Toward a General Framework for Information Fusion

Didier Dubois[1,2], Weiru Liu[2], Jianbing Ma[3], and Henri Prade[1]

[1] IRIT, CNRS & Université de Toulouse, France
[2] School of Electronics, Electrical Engineering and Computer Science, Queens University, Belfast, UK
[3] School of Design, Engineering and Computing, Bournemouth University, Bournemouth, UK

Abstract. Depending on the representation setting, different combination rules have been proposed for fusing information from distinct sources. Moreover in each setting, different sets of axioms that combination rules should satisfy have been advocated, thus justifying the existence of alternative rules (usually motivated by situations where the behavior of other rules was found unsatisfactory). These sets of axioms are usually purely considered in their own settings, without in-depth analysis of common properties essential for all the settings. This paper introduces core properties that, once properly instantiated, are meaningful in different representation settings ranging from logic to imprecise probabilities. The following representation settings are especially considered: classical set representation, possibility theory, and evidence theory, the latter encompassing the two other ones as special cases. This unified discussion of combination rules across different settings is expected to provide a fresh look on some old but basic issues in information fusion.

1 Introduction

In information fusion, each piece of information is assumed to come from a different source (measurement device or expert opinion) and the fusion is a process aiming at grasping what is known about a situation being observed. This contrasts with preference aggregation where preferences merely reflect what some agent would like the result to be, and the aggregation process is more about building compromises than finding what the true state of a situation is. The pieces of information to be fused may be inconsistent, and are often pervaded with uncertainty, which must be reflected on the result.

The information fusion problem is met in different representation settings, ranging from the merging of logical knowledge/belief bases supposed to encode the states of mind of agents about the perception of a situation ([11] in classical logic, [1] in possibilistic logic for the merging of stratified or prioritized bases), to numerical-based frameworks, such as, probability theory [20], evidence theory [17], possibility theory [10], or imprecise probability theory [22]. It is worth-noticing that all the above-mentioned settings can handle epistemic uncertainty and incomplete knowledge with the exception of probability theory that often accounts with variability and randomness, while the Bayesian approach to subjective probability yields a questionable representation of incomplete information [5]. In that respect, it is important to keep in mind the fact that, formally speaking, evidence theory encompasses both probability theory and

V. Torra et al. (Eds.): MDAI 2013, LNAI 8234, pp. 37–48, 2013.

possibility theory as particular cases; in turn, evidence theory can be seen as a particular imprecise probability system [23]; and binary-valued possibility theory is nothing but a Boolean representation for imprecise pieces of information at work in propositional or epistemic logic.

It is striking to observe that the information fusion problem until now has been discussed independently in each setting. Sometimes, specific postulates that govern fusion operations are provided [21,11,14]. Moreover in each setting, various combination rules have been advocated as behaving properly (on the basis of good properties) as opposed to the unsatisfactory behavior of other rules. In practice, we are faced with many combination rules (their number is still increasing!), and several postulate systems. It is worthwhile to provide a more unified view of the problem.

In this paper, we aim to propose common properties of fusion operators valid in any setting. They do account for various existing axiomatic systems proposed in specific settings. These properties are stated at the semantic level, rather than at the syntactic one (unlike [11]), since probabilistic settings do not have a well-established logical counterpart. Moreover, the semantical level is especially appropriate for laying bare the practical meaning of the combination rules. This provides a common ground for a rational exploration of fusion methods, despite the heterogeneity of existing frameworks. Particular instantiations of these common properties in the different settings are then considered.

The rest of the paper is organized as follows. The next section introduces eight core properties, before considering their instantiations, in Section 3 in the classical set representation and in the possibility theory setting, and in Section 4 in the context of evidence theory, in which many different combination rules have been proposed. These properties provide a basis for comparing these alternative rules.

2 Core Properties

In order to define a set of required properties that make sense in different settings ranging from logic to imprecise probability, we consider an abstract notion of information item, denoted by T, supplied by sources. Let $\Omega = \{\omega_1, ..., \omega_{|\Omega|}\}$ be a finite, non-empty set of possible worlds (e.g. the range of some unknown quantity), one of which is the true one. There are n experts/sources and the i_{th} expert/source is denoted by i. Let T_i be the information provided by i, e.g., T_i may be a basic belief assignment, a possibility distribution, or a knowledge base. $T = f(T_1, \ldots, T_n)$ denotes the fusion result using aggregation operator f over a set of information items T_i. To any information item, we associate the following features:

- The subset $\mathcal{S}(T) \subseteq \Omega$, called the *support* of T, contains the set of values considered possible by information T. It means that $\omega_i \notin \mathcal{S}(T) \iff \omega_i$ is impossible.
- Its *core* $\mathcal{C}(T) \subseteq \Omega$ contains the set of values considered fully plausible according to information T. The idea is that, by default, if information T is taken for granted, a first guess for the value of x should be an element of $\mathcal{C}(T)$. Clearly, $\mathcal{C}(T) \subseteq \mathcal{S}(T)$.
- *Internal Consistency* An information item T is said to be weakly (resp. strongly) consistent if $\mathcal{S}(T) \neq \emptyset$ (resp. $\mathcal{C}(T) \neq \emptyset$) otherwise information T is totally (resp. weakly) inconsistent. In the following, we assume $\mathcal{C}(T) \neq \emptyset$ for each source.

Strong consistency is assumed for inputs of a merging process, and weak consistency at worst for the output.

- *Mutual consistency* T and T' are said to be weakly mutually consistent when $\mathcal{S}(T) \cap \mathcal{S}(T)' \neq \emptyset$ and strongly so when $\mathcal{C}(T) \cap \mathcal{C}(T)' \neq \emptyset$.
- *Information ordering*: $T \sqsubseteq T'$ expresses that T provides at least as much information as T'. In particular, $T \sqsubseteq T'$ should imply $\mathcal{S}(T) \subseteq \mathcal{S}(T')$.
- *Plausibility ordering*: If consistent, information T induces a partial preorder $\succeq_T$ expressing relative plausibility: $\omega \succeq_T \omega'$ means that ω is at least as plausible as (or *dominates*) ω' according to T. We write $\omega \sim_T \omega'$ if $\omega \succeq_T \omega'$ and $\omega' \succeq_T \omega$. Of course, if $\omega \in \mathcal{S}(T), \omega' \notin \mathcal{S}(T)$, then $\omega \succ_T \omega'$ (ω is strictly more plausible than ω').

The *vacuous information*, expressing total ignorance is denoted by $T^\top$. Then the plausibility ordering is flat: $\mathcal{S}(T^\top) = \mathcal{C}(T^\top) = \Omega$ and $\omega \sim_{T^\top} \omega' \; \forall \omega, \omega' \in \Omega$.

The process of merging information items, supplied by sources whose reliability levels are not known, is guided by a few first principles (already in [21]):

- It is a basically symmetric process as the sources play the same role and supply information of the same kind;
- We try to use as many information items as possible in the fusion process, so as to get a result that is as precise and useful as possible. However, the result should not be arbitrarily precise, but faithful to the level of informativeness of the inputs.
- Information fusion should try to solve conflicts between sources, while neither dismissing nor favoring any of them without a reason.

These principles are implemented in the postulates listed below, called *core properties*, which are meant to be natural minimal requirements, independent of the actual representation framework.

Property 1: Unanimity.

When all sources agree on some results, then the latter should be preserved. Minimal conditions are

(c) Possibility preservation. If for all sources ω is possible, then so should the fusion result assert: if $\forall i, \omega \in \mathcal{S}(T_i)$ then $\omega \in \mathcal{S}(f(T_1, ..., T_n))$.

(d) Impossibility preservation. If all sources believe that a possible world ω is impossible, then this ω cannot become (even slightly) possible after fusion. This can be expressed as $\mathcal{S}(f(T_1, ...T_n)) \subseteq \mathcal{S}(T_1) \cup ... \cup \mathcal{S}(T_n)$.

Property 2: Informational Monotonicity.

If a set of agents provides less information than another set of *non-disagreeing* agents, then fusing the former inputs should not produce a more informative result than fusing the latter. The weakest such requirement is:

Weak Informational Monotonicity. if $\forall i, T_i \sqsubseteq T_i'$, then $f(T_1, ...T_n) \sqsubseteq f(T_1', ...T_n')$, provided that all the inputs are globally strongly mutually consistent.

Property 3: Consistency Enforcement.

This property requires that fusing individually consistent inputs should give a consistent result. At best: $\mathcal{C}(f(T_1, ...T_n)) \neq \emptyset$. At least: $\mathcal{S}(f(T_1, ...T_n)) \neq \emptyset$.

Property 4: Optimism.
In the absence of specific information about source reliability, one should assume as many sources as possible are reliable, in agreement with their observed mutual consistency. In particular: If $\cap_{i=1}^{n}\mathcal{C}(T_i) \neq \emptyset$ then $f(T_1, ..., T_n) \sqsubseteq T_i, \forall i = 1,...,n$. In general, it should be assumed that at least one source is reliable.

Property 5: Fairness. The fusion result reconcile all sources. Hence, the result of the fusion process should keep something from each input, i.e.,
$\forall i = 1,...,n, \mathcal{S}(f(T_1,...,T_n)) \cap \mathcal{S}(T_i) \neq \emptyset$.

Property 6: Insensitivity to Vacuous Information.
Sources that provide vacuous information should not affect the fusion result:
$f_n(T_1, \ldots, T_{i-1}, T^{\top}, T_{i+1}, \ldots, T_n) = f_{n-1}(T_1, \ldots, T_{i-1}, T_{i+1}, \ldots, T_n)$

Property 7: Commutativity.
Inputs from multiple sources are treated on a par, and the combination should be symmetric (up to their relative reliability).

Property 8. Minimal Commitment.
The result of the fusion should be as little informative as possible (in the sense of $\sqsubseteq$) among possible results that satisfy the other core properties.

Some comments are in order. The general core properties proposed here have counterparts in properties considered in different particular settings ; see especially [21] and also [11]. Let us further discuss each of them.

Possibility and impossibility preservation can be found in possibility theory [14] and imprecise probability [21]. It makes sense to request more than possibility preservation: plausibility preservation, replacing supports by cores [21]. The strongest form of Unanimity (Prop. 1) is *idempotence*: if $\forall i, T_i = T$, $f(T_1, ..., T_n) = T$. However, while it makes sense if sources are redundant, adopting it in all situations forbids reinforcement effects to take place when sources are independent [9]. Our Unanimity properties minimally respect the agreement between sources. A slightly more demanding requirement which leaves room for reinforcement effects can be: *Local Ordinal Unanimity*: $\forall \omega$ and ω', if ω is at least as plausible as ω', then so should it be after fusion. e.g., ω dominates ω'. Formally: if $\forall i, \omega \succeq_{T_i} \omega'$, then $\omega \succeq_{f(T_1,...,T_n)} \omega'$.

Informational Monotonicity (Prop.2), adopted as a general property in [14] should be restricted to when information items supplied by sources do not contradict each other. Indeed, if conflicting, it is always possible to make these information items less informative in such a way that they become consistent. In that case the result of the fusion may become very precise by virtue of Optimism Prop. 4, and in particular, more informative than the union of the supports of original precise conflicting items of information.

Consistency enforcement (Prop. 3) is instrumental if the result of the merging is to be useful in practice: one must extract something non-trivial, even if tentative, from available information. It is a typical requirement from the logical area [11] and a property taken for granted by numerical approaches (viz. Dempster rule of combination, but also for imprecise probabilities [21]). Still, when the representation setting is refined enough, there are gradations in consistency requirements, and Prop. 4 can be interpreted in a flexible way. For example, the re-normalisation of belief functions or possibility distributions obtained by merging is not always compulsory, even if sub-normalisation expresses a form of inconsistency.

Optimism (Prop. 4) underlies the idea of making the best of the available information: If items of information are globally consistent with each other, there is no reason to question the reliability of the sources. It is again a typical assumption in logical settings [11], but Walley [21] tries to formulate a similar property. In case of strong inconsistency, this assumption is not sustainable. Note that in the latter case (in particular if $\cap_{i=1}^{n}\mathcal{S}(T_i) = \emptyset$), and under the Impossibility Preservation property (1d), the support of the result should be at worst the union of the supports of inputs, i.e., $\mathcal{S}(f(T_1,...T_n)) \subseteq \mathcal{S}(T_1) \cup ... \cup \mathcal{S}(T_n)$, now assuming that at least one source is reliable (still a form of optimism in the presence of inconsistency). The latter requirement sounds natural for two sources only, but may be found overcautious for many sources. In particular, Optimism will lead to replace any group K of strongly consistent sources, by a single source that is more informative than and in agreement with each of them.

Fairness (Prop. 5) ensures that all input items participate to the result. At the same time, it favors no source by forbidding any input to be derived from the output result in the case of inconsistency. Note that different versions of the Fairness property can be found in the literature. In particular, a form of this property was already suggested by Walley [21] for imprecise probabilities. In the logical setting [11], the counterpart of the condition $\mathcal{S}(f(T_1,...,T_n)) \cap \mathcal{S}(T_i) \neq \emptyset$ is required to hold either for each i, or for none. The possibility that it holds for none sounds highly debatable using supports, from a knowledge fusion point of view, while it may be acceptable when fusing preferences, which is more a matter of trade-off, or when supports are changed into cores.

Insensitivity to Vacuous Information (Prop. 6) looks obvious, not to say redundant, but dispensing with it may lead to uninformative results. It appears again in the Walley postulates [21] for merging sets of probabilities. Prop. 6 implicitly admits that a non informative source is assimilated to one that does not express any opinion, and is typical of information fusion. It excludes probabilistic fusion rules like averaging, since it is sensitive to vacuous information (represented, e.g., by a uniform distribution).

Commutativity (Prop. 7) is characteristic of fusion processes as opposed to revision where prior knowledge may be altered by input information. In contrast, information fusion deals with inputs received in parallel. So, commutativity makes sense, if no information is available on the reliability of sources.

Minimal Commitment is a very important postulate that applies in many circumstances. It is central in all uncertainty theories handling incomplete information under different terminologies, including in logic-based approaches (where it is implicit). It considers as possible any state of affairs not explicitly discarded. It is called *principle of minimal specificity* in possibility theory [10], *principle of Minimal Commitment* in evidence theory [19], and it underlies the so-called *natural extension* in imprecise probability theory [22]. This is a cautious principle that is nicely counterbalanced by the Optimism postulate, and this equilibrium is sometimes useful to characterise the unicity of fusion rules: Optimism provides an upper limit to the set of possible worlds and Minimal Commitment a lower limit.

Some other properties may be required in aggregation processes, such as associativity, which makes computation more efficient, but lacking associativity is not a fatal flaw in itself (e.g., the MCS rule below), if the rule can be defined for n sources.

3 Merging Set-Valued Information: Hard Constraints

The most elementary setting one may first consider is the one of sets, whereby any information item is a subset of possible worlds, which restricts the unknown location of the true state, the simplest account of an epistemic state. Let us assume that the information items T_i are classical subsets. Then $\mathcal{S}(T_i) = T_i$, the relation $\sqsubseteq$ is set inclusion, and $\omega \succ_T \omega'$ if $\omega \in T$ and $\omega' \notin T$, while $\omega \sim_T \omega'$ if $\omega, \omega' \in T$ or $\omega, \omega' \notin T$.

If the inputs are globally consistent, i.e., if $\cap_{i=1,n} T_i \neq \emptyset$, one should have the inclusion $f(T_1,...,T_n) \subseteq \cap_{i=1}^n T_i$ by Prop. 5 (Optimism). By Possibility preservation (1c), $\cap_{i=1}^n T_i \subseteq f(T_1,...,T_n)$. Thus, $f(T_1,...,T_n) = \cap_{i=1}^n T_i$ in case of global consistency. Let us now consider the case of two inconsistent pieces of information T_1 and T_2 such that $T_1 \cap T_2 = \emptyset$. By Prop. 6 (Fairness), one should have $f(T_1,T_2) \cap T_1 \neq \emptyset$ and $f(T_1,T_2) \cap T_2 \neq \emptyset$. Moreover by Impossibility preservation (1d), one should have $f(T_1,T_2) \subseteq T_1 \cup T_2$. This leads to $f(T_1,T_2) = A_1 \cup A_2$ with $\emptyset \neq A_x \subseteq T_x$ for $x = 1, 2$. Minimal Commitment leads us to take $A_x = T_x$ for $x = 1, 2$.

This reasoning clearly extends to the case of more than two pairwise inconsistent information pieces: by Fairness, $f(T_1,...,T_n)$ should be of the form $A_1 \cup ... \cup A_n$, $\emptyset \neq A_i \subseteq T_i$ for $i = 1, \ldots, n$. Let $I \subset \{1,..,n\}$ be a maximal consistent subset (MCS) of sources, i.e., $T^I = \cap_{i\in I} T_i \neq \emptyset$ and $T^I \cap T_j = \emptyset$ if $j \notin I$. Then the partial result should be $A_j = \cap_{i\in I} T_i, \forall j \in I$ by Minimal Commitment and Optimism. Given two MCSs I and I', $T^I \cap T^{I'} = \emptyset$. Hence at most one subset I of sources is correct. Optimism dictates that at least one subset I of sources is so. We thus get the general combination rule

$$f(T_1,...,T_n) = \bigcup_{I \in MCS(\{1,\ldots,n\})} \bigcap_{i \in I} T_i \qquad (1)$$

where MCS($\{1, \ldots, n\}$) is the set of maximal consistent subsets of sources. It was first proposed by [15]. It satisfies all core properties.

This rule exhibits an apparent discontinuity when moving from a consistent situation to an inconsistent one, since shrinking two subsets that overlaps may lead from situations with more and more precise fusion results to a situation with an imprecise result. However, nothing forbids independent sources to provide information pieces having a narrow intersection. But such a precise result may sometimes become all the more debatable as its precision increases. Some approaches cope with inconsistency in fusion problems by a similarity-based enlargement of the supports and cores of information pieces [16]).

4 Possibility Theory

The possibility theory framework is a graded extension of the previous setting. Subsets are replaced by possibility distributions π, which are mappings from Ω to $[0, 1]$ that rank-order interpretations ($\omega \succeq_T \omega'$ if $\pi(\omega) \geq \pi(\omega')$). The support is $\mathcal{S}(\pi) = \{\omega | \pi(\omega) > 0\}$ and the core is $\mathcal{C}(\pi) = \{\omega | \pi(\omega) = 1\}$. A strongly consistent possibility distribution is such that $\mathcal{C}(\pi) \neq \emptyset$. The consistency degree $Cns(\pi_i, \pi_j) = \max_\omega \min(\pi_i(\omega), \pi_j(\omega))$ between two distributions ranges from 1 when there is a common ω that is fully possible, to 0 when the supports do not overlap. The information ordering is relative specificity ($\pi_i \sqsubseteq \pi_j \iff \pi_i \leq \pi_j$).

The most basic combination rules extend conjunction and disjunction, especially the Minimum rule $\min(\pi_1, ..., \pi_n)$ and the Maximum rule $\max(\pi_1, ..., \pi_n)$; other conjunctions can be t-norms t such as product instead of $\min$, which creates a reinforcement effect. The conjunctive rules do not obey the strong form of consistency enforcement.

The latter property justifies the renormalized conjunctive fusion rule (RCF) [8]

$$\hat{\bigwedge}(\pi_1, ..., \pi_n) = \frac{\bigwedge(\pi_1, ..., \pi_n)}{Cns(\pi_1, ..., \pi_n).} \tag{2}$$

It is undefined as soon as $Cns(\pi_1, ..., \pi_n) = 0$ (strong conflict). When $\bigwedge$ is product, this rule is well-known and is associative, but associativity is generally not preserved with other t-norms. This kind of fusion rule is used in logic-based merging using distances [11] instead of possibility distributions (see [1] for the connection between the two approaches). However this kind of rule cannot cope with strongly mutually inconsistent sources. We can extend the MCS rule in at least two ways:

$$MCS1(\pi_1, ..., \pi_n) = \max_{I \in MCS(\{\mathcal{C}(\pi_1), ..., \mathcal{C}(\pi_n)\})} \bigwedge_{i \in I} \pi_i \tag{3}$$

$$MCS0(\pi_1, ..., \pi_n) = \max_{I \in MCS(\{\mathcal{S}(\pi_1), ..., \mathcal{S}(\pi_n)\})} \hat{\bigwedge}_{i \in I} \pi_i \tag{4}$$

In fact, each of MCS1, MCS0 selects maximal consistent subsets in a specific way. Once, this principle chosen, the same reasoning holds as in the crisp case, and we obtain for the above three rules for merging possibility distributions:

Proposition 1. *The RCF rule (2) does not satisfy Consistency Enforcement nor Fairness (when undefined). The extended-MCS rules (3,4) satisfy all core properties.*

MCS1 is much demanding on mutual consistency of sources and yields plain disjunction if cores of π_i are disjoint. MCS0 is less demanding and more optimistic: it yields $\hat{\bigwedge}(\pi_1, ..., \pi_n)$ if all supports overlap.

Another fusion rule for possibility distribution that applies the classical MCS rule to all cuts of the input possibility distributions has been recently proposed [4]. It satisfies all basic postulates but it yields a belief function, as resulting cuts are no longer nested.

5 Evidence Theory

In evidence theory, a piece of information is modeled by a *basic belief assignment (bba)* m which is a mapping from 2^{Ω} to $[0, 1]$ such that $\sum_{A \subseteq \Omega} m(A) = 1$. A bba is consistent if $m(\emptyset) \neq 0$. A is called a focal element of m if $m(A) > 0$. Let $\mathcal{F}_m$ be the set of focal elements of m. Let $\mathcal{S}(m)$ denote the union of the focal elements: if $\mathcal{F}_m = \{A_1, .., A_n\}$, then $\mathcal{S}(m) = \bigcup_{i=1}^{n} A_i$ is the support of m. The *vacuous* bba m^{Ω} is such that $m(\Omega) = 1$.

From a bba m, two dual functions, bel and pl called *belief and plausibility functions* respectively, are defined as $bel(A) = \sum_{B \subseteq A} m(B)$, and $pl(A) = \sum_{B \cap A \neq \emptyset} m(B)$, while the commonality function q is defined by $q(A) = \sum_{A \subseteq B} m(B)$.

Evidence theory is rich enough to include as particular cases i) sets (when there is one focal element), ii) probabilities (when focal elements are singletons), and iii) possibility theory (when focal elements are nested). The *contour* function C_m of the bba m, which is the plausibility function of the singletons, $C_m(\omega) = \sum_{A \subseteq \Omega, \omega \in A} m(A)$, reduces to a possibility distribution $\pi_m = C_m$ when focal elements are nested, and then $pl(A) = \max_{\omega \in A} \pi_m(\omega)$ is a possibility measure. The contour function reduces to a probability distributions if the focal elements are singletons.

We now examine issues related to plausibility and information ordering, and inconsistency between bbas.

Plausibility Ordering. In evidence theory, from a representation point of view the contour function is a natural option for comparing possible worlds ($\omega_1 \succeq_m^{con} \omega_2$ iff $C_m(\omega_1) \geq C_m(\omega_2)$). In addition to this standard ordering, we define a more basic partial ordering relation on possible worlds induced by the bba.

Definition 1. *Let $\omega_1, \omega_2 \in \Omega$. Then ω_1 dominates ω_2 w.r.t. m, denoted by $w_1 \succeq_m^{dom} \omega_2$ iff for any $A \subseteq \Omega \setminus \{\omega_1, \omega_2\}$, $m(A \cup \{\omega_1\}) \geq m(A \cup \{\omega_2\})$.*

Proposition 2. *$\succeq_m^{dom}$ is a reflexive and transitive relation. Moreover $\omega_1 \succeq_m^{dom} \omega_1$ implies $C_m(\omega_1) \geq C_m(\omega_2)$.*

Inconsistency. The degree of inconsistency (or conflict) of two bbas m_1 and m_2 is measured by the mass received by the empty set as the result of the conjunction of m_1 and m_2 viewed as random sets: $m_{1\wedge 2}(\emptyset) = \sum_{A \cap B = \emptyset} m_1(A) m_2(B)$. It is the counterpart of $1 - Cns(\pi_1, \pi_2)$ using product instead of min. However, it has been pointed out in [12] that $m_{1\wedge 2}(\emptyset)$ is not a convincing measure of conflict, since two identical bba's usually have a non zero degree of conflict. To get a more satisfactory measure of conflict one may avoid using products $m_1(A)m_2(B)$ that presuppose source independence, and replace them by a joint mass $x(A, B)$ whose marginals are m_1 and m_2[6]. Then we can define a better inconsistency index, such that $Inc(m, m) = 0$:

$$Inc(m_1, m_2) = \inf_x \sum_{B \in \mathcal{F}_1, C \in \mathcal{F}_2 : B \cap C = \emptyset} x(B, C)$$

Note that $Inc(m_1, m_2) = 0$ whenever there exists a joint mass $x(A, B)$ whose marginals are m_1 and m_2 that assigns zero mass to all disjoint focal sets, which corresponds to saying that the two credal sets (families of probabilites) $\{P : P(A) \geq Bel_1(A), \forall A\}$ and $\{P : P(A) \geq Bel_2(A), \forall A\}$ have a non-empty intersection [2]. So we can call this index one of probabilistic consistency. Its calculation requires the use of linear programming. It is easy to see that $Inc(m_1, m_2) \leq m_{1\wedge 2}(\emptyset)$.

Alternatively we can adopt definitions that do not rely on numerical values of bba's: two mass functions m and m' with focal sets $\mathcal{F}$ and $\mathcal{F}'$ are said to be

– *Weakly mutually consistent* if $\exists E \in \mathcal{F}, E' \in \mathcal{F}' : E \cap E' \neq \emptyset$ (note that it implies that $m_{1\wedge 2}(\emptyset) < 1$, hence $Inc(m_1, m_2) < 1$ as well)
– *Strongly (or logically [3]) mutually consistent* if $\forall E \in \mathcal{F}, \forall E' \in \mathcal{F}' : E \cap E' \neq \emptyset$ (note that it does imply that $Inc(m_1, m_2) = m_{1\wedge 2}(\emptyset) = 0$).

Information Ordering. In the literature, different information orderings in evidence theory have been proposed for comparing the information contents of bba's (see e.g. [7]). We here only consider the one that can be expressed in terms of mass functions, and echoes the above inconsistency index. It is the strongest information ordering among those previously introduced in the literature.

Definition 2 (Specialization). *Let m_1 and m_2 be two bbas over Ω, m_1 is a specialization of m_2 (denoted by $m_1 \sqsubseteq_s m_2$) if and only if there exists a joint mass $x(A,B)$ whose marginals are m_1 and m_2 , such that $x(A,B) = 0$ whenever $A \not\subseteq B, A \in \mathcal{F}_1, B \in \mathcal{F}_2$.*

We are in a position to propose one possible instantiation of the basic fusion postulates, for two sources here, denoting by m_{12} the result:

1. **Unanimity** Possibility and impossibility preservation.
2. **Weak Information Monotonicity**: If m_1 and m_2 are strongly consistent, and moreover $m_1 \sqsubseteq_s m'_1$, $m_2 \sqsubseteq_s m'_2$ then $m_{12} \sqsubseteq_s m'_{12}$
3. **Consistency enforcement**:
   - $\sum_{E \subseteq S} m_{12}(E) = 1$ (strong version)
   - $\sum_{E \subseteq S} m_{12}(E) > 0$ (weak version)
4. **Optimism**
   - If m_1 and m_2 are strongly mutually consistent, then $m_{12} \sqsubseteq_s m_i, i = 1, 2$.
   - There exists a joint bba $x(\cdot,\cdot)$ whose marginals are m_1 and m_2, such that $m_{12} \sqsubseteq_s m_1 \oplus m_2$, with $m_1 \oplus m_2(E) = \sum_{F,G:E=F \cup G} x(F,G)$.
5. **Fairness**: Each m_i should be weakly consistent with m_{12}.
6. **Insensitivity to Vacuous Information**: If $m_1(\Omega) = 1$ then $m_{12} = m_2$
7. **Symmetry**: $m_{12} = m_{21}$
8. **Minimal Commitment**: m_{12} should be minimally specific for specialization.

5.1 Checking Some Existing Combination Rules

Several rules have been proposed in evidence theory for merging information, apart from the well-known Dempster's rule of combination. We first focus on the main rules.

$$m_{Dem}(C) = \frac{\sum_{A,B:A \cap B = C} m_1(A)m_2(B)}{1 - \sum_{A,B:A \cap B = \emptyset} m_1(A)m_2(B)} \text{ (Dempster's rule)} \quad (5)$$

$$m_{Sm}(C) = \sum_{A,B \subseteq \Omega, A \cap B = C} m_1(A)m_2(B) \text{ (Smets' rule)} \quad (6)$$

$$m_{Ya}(C) = \begin{cases} \sum_{A,B:A \cap B = C} m_1(A)m_2(B) \text{ if } C \neq \Omega \text{ (Yager's rule)} \\ m_1(\Omega)m_2(\Omega) + \sum_{A \cap B = \emptyset} m_1(A)m_2(B) \text{ if } C = \Omega \end{cases} \quad (7)$$

$$m_{DP}(C) = \sum_{A,B:A \cap B = C} m_1(A)m_2(B) + \sum_{A,B:A \cup B = C, A \cap B = \emptyset} m_1(A)m_2(B).[8] \quad (8)$$

All four fusion rules presuppose independence between sources, as an additional assumption, which enforces the choice of $x(\cdot,\cdot) = m_1(\cdot) \cdot m_2(\cdot)$. It reduces the scope of the Minimal Commitment axiom to the choice of a set-theoretic combination for focal sets. The main difference between Dempster's rule and the three other rules respectively proposed in [19] (see also [18]) [24], [8] concern the way the mass $(m_1 \otimes m_2)(\emptyset) = \sum_{A,B:A\cap B=\emptyset} m_1(A)m_2(B)$ is re-allocated. In Dempster's rule, the renormalization by division enforces strong consistency of the result, when the two bba's are weakly mutually consistent (otherwise the operation is not defined). Smets's rule simply keeps this mass on $\emptyset$, whilst Yager's rule assigns it to Ω.

All four fusion rules coincide with each other if the two bba's are strongly consistent. Then all postulates are satisfied. When $\sum_{A\cap B=\emptyset} m_1(A)m_2(B) = 1$, m_{Dem} is not defined due to a total conflict between the sources, which violates the Consistency Enforcement postulate, like for the normalized conjunctive rule of possibility theory. When the two bba's are weakly mutually consistent, the result is consistent since $m_{Dem}(\emptyset) = 0$. Dempster's rule of combination is over-optimistic in case of weak consistency; it may fail to satisfy the second Optimism condition, due to renormalization (it would satisfy it if we replace it by the weaker condition $\mathcal{S}(m_{12}) \subseteq \mathcal{S}(m_1) \cup \mathcal{S}(m_2)$).

In Smets rule, the mass assigned to the empty set $m_S(\emptyset)$ may be different from 0. Smets rule does not respect the consistency enforcement principle, even if it is always defined, since it may deliver the plain empty set in case m_1 and m_2 are strongly inconsistent. Like Dempster rule of combination, Smets' rule is purely conjunctive, hence does not behave in agreement with the postulates in case of partial mutual inconsistency. The Fairness axiom formally fails with this fusion rule like for Dempster's, because it is not compatible with the failure of the consistency enforcement postulate.

Yager's rule is similar to Smets' rule except that $(m_1 \otimes m_2)(\emptyset)$ is added to $m_{Ya}(\Omega)$ instead of leaving it in $m_{Ya}(\emptyset)$, just changing conflict into ignorance (a form of renormalization). It does not respect Unanimity, nor Optimism and in particular impossibility preservation is clearly violated. In fact, this rule is far too cautious in the presence of conflicts.

Three of the four above rules are conjunctive, while the last one, proposed in [8] extends the basic fusion rule (1) for sets to belief functions (hence it is a special case of the MSC rule). It is a hybrid rule, like Yager's, that contains both conjunctive and disjunctive elements. It is more informative than Yager's. This fusion rule satisfies all fusion postulates like the MCS fusion rule for two sets, which it generalizes.

Dempster rule and Smets rule are associative, while the others are not. However, Dubois and Prade combination rule can be readily extended to $n > 2$ sources using the MCS rule on all n-tuples of focal sets.

We may complement Unanimity with Local Ordinal Unanimity with respect to dominance ordering: for two possible worlds ω and ω', $\omega \succeq_1^{dom} \omega'$ and $\omega \succeq_2^{dom} \omega'$ then $\omega \succeq_{12}^{dom} \omega'$. Indeed, we can prove the following:

Proposition 3. *Smets, Yager and Dempster combination rules obey Local Ordinal Unanimity with respect to the dominance ordering.*

It is still unclear whether this result holds for the 4th fusion rule. The above results are summarized by the following Table 5.1 (all above rules are symmetric).

rule/Prop	Una	Mono	Cons	Opti	Fair	Vacuous	Min-Com
Dempster	Yes^1	Yes	Strong^1	No^3	Yes^1	Yes	No^3
Smets	Yes^2	Yes	No	Yes	No	Yes	Yes^1
Yager	No	Yes	Strong	No	Yes	Yes	Yes
DP	Yes	Yes	Strong	Yes	Yes	Yes	Yes

- 1. Only when there is no strong global inconsistency
- 2. Trivially in case of strong global inconsistency
- 3. Overoptimistic in case of weak inconsistency

All the fusion rules considered above assume source independence but can be extended by replacing the product of bba's $m_1(F)m_2(G)$ by a suitably chosen joint mass function $x(E, F)$ whose marginals are m_1 and m_2 [3]. The main difference is that we can replace strong consistency by probabilistic consistency, that is all four fusion rules would coincide with $m_{12}(E) = \sum_{E=F\cap G} x(E, F)$ if m_1 and m_2 are mutually consistent in the sense that $Inc(m_1, m_2) = 0$. However there may be several minimally specific fusion rules, some of which are idempotent [3], if we leave the choice of $x(E, F)$ open.

6 Concluding Remarks

In this paper, we have provided a general framework for analyzing fusion operators proposed in different settings, in a unified way. Due to space limitation, we have concentrated the presentation on three types of representation using classical sets, possibility theory, and evidence theory respectively, considering only a representative sampling of operators. It is clear that the analysis may be applied more systematically, as well as to other settings, whether numerical (such as imprecise probabilities [22,21]), ordinal[13] or yet logical [11]. The latter case comes down to viewing the set of models of a knowledge base K as the core $\mathcal{C}(T_K)$ of the corresponding information item T_K. We did not discuss the case of single probability distributions as they only support weighted arithmetic means [20], which violates Insensitivity to Vacuous Information (assuming the latter is expressed by uniform probability distributions). When distinct, they always conflict, but taking their convex hull satisfies all postulates [21]. Beyond our core properties, that are usually completely intuitive, and should be satisfied by any reasonable fusion rule, other less universal properties, may make sense in specific contexts. For instance a discontinuous fusion rule in a continuous setting is questionable (e.g. Dempster's rule is oversensitive to small changes of input values). Some properties are useful in some situations but not possessed by many rules (e.g., idempotency when sources are redundant). Moreover, when the representation setting becomes richer, more options are available for expressing the properties with various strengths. Adapting the basic postulates to prioritized merging is another line for further work.

References

1. Benferhat, S., Dubois, D., Kaci, S., Prade, H.: Possibilistic merging and distance-based fusion of propositional information. Annals of Mathematics and Artificial Intelligence 34(1-3), 217–252 (2002)

2. Chateauneuf, A.: Combination of compatible belief functions and relation of specificity. In: Advances in the Dempster-Shafer Theory of Evidence, pp. 97–114. John Wiley & Sons, Inc., New York (1994)
3. Destercke, S., Dubois, D.: Idempotent conjunctive combination of belief functions: Extending the minimum rule of possibility theory. Information Sciences 181(18), 3925–3945 (2011)
4. Destercke, S., Dubois, D., Chojnacki, E.: Possibilistic information fusion using maximal coherent subsets. IEEE Tr. Fuzzy Systems 17, 79–92 (2009)
5. Dubois, D.: The role of epistemic uncertainty in risk analysis. In: Deshpande, A., Hunter, A. (eds.) SUM 2010. LNCS, vol. 6379, pp. 11–15. Springer, Heidelberg (2010)
6. Dubois, D., Prade, H.: On the unicity of Dempster rule of combination. Int. J. of Intelligent Systems 1, 133–142 (1986)
7. Dubois, D., Prade, H.: A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets. Int. J. General Systems 12 (1986)
8. Dubois, D., Prade, H.: Representation and combination of uncertainty with belief functions and possibility measures. Computational Intelligence 4, 244–264 (1988)
9. Dubois, D., Prade, H.: Possibility theory and data fusion in poorly informed environments. Control Engineering Practice 2, 811–823 (1994)
10. Dubois, D., Prade, H.: Possibility theory: qualitative and quantitative aspects. In: Handbook of Defeasible Reasoning and Uncertainty Management Systems, vol. 1, pp. 169–226. Kluwer (1998)
11. Konieczny, S., Pino Pérez, R.: Logic based merging. J. Philosophical Logic 40(2), 239–270 (2011)
12. Liu, W.: Analyzing the degree of conflict among belief functions. Artificial Intelligence 170, 909–924 (2006)
13. Maynard-Reid, P., Lehmann, D.: Representing and aggregating conflicting beliefs. In: Proc. of Int. Conf on Principles of Knowledge Representation and Reasoning (KR 2000), pp. 153–164 (2000)
14. Oussalah, M.: Study of some algebraic properties of adaptative combination rules. Fuzzy Sets and Systems 114, 391–409 (2000)
15. Rescher, N., Manor, R.: On inference from inconsistent premises. Theory and Decision 1, 179–219 (1970)
16. Schockaert, S., Prade, H.: An inconsistency-tolerant approach to information merging based on proposition relaxation. In: Proc. 24th AAAI Conf. on Artificial Intelligence, pp. 363–368 (2010)
17. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press (1976)
18. Smets, P.: Data fusion in the transferable belief model. In: Proc. of Intern. Conf. on Information Fusion, Paris (2000)
19. Smets, P., Kennes, R.: The transferable belief model. Artificial Intelligence 66(2), 191–234 (1994)
20. Wagner, C., Lehrer, K.: Rational Consensus in Science and Society. D. Reidel (1981)
21. Walley, P.: The elicitation and aggregation of beliefs. Technical Report 23, University of Warwick, UK (1982)
22. Walley, P.: Statistical Reasoning with Imprecise Probabilities. Chapman and Hall (1991)
23. Walley, P.: Measures of uncertainty in expert systems. Artificial Intelligence 83, 1–58 (1996)
24. Yager, R.: On the Dempster-Shafer framework and new combination rules. Information Sciences 41, 93–138 (1987)