

The cautious rule of combination for belief functions and some extensions

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

Abstract - *Dempster's rule plays a central role in the theory of belief functions. However, it assumes the items of evidence combined 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 non distinct items of evidence. Extensions based on triangular norms (some of which allow to define operators whose behavior is intermediate between the Dempster's rule and the cautious rule) are also introduced.*

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

1 Introduction

Dempster's rule of combination [1, 14] is known to play a pivotal role in the theory of belief functions, together with its unnormalized version (hereafter referred to as the TBM conjunctive rule) introduced by Smets in the Transferable Belief Model (TBM) [16]. Justifications for the origins and unicity of these rules have been provided by several authors [2, 16, 11, 10]. 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 the conjunctive rules of combination seem to be their lack of robustness with respect to conflicting evidence (a criticism which mainly applies to the normalized 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 [24, 5, 13]. The disjunctive rule of combination [3, 17] is both associative and more robust than Dempster's rule in the presence of conflicting evidence, but it is often too conservative to be useful in practice. It may also be argued that problems with Dempster's rule (and, to a lesser extent, with its unnormalized version) 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 [8].

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 [1], 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" [1]. 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., [2, 15]). However, it is often the case that, although two items of evidence (such as, e.g., opinions expressed by two experts sharing certain experiences, or observations of correlated random quantities) can clearly not be regarded as distinct, the interaction between them is only partially known and, in many cases, almost impossible to describe.

In such a common situation, it would be extremely helpful to have a combination rule that would not rely on the distinctness assumption. A natural requirement for such a rule is idempotence. The arithmetic mean does possess this property, but it is not associative, another requirement often regarded as essential. Whereas an associative and idempotent conjunctive operator (the minimum t-norm) exists in possibility theory, no equivalent existed until now in the theory of belief functions. Such an operator is proposed in this paper, building on some ideas suggested to the author by the late Philippe Smets [21].

The rest of this paper is organized as follows. The underlying fundamental concepts, including the canonical decomposition of a belief function, will first be recalled in Section 2. The cautious rule of combination will then be introduced in Section 3, and some extensions applying to separable and nonseparable belief functions will be described in Section 4. Section 5 will conclude the paper.

2 Fundamental concepts

2.1 Basic definitions and notations

In this paper, the TBM [22, 19] is accepted as a model of uncertainty. An agent's state of belief expressed on a finite frame of discernment $\Omega = \{\omega_1, \dots, \omega_K\}$ 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);
- 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, including Ω ;
- categorical if it is simple and dogmatic.

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, plausibility and commonality functions, defined, respectively, as:

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

$$pl(A) = \sum_{B \cap A \neq \emptyset} m(B), \quad (2)$$

and

$$q(A) = \sum_{B \supseteq A} m(B), \quad (3)$$

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

$$m(A) = \sum_{B \supseteq A} (-1)^{|B|-|A|} q(B), \quad \forall A \subseteq \Omega. \quad (4)$$

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

$$m_{1 \odot 2}(A) = \sum_{B \cap C = A} m_1(B) m_2(C), \quad \forall A \subseteq \Omega, \quad (5)$$

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

$$m_{1 \oplus 2}(A) = \begin{cases} 0 & \text{if } A = \emptyset \\ \frac{m_{1 \odot 2}(A)}{1 - m_{1 \odot 2}(\emptyset)} & \text{otherwise.} \end{cases} \quad (6)$$

Both rules are commutative, associative, and admit a unique neutral element: the vacuous BBA. Let A^{w_1} and A^{w_2} be two SBBA with the same focal element $A \neq \Omega$. The result of their \odot -combination is the SBBA $A^{w_1 w_2}$. The \oplus operator yields the same result as long as $A \neq \emptyset$.

2.2 Canonical decomposition of a belief function

Shafer [14, Chapter 4] defined a (normal) separable BBA as a the result of the \oplus combination of SBBA. The normality condition can easily be dropped. For every separable BBA (normal or not), one has:

$$m = \bigoplus_{A \subsetneq \Omega} A^{w(A)}, \quad (7)$$

with $w(A) \in [0, 1]$ for all $A \subsetneq \Omega$. This representation is unique if all focal elements are different and m is non dogmatic. Shafer called this representation the canonical representation of m .

The above representation was extended to any non dogmatic BBA by Smets [18]. The key to such a generalization is the notion of generalized simple BBA (GS-BBA), defined as a function μ from 2^Ω to \mathbb{R} verifying

$$\mu(A) = 1 - w, \quad (8)$$

$$\mu(\Omega) = w, \quad (9)$$

$$\mu(B) = 0 \quad \forall B \in 2^\Omega \setminus \{A, \Omega\}, \quad (10)$$

for some $A \neq \Omega$ and some $w \in [0, +\infty)$. Any GS-BBA 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. 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 GSBBA:

$$m = \bigodot_{A \subsetneq \Omega} A^{w(A)}, \quad (11)$$

with $w(A) \in [0, +\infty)$ for all $A \subsetneq \Omega$. This is clearly an extension of (7). The weights $w(A)$ for every $A \in 2^\Omega \setminus \{\Omega\}$ can be obtained from the following formula:

$$w(A) = \prod_{B \supseteq A} q(B)^{(-1)^{|B|-|A|+1}}, \quad (12)$$

$$= \begin{cases} \frac{\prod_{B \supseteq A, |B| \notin 2\mathbb{N}} q(B)}{\prod_{B \supseteq A, |B| \in 2\mathbb{N}} q(B)} & \text{if } |A| \in 2\mathbb{N} \\ \frac{\prod_{B \supseteq A, |B| \in 2\mathbb{N}} q(B)}{\prod_{B \supseteq A, |B| \notin 2\mathbb{N}} q(B)} & \text{otherwise,} \end{cases} \quad (13)$$

where $2\mathbb{N}$ denotes the set of even natural numbers. Eq. (12) can be equivalently written

$$\ln w(A) = - \sum_{B \supseteq A} (-1)^{|B|-|A|} \ln q(B). \quad (14)$$

One notices the similarity with (4). Hence, any procedure for transforming q to m (such as the Fast Möbius Transform [9] or matrix multiplication [20]) can be used to compute $\ln w$ from $-\ln q$.

Example 1 Let $\Omega = \{a, b, c\}$ be a frame of discernment. The commonality and weight functions for a BBA m on Ω are shown in Table 1. We can see that m is not separable, since we have $w(\{b\}) > 1$.

The function $w : 2^\Omega \rightarrow [0, +\infty)$ (hereafter referred to as the weight function) is thus yet another equivalent representation of any non dogmatic BBA. It can be extended to a dogmatic BBA m by discounting it with some discount rate ϵ and letting ϵ tend towards

Table 1: A non separable 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	

0 [18]. 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 is in fact $m(\{Heads\}) = 0.5(1 - \epsilon)$, $m(\{Tails\}) = 0.5(1 - \epsilon)$ and $m(\Omega) = \epsilon$ for some small $\epsilon > 0$.

Note that any non normal BBA $m = \bigoplus_{A \subseteq \Omega} A^{w(A)}$ may be transformed into a normal BBA m^* by combining it with \emptyset^k with $k = (1 - m(\emptyset))^{-1}$. Hence, the weight function w^* associated to m^* is identical to w , except for the weight assigned to \emptyset . We can write equivalently

$$m^* = \bigoplus_{\emptyset \neq A \subseteq \Omega} A^{w(A)}. \quad (15)$$

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:

$$\begin{aligned} m_1 \circledast m_2 &= \left(\bigoplus_{A \subseteq \Omega} A^{w_1(A)} \right) \circledast \left(\bigoplus_{A \subseteq \Omega} A^{w_2(A)} \right) \\ &= \bigoplus_{A \subseteq \Omega} A^{w_1(A)w_2(A)}. \end{aligned} \quad (17)$$

We can thus write, with obvious notations: $w_1 \circledast w_2 = w_1 w_2$ (a similar property is known to exist for the commonalities: $q_1 \circledast q_2 = q_1 q_2$). Similarly, using (15), it is easy to see that

$$m_1 \oplus m_2 = \bigoplus_{\emptyset \neq A \subseteq \Omega} A^{w_1(A)w_2(A)}. \quad (18)$$

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 [17], 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 to compare belief functions according to their information content. Three such partial orderings, generalizing set inclusion, were proposed in the 1980's by Yager [23] and Dubois and Prade [3]; 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 $S(A, B)$, $A, B \in 2^\Omega$ such that

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

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

and

$$m_1(A) = \sum_{B \subseteq \Omega} S(A, B)m_2(B), \quad \forall A \subseteq \Omega. \quad (19)$$

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 [11, 20], and m_1 is said to be a specialization of m_2 .

As shown in [3], 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 [3] from a set-theoretical perspective, and in [6] 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 [11]. 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 $m_1 = m \circledast m_2$. As shown in [11], this is a stronger condition than specialization, i.e., we have $m_1 \sqsubseteq_d m_2 \Rightarrow m_1 \sqsubseteq_s m_2$: if $m_1 = m \circledast m_2$, then there is a specialization matrix S_m defined as function of m , called a Dempsterian specialization matrix, allowing to compute m_1 from m_2 using relation (19).

Finally, we can think of a last 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 $A \in 2^\Omega \setminus \Omega$. Because of (17), this is equivalent to the existence of a *separable* BBA m such that $m_1 = m \circledast m_2$. 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 the \sqsubseteq_w , the requirement that m be 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 separable. Furthermore, we will see that w -ordering may be a simpler and more convenient notion, for some purposes, than s -ordering.

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

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

and all these implications are strict.

Using these definitions, 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 $x \in \{pl, q, s, d, w\}$. In the absence of decisive arguments in favor of one ordering or another, the choice of x is usually guided by technical considerations. For instance, q -ordering is adopted in [6] to derive the expression of the q -least committed BBA with given pignistic probability function [22]. In the following section, the same principle is used to derive a rule of combination, using partial ordering \sqsubseteq_w .

3 The cautious rule

3.1 Derivation from the LCP

Just as relations \sqsubseteq may be seen as generalizing set inclusion, it is possible to conceive conjunctive combinations rules generalizing set intersection, by reasoning as follows. Assume that we have two sources of information, one of which indicates that the 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 be consider as certain the fact 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 provides BBAs m_1 and m_2 . If they both are trusted, then the agent's state of belief, after receiving these two pieces of information, should 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 $m' \sqsubseteq m$, 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 [6], 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 $A \subseteq \Omega$, where \wedge denotes the minimum operator, is the x -least committed element in the set $\mathcal{S}_x(m_1) \cap \mathcal{S}_x(m_2)$, for $x \in \{pl, q, s\}$. 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 order does yield an interesting solution, as shown by the following proposition.

Proposition 1 *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:*

$$w_{1 \circledast 2}(A) = w_1(A) \wedge w_2(A), \quad A \in 2^\Omega \setminus \Omega. \quad (20)$$

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 $A \in 2^\Omega \setminus \Omega$. 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 $A \in 2^\Omega \setminus \Omega$. Let us define the following weight function: $w_{1 \circledast 2}(A) = w_1(A) \wedge w_2(A)$, for all $A \in 2^\Omega \setminus \Omega$. We must show that this weight function is valid, i.e., it corresponds to a BBA. For that, it is sufficient to remark that

$$w_{1 \circledast 2}(A) = w_1(A) \cdot \frac{w_1(A) \wedge w_2(A)}{w_1(A)} \\ = w_1(A) \cdot \left(1 \wedge \frac{w_1(A)}{w_2(A)} \right).$$

Consequently, $w_{1 \circledast 2}$ is the weight function of a BBA $m_{1 \circledast 2}$ obtained by computing the \circledast -combination of m_1 with a separable bba with weight function $1 \wedge w_1/w_2$. \square

Equation (20) defines a new rule which can be formally defined as follows.

Definition 1 (Cautious rule) *Let m_1 and m_2 be two non dogmatic BBAs. Their combination using the cautious rule is noted $m_{1 \circledast 2} = m_1 \circledast m_2$. It is defined as the BBA with the following weight function:*

$$w_{1 \circledast 2}(A) = w_1(A) \wedge w_2(A), \quad A \in 2^\Omega \setminus \Omega. \quad (21)$$

We thus have

$$m_{1 \circledast 2} = \bigoplus_{A \subseteq \Omega} A^{w_1(A) \wedge w_2(A)}. \quad (22)$$

Note that this rule happens to generalize a method proposed by Kennes [9] for combining 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 a much wider class of belief functions.

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 (3);

- Compute the weight functions w_1 and w_2 using (12);
- Compute $m_1 \circledast m_2 = m_1 \circledast m_2$ as the \circledast combination of GSBBA's $A^{w_1(A) \wedge w_2(A)}$, for all $A \subset \Omega$ such that $w_1 \wedge w_2(A) \neq 1$.

Example 2 Tables 3 shows the weight functions of the two BBAs m_1 and m_2 of Table 2, a well as the combined weight function $w_1 \circledast m_2$ and BBA $m_1 \circledast m_2$. In this case, $m_1 \circledast m_2$ is obtained as the TBM conjunctive combination of three SBBA's: $\{b\}^{0.7}$, $\{a, b\}^{2/5}$ and $\{b, c\}^{2/7}$. The resulting BBA is shown in the last column of Table 3. In contrast, the TBM conjunctive rule yields the following result in that case:

$$m_1 \circledast m_2(\{b\}) = 0.42, \quad m_1 \circledast m_2(\{a, b\}) = 0.09,$$

$$m_1 \circledast m_2(\{b, c\}) = 0.43, \quad m_1 \circledast m_2(\Omega) = 0.06.$$

Table 2: Two BBAs and their commonality functions.

A	$m_1(A)$	$q_1(A)$	$m_2(A)$	$q_2(A)$
\emptyset	0	1	0	1
$\{a\}$	0	0.5	0	0.3
$\{b\}$	0	1	0.3	1
$\{a, b\}$	0.3	0.5	0	0.3
$\{c\}$	0	0.7	0	0.7
$\{a, c\}$	0	0.2	0	0.3
$\{b, c\}$	0.5	0.7	0.4	0.7
Ω	0.2	0.2	0.3	0.3

Table 3: Combination of the BBAs of Table 2 using the cautious rule.

A	$w_1(A)$	$w_2(A)$	$w_1 \circledast m_2(A)$	$m_1 \circledast m_2(A)$
\emptyset	1	1	1	0
$\{a\}$	1	1	1	0
$\{b\}$	7/4	0.7	0.7	0.6
$\{a, b\}$	2/5	1	2/5	0.12
$\{c\}$	1	1	1	0
$\{a, c\}$	1	1	1	0
$\{b, c\}$	2/7	3/7	2/7	0.2
Ω				0.08

3.2 Properties

The main properties of the cautious rule result directly from corresponding properties of the minimum operator:

commutativity: for all m_1 and m_2 , $m_1 \circledast m_2 = m_2 \circledast m_1$;

associativity: for all m_1 , m_2 and m_3 , $m_1 \circledast (m_2 \circledast m_3) = (m_1 \circledast m_2) \circledast m_3$;

idempotence: for all m , $m \circledast m = m$;

distributivity of \circledast with respect to \circledast : for all m_1 , m_2 and m_3 ,

$$m_1 \circledast (m_2 \circledast m_3) = (m_1 \circledast m_2) \circledast (m_1 \circledast m_3).$$

This last property becomes obvious when expressed using weight functions: $w_1(w_2 \wedge w_3) = (w_1 w_2) \wedge (w_1 w_3)$. It is actually quite important, as it explains why the cautious rule can be considered to be more relevant than the TBM conjunctive rule \circledast when combining non distinct items of evidence: if two sources provide BBAs $m_1 \circledast m_2$ and $m_1 \circledast m_3$ having some evidence m_1 in common, the shared evidence is not counted twice.

Other properties of the cautious rule are more difficult to interpret. For instance, we have $m \circledast m_\Omega = m$ for any separable BBA m , but $m \circledast m_\Omega \neq m$ when m is not separable (since we then have $w(A) > 1$ for some A). Consequently, the vacuous BBA m_Ω is not a neutral element for \circledast . In fact, any non dogmatic bba can be decomposed into a confidence component (corresponding to the weights $w(A) \leq 1$) and a diffidence component (corresponding to the weights $w(A) > 1$) [18]. The combination of a BBA m with the vacuous BBA has the effect of erasing the diffidence component. The interpretation of this property is not clear right now. In general, combination with a separable BBA using \circledast transforms any non separable BBA into a separable one.

4 Extensions based on t-norms

As we have seen, the cautious rule is based on the *minimum* of the weights obtained from the canonical decomposition of belief functions, whereas the TBM conjunctive rules is based on the *product* of these weights. In the case where the weights belong to the $[0, 1]$ interval, the minimum and the product may be seen as two particular triangular norms, or t-norms for short (see, e.g., [12] for a recent survey on this topic). We recall that a t-norm is a commutative and associative binary operator \top on the unit interval satisfying the boundary condition

$$x \top 1 = x, \quad \forall x \in [0, 1], \quad (23)$$

and the monotonicity axiom

$$y \leq z \Rightarrow x \top y \leq x \top z, \quad \forall x, y, z \in [0, 1]. \quad (24)$$

Furthermore, a t-norm is said to be positive iff $x, y > 0$ implies $x \top y > 0$. It is known that the minimum is the only idempotent t-norm ($x \top x = x$, for all $x \in [0, 1]$), and it is the largest t-norm, i.e., we have $x \top y \leq x \wedge y$ for all t-norm \top and all $x, y \in [0, 1]$.

By using the concept of t-norm, it can be shown that the cautious and TBM conjunctive rules actually belong to general families of rules sharing interesting properties, including associativity.

4.1 Separable BBAs

If we consider only operations on *separable BBAs* (verifying $w(A) \leq 1$ for all $A \subset \Omega$), then new commutative and associative conjunctive rules may be obtained by replacing the minimum in (21) by any other t-norm \top . The TBM conjunctive rule is recovered when \top is the product. Condition (23) guarantees that the

vacuous BBA is the neutral element, and axiom (24) implies that combining more informative BBAs (according to the \sqsubseteq_w ordering) yields a more informative result, which seems indeed reasonable. Positivity of the t-norm seems also required to avoid generating dogmatic BBAs (with at least one zero weight $w(A)$) when combining non dogmatic ones. The cautious rule is the only one to be idempotent in this class of conjunctive operations, and it is the most conservative (the minimum being the largest t-norm). Of particular interest are parameterized families of t-norms, including the product and the minimum as special cases. For instance, the Dubois-Prade family is defined as

$$x \top_{\gamma}^{DP} y = \frac{xy}{\max(x, y, \gamma)} \quad (25)$$

for $\gamma \in [0, 1]$. The minimum and the product are recovered for $\gamma = 0$ and $\gamma = 1$, respectively. Using such a family of t-norms defines a family of operations on separable BBAs ranging from the cautious rule to the TBM conjunctive rule.

Example 3 Table 5 shows the result of the combinations of the two separable BBAs of Table 4 using the generalized cautious rule, with t-norms in the Dubois-Prade family, for different values of γ . We can see setting parameter γ to a value between 0 and 1 yields a result somewhere between the TBM conjunctive rule and the cautious rule.

Table 4: Two separable BBAs and their weight functions.

A	$w_1(A)$	$m_1(A)$	$w_2(A)$	$m_2(A)$
\emptyset	1	0	1	0.30
$\{a\}$	0.5	0.56	1	0.06
$\{b\}$	1	0	0.5	0.20
$\{a, b\}$	0.8	0.04	0.8	0.04
$\{c\}$	1	0	1	0
$\{a, c\}$	0.4	0.24	0.4	0.24
$\{b, c\}$	1	0	1	0
Ω	-	0.16	-	0.16

Table 5: BBAs obtained by combining the separable BBAs of Table 4 using the generalized cautious rule, with t-norms in the Dubois-Prade family, for different values of γ . The cautious and TBM conjunctive rules are recovered for $\gamma = 0$ and $\gamma = 1$, respectively.

A	$\gamma = 0$	$\gamma = 0.5$	$\gamma = 0.8$	$\gamma = 1$
\emptyset	0.4000	0.4200	0.4500	0.4600
$\{a\}$	0.2800	0.2840	0.2900	0.3256
$\{b\}$	0.1000	0.0800	0.0500	0.0400
$\{a, b\}$	0.0200	0.0160	0.0100	0.0144
$\{c\}$	0	0	0	0
$\{a, c\}$	0.1200	0.1360	0.1600	0.1344
$\{b, c\}$	0	0	0	0
Ω	0.0800	0.0640	0.0400	0.0256

4.2 Non Separable BBAs

Up to now, we have only considered t-norm-based extensions of the cautious rule applicable to the combination of separable belief functions. However, we need rules applicable to a wider class of belief functions, including, at least, all non dogmatic belief functions. In the case of non separable BBAs, we may have $w(A) > 1$ for some $A \subset \Omega$. Consequently, triangular norms, which are defined on $[0, 1]^2$, have to be extended from $[0, 1]^2$ to $[0, +\infty)^2$, while maintaining commutativity, associativity, as well as monotonicity:

$$y \leq z \Rightarrow x \top y \leq x \top z, \forall x, y, z > 0, \quad (26)$$

and positivity:

$$\forall x, y > 0, \quad x \top y = 0 \Rightarrow x = 0 \text{ or } y = 0. \quad (27)$$

Additionally, the following condition is a sufficient condition for the result of the combination to be a belief function:

$$x \top y \leq x \wedge y, \quad \forall x, y > 0. \quad (28)$$

Boundary condition (23) does not seem essential as it is not satisfied by the minimum in $[0, +\infty)$.

Propositions 2 and 3 below indicate two ways of extending t-norms to $[0, +\infty)$ while maintaining these properties. These “extended t-norms” may be used to replace the minimum in (21), thus defining extended cautious rules applicable to arbitrary non dogmatic BBAs.

Proposition 2 Let \top be a positive t-norm, and $\gamma \in (0, 1]$. Then, the operator \top_{γ} defined by

$$x \top_{\gamma} y = \begin{cases} x \top y & \text{if } x, y \leq \gamma \\ x \wedge y & \text{otherwise,} \end{cases} \quad (29)$$

for all $x, y \geq 0$, is commutative, associative, and verifies conditions (26)-(28).

Proof. Commutativity, positivity, and property (28) are satisfied by definition. We only need to check associativity and monotonicity. Concerning associativity, we observe that, for $x, y, z \leq \gamma$, we have $(x \top y) \top z = x \top (y \top z)$ as a consequence of the associativity of \top . If at most one of the values x, y and z is strictly greater than γ , then

$$(x \top_{\gamma} y) \top_{\gamma} z = x \wedge y \wedge z = x \top_{\gamma} (y \top_{\gamma} z).$$

Now, assume that $x, y \leq \gamma$ and $z > \gamma$. We have

$$\begin{aligned} (x \top_{\gamma} y) \top_{\gamma} z &= (x \top_{\gamma} y) \wedge z \\ &= x \top_{\gamma} y \\ &= x \top_{\gamma} (y \wedge z) \\ &= x \top_{\gamma} (y \top_{\gamma} z). \end{aligned}$$

The two remaining cases are completely analogous. For monotonicity, fix $x, y, z \geq 0$ with $y \leq z$. If $x > \gamma$, or if $x \leq \gamma$ and y and z are both either greater, or smaller than γ , then $x \top_{\gamma} y \leq x \top_{\gamma} z$ results from the monotonicity of \top and \wedge . If $x, y \leq \gamma$ and $z > \gamma$, we

have $x \top_{\gamma} y = x \top y$ and $x \top_{\gamma} z = x \wedge z = x = x \top_{\gamma} 1$. Consequently, $x \top_{\gamma} y \leq x \top_{\gamma} z$, as a consequence of the monotonicity of \top . \square

Note that, by setting $\gamma = 1$ and $\top = \text{prod}$, we obtain a rule which coincides with the TBM conjunctive rule when applied to separable BBAs. However, the two rules yield different results for non separable BBAs. It would be interesting to define a parameterized family of rules containing both the cautious and TBM conjunctive rules, but we have not been able to find such a rule up to now.

The following proposition gives yet another method for extending any t-norm to $[0, +\infty) \times [0, +\infty)$ while maintaining the desired properties.

Proposition 3 *Let \top be a positive t-norm, and g a strictly increasing function from $[0, +\infty)$ to $[0, 1]$ such that $g(0)=0$. Then the operator \top_g defined by*

$$x \top_g y = g^{-1} [g(x) \top g(y)], \quad \forall x, y \geq 0 \quad (30)$$

is commutative, associative, and verifies conditions (26)-(28).

Proof. Commutativity is verified by definition. For positivity, we have

$$\begin{aligned} x \top_g y = 0 &\Rightarrow g^{-1} [g(x) \top g(y)] = 0 \\ &\Rightarrow g(x) \top g(y) = g(0) = 0 \\ &\Rightarrow g(x) = 0 \text{ or } g(y) = 0 \\ &\Rightarrow x = 0 \text{ or } y = 0. \end{aligned}$$

For property (28), we observe that $g(x) \top g(y) \leq g(x) \wedge g(y)$, for all $x, y \geq 0$. Consequently

$$\begin{aligned} x \top_g y &= g^{-1} [g(x) \top g(y)] \\ &\leq g^{-1} [g(x) \wedge g(y)] = g^{-1} (g(x)) \wedge g^{-1} (g(y)) = x \wedge y. \end{aligned}$$

For associativity, we have, for all $x, y, z \geq 0$,

$$\begin{aligned} g(x \top_g (y \top_g z)) &= g(x) \top g(y \top_g z) \\ &= g(x) \top (g(y) \top g(z)) \\ &= (g(x) \top g(y)) \top g(z) \\ &= g(x \top_g y) \top g(z) \\ &= g((x \top_g y) \top_g z). \end{aligned}$$

Lastly, we have, for all $x, y, z \geq 0$ such that $y \leq z$,

$$x \top_g y = g^{-1} [g(x) \top g(y)] \leq g^{-1} [g(x) \top g(z)] = x \top_g z,$$

hence the monotonicity of \top_g . \square

When $\top = \wedge$, then $\top_g = \wedge$ whatever the choice of g . Hence, one can define a family of operators on belief functions containing the cautious rule as a particular case by considering a family of t-norms containing the minimum, and a function g .

One may object that these new t-norm based are only weakly justified, apart from the properties outlined above. However, we may remark that the situation is the same as that encountered in Possibility theory [4], 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. In contrast, Dempster-Shafer theory has sometimes been criticized for its lack of flexibility in the choice of combination operator [2], a criticism which, in light of the new results presented in this paper, appears to be unjustified.

It remains to be seen how such t-norm-based operators could be used in practice. For instance, when combining separable belief functions, parameter γ in the Dubois-Prade family (or s in the Frank family) can perhaps be related to some subjective judgement regarding the distinctness of the items of evidence combined, a value between 0 and 1 leading to the choice of a combination operator situated somewhere between the conjunctive and cautious rules. Another possibility is to learn the parameter from data, in the same way as discount rates are learnt in the expert tuning method [7].

5 Conclusion

The cautious rule of combination, a new commutative, associative and idempotent operator for belief functions, has been introduced. The idempotence of this operator (and its normalized version which can easily be defined) makes them useful alternatives to the so far ubiquitous Dempster's rule when the information sources are not independent.

Whereas the behavior of the cautious rule is clear in the case of separable belief function, its properties need to be further studied and interpreted in the more general case of non dogmatic belief function. In particular, the vacuous belief function is no longer a neutral element in that case, a property that may perhaps be better understood in the light of the notions of confidence and diffidence components of a belief function introduced in [18].

The cautious rule has also been shown to be a member of a much wider family of commutative and associative operators based on triangular norms. These findings indicate that, contrary to a so far widely accepted opinion [2], the richness of potential combination operators is not lower in the theory of belief functions than it is in possibility theory.

Acknowledgment

This work is dedicated to the memory of Professor Philippe Smets, who inspired many of the ideas expressed in this paper.

References

- [1] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Math-*

- ematical Statistics*, 38:325–339, 1967.
- [2] D. Dubois and H. Prade. On the unicity of Dempster’s rule of combination. *International Journal of Intelligent Systems*, 1:133–142, 1986.
- [3] D. Dubois and H. Prade. A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets. *International Journal of General Systems*, 12:193–226, 1986.
- [4] D. Dubois and H. Prade. *Possibility Theory: An approach to computerized processing of uncertainty*. Plenum Press, New-York, 1988.
- [5] D. Dubois and H. Prade. Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence*, 4:244–264, 1988.
- [6] D. Dubois, H. Prade, and Ph. Smets. New semantics for quantitative possibility theory. In *2nd International Symposium on Imprecise Probabilities and their Applications*, Ithaca, NY, 2001.
- [7] 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.
- [8] 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.
- [9] 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, *Uncertainty in AI 91*, pages 174–181. Morgan Kaufman, San Mateo, CA, 1991.
- [10] F. Klawonn and E. Schwecke. On the axiomatic justification of dempster’s rule of combination. *International Journal of Intelligent Systems*, 7:469–478, 1992.
- [11] 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.
- [12] E. P. Klement, R. Mesiar, and E. Pap. *Triangular norms*. Kluwer Academic Publishers, Dordrecht, 2000.
- [13] E. Lefèvre, O. Colot, and P. Vannoorenberghe. Belief function combination and conflict management. *Information Fusion*, 3(2):149–162, 2002.
- [14] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, Princeton, N.J., 1976.
- [15] Ph. Smets. Combining non distinct evidences. In *Proc. of NAFIPS’86*, pages 544–549, New Orleans, USA, 1986.
- [16] Ph. Smets. The combination of evidence in the Transferable Belief Model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5):447–458, 1990.
- [17] Ph. Smets. Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem. *International Journal of Approximate Reasoning*, 9:1–35, 1993.
- [18] 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.
- [19] 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.
- [20] Ph. Smets. The application of the matrix calculus to belief functions. *International Journal of Approximate Reasoning*, 31(1–2):1–30, 2002.
- [21] Ph. Smets. Personal communication, 2005.
- [22] Ph. Smets and R. Kennes. The Transferable Belief Model. *Artificial Intelligence*, 66:191–243, 1994.
- [23] R. R. Yager. The entailment principle for Dempster-Shafer granules. *Int. J. of Intelligent Systems*, 1:247–262, 1986.
- [24] R. R. Yager. On the Dempster-Shafer framework and new combination rules. *Information Sciences*, 41:93–137, 1987.