap}_{i=1}^{d} m^\Omega[x_i] & (75) \\ &= \textcircled{\cap}_{k=1}^{K} \overline{\{\omega_k\}}^{\prod_{i=1}^{d} pl_k(x_i)}. & (76) \end{aligned}$$

If conditional independence is not assumed, then the cautious rule may be more appropriate. We then have:

$$m^\Omega[x_1, \ldots, x_d] = \textcircled{\wedge}_{i=1}^{d} m^\Omega[x_i] \tag{77}$$

$$= \textcircled{\cap}_{k=1}^{K} \overline{\{\omega_k\}}^{\bigwedge_{i=1}^{d} pl_k(x_i)}. \tag{78}$$

Note that the possibility distributions $pl_k(x_i)$ are combined using the product t-norm in (76), whereas they are combined using the minimum t-norm in (78). Using a generalized cautious rule $\circledast_{\top,\perp}$ such as introduced in Section 5.1 amounts to combining the $pl_k(x_i)$ using t-norm $\top$. If $\top$ is chosen in the Frank family, then the cautious rule is recovered for $s = 0$, whereas the TBM conjunctive rule is recovered for $s = 1$. Choosing s between 0 and 1 results in a combination rule somewhere between these two extremes.

6.2 Experimental results

Numerical simulations were performed for a particular instance of the above problem, with $K = 2$ classes and $d = 10$ features. The conditional distribution of feature vector $(X_1, \ldots, X_d)$ in class ω_k was assumed to be multivariate normal with mean $\mu_1 = (0, \ldots, 0)$ in class ω_1 and $\mu_2 = (1, \ldots, 1)$ in class ω_2, and with common variance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho & \rho & \ldots & \rho & 0 \\ \rho & 1 & \rho & \ldots & \rho & 0 \\ \rho & \rho & \ddots & & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \rho & \vdots \\ \rho & \rho & \ldots & \rho & 1 & 0 \\ 0 & 0 & \ldots & \ldots & 0 & 1 \end{pmatrix},$$

with $\rho \in [0, 1]$. Conditionally on each class, the last feature X_{10} was thus assumed to be independent from all other features, whereas the correlation coefficient between any two features X_i and X_j, $i, j \in \{1, \ldots, 9\}$ was equal to ρ.

For any observed feature vector $x = (x_1, \ldots, x_{10})$, a decision was computed as follows:

- The plausibilities $pl_k(x_i)$ were computed using (73), for each i and each k;
- The BBAs $m^\Omega[x_i]$ on Ω given feature X_i were computed using (74);
- The combined BBA $m^\Omega[x_1, \ldots, x_{10}]$ was computed using the conjunctive rule using (76), the cautious rule using (78), or the generalized cautious rule $\circledast_{\top,\perp}$ with $\top$ equal to the Frank t-norm with $s = 0.01$;
- The pignistic transformation (18) was applied to $m^\Omega[x_1, \ldots, x_{10}]$, and the pignistic probability of class ω_1 was compared to some threshold.

The above procedure was repeated for $n = 5000$ test vectors in each class, for $\rho = 0$, $\rho = 0.5$ and $\rho = 0.9$. The false positive rate (proportion of test vectors from class ω_1 wrongly classified as ω_2) and the true positive rate (proportion of test vectors from class ω_2 correctly classified as ω_2) were estimated for each combination rule and each value of ρ. The corresponding ROC curves (plot of the true positive rate as a function of the false positive rate) are shown in Figure 2. In this representation, a higher curve corresponds to higher performance (higher true positive rate for any false positive rate).

Figure 2: ROC curves for the TBM conjunctive, cautious and generalized cautious rules, for $\rho = 0$ (top), $\rho = 0.5$ (middle) and $\rho = 0.9$ (bottom).

As expected, the TBM conjunctive rule achieves higher performance in the case of independent features. However, it is outperformed by the cautious rule when features are no longer independent. The generalized cautious rule with the Frank t-norm for $s = 0.01$ has intermediate performances in all three situations. This is an experimental verification of the validity of the cautious rule in the case of non distinct evidence.

7 Conclusion

Two new commutative, associative and idempotent operators for belief functions have been introduced. The cautious conjunctive rule $\owedge$ has been derived from the Least Commitment Principle with a suitable informational ordering: the $\owedge$-combination of two non dogmatic BBAs m_1 and m_2 has been defined as the least committed BBA according to the $\sqsubseteq_w$ ordering, among those which are more committed than m_1 and m_2, according to the same ordering. Symetrically, the combination of two subnormal BBAs m_1 and m_2 using the bold disjunctive rule $\ovee$ has been defined as the most committed BBA according to $\sqsubseteq_v$, an ordering dual to $\sqsubseteq_w$, among BBAs that are less committed than m_1 and m_2.

Contrary to the TBM conjunctive and disjunctive rules $\ocap$ and $\ocup$, these two operators do not require the assumption of independence, or distinctness of the information sources from which BBAs are derived. Independently from this distinctness assumption, conjunctive operators $\ocap$ and $\owedge$ are appropriate when all sources are believed to be reliable, whereas disjunctive operators should be used when one only assumes that at least one of the sources is reliable. The choice of one operator among $\ocap$, $\ocup$, $\owedge$ and $\ovee$ thus depends on assumptions regarding both the distinctness and reliability of the sources, as summarized in the following table:

sources	all reliable	at least one reliable
distinct	$\ocap$	$\ocup$
non distinct	$\owedge$	$\ovee$

The cautious and bold rules have also been shown to belong to infinite families of conjunctive and disjunctive operators based on t-norms and t-conorms. Using parametrized families of t-norms and t-conorms, corresponding families of conjunctive and disjunctive operators can be defined. All these operators are commutative and associative, but only $\owedge$ and $\ovee$ are idempotent. Although these operators do not appear to be as well justified as the cautious and bold rules, they may be useful in classification or information fusion applications where the behavior of a combination rule can be tuned to optimize a given performance measure [43, 5, 17]. In any case, it appears that, contrary to a so far widely accepted opinion [9], the richness of potential combination operators is not lower in the theory of belief functions than it is in possibility theory, which opens new perspectives for applying belief functions theory to information fusion problems.

Acknowledgment

This work is dedicated to the memory of Professor Philippe Smets, who inspired most of the ideas expressed in this paper. The author thanks the anonymous referees for their careful reading of the manuscript and their constructive comments that helped to improve this paper. Thanks also to Frédéric Pichon for stimulating discussions on the cautious rule and the canonical decompositions.

References

[1] M. E. G. V. Cattaneo. Combining belief functions issued from dependent sources. In J. M. Bernard, T. Seidenfeld, and M. Zaffalon, editors, *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications (ISIPTA'03)*, pages 133–147, Lugano, Switzerland, 2003. Carleton Scientific.

[2] F. Delmotte and Ph. Smets. Target identification based on the Transferable Belief Model interpretation of Dempster-Shafer model. *IEEE Transactions on Systems, Man and Cybernetics A*, 34(4):457–471, 2004.

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

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

[5] T. Denœux. A neural network classifier based on Dempster-Shafer theory. *IEEE Trans. on Systems, Man and Cybernetics A*, 30(2):131–150, 2000.

[6] T. Denœux. The cautious rule of combination for belief functions and some extensions. In *Proceedings of the 9th International Conference on Information Fusion*, Florence (Italy), July 2006. Paper #114.

[7] T. Denœux and P. Smets. Classification using belief functions: the relationship between the case-based and model-based approaches. *IEEE Transactions on Systems, Man and Cybernetics B*, 36(6):1395–1406, 2006.

[8] D. Dubois and H. Prade. On several representations of an uncertainty body of evidence. In M. M. Gupta and E. Sanchez, editors, *Fuzzy Information and Decision Processes*, pages 167–181. North-Holland, New-York, 1982.

[9] D. Dubois and H. Prade. On the unicity of Dempster's rule of combination. *International Journal of Intelligent Systems*, 1:133–142, 1986.

[10] D. Dubois and H. Prade. A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets. *International Journal of General Systems*, 12(3):193–226, 1986.

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

[12] D. Dubois and H. Prade. Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence*, 4:244–264, 1988.

[13] D. Dubois, H. Prade, and Ph. Smets. New semantics for quantitative possibility theory. In S. Benferhat and Ph. Besnard, editors, *Proc. of the 6th European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU 2001)*, pages 410–421, Toulouse, France, 2001. Springer-Verlag.

[14] D. Dubois, H. Prade, and R. Yager. Merging fuzzy information. In J. C. Bezdek, D. Dubois, and H. Prade, editors, *Fuzzy sets in approximate reasoning and information systems*, pages 335–401. Kluwer Academic Publishers, Boston, 1999.

[15] D. Dubois and R. R. Yager. Fuzzy set connectives as combination of belief structures. *Information Sciences*, 66:245–275, 1992.

[16] Z. Elouedi and K. Mellouli. Pooling dependent expert opinions using the theory of evidence. In *Proc. of the Seventh Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 98)*, volume 1, pages 32–39, Paris, France, July 1998.

[17] Z. Elouedi, K. Mellouli, and Ph. Smets. Assessing sensor reliability for multisensor data fusion within the Transferable Belief Model. *IEEE Transactions on Systems, Man and Cybernetics B*, 34(1):782–787, 2004.

[18] R. Haenni. Are alternatives to Dempster's rule of combination real alternatives?: Comments on "about the belief function combination and the conflict management problem" – Lefèvre et al. *Information Fusion*, 3(3):237–239, 2002.

[19] A.-L. Jousselme, D. Grenier, and E. Bossé. A new distance between two bodies of evidence. *Information Fusion*, 2(2):91–101, 2001.

[20] R. Kennes. Evidential reasoning in a categorical perspective: conjunction and disjunction of belief functions. In B. D'Ambrosio, P. Smets, and P. P. Bonissone, editors, *Proceedings of the 7th Conference on Uncertainty in Artificial Intelligence*, pages 174–181. Morgan Kaufman, San Mateo, CA, 1991.

[21] R. Kennes and Ph. Smets. Computational aspects of the Möbius transformation. In P.P. Bonissone, M. Henrion, L. N. Kanal, and J. F. Lemmer, editors, *Proceedings of the 6th Conference on Uncertainty in Artificial Intelligence*, pages 401–416. Elsevier Science Publishers, 1990.

[22] F. Klawonn and E. Schwecke. On the axiomatic justification of Dempster's rule of combination. *International Journal of Intelligent Systems*, 7:469–478, 1992.

[23] F. Klawonn and Ph. Smets. The dynamic of belief in the transferable belief model and specialization-generalization matrices. In D. Dubois et al., editor, *Proc. of the 8th conference on Uncertainty in Artificial Intelligence*, pages 130–137. Morgan Kaufmann, San Mateo, CA, 1992.

[24] E. P. Klement, R. Mesiar, and E. Pap. *Triangular norms*. Kluwer Academic Publishers, Dordrecht, 2000.

[25] G. J. Klir and M. J. Wierman. *Uncertainty-Based Information. Elements of Generalized Information Theory*. Springer-Verlag, New-York, 1998.

[26] E. Lefèvre, O. Colot, and P. Vannoorenberghe. Belief function combination and conflict management. *Information Fusion*, 3(2):149–162, 2002.

[27] X. N. Ling and W. G. Rudd. Combining opinions from several experts. *Applied Artificial Intelligence*, 3:439–452, 1989.

[28] K. Sentz and S. Ferson. Combination of evidence in Dempster-Shafer Theory. Technical report, SANDIA National Laboratories, 2002.

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

[30] Ph. Smets. Combining non distinct evidences. In *Int. Conf. of the North American Fuzzy Information Processing Society (NAFIPS'86)*, pages 544–549, New Orleans, USA, 1986.

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

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

[33] Ph. Smets. The canonical decomposition of a weighted belief. In *Int. Joint Conf. on Artificial Intelligence*, pages 1896–1901, San Mateo, Ca, 1995. Morgan Kaufman.

[34] Ph. Smets. The Transferable Belief Model for quantified belief representation. In D. M. Gabbay and Ph. Smets, editors, *Handbook of Defeasible reasoning and uncertainty management systems*, volume 1, pages 267–301. Kluwer Academic Publishers, Dordrecht, 1998.

[35] Ph. Smets. The application of the matrix calculus to belief functions. *International Journal of Approximate Reasoning*, 31(1–2):1–30, 2002.

[36] Ph. Smets. Personal communication, 2005.

[37] Ph. Smets. Belief functions on real numbers. *International Journal of Approximate Reasoning*, 40(3):181–223, 2005.

[38] Ph. Smets. Analyzing the combination of conflicting belief functions. *Information Fusion (In press)*, 2007. doi:10.1016/j.inffus.2006.04.003.

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

[40] R. R. Yager. The entailment principle for Dempster-Shafer granules. *Int. J. of Intelligent Systems*, 1:247–262, 1986.

[41] R. R. Yager. On the Dempster-Shafer framework and new combination rules. *Information Sciences*, 41:93–137, 1987.

[42] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.

[43] L. M. Zouhal and T. Denœux. An evidence-theoretic k-NN rule with parameter optimization. *IEEE Trans. on Systems, Man and Cybernetics C*, 28(2):263–271, 1998.