The basic principles of uncertain information fusion. An organised review of merging rules in different representation frameworks.

Didier Dubois, Weiru Liu, Jianbing Ma, Henri Prade

The basic principles of uncertain information fusion. An organised review of merging rules in different representation frameworks

Didier Dubois[a,b,*], Weiru Liu[b], Jianbing Ma[c], Henri Prade[a]

[a] IRIT, CNRS & Université de Toulouse, 118 Route de Narbonne, 31062 Toulouse Cedex 09, France
[b] School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast BT7 1NN, UK
[c] School of Computing, Electronics and Maths, Coventry University, Coventry, CV1 5FB, UK

Keywords:
Information fusion
Knowledge-based merging
Evidence theory
Combination rules
Plausibility orderings
Possibility theory
Imprecise probability

A B S T R A C T

We propose and advocate basic principles for the fusion of incomplete or uncertain information items, that should apply regardless of the formalism adopted for representing pieces of information coming from several sources. This formalism can be based on sets, logic, partial orders, possibility theory, belief functions or imprecise probabilities. We propose a general notion of information item representing incomplete or uncertain information about the values of an entity of interest. It is supposed to rank such values in terms of relative plausibility, and explicitly point out impossible values. Basic issues affecting the results of the fusion process, such as relative information content and consistency of information items, as well as their mutual consistency, are discussed. For each representation setting, we present fusion rules that obey our principles, and compare them to postulates specific to the representation proposed in the past. In the crudest (Boolean) representation setting (using a set of possible values), we show that the understanding of the set in terms of most plausible values, or in terms of non-impossible ones matters for choosing a relevant fusion rule. Especially, in the latter case our principles justify the method of maximal consistent subsets, while the former is related to the fusion of logical bases. Then we consider several formal settings for incomplete or uncertain information items, where our postulates are instantiated: plausibility orderings, qualitative and quantitative possibility distributions, belief functions and convex sets of probabilities. The aim of this paper is to provide a unified picture of fusion rules across various uncertainty representation settings.

1. Introduction

Information fusion is a specific aggregation process which aims to extract truthful knowledge from incomplete or uncertain information coming from various sources [15]. This topic is relevant in many areas: expert opinion fusion in risk analysis [24], image fusion in computer vision [13,14], sensor fusion in robotics [1,61,86], database merging [18,21], target recognition [78], logic [67,68] and so forth. Historically the problem is very old. It lies at the origin of probability theory whose pioneers in the XVIIth century were concerned by merging unreliable testimonies at courts of law [98]. Then, this problem fell into oblivion with the development of statistics in the late XVIIIth century. It was revived in the late XXth century in connection with the widespread use of computers, and the necessity of dealing with large amounts of data coming from different sources, as well as the renewed interest toward process human-originated information, and the construction of autonomous artefacts that sense their environment and reason with uncertain and inconsistent inputs.

Information fusion is often related to the issue of uncertainty modelling. Indeed, sources often provide incomplete or unreliable information, and even if such pieces of information are precise, the fact that they come from several sources often results in conflicts to be solved, as inconsistency threatens in such an environment. The presence of incomplete, unreliable and inconsistent information leads to uncertainty, and the necessity of coping with it, so as make the best of what is available, while discarding the wrong. This is the role of information fusion.

There are many approaches and formats to model information, and several uncertainty theories [51]. The fusion problem in the presence of uncertain or incomplete information has been discussed in each of these settings almost independently of the

* Corresponding author at: IRIT, CNRS and Université de Toulouse, 118 Route de Narbonne, 31062 Toulouse Cedex 09, France.Tel.: +33 561556331.
E-mail address: dubois@irit.fr (D. Dubois).

other ones [49,80,83,101]. Sometimes, dedicated principles have been stated in order to characterise the specific features of the fusion process in the language of each particular formal setting [69,73,87,109]. Several fusion strategies exist according to the various settings. These strategies share some commonalities but may differ from each other in some aspects due to their specific representation formats (for instance, symbolic vs. numerical).

This paper takes an inclusive view of the current available properties from different theories and investigates the common laws that *must be* followed by these fusion strategies[1]. We argue that some properties are mandatory and some are facultative only. The latter can be useful in certain circumstances, or in order to speed up computation time. It is interesting to notice that although each requested property looks intuitively reasonable on its own, they can be inconsistent when put together. This happens in the problem of merging preferences from several individuals modelled by complete preorderings (Arrow impossibility theorem, see the discussion in [22]). However, the basic mandatory properties of information fusion we propose are globally consistent.

The aim of the paper is to lay bare the specific nature of the information fusion problem. This general analysis yields a better understanding of what fusion is about and how an optimal fusion strategy (operator) can be designed. In particular, information fusion differs from preference aggregation, whose aim is to find a good compromise between several parties. Noticeably, while the result of information fusion should be consistent with what reliable sources bring about, a good compromise in a multiagent choice problem may turn out to be some proposal no party proposed in the first stand. So while they share some properties and methods, we claim that information fusion and preference aggregation do not obey exactly the same principles.

We also wish to show the deep unity of information fusion methods, beyond the particulars of each representation setting. To this aim, we look at special characteristics of each theory and what becomes of fusion principles, what are the fusion rules in agreement with these principles. We will check whether known fusion rules in each theory comply with general postulates of information fusion. We explain how these basic properties can be written in different representation settings ranging from set-based and logic-based representations to possibility theory, belief function theory and imprecise probabilities. These comparisons demonstrate that the proposed basic properties truly reflect the nature of fusion in different settings.

The rest of the paper is organised as follows. The next section presents general features of what can be called an information item. Such features can be extracted from information items in each representation framework. Section 3 presents basic principles of information fusion that apply to information items and discuss their relevance. Some additional and facultative principles are discussed. The problem of merging information is carefully distinguished from the one of preference aggregation. Section 4 instantiates our principles on the crudest representation of an information item, as a set of possible values. When such a set basically excludes impossible values, we show that our setting characterises the method of maximal consistent subsets. The case of merging propositional belief bases, for which a set of postulates, due to Konieczny and Pino-Perez [68], exists, is then discussed. We compare them to our fusion principles, and show that the corresponding Boolean information items in our sense correspond to subsets of most plausible values. The next section discusses the fusion of information items represented by plausibility rankings of possible values, going from ordinal representations to numerical ones in terms of possibility distributions. Again, we compare our instantiated principles with existing proposals, and provide examples of rational fusion rules in our sense. Finally the last section discusses representations that blend set-based and probabilistic formalisms, and account for incomplete information, such as belief functions and imprecise probabilities. We instantiate our principles in each setting, and study the property of known rules for merging belief functions. We also analyse postulates for merging imprecise probabilities proposed by Peter Walley [109] in the light of our general approach.

[1] Preliminary and partial versions of this paper were presented in two conferences [37,38].

2. A general setting for representing information items

We call what sources of information provide to an end-user *information items* pertaining to some uncertain entity. An information item is understood as a statement, possibly tainted with uncertainty, forwarded by some source, and describing what the current state of affairs is. In order to define a set of requirements that make sense in different representation settings ranging from logic to imprecise probability, we need to describe several features of an information item, that we consider essential.

Consider a non-empty set of possible worlds or state descriptions or alternatives, one of which is the true one, denoted by $W = \{w_1, \ldots, w_{|W|}\}$ (it will often be the range of some unknown precise entity denoted by x). For simplicity, we restrict ourselves to a finite setting. We assume that there are n agents/sources (sensors, experts, etc.) and the ith one is denoted by e_i. Let T_i denote the information item provided by agent e_i about x. For example T_i can be a set, a probability or a possibility distribution [42], or an ordinal conditional function [104] or a knowledge base.

In this paper, we do not discuss the fusion of precise set-valued entities, such as multisets [17], where sets represent complex entities made of the conjunction of several, possibly identical elements, representing hierarchical data structures [19], or related tuples in relational databases. Such multiset fusion problems can be found when cleaning databases containing duplicate data [18] or for the summarisation of documents. On the contrary, sets used in the representation of uncertain items of information contain mutually exclusive values[2].

Here, an information item indicates which values or states of affairs in W are plausible, and which ones are not, for the uncertain entity or parameter x, according to a source. In that sense an information item is completely attached to the source that supplies it and is not an objective description of the state of affairs. It is a representation of knowledge that is likely to be modified by additional information. An information item T will then be characterised by several features:

- Its *support* $\mathcal{S}(T) \subseteq W$, contains the set of values of x considered not impossible according to information T. Namely, $w \notin \mathcal{S}(T)$ if and only if the value w is considered impossible for the source offering T. One may see $\mathcal{S}(T)$ as a kind of integrity constraint attached to T. If $\mathcal{S}(T) = \emptyset$ then information T is said to be strongly inconsistent. The condition $\mathcal{S}(T) \neq \emptyset$ is a weak form of (internal) consistency.
- Its *core* $\mathcal{C}(T) \subseteq W$, contains the set of values considered fully plausible according to information T. One may see $\mathcal{C}(T)$ as the

[2] Note that if x is a set-valued attribute, we do not consider the fusion of such precise set-values, e.g. $x = A$. But our approach encompasses the case of incomplete information for set-valued attributes [46]. For instance, if x is the precise time interval when the museum is open, a piece of information like "the museum is open from 9 to 12 h" is imprecise in the sense that what we know from it is that $[9, 12] \subseteq x$. If another source claims that "the museum is open from 14 to 17 h." we may conclude that the museum is open from 9 to 12 h *and* from 14 to 17 h, which here is modelled by $[9, 12] \cup [14, 17] \subseteq x$. However this disjunction is actually obtained by the conjunctive fusion of two sets of time spans, namely $\{A: [9, 12] \subseteq A\} \cap \{A: [14, 17] \subseteq A\}$.

plausibility set attached to 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, it follows that $\mathcal{C}(T) \subseteq \mathcal{S}(T)$, as the most plausible elements are to be found in the support. An information item such that $\mathcal{C}(T) \neq \emptyset$ is said to be *strongly consistent*. This strong form of internal consistency can be requested for inputs of a merging process, but not necessarily so for its result. In the following, we assume each source provides strongly consistent information.

- Its induced *plausibility ordering*: If consistent, information T induces a partial preorder $\succeq_T$ (reflexive and transitive) on possible values, expressing relative plausibility[3]: $w \succeq_T w'$ means that w is at least as plausible as (or *dominates*) w' according to T. We write $w \sim_T w'$ if $w \succeq_T w'$ and $w' \succeq_T w$. The concept of plausibility ordering corresponds to the idea of potential surprise already discussed by Shackle [96], namely a state of affairs is all the more implausible as its presence is more surprising. Of course, the plausibility ordering should agree with $\mathcal{C}$ and $\mathcal{S}$, i.e., if $w \in \mathcal{S}(T), w' \notin \mathcal{S}(T)$, then $w \succ_T w'$ (w is strictly more plausible than w'). Likewise if $w \in \mathcal{C}(T), w' \notin \mathcal{C}(T)$.

The triple $(\mathcal{S}(T), \mathcal{C}(T), \succeq_T)$ is not redundant. Indeed, if only $\succeq_T$ is known, we still do not know if the least worlds according to $\succeq_T$ are possible or not, nor if the best worlds according to $\succeq_T$ are fully plausible or not. So, while $\succeq_T$ provides relative information, the sets $\mathcal{S}(T), \mathcal{C}(T)$ respectively point out impossible and fully plausible worlds according to each source.

Let us give a number of examples of formats for representing information items:

- *Sets*: T is of the form of a non-empty subset E_T of W, representing all the mutually exclusive possible values for x according to the information source. The set E_T is often called a disjunctive set, representing the information possessed by an agent (an epistemic state). Then $\mathcal{S}(T) = \mathcal{C}(T) = E_T$. And $w_1 \sim_T w_2$ when w_1, w_2 both belong to E_T or both belong to its complement. Moreover, $w_1 \succ_T w_2$, $\forall w_1 \in E_T$, and $\forall w_2 \notin E_T$. An alternative interpretation of E_T is a set of plausible values, and then $\mathcal{S}(T) = W$ and $\mathcal{C}(T) = E_T$. This latter view is sometimes used in belief revision and the related approaches to fusion, as we shall see.
  Important special cases of set-valued information are
  - *Vacuous information*, expressing total ignorance is denoted by $T^\top$. Then $\mathcal{S}(T^\top) = \mathcal{C}(T^\top) = W$ and the plausibility ordering is flat in the sense that $w \sim_{T^\top} w'$ $\forall w, w' \in W$.
  - *Complete knowledge* expressing that the actual world is known to be w is denoted by T^w: then $\mathcal{S}(T^w) = \mathcal{C}(T^w) = \{w\}$.

  Note that T can take the form of a propositional knowledge base K [68]; then W is the set of interpretations of a propositional language $\mathcal{L}$. Then $E_T = [K]$ the set of models of K. Alternatively, if T represents information about a numerical parameter, it may take the form of an interval $[a, b]$ on the real line [4]. This case is studied in more details in Section 4
- *Plausibility relations*: we call an information item *ordinal* if it consists of the triple $(\mathcal{S}(T), \mathcal{C}(T), \succeq_T)$. If only the plausibility ordering is provided, one may accept that by default, the maximal elements according to $\succeq_T$ form the core of T. However, we can find examples of information items for which no world is fully plausible. For instance, if $\succeq_T$ stems from a probability distribution p, the most probable situation may not be very probable: it is clear that $\{w : p(w) = 1\} = \emptyset$, generally, since $\sum_{w \in W} p(w) = 1$. Likewise, by default we can assume the support of $\succeq_T$ is W itself unless otherwise specified. This case is studied in more details in Section 5. This format encompasses the previous one when $\succeq_T$ is complete and induces only two levels.
- A *possibility distribution* [50], namely a mapping π_T: $W \to L$ where L is a totally ordered set of plausibility levels, its bottom 0 encoding impossibility, and its top 1 encoding full plausibility. The existence of a scale L is the key difference between this format and the one of plausibility relations, where we only have $\succeq_T$, not π_T. A possibility distribution π_T is then more expressive than the plausibility ordering it induces, as the use of scale L enables the user to say that some situation is fully plausible (and not only the most plausible) or some other is impossible (and not only the least plausible). More numerical settings for possibility distributions can be used. Then $\mathcal{S}(T) = \{w : \pi(w) > 0\}$ is the support and $\mathcal{C}(T) = \{w : \pi(w) = 1\}$ is the core in the sense of fuzzy sets, viewing π_T as a membership function of a fuzzy set. The plausibility ordering $\succeq_T$ is induced by π_T. Note that the possibility scale L can be numerical or not. In the most qualitative situation, it could be a finite chain of symbolic levels. In contrast, we can let $L = [0, 1]$ and use numerical degrees of possibility (often interpreted as upper probability bounds [50]). Alternatively, (im)possibility levels can be encoded by integers, as done by Spohn [104]. However, in that case the scale is one of implausibility, namely a mapping $\kappa_T : W \to \mathbb{N}$ such as $\kappa_T(w) = 0$ for normal situations, and w is all the less plausible as $\kappa_T(w)$ is greater (then one may let $\pi_T(w) = k^{-\kappa_T(w)}$ for some integer $k > 1$ [50]). This case is studied in more details in Section 6.
- A *mass assignment* m_T that defines belief and plausibility functions in Shafer's theory of evidence [97]. A mass assignment is formally a random set, i.e., a probability distribution over possible choices of epistemic states, $m_T(E)$ being the probability that the best epistemic state representing T is E. This representation is more general than a mere probabilistic representation; for the latter, $m_T(E) > 0$ only if E is a singleton. One choice of the induced triple $(\mathcal{S}(T), \mathcal{C}(T), \succeq_T)$ can rely on the so-called contour function (plausibility of singletons, understood as the probability of hitting them by sets E). This case is studied in more details in Section 7.
- A *convex set of probability measures* [109]: it may represent either a set of possible probabilistic information items (it is a second order 0–1 possibility distribution) or the state of belief of an agent described via desirable gambles (see Section 8.1 for details). The triple $(\mathcal{S}(T), \mathcal{C}(T), \succeq_T)$ is then for instance derived from the upper probability of singletons. This more complex case is studied in some details in Section 8.

Note that these frameworks are listed in increasing order of expressiveness, so that any information item expressible in one setting can be encoded in the settings further down (possibly adding some extra information, for instance if we encode a plausibility relation in the form of a possibility distribution).

Finally we must be able to compare two items of information in terms of their relative informativeness and their mutual consistency.

- *Information ordering* denoted by $\sqsubseteq$: it is a partial preorder relation (reflexive, transitive) on information items. $T \sqsubseteq T'$ expresses that T provides at least as much information as T' (said otherwise: is more precise, more specific). In particular, the information ordering is defined such that $T \sqsubseteq T'$ implies $\mathcal{S}(T) \subseteq \mathcal{S}(T')$ and $\mathcal{C}(T) \subseteq \mathcal{C}(T')$. It makes full sense if T is strongly consistent ($\mathcal{C}(T) \neq \emptyset$), and its meaning becomes trival if T is strongly inconsistent. In the case of set-based representations, this ordering coincides with set-inclusion, and in possibility theory, fuzzy

[3] In some settings, there may exist several candidates for $\succeq_T$, like the setting of belief functions, see Section 7. In some cases, the plausibility ordering may be partial.

[4] Then, W is no longer finite; however, our setting can be extended to the infinite case.

set inclusion. It is less obvious for ordinal plausibility representations (see Section 5) and belief functions (see Section 7) as there are several options. The information ordering is also more difficult to define between pieces of information having empty cores as it denotes some internal inconsistency that may override the notion of informativeness.

- *Imprecision index*: this is a measure $II(T)$ of how much information is contained in an information item T. If T reduces to a set of possible values E_T, and then it can be the cardinality of E_T in the finite set setting (or its logarithm). More generally, it could be some index of non-specificity for possibility measures or belief functions [114,115].

Note that $T \sqsubseteq T'$ conveys more meaning than simply saying that T' is more imprecise than T. The latter could be expressed by comparing imprecision indices as $II(T') \geq II(T)$. Actually, $T \sqsubseteq T'$ also means that T' can be derived from T: the relation $\sqsubseteq$ should be viewed as a (generalised) entailment relation as well, while if $II(T') \geq II(T)$, nothing forbids T and T' from being totally inconsistent with each other.

- *Mutual consistency*: two items of information T and T' will be called *weakly mutually consistent* if their supports overlap ($\mathcal{S}(T) \cap \mathcal{S}(T') \neq \emptyset$), and *strongly mutually consistent* if their cores overlap ($\mathcal{C}(T) \cap \mathcal{C}(T') \neq \emptyset$). However, the latter property is violated by information items having empty cores as they already display a form of internal inconsistency.

We can give a number of examples of situations where such kinds of information item appear naturally:

- In the merging of expert opinions, experts provide knowledge about parameters of components of a complex system (for instance, failure rate of a pump in a nuclear power plant), in the form of an uncertainty distribution that can be a subjective probability distribution [24], or yet a likelihood function [57], or a possibility distribution [94].
- The problem of syntax-independent merging of logical databases [69] comes down to merging their sets of models.
- In sensor fusion, information provided is often modelled by random sets that account for reliability coefficients [61,86]. Namely, sensor readings can be mapped to a set of decision hypotheses modelled by mass functions [97].

3. General postulates of information fusion

Let $\mathcal{T}$ be the set of possible information items of a certain format. It is assumed that the input information items are strongly internally consistent ($\mathcal{C}(T_i) \neq \emptyset$). A n-ary *fusion operation* f_n is a mapping from $\mathcal{T}^n$ to $\mathcal{T}$, that operates the merging process:

$$T = f_n(T_1, \ldots, T_n)$$

denotes the result of the fusion of a set of information items T_i. Following the terminology in [5], a *fusion operator* is a collection of fusion operations f_n, $n \in \mathbf{N}$, $n \geq 1$ for all arities. By convention $f_1(T) = T$. When this is not ambiguous we shall replace f_n by f.

The process of merging information items, supplied by sources whose reliability levels are not known to differ from one another, is guided by a few general principles, already proposed in [45,49] that we shall formalise in the following:

- It is a basically symmetric process as the sources play the same role and supply information of the same kind;
- In the fusion process, we consider as many information items as possible as reliable, so as to get a result that is as precise and useful as possible, however not arbitrarily precise, if there is not enough information to be precise. The result should be faithful to the level of informativeness of the inputs.
- Information fusion should solve conflicts between sources, while neither dismissing nor favouring any of them without a reason.

A prototypical example of a fusion situation can be the following. Suppose we have three witnesses i, each of whom provides a piece of information T_i. Suppose that T_2 and T_3 are compatible, but that the truth of T_1 is incompatible with T_2 and T_3. In case all sources inform on the same issue, and are considered equally relevant, an intuitively natural way of making the best of such information is to consider the true situation to be in agreement with either the part of the information common to witnesses 2 and 3, or with the information provided by witness 1. This is achieved by a conjunctive combination of information items T_2 and T_3 followed by a disjunctive merging of the result with information item T_1. It respects symmetry, does not dismiss any of the witnesses, and it is the most precise conclusion one may legitimately draw. The purpose of this paper is to provide postulates embodying the above principles underlying the fusion process at work in such a kind of situation, and to instantiate them in different formal settings where the pieces of information T_i can take various formats recalled in the previous section. In each case, we lay bare what is the main fusion operation that respects these principles.

A fusion operation with such an agenda was called *arbitration* by Revesz [93] and taken over by Liberatore and Schaerf [73], in the set-theoretic or logical framework. These principles are implemented in the postulates listed below, called *basic properties*, which are meant to be natural minimal requirements, independent of the actual representation framework.

3.1. Basic properties

The postulates we consider essential and that any information fusion process should satisfy are as follows:

Property 1: Unanimity

The result of the fusion should propose values on which all sources agree and reject those values rejected by all sources. Formally it reads:

(a) *Possibility preservation*: if all sources consider some world is possible, then so should the result of the fusion. It means in particular that

$$\bigcap_{i=1}^{n} \mathcal{S}(T_i) \subseteq \mathcal{S}(f(T_1, \ldots, T_n)).$$

(b) *Impossibility preservation*: if all sources believe that some world is impossible, then this world is considered impossible after fusion. Mathematically, this can be expressed as

$$\mathcal{S}(f(T_1, \ldots, T_n)) \subseteq \mathcal{S}(T_1) \cup \cdots \cup \mathcal{S}(T_n).$$

Property 2: Information monotonicity

When a set of agents provides less information than another set of *non-disagreeing* agents, then fusing the former set of information items should not produce a result that is more informative than fusing the latter set of information items. Formally, it reads:

If $\forall i, T_i \sqsubseteq T_i'$, then $f(T_1, \ldots, T_n) \sqsubseteq f(T_1', \ldots, T_n')$, provided that inputs are globally strongly consistent ($\mathcal{C}(T_1) \cap \cdots \cap \mathcal{C}(T_n) \neq \emptyset$).

Property 3: Consistency enforcement

This property requires that fusing individually consistent inputs should give a consistent result. In particular, at the very least one should require that

$$\mathcal{S}(f(T_1, \ldots, T_n)) \neq \emptyset.$$

Property 4: Optimism

In the absence of specific information about source reliability, one should assume that as many sources as possible are reliable, in agreement with their observed mutual consistency. In particular,

- If all the inputs are mutually consistent, then the fusion should preserve the information supported by every input:
  If $\bigcap_i \mathcal{C}(T_i) \neq \emptyset$ then $f(T_1, \ldots, T_n) \sqsubseteq T_i, \forall i = 1, \ldots, n$.
- If all the inputs are mutually inconsistent, it should be assumed that at least one source is reliable.

More generally, this basic property comes down to assuming that any group of consistent sources is potentially reliable, and at least one of this group is truthful.

Property 5: Fairness

The result of the fusion process should keep something from each input, i.e., $\forall i = 1, \ldots, n, \mathcal{S}(f(T_1, \ldots, T_n)) \cap \mathcal{S}(T_i) \neq \emptyset$.

Property 6: Insensitivity to vacuous information

Sources that provide vacuous information should not affect the result of fusion. That is, $f_n(T_1, \ldots, T_{i-1}, T^\top, T_{i+1}, \ldots, T_n) = f_{n-1}(T_1, \ldots, T_{i-1}, T_{i+1}, \ldots, T_n)$.

Property 7: Commutativity

Inputs from multiple sources are treated on a par, and the combination should be symmetric. This is represented as $f(T_1, \ldots, T_n) = f(T_{i_1}, \ldots, T_{i_n})$ for any permutations of indices.

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 basic properties.

The basic properties proposed here in generality have counterparts in properties considered in some particular settings; see especially [109] for imprecise probability, [87] for possibility theory and [68] for knowledge bases. We shall compare their proposals with the above more general one in the sections devoted to these specific frameworks.

Besides, note that some of the above principles are expressed using the supports of information items, and some others use their cores. The choice was guided by the concern to make each property as little demanding as possible, while still meaningful. This choice can of course be debated, and some of these postulates can be written using cores only, but their strength and possibly their intuitive nature are then altered. However, some of these axioms as stated above (1, 3, 5) trivialise when merging information items whose support is W. This is discussed below.

3.2. Arguing for the basic postulates and some variants

In the following we provide the rationale of the above postulates and discuss possible variants for them.

3.2.1. Unanimity

The corresponding basic postulate is the weakest form of unanimity one may require: accepting what is unanimously possible, and rejecting what is unanimously impossible. This property admits of variants of various strength. First, one might replace supports by cores. Then, Property 1a means that the result should consider as fully plausible at least all worlds judged fully plausible by all sources (plausibility preservation). The core-based counterpart to Property 1b is rather demanding and more debatable, as the most plausible worlds after fusion could well be among worlds that are not considered fully plausible by some source.

A natural, often found, form of unanimity is:

Idempotence: $\forall i, T_i = T,\ f(T_1, \ldots, T_n) = T$,

However, adopting it in all situations forbids reinforcement effects to take place in case sources can be assumed to be independent. Idempotence could be adopted if it is not known whether the sources are independent or not [49]. If sources are independent, one expects possible worlds judged somewhat implausible by many sources to be more implausible globally.

The basic postulate takes a form that leaves room to reinforcement effects, while minimally respecting the agreement between sources. It trivially implies that if all sources supply empty information, the result of the fusion will be empty as well. Likewise if all information item supports are the same, the result of the fusion will have the same support. For instance, if all sources claim the only possible world is w, then so is the global result.

Somewhat stronger than our postulate, yet weaker than idempotence is the following axiom that is used in social choice:

Ordinal unanimity If $\forall i, \geq_{T_i} = \geq_T$, then $\geq_{f(T_1,\ldots,T_n)} = \geq_T$.

Ordinal unanimity can be restricted to each pair of worlds (w, w'):

Local ordinal unanimity
$\forall w, w'$, if $\forall i, w \geq_{T_i} w'$, then $w \geq_{f(T_1,\ldots,T_n)} w'$.

Ordinal unanimity is a global notion that is weaker than local ordinal unanimity since the global form only constrains the result when all information items T_i generate the same plausibility orderings, while the latter property only applies to the part of $W \times W$ where all ordering relations coincide. Local Ordinal Unanimity is a special case of the so-called Arbitration property used in knowledge base merging [69][5], a variant of which can be written here in the case of n sources:

Arbitration if $\forall i = 1, \ldots, n, w \geq_{T_i} w_i$, and $\forall i, j = 1, \ldots, n$, $w_i \sim_{f(T_1,\ldots,T_n)} w_j$, then $\forall i = 1, \ldots, n, w \geq_{f(T_1,\ldots,T_n)} w_i$.

3.2.2. Information monotonicity

This basic property 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 mutually consistent. In that case the result of the fusion may become artificially very precise, by virtue of the Optimism postulate, and in particular, more precise than the union of the supports of original conflicting items of information (as the intersection of enlarged supports is performed).

One can strengthen this postulate by requesting the preservation of strict relations:

Strict information monotonicity If $\forall i, T_i \sqsubseteq T'_i$, and $\exists j, T_j \sqsubset T'_j$ then $f(T_1, \ldots, T_n) \sqsubset f(T'_1, \ldots, T'_n)$, whenever $\mathcal{C}(T_1) \cap \cdots \cap \mathcal{C}(T_n) \neq \emptyset$.

This is generally too demanding in purely Boolean representation settings. Even set-intersection and set-union violate it. However, it makes more sense in numerical representation settings.

3.2.3. Consistency enforcement

This postulate is instrumental if the result of the fusion is to be useful in practice: one must extract something meaningful and non-trivial, even if tentative, from the available information, even if sources contradict one another. However, when the representation framework is sufficiently refined, there are gradations in consistency requirements, and this property can be interpreted in a

[5] This name is borrowed from [93], and [73], but the Arbitration property here seems to be only loosely related to the notion of arbitration operation in the sense of [73].

more flexible way. For instance, re-normalisation of belief functions or possibility distributions obtained by fusion is not always compulsory [99], even if sub-normalisation expresses a form of inconsistency. Likewise, in the symbolic setting, where knowledge is expressed by means of logical formulas, one may relax this assumption by adopting a paraconsistent approach whereby each formula is either supported, denied, unknown or conflicting with respect to a set of sources (as for instance in the approach by Belnap [6]).

A more demanding variant of this postulate is obtained replacing support by core. Then the enforcement of weak consistency is replaced by a requirement of strong consistency of the result.

3.2.4. Optimism

This postulate underlies the idea of making the best of the available information. If items of information are consistent and no other information is otherwise available, there is no reason to question the reliability of the sources. This means that if all the inputs are globally consistent with one another, then the information provided by each source should be preserved, i.e., $f(T_1, \ldots, T_n)) \sqsubseteq T_i, \forall i = 1, \ldots, n$, or at least $\mathcal{S}(f(T_1, \ldots, T_n)) \subseteq \mathcal{S}(T_i), \forall i = 1, \ldots, n$. In that case, we assume that they all supply correct information, so that the result should be more informative than, and in agreement with each original information item, which is clearly an optimistic attitude. This postulate is implicitly at work in belief revision [56] as well, since in the AGM axioms[6] it is assumed that when the new information does not contradict the prior one, the revision comes down to an expansion [2], which is, in our sense, an optimistic fusion operation [32].

In case of inconsistent sources, this formal requirement is no longer sustainable. Note when inputs are globally inconsistent (in particular, strongly so: $\bigcap_i \mathcal{S}(T_i) = \emptyset$), and we accept Impossibility Preservation property 1b, then the support of the result should be contained in the union of the supports of inputs, i.e., $\mathcal{S}(f(T_1, \ldots, T_n)) \subseteq \mathcal{S}(T_1) \cup \cdots \cup \mathcal{S}(T_n)$. This makes sense provided that at least one source is supposed to be reliable (still a form of optimism). Requiring equality in the latter inclusion would be a very cautious requirement (assuming that only one source is reliable). It sounds natural for two sources only, but may be found overcautious in the case of many sources [45]. So one usually expects a strict inclusion $\mathcal{S}(f(T_1, \ldots, T_n)) \subset \mathcal{S}(T_1) \cup \cdots \cup \mathcal{S}(T_n)$. More specifically, one may expect that *for each subset* I of mutually consistent sources, there is a piece of information $T_I \sqsubseteq T_i$, $\forall i \in I$ such that $T_I \sqsubseteq f(T_1, \ldots, T_n)$, and that (this is where optimism comes in) $f(T_1, \ldots, T_n)$ should be the most specific output satisfying this condition. One is led to choose I as a maximal set of consistent sources, so as to select T_I as informative as possible (although Minimal Commitment will prevent an arbitrarily precise choice). Of course, there are several possible choices of maximal subsets I of consistent sources.

3.2.5. Fairness

It ensures that all input items participate to the result. In particular, when inputs are globally inconsistent (especially, $\bigcap_i \mathcal{S}(T_i) = \emptyset$), the fusion result treats all sources on a par. For instance, if T_1 is inconsistent with T_2 and T_3 that are mutually consistent, then having $\mathcal{S}(f(T_1, \ldots, T_n)) = \mathcal{S}(T_2) \cap \mathcal{S}(T_3)$ is optimistic but it is unfair to T_1.

Fairness also implies no source is privileged in the following sense:

Proposition 1. If the Fairness axiom is satisfied the following property holds:

No Favouritism: the fusion result never implies any single input inconsistent with some of the other inputs: it does not hold that $f(T_1, \ldots, T_n) \sqsubseteq T_i$ for any T_i such that $\exists j\ \mathcal{S}(T_j) \cap \mathcal{S}(T_i) = \emptyset$,

Proof Due to Fairness, if $\mathcal{S}(T_j) \cap \mathcal{S}(T_i) = \emptyset$ then $\exists A_i \subseteq \mathcal{S}(T_i), A_j \subseteq \mathcal{S}(T_j)$, non-empty sets such that $A_i \cup A_j \subset \mathcal{S}(f(T_1, \ldots, T_n)$. As a consequence $\mathcal{S}(f(T_1, \ldots, T_n) \not\subseteq \mathcal{S}(T_i)$, and so $f(T_1, \ldots, T_n) \not\sqsubseteq T_i$.

So this axiom favours no source by preventing any input information item from being the global output result in case of inconsistency. Note that different versions of the idea of fairness can be found in the literature. In particular, in [67,68], where information items are consistent knowledge bases with sets of models E_i, they propose the condition that $f(E_1, \ldots, E_n) \cap E_i \neq \emptyset$ either holds for each i, or for none. The possibility that it holds for none (that is, the result of the fusion may contradict all information items in case of conflict) is a matter of debate from a knowledge fusion point of view; it may be acceptable when fusing preferences, which is a matter of building a compromise, and also if the sets E_i correspond to cores of information items T_i; but it sounds strange if they correspond to supports. In the latter case, as we assume no information about reliability of sources, we take it for granted that the final result should keep some memory of all sources. Replacing supports by cores in our fairness condition is more demanding and may sound questionable.

One way to strengthen the Fairness axiom is to combine it with Optimism and to require that the partial information from each source retained in the final result be common to as many sources as possible:

Optimistic fairness: For any subset I of consistent sources, $\mathcal{S}(f(T_1, \ldots, T_n)) \cap \bigcap_{i \in I} \mathcal{S}(T_i) \neq \emptyset$.

This condition will improve the informativeness of the result, as it will enforce values considered possible in the result to be in agreement with maximal consistent subsets of sources.

3.2.6. Insensitivity to vacuous information

This one looks obvious, even redundant, but dispensing with it may lead to overly uninformative results. In fact, this postulate implicitly admits that a non-informative source is useless and irrelevant, and is assimilated to one that does not express any opinion. In other words this postulate does not allow for interpreting the input as meta-information, like a source declaring that one cannot know anything more informative than what is declared. There is no "contraction" [7] effect allowed by acquiring poor information. It is also the only relationship explicitly requested at this level of generality, between fusion functions of different arities. This postulate is typical of information fusion, and excludes fusion rules like some forms of averaging that are always sensitive to vacuous information (if represented by, for instance uniform distributions). Of course, in some uncertainty theories, averaging is built-in, and is useful (e.g., in probability or belief function theories). But it arguably addresses other tasks than the one considered here (like estimation, where independence assumptions are needed, and precise observations are available, or the explicit discounting of sources [97], which is a form of contraction [2]).

3.2.7. Commutativity

This is really characteristic of fusion processes as opposed to revision. Revision is about how input information should alter prior knowledge. This process is fundamentally asymmetric: generally, priority is given to the input information and the process is driven by the minimal change principle [2,32] (the prior information is minimally changed so as to accept the input information). On the

[6] named from Alchourrón, Gärdenfors and Makinson [2].

[7] In the sense of belief revision theory [2].

contrary, the kind of fusion process we deal with here has to do with information items obtained in parallel. So, commutativity makes sense, if no information is available on the reliability of sources. One obvious objection against commutativity is that information items are often not equally reliable. However, a natural way of handling an unreliable information item is to use the discounting method [97] in order to get a reliable but less informative information item. Typically, assume a set-valued information item of the form $x \in T$ is reliable with probability p. Then this is equivalent to an information item T_p in the form of a random set granting mass p to T and mass $1-p$ to the whole set W. Then the asymmetric merging of unreliable set-like information items comes down to the symmetric merging of mass assignments in evidence theory (in fact, possibility distributions π such that $\pi(w) = 1$, for $x \in T$, and $1-p$ otherwise). However, we do not consider prioritised merging where information coming from unreliable sources is discarded if inconsistent with information coming from more reliable ones. This topic is discussed in [26] for logical databases. The framework of prioritised merging can encompass both fusion and revision.

3.2.8. Minimal commitment

This is a very important postulate that applies in many circumstances. It comes down to saying we should never express more information than the one that is actually available. It appears in all uncertainty theories in a specific form as we shall see later, including in logic-based approaches. It is in some sense the converse of the closed world assumption where any statement not explicitly formulated is considered to be false. Here we consider possible any state of affairs not explicitly discarded. 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.

An important consequence of Optimism along with some of the other postulates can be asserted:

Proposition 2. *Suppose* $\bigcap_i \mathcal{C}(T_i) \neq \emptyset$. *If a fusion operation* f *satisfies Optimism and any of Possibility Preservation 1a or Minimal Commitment, then for globally consistent information items* $T_i, i = 1, \ldots, n$, *we have* $\mathcal{S}(f(T_1, \ldots, T_n)) = \bigcap_i \mathcal{S}(T_i)$.

Proof. From Optimism we have $\mathcal{S}(f(T_1, \ldots, T_n)) \subseteq \bigcap_i \mathcal{S}(T_i)$. From Minimal Commitment there is no other reason to discard more possible worlds. Alternatively, Possibility Preservation ensures $\bigcap_i \mathcal{S}(T_i) \subseteq \mathcal{S}(f(T_1, \ldots, T_n))$, hence in either case, we get $\mathcal{S}(f(T_1, \ldots, T_n)) = \bigcap_i \mathcal{S}(T_i)$.

Note that, in some representation settings, other postulates than Optimism may further restrict the set of possible worlds.

3.3. Facultative additional properties of fusion operations

Some other properties are often either required or implicit in information fusion. But they turn out to be debatable in some situations.

3.3.1. Universality

It has two complementary aspects:

Unrestricted domain: $\forall T_i \in \mathcal{T}, \exists T \in \mathcal{T}, T = f(T_1, \ldots, T_n)$.
Attainability: $\forall T \in \mathcal{T}, \exists T_1, \ldots, T_n \in \mathcal{T}$, such that $T = f(T_1, \ldots, T_n)$.

Universality is one used in social choice, but it may apply to any aggregation problem. Indeed, Unrestricted Domain claims we must be able to get a result whatever items of information are merged. Moreover, Attainability says that no item of information should be excluded as a possible result.

Strictly speaking, Universality is in fact a consequence of our postulates since:

- Attainability is trivially implied by Insensitivity to Vacuous Information: $\forall T \in \mathcal{T}, T = f(T, T^{\top})$.
- Consistency Enforcement is a strong version of Unrestricted Domain whereby the combined result should not only exist, but be consistent.

However, since we define the fusion operation as a mapping from $\mathcal{T}^n$ to $\mathcal{T}$, it requires the result be expressed in the same kind of representation setting as the inputs. This feature introduces a constraint on the possible fusion rules, that may be damaging in some situations: it is a closure requirement (namely all results should lie in the class $\mathcal{T}$). For instance, merging knowledge bases should yield a single knowledge base (not a subset thereof), merging possibility distributions should yield a possibility distribution (not a more general object like a belief function structure), etc. So this property could be relaxed to account for the possibility of increasing the level of generality of the obtained result (allowing a larger class of operations that yield results outside $\mathcal{T}$, the most general being convex probability sets), especially in the case of conflicting inputs, or when the fact of forcing the result to be in $\mathcal{T}$ rules out certain modes of fusion that sound natural in other respects. For instance, merging probability measures into another one requires a weighted average [81], which rules out conjunctive and disjunctive modes of fusion (that yield belief functions) [45]. So we may in the following admit that in some situations, the result of the fusion is not necessarily a member of the specific class $\mathcal{T}$ where inputs lie, but may lie in a larger class.

3.3.2. Non-sensitivity

When merging consistent items, increasing (resp. decreasing) the informativeness of one of them *slightly* should result in a *slightly* more (resp. less) informative result. This is a property of robustness of the aggregation operation, that sounds natural for numerical aggregation schemes. It can be expressed as follows:

For all n-tuples of globally consistent information items $T_i, i = 1, \ldots, n$, $\exists k > 0$, such that $\forall j = 1, \ldots, n$ if $T_j \sqsubseteq T'_j$, then

$$d(f(T_1, \ldots, T_j, \ldots, T_n), f(T_1, \ldots, T_{j-1}, T'_j, T_{j+1}, \ldots, T_n)) \leq k \cdot d(T_j, T'_j)$$

for some informational distance d between pieces of information.

Under the same consistency assumptions, Non-Sensitivity as formulated above is stronger than the mere continuity of the aggregation operation. Semantically, this property requires that fusion should not be over-reactive to small informational changes. Note that again this property may not make sense for conflicting inputs, as they may become consistent in case of a small relaxation, thus possibly resulting in a dramatic change of the result, if Consistency Enforcement and Fairness are respected. This is the case in a purely Boolean setting. In numerical settings, conjunctive rules like Dempster rule of combination for belief functions are even discontinuous in the presence of severely (but not fully) conflicting information [45].

3.3.3. Associativity

$f(f(T_1, T_2), T_3) = f(T_1, f(T_2, T_3))$. This property is useful to facilitate the computation of the fusion process, but it has no other motivation pertaining to the nature of the fusion process. If a fusion operation can be associative, so much the worth. However, the lack of associativity is not a fatal flaw (e.g., averaging operations are not), if the fusion operation can be defined for all arities.

3.3.4. Independence of irrelevant alternatives

(IIR): It means that the resulting relative plausibility between w and w' only depends on the relative plausibility between w and w' declared by the sources. This is yet another property that comes

from social choice. It is satisfied by the union and intersection of sets or fuzzy sets. However, in some settings, especially the one of belief functions, the relative plausibility between two possible worlds after fusion is influenced by other factors not just the relative plausibilities of the two worlds in the original information items. The same is true in general when pieces of information are conflicting. So this property cannot be within the set of basic postulates. Nevertheless note that the Local Ordinal Unanimity property underlies a form of IIR.

3.3.5. Majority

Consider a countable set of non-vacuous information items $\{T_1, T_2, \ldots\}$ such that $T_i = T \neq T_1, \forall i > 1$. Then $\exists n > 2$, $f_n(T_1, \ldots, T_n) = T$.

This property, which sounds very natural in the case of voting processes, may also sound fine for fusion processes if n is large enough. It however implicitly presupposes that sources are independent and identical (similar to the statistical i.i.d.[8] assumption), so that when n increases, the source that proposes something different from other ones appears more and more like an outlier, and can be dismissed. Otherwise in the case of not provably independent sources, this property sounds debatable. And in fact, it is not credible that in a real information fusion problem, there exists a large number of independent sources. If many sources are available, it is very likely that some of them will be redundant (for instance, it is hard to find 100 experts agreeing on some issue of interest while having strictly different backgrounds). In the case of voters, they are *legally* considered independent (even if they have read the same newspapers), so that the majority rule is used for preference aggregation (see next subsection) to serve the purpose of democracy. But it does not sound like a basic universal postulate for information fusion. Finally it contradicts the Fairness axiom, as $f_n(T_1, \ldots, T_n)$ eventually does not take T_1 into account any more. Again, Majority presupposes a "the more, the more likely" assumption, which we do not regard as universal unless one can be sure about the independence of sources.

3.4. Information fusion vs preference aggregation

Information fusion takes a set of imperfect inputs (imprecise and uncertain) from different sources and produces a single output which should best reflect what is known about the true state of the world. In other words, in this paper, information items model what the world is supposed to be. In the case of preference aggregation, the items of information reflect how the world *should* be, according to sources that can represent individuals (in voting theory) or criteria (in multifactorial evaluation) [107]. In a nutshell, the aim of preference aggregation is to find a compromise between antagonistic options, while the aim of information fusion is to find the truth. This difference of perspective implies that methods that make sense for preference aggregation may sound debatable for information fusion.

For example, if one source states that *the room is hot* while the other states *the room is cold* (e.g. sources are thermometers), then the role of merging or fusion is to resolve such inconsistency, possibly by using some other kind of information such as reliability of sources, in order to find out a correct range of temperature. If the two words *hot* and *cold* are viewed as totally incompatible despite their fuzzy nature, the outcome is unlikely to be *warm*, a result that is not reported by any source. However, if when deciding a duration for holidays, the husband prefers a short break of one week and the wife prefers a long vacation period of 4 weeks, then a trade-off could be a two week holiday, a trade-off originally not suggested by any of the sources but plausibly acceptable by both of them, eventually. Therefore, information fusion or knowledge merging should focus on outcome(s) supported by some sources and not on possible worlds not suggested by any, whilst preference aggregation tries to find a compromise that maximises the satisfaction of most parties involved (or in other words, minimise some kind of distance between the merged result and individual information items), even if not previously suggested.

Nevertheless the tools available for information fusion and preference aggregation clearly overlap, and so do the postulates that delimit rational aggregation methods. For instance, the famous Arrow axioms of voting apply to the fusion of total orderings and some of them are similar to some of our information fusion postulates [22]. For instance, as explained above, Unrestricted Domain, Attainability, and (ordinal) Unanimity are social choice axioms that make sense for ordinal information fusion. Independence of Irrelevant Alternatives may be hard to sustain for information fusion in the face of inconsistency. But Non-Dictatorship is implied by the Fairness axiom.

As can be seen, our framework for information fusion is not the same as the framework for social choice and voting despite the relationships existing between some of their postulates. In particular, the Optimism, Minimal Commitment, and Information Monotonicity postulates of information fusion have no counterpart in voting theory.

3.5. Basic information aggregation modes

Based on the definition of information items T_i and their informational ordering, several aggregation modes can be defined, that will be instrumental for constructing fusion operations:

1. *Conjunctive operators:* a *conjunctive operator* is a commutative function $c_n : \mathcal{T}^n \to \mathcal{T}$ such that $c_n(T_1, \ldots, T_n) \sqsubseteq T_i, \forall i = 1, \ldots, n$; a *cautious* conjunctive operator is a minimally committed conjunctive operator, that is one such that $c_n(T_1, \ldots, T_n)$ is maximal for $\sqsubseteq$.
2. *Disjunctive operators:* a *disjunctive operator* is a commutative function $\delta_n : \mathcal{T}^n \to \mathcal{T}$ such that $T_i \sqsubseteq \delta_n(T_1, \ldots, T_n), \forall i = 1, \ldots, n$; an *optimistic* disjunctive operator is a maximally committed disjunctive operator, that is one such that $\delta_n(T_1, \ldots, T_n)$ is minimal for $\sqsubseteq$.

Conjunctive operators can be used for information fusion when the inputs are not mutually inconsistent (overlapping supports). Otherwise, they may fail the Consistency Property, hence Fairness, and Minimal Commitment.

Proposition 3. *When inputs are not strongly globally inconsistent (overlapping supports), cautious conjunction operators obey all basic properties of fusion.*

Proof. We use the fact that $T_1 \sqsubseteq T_2$ imply the similar inclusions between supports and between cores.

- For possibility preservation, note that the cautiousness assumption implies that $\mathcal{S}(c_n(T_1, \ldots, T_n)) = \bigcap_i \mathcal{S}(T_i)$ (otherwise $c_n(T_1, \ldots, T_n)$ would not be maximal for $\sqsubseteq$). It also implies Fairness. Impossibility preservation is obvious.
- Information monotonicity holds as if $T_i \sqsubseteq T_i'$ the condition $c_n(T_1, \ldots, T_n) \sqsubseteq T_i$ is more demanding than $c_n(T_1, \ldots, T_n) \sqsubseteq T_i'$, so that there is a minimally committed result $c_n(T_1, \ldots, T_i', \ldots, T_n)$ that is equal to or less informative than $c_n(T_1, \ldots, T_n)$.
- Consistency Enforcement: by construction, since $\mathcal{S}(c_n(T_1, \ldots, T_n)) = \bigcap_i \mathcal{S}(T_i) \neq \emptyset$. It will hold in the strong sense if the cores intersect.

[8] Independent and identically distributed random variables.

- Optimism, Commutativity, Minimal commitment, by construction.
- Insensitivity to vacuous information is obvious since the condition $c_n(T_1, \ldots, T_n) \sqsubseteq T_i = T^\top$ is a tautology. And so $c_n(T_1, \ldots, T_n) = c_{n-1}(T_1, \ldots, T_{i-1}, T_{i+1}, \ldots, T_n)$. □

Note that if the supports of inputs are globally inconsistent, then Consistency Enforcement fails for cautious conjunctive operations, hence Fairness and Minimal Commitment as well. Disjunctive operators are not optimistic (even when they are called optimistic) when the inputs are mutually consistent (overlapping cores).

Adaptive aggregation schemes can then be devised. In the case of two inputs:

- A binary *cautious adaptive aggregation operation ac* is defined by means of a cautious conjunctive aggregation c_2 and an optimistic disjunctive one δ_2:

$$ac(T_1, T_2) = \begin{cases} c_2(T_1, T_2) \text{ if } \mathcal{C}(T_1) \cap \mathcal{C}(T_2) \neq \emptyset; \\ \delta_2(T_1, T_2) \text{ otherwise.} \end{cases} \quad (1)$$

- A binary *bold adaptive aggregation operation ab* is defined by

$$ab(T_1, T_2) = \begin{cases} c_2(T_1, T_2) \text{ if } \mathcal{S}(T_1) \cap \mathcal{S}(T_2) \neq \emptyset; \\ \delta_2(T_1, T_2) \text{ otherwise.} \end{cases} \quad (2)$$

The difference between the two adaptive aggregations is the condition under which the conjunctive aggregation is applied. In the cautious case, the conjunction is used only if the inputs are strongly mutually consistent.

Proposition 4. *Cautious and bold adaptive aggregation operations satisfy the eight basic properties. A cautious one satisfies strong versions of Consistency Enforcement and Fairness (using cores).*

Proof.

- For Unanimity, if the two inputs are strongly mutually consistent, it follows from Proposition 3, as both operations reduce to cautious conjunctive ones. If the inputs are weakly mutually consistent or inconsistent, Unanimity is obvious for the cautious adaptive aggregation (since disjunctions preserve possibility and impossibility). For the bold one, if the inputs are weakly mutually consistent it reduces to a conjunction, but then $\mathcal{S}(c_n(T_1, \ldots, T_n)) = \bigcap_i \mathcal{S}(T_i)$ holds to respect maximality w.r.t. $\sqsubseteq$.
- Information monotonicity follows from Proposition 3, Consistency Enforcement is obvious and is even strong for the cautious adaptive operation.
- Optimism of the cautious adaptive aggregation is built-in if strong Consistency Enforcement is required. For the bold one, it is even more optimistic at the expense of getting a result with empty core (if cores are disjoint, then $c_2(T_1, T_2)$ has an empty core).
- Insensitivity for Vacuous information follows from the fact that $ac(T_1, T^\top) = ab(T_1, T^\top) = c_2(T_1, T^\top)$ and by Proposition 3.
- Fairness, Commutativity, Minimal commitment are obvious by construction. For the cautious adaptive aggregation, $\forall i = 1, \ldots, n, \mathcal{C}(ac(T_1, \ldots, T_n)) \cap \mathcal{C}(T_i) \neq \emptyset$ (when $ac = c_2$ this is obvious, and otherwise it is a disjunctive aggregation whose core includes all cores of all T_i's). So this is a strong form of fairness.

These results are summarised by Table 1. The case of n-ary adaptive fusion rules is considered later in the paper.

Table 1
Properties of general conjunctive, disjunctive and adaptive operations.

Properties	Cautious conjunction	Optimistic disjunction	Cautious adaptive	Bold adaptive
Unanimity	Yes	Yes	Yes	Yes
Information monotonicity	Yes	Yes	Yes	Yes
Consistency enforcement	No	Yes	Yes strong	Yes
Optimism	Yes	No	Yes	Yes
Fairness	No	Yes	Yes	Yes
Insensitivity for vacuous	Yes	No	Yes	Yes
Commutativity	Yes	Yes	Yes	Yes
Minimal commitment	No	No	Yes	Yes

4. Merging set-valued and Boolean information

Having proposed a set of general postulates for information fusion, the next step is to demonstrate the existence of fusion rules that obey these postulates. This question can be posed in the various settings that can be envisaged for representing information items. Moreover, we should compare our set of postulates with existing proposals in more specialised settings. The most elementary setting one may first consider is the one of sets, whereby any information item is a subset of possible worlds, one of which being the actual world, the simplest account of an epistemic state. This setting is important because it is the simplest information representation framework and also has connections with knowledge base merging in Boolean logic, where fusion postulates were proposed by Konieczny and Pino-Perez [68,69].

4.1. Merging set-valued information: hard constraints

Let us assume that the information items T_i are defined by classical subsets $E_i \subseteq W$ representing plain epistemic states, so that here $\mathcal{T} = 2^W \setminus \{\emptyset\}$. In this subsection, the triple $(\mathcal{S}(T), \mathcal{C}(T), \succeq_T)$ is defined as follows:

- Core and Support coincide: $\mathcal{C}(T) = \mathcal{S}(T) = E \subseteq W$.
- Plausibility ordering induced by T: $w \succ_T w'$ if $w \in E$ and $w' \notin E$, while $w \sim_T w'$ if $w, w' \in E$ or $w, w' \notin E$.

Note that this choice is not unique. One could also decide that $\mathcal{S}(T) = W$ (as in the next subsection). Instead we study here the case where any world $w \notin E$ is considered impossible (for instance, the fusion of integrity constraints).

It is clear that $\mathcal{T}$ is consistent if and only if E is not an empty set, and the information ordering relation $\sqsubseteq$ is set inclusion $\subseteq$. Then all the basic properties can be naturally adapted to the Boolean representation. In order to characterise a canonical fusion rule, though, only three axioms among the eight ones are needed to ensure uniqueness in the case of two sources [32]:

- *Optimism:* If $E_2 \cap E_1 \neq \emptyset$, then both $f(E_1, E_2) \subseteq E_1$ and $f(E_1, E_2) \subseteq E_2$ hold.
- *Unanimity*: $E_2 \cap E_1 \subseteq f(E_1, E_2) \subseteq E_2 \cup E_1$.
- *Minimal commitment*: $f(E_1, E_2)$ is the largest subset of possible worlds obeying Optimism and Unanimity.

Proposition 5. *If there are two sources, the only fusion rule that satisfies Optimism, Unanimity and Minimal Commitment is*

$$f_2(E_1, E_2) = \begin{cases} E_2 \cap E_1 \text{if } E_2 \cap E_1 \neq \emptyset \\ E_2 \cup E_1 \text{otherwise.} \end{cases} \quad (3)$$

Proof. in the consistent case, Optimism and Unamimity imply $f_2(E_1, E_2) = E_1 \cap E_2$, and otherwise, Minimal commitment and Unanimity imply $f_2(E_1, E_2) = E_1 \cup E_2$. □

It is clear that the above three axioms imply the other five ones for two sources: The intersection operation is information-monotonic, the fusion rule always yields a consistent outcome if inputs are consistent. Fairness obviously holds as the result clearly keeps track of the two information items. The above fusion rule is commutative, and is not sensitive to vacuous information items (as

the latter are always less informative than the other items in the sense of relation $\sqsubset$). It is a first example of adaptive aggregation. This framework leaves room for just one such aggregation (first advocated in [43]).

The rationale for such a fusion rule is clear: in case of consistent sources, we assume both are reliable and increase the precision accordingly; if they are mutually inconsistent, the fusion rule does not take sides and remains optimistic by assuming one source at least yields correct information. This way of tackling inconsistency is similar to variable forgetting in logical knowledge bases [70] (if W contains two elements only, $E_1 \cup E_2$ is vacuous).

This rule exhibits a discontinuous behaviour when moving from a consistent situation to an inconsistent one, since the less the two subsets overlap, the more precise the result of fusion, until a total contradiction appears and then the result suddenly becomes imprecise. However, nothing forbids independent sources to provide information items having a narrow intersection. In that respect, the set representation is too rigid, and it pleads for a more flexible representation setting where inconsistency can be a matter of degree (some approaches get rid of inconsistency in fusion problems by a similarity-based enlargement of the sets of interpretations of information items [95]).

An obvious consequence of Proposition 5 this result is that the Associativity postulate is inconsistent with the eight information fusion postulates. Indeed the canonical fusion rule they characterise in the set-valued setting is clearly not associative. For instance, suppose three sources providing $E_i, i = 1, 2, 3$ where $E_1 \cap (E_2 \cup E_3) = \emptyset$ while E_2 and E_3 are consistent. Then

- $f_2(E_1, f_2(E_2, E_3)) = f_2(E_1, E_2 \cap E_3) = E_1 \cup (E_2 \cap E_3)$
- $f_2(f_2(E_1, E_2), E_3)) = f_2(E_1 \cup E_2), E_3) = E_2 \cap E_3$

The origin of this incompatibility between Associativity and the eight postulates of information fusion lies in the Consistency Enforcement property, which implies the Unrestricted Domain postulate. The former is never satisfied by known fusion rules. For instance, Dempster rule of combination for belief functions is associative but is not defined for totally inconsistent belief functions (see Section 7). The lack of associativity forces us to directly define the n-ary counterpart of the basic fusion rule.

If the inputs are globally consistent, i.e., if $\bigcap_{i=1}^{n} E_i \neq \emptyset$, Optimism and Possibility Preservation imply $f_n(E_1, \ldots, E_n) = \bigcap_i E_i$. This is a direct consequence of Proposition 2. Let $I \subset \{1, \ldots, n\}$ be a maximal consistent subset (MCS) of information items, i.e., $T^I = \bigcap_{i \in I} E_i \neq \emptyset$ and $T^{I \cup \{j\}} = \emptyset, \forall j \notin I$. Let T^I and $T^{I'}$ be the results of the conjunctive combination of the information items given by two MCSs I and I', then $T^I \cap T^{I'} = \emptyset$, by construction.

There is a generalisation of the two-source combination rule (3) characterised by our postulates, namely the maximal-consistent subset fusion rule (MCS):

$$f_n^{MCS}(E_1, \ldots, E_n) = \bigcup_{I \in MCS(\{1,\ldots,n\})} \bigcap_{i \in I} E_i \tag{4}$$

where MCS({1,..., n}) is the set of maximal consistent subsets of sources. It was first proposed by Rescher and Manor already in 1970 [92]. This operator extends (3) to n sources. Indeed, in the case of 2 sources, $MCS(\{1, 2\}) = \{\{1, 2\}\}$ (hence the conjunction rule) or $\{\{1\}, \{2\}\}$ (hence the disjunction rule). We can prove that our axiomatic framework for set-fusion characterises the MCS fusion rule, if we take an optimistic view of the Fairness axiom:

Proposition 6. *Suppose the information items take the form of sets. A fusion rule satisfies Consistency Enforcement, Optimism, Optimistic Fairness, and Minimal Commitment if and only if it is the MCS fusion rule.*

Proof. Due to Consistency Enforcement and Fairness postulate, $f_n(E_1, \ldots, E_n) = A_1 \cup \cdots \cup A_n$ for non-empty sets $A_i \subseteq E_i$. Now due to Optimism and Optimistic Fairness, if I is a maximal consistent subset, one should select $A_i = A^I \subseteq E_i, i \in I$; it means that no set of sources strictly containing I can be assumed to be simultaneously reliable, but nothing prevents the set I from containing reliable sources only. As a consequence $f_n(E_1, \ldots, E_n) = \bigcup_{I \in MCS(\{1,\ldots,n\})} A^I$. Now due to minimal commitment, there is no reason to choose $A^I \subset \bigcap_{i \in I} E_i$, as this choice would be arbitrary. Hence $A^I = \bigcap_{i \in I} E_i$. So this is the MCS fusion rule. □

It is obvious MCS satisfies the other basic properties 1, 2, 6, 7. It is clear that the MCS fusion rule does not take into account the number of sources supplying the same information item. The duplication of sources does not affect the result of the MCS rule. It satisfies the following property [68]:

Majority-Insensitivity: if $E_1 = E_2 = \cdots = E_n = E$, then $f_{n+1}(E_0, E_1, \ldots, E_n) = f_2(E_0, E), \forall n > 1$.

Another way to circumvent the opposite drawbacks of conjunctive and disjunctive fusion modes (the former leading to contradiction as more sources are involved and the latter leading to uninformativeness), one may use so-called k-quota merging operators. A k-quota merging operator selects a subset of k sources that are jointly consistent and performs a conjunctive operation over them. If there are several such groups of k sources one may perform the union of the partial results. Using the principle of optimism, one is led to use a value of k that is as large as possible. It then comes down to selecting the subset of most numerous maximal consistent subsets of sources. Fusion rules of this kind can be found in [52,54,91]. They are intermediary between conjunctive and disjunctive fusion modes and are not majority-insensitive. This quota-merging rule satisfies most basic postulates, but it fails the fairness axiom as the information from sources other than the k selected ones is discarded; one may consider it also fails the minimal specificity requirement due to the selection process that picks the largest group of jointly consistent sources, which may sound questionable unless sources are considered independent.

4.2. The axioms of arbitration

It is interesting to compare the set-based instantiations of our general postulates to the axioms of arbitration after Liberatore and Schaerf [73]. They consider five postulates (here couched in our terminology), applied to two sets of possible worlds (however possibly empty), as basic:

- A closure assumption: $f(E_1, E_2)$ is a set (implicit in [73]).
- Commutativity.
- Possibility preservation, which here reads: $E_1 \cap E_2 \subseteq f(E_1, E_2)$.
- Optimism (as restored in this subsection).
- Consistency, expressed as $f(E_1, E_2) = \emptyset$ if and only if $E_1 = E_2 = \emptyset$.

The latter are four of our basic postulates and the first one is taken for granted by definition of the fusion operation. They propose additional postulates as follows:

- Disjunctive decomposability: $f(E_1, E_2 \cup E_3) = f(E_1, E_2)$ or $f(E_1, E_3)$ or yet $f(E_1, E_2) \cup f(E_1, E_3)$.
- Impossibility preservation, which here reads $f(E_1, E_2) \subseteq E_1 \cap E_2$
- Fairness: if $E_1 \neq \emptyset$, $f(E_1, E_2) \cap E_1 \neq \emptyset$.

Clearly the two last of these additional axioms are among our basic postulates. In the set-theoretic representation, these postulates are not independent, as shown in [73]; for instance Impossibility Preservation and Fairness are consequence of the other six ones. Moreover, they show that an arbitration operation is of the form $f(E_1, E_2) = A \cup B$ for some subsets $A \subseteq E_1$, $B \subseteq E_2$, and $f(E_1, E_2) = E_1 \cap E_2$ if the inputs are consistent. Note that if

$E_1 \cap (E_2 \cup E_3) \neq \emptyset$ then $f(E_1, E_2 \cup E_3) = E_1 \cap (E_2 \cup E_3)$ (due to Optimism and Possibility Preservation), which is of the form prescribed by Disjunctive decomposability.

Although written for two inputs, the above axioms can easily be written for n inputs but for Disjunctive decomposability. It is also clear that in the case of n inputs, arbitration takes the form $\bigcup_{I \in MCS(\{1,\ldots,n\})} A^I$, where $A^I \subset \bigcup_{i \in I} E_i$ (see the proof of Proposition 6).

The Liberatore and Schaerf postulates do not include Information Monotonicity, Insensitivity to Vacuous Information (which is a trivial property of arbitration), nor Minimal Commitment. It is clear that the only minimally committed arbitration operation is our adaptive operation (3). Note that it is Minimal Commitment that ensures the unicity of our fusion rule (3). This cautiousness assumption is absent from [73], in which it is assumed (like in the AGM revision setting) that the inputs E_i underlie more information than what is explicitly represented, under the form of weak plausibility orders $\geq_i$, so that $f(E_1, E_2) = A \cup B$ where A (resp. B) is formed by the maximally plausible elements of E_1 for $\geq_2$ (resp. E_2 for $\geq_1$). Liberatore and Schaerf [73] also mention Information Monotonicity, albeit without the restriction to consistent inputs, thus getting trivialisation results.

These results shed light on the rationale of our axiomatic setting. Our view of fusion is more a matter of displaying a cautious but useful synthesis of the information provided by several sources, than one of making a final decision that would be the result of a choice between sources. Moreover, it implicitly assumes that if several sources deliver the same hard constraint, it means they may be dependent. As argued above, this assumption is sometimes all the more plausible as the number of sources is high if sources are experts.

When sources provide observations resulting from processes *designed* to be independent, as in statistics, repeated information matters. For instance, one can count the number of sources that state $E_i = E$, say n_E. In the latter case one would get a belief function with focal sets E_i, and a mass function $m(E) = \frac{n_E}{n}$, which is a kind of fusion rule which does not respect the closure property, since then $f_n(E_1, \ldots, E_n) = m \notin \mathcal{T}$. Note that the contour function $\pi_m(s) = \sum_{s \in E} m(E)$ consists in averaging the characteristic functions of the E_i, i.e., $\pi_m(s) = \frac{|\{i : s \in E_i\}|}{n}$, which is intermediary between conjunction and disjunction. It gives a fuzzy set as the result of the fusion. However, if the result of the fusion of sets should be a set, there is hardly any way of expressing a reinforcement effect due to identical and independent information items.

The MCS rule is in this sense a cautious combination rule. The fact that we retrieve a well-known, probably the oldest approach to handling inconsistency in logic-based representations comforts the idea that these postulates are natural. However, it prompts us to compare this setting with the one of knowledge-based merging, that seems to argue against fusion methods obeying Impossibility Preservation.

4.3. Logic-based merging: the fusion of plausible sets

Knowledge base merging addresses how a set of logical formulae in propositional logic obtained from different agents should be merged in order to obtain a consistent set of formulae.

A commonly accepted set of postulates for judging a logic-based merging operator was proposed in [68]. We consider a propositional language $\mathcal{L}_{\mathcal{PS}}$ defined on a finite, non-empty set $\mathcal{P}$ of propositional atoms, which are denoted by p, q, r etc. A proposition (or formula) ϕ is constructed by propositional atoms with logical connectives $\neg$, $\wedge$, $\vee$, $\rightarrow$ in the standard way. An interpretation w (or possible world) is a function that maps $\mathcal{P}$ onto the set $\{0, 1\}$. The set of all possible interpretations on $\mathcal{P}$ is denoted by W. Function w can be extended to any propositional sentence in $\mathcal{L}_{\mathcal{P}}$ in the usual way, $w : \mathcal{L}_{\mathcal{P}} \rightarrow \{0, 1\}$. An interpretation w is a model of (or satisfies) ϕ if and only if $w(\phi) = 1$, denoted by $w \models \phi$. We use $Mod(\phi)$ to denote the set of models for ϕ.

A *(flat) knowledge base* K is a finite set of propositions. K can be equivalently expressed as a formula ϕ consisting of the conjunction of formulas in K. A knowledge profile is a multi-set of propositional formulas $\mathcal{E} = \{\phi_1, \phi_2, \ldots, \phi_n\}$. The conjunction of formulas in $\mathcal{E}$ is $\bigwedge \mathcal{E} = \phi_1 \wedge \cdots \wedge \phi_n$. $\mathcal{E}$ is called consistent if and only if $\bigwedge \mathcal{E}$ is consistent. $\mathcal{E}_1 \leftrightarrow \mathcal{E}_2$ denotes that there is a bijection g from $\mathcal{E}_1 = \{\phi_1^1, \ldots, \phi_n^1\}$ to $\mathcal{E}_2 = \{\phi_1^2, \ldots, \phi_n^2\}$ such that $\vdash f(\phi) \leftrightarrow \phi$. We denote by $\mathcal{E}_1 \bigsqcup \mathcal{E}_2$ the multi-set union of $\mathcal{E}_1$ and $\mathcal{E}_2$.

An operator Δ is a mapping from knowledge profiles to knowledge bases, and Δ is a *KP-merging operator* (after Koniecny and Pino-Pérez [68]) if and only if it satisfies the following postulates [68].

A1 $\Delta(\mathcal{E})$ is consistent.
A2 If $\mathcal{E}$ is consistent, then $\Delta(\mathcal{E}) = \bigwedge \mathcal{E}$.
A3 If $\mathcal{E}_1 \leftrightarrow \mathcal{E}_2$, then $\vdash \Delta(\mathcal{E}_1) \leftrightarrow \Delta(\mathcal{E}_2)$.
A4 If $\phi_1 \wedge \phi_2$ is not consistent, then $\Delta(\{\phi_1, \phi_2\}) \nvdash \phi_1$.
A5 $\Delta(\mathcal{E}_1) \wedge \Delta(\mathcal{E}_2) \vdash \Delta(\mathcal{E}_1 \bigsqcup \mathcal{E}_2)$.
A6 If $\Delta(\mathcal{E}_1) \wedge \Delta(\mathcal{E}_2)$ is consistent, then $\Delta(\mathcal{E}_1 \bigsqcup \mathcal{E}_2) \vdash \Delta(\mathcal{E}_1) \wedge \Delta(\mathcal{E}_2)$.

These postulates are described at the syntax-level, e.g., formula-level, whilst the common properties we proposed above are represented at the semantic level, e.g., possible worlds level. However, since axiom A3 assumes syntax independence, these postulates can be written at the semantic level in terms of merging sets of possible worlds, which is the representation setting of this section, without loss of content. Let $E_i = Mod(\phi_i)$ be the set of models of ϕ_i and we use notation f^Δ instead of Δ to denote a fusion operation obeying the KP axioms.

Then the KP postulates (minus A3, trivially true) can be equivalently stated as:

A1* $f^\Delta(E_1, \ldots, E_n) \neq \emptyset$.
A2* If $E_1 \cap \cdots \cap E_n \neq \emptyset$, then $f^\Delta(E_1, \ldots, E_n) = E_1 \cap \cdots \cap E_n$.
A4* If $E_1 \cap E_2 = \emptyset$, then $f^\Delta(E_1, E_2) \subsetneq E_1$.
A5* $f_n^\Delta(E_1, \ldots, E_n) \cap f_m^\Delta(E'_1, \ldots, E'_m) \subseteq f_{m+n}^\Delta(E_1, \ldots, E_n, E'_1, \ldots, E'_m)$.
A6* If $f_n^\Delta(E_1, \ldots, E_n) \cap f_m^\Delta(E'_1, \ldots, E'_m) \neq \emptyset$, then $f_{m+n}^\Delta(E_1, \ldots, E_n, E'_1, \ldots, E'_m) \subseteq f_n^\Delta(E_1, \ldots, E_n) \cap f_m^\Delta(E'_1, \ldots, E'_m)$.

A1*–A6* can be interpreted as follows in the light of our setting. A1* is the Consistency Enforcement axiom. A2* is implied by our information fusion postulates, after Proposition 2. It is a consequence of Optimism and Possibility Preservation: it presupposes that consistent sources are reliable. A4* is a consequence of our Fairness axiom (Proposition 1). It does not prevent the result from being inconsistent with ϕ_1 and ϕ_2, should they be inconsistent. This is the main difference with arbitration operators. A5* and A6* together state that if it is possible to find two subgroups which agree on at least one possible world, then the result of global fusion will contain exactly those possible worlds the two subgroups agree on.

The KP framework is thus different from ours. The axioms A1 to A4 have set-valued counterparts that are consequences of our setting, but A5* and A6* are not presupposed by our framework. In fact, as the MCS rule is characteristic of our postulates, and since it satisfies the Majority-Insensitivity property, it follows that it does not satisfy A5* and A6*. Indeed Konieczny and Pino-Perez [68] indicate that the Majority-Insensitivity property is incompatible with their axiomatic setting A1–A6 (and they propose a weak version of it).

So, how do they define merging operators? In fact, the use of A5 and A6 forces them into a more expressive framework than the

one of sets, namely, the one of partially ordered epistemic states. A subset of possible worlds viewed as an epistemic state E_i is interpreted as a so called *syncretic assignment*, which is a partial plausibility order $\succeq_i$ on W attached to E_i such that all maximal elements for this plausibility order are those of E_i. In fact, we are back to the triple $T_i = (\mathcal{S}(T_i), \mathcal{C}(T_i), \succeq_i)$ where conditions are expressed in our setting as follows:

- $\mathcal{C}(T_i) = E_i = \max_{\succeq_i} W$. In particular $\mathcal{C}(T_i)$ is never empty.
- $\mathcal{S}(T_i) = W$

The resulting partial plausibility ordering $\succeq_T$ is obtained by merging the local preorders $\succeq_i$ and the epistemic state E resulting from the fusion is $E = \mathcal{C}(T) \neq \emptyset$, such that

- If $\mathcal{C}(T_1) \cap \mathcal{C}(T_2) = \emptyset$ then $\mathcal{C}(T) \not\subseteq \mathcal{C}(T_i), i = 1, 2$ (this is reflecting A4); one may have that $\mathcal{C}(T) \cap (\mathcal{C}(T_1) \cup \mathcal{C}(T_2)) = \emptyset$.
- If $\mathcal{C}(T) \cap \mathcal{C}(T') \neq \emptyset$, where T and T' are the results of merging the profiles $\mathcal{E}$ and $\mathcal{E}'$, then $\mathcal{C}(\Delta(\mathcal{E}, \mathcal{E}')) = \mathcal{C}(T) \cap \mathcal{C}(T')$ (this is reflecting A5 and A6)

The authors prove that any fusion operation satisfying A1–A6 can be obtained by means of merging syncretic assignments associated to the sets E_i and the result of the fusion is the core of the obtained global partial ordering $\succeq_T$.

Note that Axioms A5 and A6 are instrumental in proving the existence of syncretic assignments. These axioms (like those in the AGM theory of revision [2]) are directly inspired by choice function theory (see [16] for such a connexion) which defines axioms for the selection of best items from subsets of options, such that there exists a preference relation that can justify these choices. One may argue that the role of such axioms is more that of justifying a particular representation of plausibility via orderings, than expressing key-properties of information fusion (A5 and A6 are also at work as postulates 7 and 8 in the AGM axiomatic setting for revision). In any case, the merging of sets according to Koniecny and Pino-Perez uses a richer framework than the one of sets. But the plausibility ordering does not appear in the KP axioms.

Since the knowledge merging in the KP style uses information items of the form $(\mathcal{S}(T_i), \mathcal{C}(T_i), \succeq_i) = (W, E_i, \succeq_i)$, it is clear that the subset of possible worlds E_i in the KP setting is then viewed as a most plausible set and no longer as a hard constraint. The absence of any impossible world in the KP setting makes some of our axioms of information fusion trivially satisfied under such inputs (for instance, Properties 1, 5). In order to account for this interpretation of E_i in our setting, we could revise some of our axioms as follows:

- *Unanimity of Plausibility*:
  - **(a)** $\bigcap_{i=1}^{n} \mathcal{C}(T_i) \subseteq \mathcal{C}(f(T_1, \ldots, T_n))$.
  - **(b)** $\mathcal{C}(f(T_1, \ldots, T_n)) \subseteq \mathcal{C}(T_1) \cup \cdots \cup \mathcal{C}(T_n)$.
- *Strong Consistency Enforcement*: $\mathcal{C}(f(T_1, \ldots, T_n)) \neq \emptyset$.
- *Fairness* : $\forall i = 1, \ldots, n, \mathcal{C}(f(T_1, \ldots, T_n)) \cap \mathcal{C}(T_i) \neq \emptyset$.

Interestingly, put in this form, these axioms no longer sound as compelling as their counterparts in terms of support. While part (a) of Unanimity is sanctioned by the KP axiom A2 and is hard to challenge, part (b) looks very demanding. Indeed, one may sometimes find it intuitively satisfactory if the most plausible worlds after fusion do not lie in $\mathcal{C}(T_1) \cup \cdots \cup \mathcal{C}(T_n)$ (for instance if the most plausible worlds for one source are all found to be very implausible for another one, while there are worlds that are reasonably plausible for all sources). As a consequence the above form of the Fairness axiom also sounds too demanding. The Strong Consistency Enforcement is also endorsed by the KP framework (this is A1) but it may look too strong if one would like to leave room for partial inconsistency.

However, it is clear that these postulates along with the five other ones would yield the MCS rule on plausible sets, which, in the Boolean setting, may look less plausible than the MCS rule on hard constraints. In particular, it rules out the kind of trade-off fusion rules envisaged by Konieczny and Pino-Perez, where the core of the result of the fusion may be disjoint from the cores of the inputs.

In order to lay bare concrete examples of merging rules, these authors appeal to numerical encodings of their syncretic assignments, using a syntax-based Hamming distance between logical interpretations. In other words, they do not merge sets of models, but fuzzy sets thereof, where the degrees of membership of an interpretation w to a formula ϕ_i reflects the minimal Hamming distance to a model of ϕ_i. This way of envisaging knowledge-based fusion is explained in more details by Benferhat et al. [7], who show it comes down to merging possibilistic logic bases (see Section 6.3).

The above discussion also lays bare the difference of perspectives between the fusion of hard constraints and knowledge-base merging: the idea of Konieczny and Pino-Perez is to explain the fusion of plain epistemic states, understood as a set of plausible worlds, by the existence of underlying partial orderings or numerical plausibility degrees (obtained by distances), based on axioms that only use plausible sets attached to these orderings. In [67] the same authors use both hard (integrity) constraints and belief sets referring to plausible worlds, and try to extend both the AGM revision and knowledge-based merging. However, they do not envisage the merging of integrity constraints discussed in the previous section. The belief revision and merging literature takes an external point of view on cognitive processes under study. The underlying ordered structures are here a consequence of the merging postulates, but they do not appear explicitly in the axioms and they are not observable from the outside. On the contrary, our approach is to construct fusion rules that only rely on what is *explicitly supplied* by sources. In the sequel we consider the counterpart of our fusion postulates for ranked models, that can be expressed by means of total orders of possible worlds or by their encodings on a plausibility scale.

5. Merging of ranked epistemic states: ordinal setting

In this section we study how to adapt the eight postulates of information fusion to refined epistemic states, whereby some possible worlds are more plausible than others. The various representations of likelihood where the whole information is contained in the plausibility ordering over possible worlds can be captured under the umbrella of possibility theory. In the purely ordinal case comparative possibility was introduced by Lewis [72] and independently by Dubois [31]. From the late 1980's onward, ordinal plausibility orderings are found in the theory of belief revision after Grove [58], Gärdenfors [56], Katsuno and Mendelzon [63]. In the later works, though, plausibility orderings are the result of the axiom systems adopted for belief revisions, and do not stand as primitive explicit available data.

5.1. Ordinal representations

In this subsection, we suppose information items T are explicitly defined by complete preorders $\succeq_T$ on W, following Lewis, Dubois and Grove. They should be understood as plausibility rankings on interpretations or states, and can be equivalently described by well-ordered partitions $\mathcal{P}_T = \{A_1, \ldots, A_k\}$ such that

$$\forall w_i \in A_i, w_j \in A_j, w_i \succ_T w_j \iff i < j$$

means that w_i is a more plausible world than w_j. Internal consistency of such information items can be interpreted by the acyclicity of the strict relation $\succ_T$, which holds if we start with complete preorders. This is a strong form of internal consistency.

Let us denote by R_T the subset of $W \times W$ corresponding to $\succeq_T$. In contrast to strong internal consistency, strong internal inconsistency would correspond to $R_T = \emptyset$ (all pairs (w_1, w_2) consist of incomparable states), which is made impossible by assuming complete preorders. The conventions coming from knowledge-based fusion define the core $\mathcal{C}(T)$ as $\max_{\succeq_T} W$ irrespective of how much plausible are the most plausible worlds. The support is the whole set of states, none of which is assumed to be impossible.

As seen above, the KP knowledge-based merging approach underlies plausibility orderings $\succeq_i$ that are not explicitly used in the postulates, but whose existence is a consequence of them. Here, we assume these are explicitly handled as complete preorders (or rankings). KP axioms only bear on the plausible sets $E_i = \mathcal{C}(T_i)$ induced by abstract information items. It comes down to a kind of renormalisation, bringing the most plausible worlds for each $\succeq_i$ at the same plausibility level to make them minimally commensurate. These plausible sets are considered to be the "visible part" of the information items.

For instance, in the KP approach, the information ordering $T_1 \sqsubseteq T_2$ between information items is expressed by inclusion between plausible sets $\mathcal{C}_1(T) \subseteq \mathcal{C}_2(T)$, irrespective of the actual underlying rankings. One may think of more demanding notions of entailment between possibility rankings. Here are two of them [9]:

- *Refinement ordering:* $\succeq_1 \sqsubseteq \succeq_2$ if and only if $\mathcal{P}_1$ refines $\mathcal{P}_2$;
- *Specificity ordering* [50,88], letting k and ℓ be the number of elements of the partitions induced by $\succeq_1$ and $\succeq_2$ respectively: $\succeq_1 \sqsubseteq_s \succeq_2$ if and only if $\forall i = 1, \ldots, \min(k, \ell), \bigcup_{j=1\ldots i} A_j^1 \subseteq \bigcup_{j=1\ldots i} A_j^2$

The refinement ordering introduces a natural mapping between elements of the well-ordered partitions associated to ordinal information items, as the equivalence classes of the coarser partition are unions of elements of the finer one.

The specificity ordering (put forward by Pearl [88] as favouring the most compact ranking) is weaker than refinement but it introduces a systematic commensurateness assumption between the weak order relations, whereby A_j^1 is mapped to A_j^2, for $j = 1, \ldots, \min(k, \ell)$, namely it presupposes that in each pair (A_j^1, A_j^2) the sets are equally likely. However, there exists a dual notion presupposing that A_k^1 and A_ℓ^2 are equally unlikely, and A_{k-j}^1 and $A_{\ell-j}^2$ as well, for $j = 1, \ldots, \min(k, \ell) - 1$. The choice between these two ordering comparison methods may look sowewhat arbitrary. So, in the purely ordinal setting, the refinement ordering looks like the least controversial as not involving commensurateness. Nevertheless the specificity ordering is used in ordinal approaches to non-monotonic reasoning like Pearl's system Z [88], or the semantic account of preferential inference in terms of linear possibility distributions [9], as it fits the natural informational ordering of possibility theory [50].

Likewise, in the KP setting, two information items T_1 and T_2 are said to be consistent if $\mathcal{C}(T_1) \cap \mathcal{C}(T_2) \neq \emptyset$. In contrast, a much more demanding form of consistency between rankings, that does not involve a commensurateness assumption, is as follows:

$\succeq_1$ is *strongly consistent* with $\succeq_2$ if and only if $\nexists w_1, w_2, w_1 \succ_1 w_2, w_2 \succ_2 w_1$.

It means that there is no preference reversals between T_1 and T_2 (this is often called comonotonicity). At the opposite, a very weak consistency requirement can be expressed by the existence of at least one pair of possible worlds for which the two sources agree:

$\succeq_1$ is *weakly consistent* with $\succeq_2$ if and only if $\exists w_1, w_2, w_1 \succeq_1 w_2, w_1 \succeq_2 w_2$.

It is weaker than the consistency condition in the KP setting, since the latter is weak consistency enforced to $w_1 \in \mathcal{C}(T^1) \cap \mathcal{C}(T^2)$.

In voting theory, well-known axioms have been proposed for aggregating complete preorderings into complete preorderings. Some have been already mentioned in the previous section: universality (UNI), local ordinal unanimity (LOU) and independence of irrelevant alternatives (IIR). Of course, together with non-dictatorship they lead to the celebrated Arrow's impossibility theorem from voting theory [3], saying that dictatorship is then the only possible fusion rule, namely the result of merging several complete preorders is necessarily one of them. Lehmann and Maynard-Reid [79] try to bypass the impossibility result by relaxing the complete preordering representation. They use modular relations, where transitivity is changed into modularity: if $w \succeq w'$ then $w \succeq z$ or $z \succeq w'$. Such relations also lead to well-ordered partitions, but they are no longer acyclic, and allow circuits inside equivalence classes. In the case of sources treated on a par, Lehmann and Maynard-Reid propose as the basic rational fusion rule the set theoretic union of the strict part of (modular) relations $\succeq_i$. The obtained relation is modular but not necessarily transitive.

However, it is not clear why we should adopt the axioms of voting theory for information fusion. Clearly, voting theory uses no concept of information ordering, nor any form of minimal commitment. The IIR property is not always satisfied by the fusion rules in the set-based case (e.g., when sources are inconsistent), as clear from the role of MCS fusion rule in the previous section. However, it sounds natural to rely on plausibility orderings for representing information items and to share some axioms with voting theory. For instance, as seen earlier, Universality is a consequence of our framework.

5.2. The basic postulates for the fusion of plausibility orderings

So, let us try to instantiate the eight postulates on the fusion of complete preorders $\succeq_i, i = 1, n$. The use of complete preorders corresponds to the idea that input items of information should be strongly internally consistent. We denote by $\sim_i$ the equivalence relation associated with $\succeq_i$ (that is, $w_1 \sim_i w_2 \iff w_1 \succeq_i w_2$ and $w_2 \succeq_i w_1$). In order to protect the approach against Arrow's impossibility theorem, we shall assume the result of the fusion need not be a complete preorder: some states may be incomparable in the final result (denoted by $w_1 \nsim_i w_2$). This is the price paid for not making any commensurateness assumption. In this setting our axioms can take the following form:

1. *Local ordinal unanimity.* As defined earlier: If $\forall i, w \succeq_i w'$, then $w \succeq w'$.
2. *Information monotonicity.* If $\succeq_i^1$ refines $\succeq_i^2, \forall i = 1, n$ then the result of aggregating $\succeq_i^1$ should refine the result of aggregating $\succeq_i^2$, if the relations $\succeq_i^1$ are strongly mutually consistent.
3. *Consistency enforcement.* The result must be at least weakly consistent.
4. *Optimism.* If $w_1 \succ_i w_2$ for one source i, then $w_1 \succ w_2$ if $w_2 \succ_j w_1$ for no other source j.
5. *Fairness.* If after fusion $w_2 \succeq w_1$, then $\exists i : w_2 \succeq_i w_1$.
6. *Insensitivity to vacuous information.* If $R_i = W \times W$ then source i does not influence the result of the fusion.
7. *Commutativity.*
8. *Minimal commitment.* The obtained plausibility relation should be the least refined among those that satisfy the above postulates.

Regarding the unanimity postulate, it seems that it is very natural to conclude that w is more plausible than w' if all sources say so, and likewise for equally plausible. Note that this is true regardless of other alternatives. Property 3 copes with the impossibility theorem of Arrow to some extent by allowing some incomparabilities to take place in the resulting plausibility relation. However, in the case of two sources providing opposite linear rankings

of states, all options are incomparable, and there is no aggregation method obeying the 8 postulates. The above form of the optimism postulate tries to exploit all reasons to consider one state more plausible than another. The fairness property as formulated here is minimally demanding: the final comparability between any two states w_1 and w_2 corresponds to the opinion of at least one source. Property 6 is not acceptable in the scope of voting theory and Property 8 makes no sense in that setting.

The fusion rule enforced by all required properties but Property 3 can be summarised by the following procedure for each pair (w_1, w_2).

- If $\forall i, w_1 \succeq_i w_2$, and $\nexists j, w_1 \succ_i w_2$, then $w_1 \succeq w_2$ (unanimity).
- If $\forall i, w_1 \succeq_i w_2$, and $\exists j, w_1 \succ_j w_2$, then $w_1 \succ w_2$ (optimism).
- Otherwise $w_1 \sim w_2$ (incomparability).

This fusion rule is just the Pareto vector ordering (PAR) applied to all n complete preorders. It is easy to check that other postulates are satisfied. Information monotonicity holds because making one preference relation coarser will not create any inconsistency with other ones if the more refined relation was strongly consistent with the other ones. Fairness hold as at worst $\forall i, \exists w_1, w_2$ such that w_1 and w_2 will be incomparable, hence not contradicting $w_1 \succ_i w_2$. Insensivity to vacuous information is here insensivity to full indifference, that cannot affect unanimity and optimism. Minimal commitment is due to the fact that this fusion rule does not solve outright conflicts and proposes incomparability in this case. Note that the resulting relation may be very partial but it is transitive. In the case where the sources are not even globally weakly consistent, that is $\nexists w_1, w_2, \forall i, w_1 \succeq_i w_2$, then $w_1 \sim w_2, \forall w_1, w_2 \in W$, which contradicts Property 3.

Other combination rules for partial orders exist in the literature. For instance

- Lehmann and Maynard-Reid [79] suggest to compute the union of the strict parts of $\succeq_i$: in terms of combining subsets of $W \times W$, the strict part $\succ_{MR}$ of the result reads
$$R^{\succ}_{MR} = \bigcup_{i=1}^{n} (R_i \cap (R_i^{-1})^c),$$
where $(w_1, w_2) \in R_i^{-1}$ if and only if $w_2 \succeq_i w_1$, and $R_i \cap (R_i^{-1})^c$ represents the strict part of $\succeq_i$.
- The likely dominance rule [33]:
$w_1 \succ_{LD} w_2 \iff [w_1 > w_2] >_\iota [w_2 > w_1]$;
$w_1 \sim_{LD} w_2 \iff [w_1 \geq w_2] =_\iota [w_2 \geq w_1]$,
where $[w_1 > w_2] = \{i : w_1 \succ_i w_2\}$, and $\geq_\iota$ is a partial preordering on subsets I, J of sources, representing their relative reliability.

Lehmann and Maynard-Reid [79] propose their fusion rule for merging epistemic states represented by transitive and modular relations. This is a looser framework than complete preorderings. The strict part of such relations still rank-orders a partition. Elements of a partition may be either mutually indifferent or conflicting. The result of the fusion may fail to be transitive and the authors suggest to compute its transitive closure. The likely dominance rule is the one that remains in decision under non-commensurate uncertainty and utility, when the first five Savage axioms of decision theory (see [51] Chapter 9) are preserved, but for the completeness and transitivity of the preference over acts.

We can show the following result:

Proposition 7. *When merging complete preorderings, the Pareto fusion rule is equivalent to the intersection $\bigcap_{i=1}^{n} R_i$ of $\succeq_i$. Moreover, its strict part is equivalent to*

1. *The union of the strict parts of $\succeq_i$: $R_{PAR} = (\bigcup_{i=1}^{n} R_i \cap (R_i^{-1})^c)^c$*
2. *The likely dominance rule when groups of sources have the same reliability: $I >_\iota J$ if and only if $I \neq \emptyset = J$, and $I =_\iota J$ otherwise.*

Proof. Suppose $w_1 \succeq_{PAR} w_2$ by the Pareto fusion rule. It is equivalent to $w_1 \succeq_i w_2, \forall i$. It is equivalent to $(w_1, w_2) \in \bigcap_{i=1}^{n} R_i$. Now $w_1 \succ_{PAR} w_2$, if, moreover, there is some source i with $w_1 \succ_i w_2$, which corresponds to the disjunction of strict parts of $\succeq_i$. The case of incomparability $w_1 \sim_{PAR} w_2$ corresponds to when $w_1 \succ_i w_2$ and $w_2 \succ_j w_1$ for some i, j, w_1, w_2, whereby both $w_1 \succ_{MR} w_2$ and $w_2 \succ_{MR} w_1$ hold. For the third property, note that under the proposed assumptions, $w_1 \succ_{LD} w_2 \iff [w_2 > w_1] = \emptyset$ and $[w_1 > w_2] \neq \emptyset$; It means $w_1 \succ_i w_2$ for some i and $w_1 \succeq_j w_2$ for $j \neq i$. This is $w_1 \succ_{PAR} w_2$. □

In the case where the obtained relation is strongly inconsistent, one way to escape the trival result of a strongly inconsistent result is to attach priorities to sources, as done by Maynard-Reid and Shoham [80]. The likely dominance rule does the same by attaching priorities to groups of sources, but the resulting relation may include cycles [33]. Alternatively, we may resort to a counterpart of a MCS fusion rule. It consists of the following steps

- Find all the maximal subsets of weakly consistent sources for the Pareto fusion rule (i.e., maximal subsets I_k such that $R_{PAR(\{\succeq_i : i \in I_k\})} \neq \emptyset$).
- $MCS(\succeq_1, \ldots, \succeq_n) = \bigcup_k R_{PAR(\{\succeq_i : i \in I_k\})}$

This technique could avoid violating Property 3, but may not yield a transitive relation any longer.

In summary, the issue of merging epistemic states represented by plausibility relations on states of affairs has received only little attention, compared to the same problem where the relations represent preferences of individuals in a group. Preliminary results suggest that one cannot just apply results from voting theory to the setting of belief merging, as the working assumptions in both areas seem to differ, even if they overlap. There may be specific approaches to belief merging likely to circumvent impossibility results inherited from the voting area, that may be of interest only for belief merging.

In particular, one may replace the notion of refinement between ordinal epistemic states with the specificity ordering. However, as discussed earlier, it comes down to restoring commensurateness between sources that is absent from the purely ordinal view. The specificity-based approach is more naturally addressed in the framework of possibility theory, using a common scale for sources of information.

6. Merging in possibility theory

Possibility theory is one of the main theories for reasoning under uncertainty due to incomplete information [44]. It is a flexible framework for merging information because set-based fusion modes can be directly extended to fuzzy sets representing items of incomplete information in a gradual way. An extensive overview of fusion methods in possibility theory appears in [52].

Possibility theory has several variants [50], some being qualitative (ordinal, as in the previous subsection, or classificatory), some being numerical such as the theory of kappa functions by Spohn [104,106]. Instead of using a purely relative notion of plausibility, it may be convenient to use a plausibility scale L, in the form of a bounded totally ordered set. The primary concept in scaled possibility theory is the *possibility distribution* π on W, a mapping with domain W taking values on L. It rank-orders interpretations in terms of plausibility ($w \succeq_T w'$ if $\pi(w) \geq \pi(w')$, albeit using absolute ratings). Moreover, it also enables landmark possibility values to be defined: the bottom 0 of L refers to complete impossibility while 1 refers to full plausibility (no impediment to the realisation of an event). In particular, the normalisation condition for possibility distributions is that $\pi(w) = 1$ for some $w \in W$.

Pioneers of possibility theory are the English economist Shackle [96] who interpreted plausibility in terms of lack of surprise, the

philosopher Lewis [72], who introduced comparative possibility relations in order to provide a semantics to his logic of counterfactuals, and the professor of electrical engineering Zadeh [118], who related degrees of possibility and membership functions of fuzzy sets in the setting of natural language understanding, for the purpose of modelling the extension of gradual predicates. More recently, numerical degrees of possibility were understood as upper bounds of probability degrees [48]. Independently, Spohn [104] introduced an integer-valued theory of plausibility, where higher integers reflect more implausible situations and are viewed as exponents of infinitesimal probabilities [105].

6.1. Plausibility scales

All representations of epistemic states by possibility distributions do not have the same expressive power since the plausibility scale can be numerical or not. In fact we can distinguish between several representation settings according to the expressiveness of the scale used. There are four kinds of scales one may envisage in increasing order of expressiveness: qualitative (finite or not), integer-valued, or real-valued [10].

1. The qualitative finite or classificatory setting (QUALFI for short) with possibility degrees lying in a finite totally ordered scale: $L = \{\alpha_0 = 1 > \alpha_1 > \cdots > \alpha_{m-1} > \alpha_m = 0\}$. This setting is used in possibilistic logic [36]. The values α_i can be viewed as just denoting ranked class names, from the class *fully plausible* (α_0) to the class *impossible*.
2. The dense ordinal setting (DORD for short) using $L = [0, 1]$, seen as an ordinal scale. In this case, the possibility distribution π is defined up to any monotone increasing transformation $f : [0, 1] \to [0, 1], f(0) = 0, f(1) = 1$. This setting is also used in possibilistic logic [36].
3. The denumerable setting (DENUM for short), using a scale made of successive powers $L = \{\alpha^0 = 1 > \alpha^1 > \cdots > \alpha^i > \cdots 0\}$, for some $\alpha \in (0, 1)$[9]. This scale is quite expressive as it is equipped with semi-group operations min, max, product, and also division. This is isomorphic to the use of integers in so-called κ-functions by Spohn [104]. It is used in belief merging by Chopra et al. [22].
4. The dense absolute setting (DABS for short) where $L = [0, 1]$, seen as a genuine numerical scale equipped with product. In this case, a possibility measure can be viewed as special case of a Shafer [97] plausibility function, actually a consonant plausibility function, and $1 - \pi$ a potential surprise function in the sense of Shackle [96].

The possibility theory framework is a graded extension of the set-theoretic setting. An item of information T expressed in the form of a possibility distribution π has support $\mathcal{S}(\pi) = \{w : \pi(w) > 0\}$ and core $\mathcal{C}(\pi) = \{w : \pi(w) = 1\}$ (which is not empty as long as the possibility distribution is normalised), and naturally defines a plausibility ordering on W. The information ordering between possibility distributions is relative specificity ($\pi_i \sqsubseteq \pi_j \Leftrightarrow \pi_i \leq \pi_j$), which differs from the refinement ordering in the previous section for comparative possibilities. This is a major difference with the ordinal setting, as here possibility distributions are commensurate.

Like sets, uncertain inputs represented by possibility distributions can be combined by means of aggregation operators in two main modes: conjunctive or disjunctive [49] with the same rationale on the assumed reliability of sources. According to the nature of the plausibility scale, basic operations that are instrumental for fusion may differ. Conjunctions and disjunctions are defined as follows:

[9] As usual α^i stands for the ith power of α.

Definition 1. A conjunction on L is a binary mapping $*$: $L \times L \mapsto L$ such that

1. If $x \leq y$ then $x*z \leq y*z$;
2. If $x \leq y$ then $z*x \leq z*y$;
3. $0 * 0 = 0 * 1 = 1 * 0 = 0$ and $1 * 1 = 1$;
4. $*$ is commutative.

It is a monotonic binary operation that coincides with a Boolean conjunction. A disjunction operation $\oplus$ is defined likewise replacing the third condition by $0 \oplus 0 = 0; 0 \oplus 1 = 1 \oplus 0 = 1 \oplus 1 = 1$. As previously, conjunctive operators are instrumental to combine (especially strongly) consistent inputs while disjunctive operators are useful for the combination of inconsistent uncertain inputs.

In the QUALFI setting the most usual conjunctions and disjunctions are min and max. They satisfy the De Morgan laws with respect to the negation operation $\nu(\alpha_i) = \alpha_{m-i}$. They are associative. Due to the qualitative nature of this scale we restrict to such connectives. The DORD setting only allows for such connectives as well. Operations min and max are also exchanged via any decreasing and involutive mapping ν: $[0, 1] \to [0, 1]$ as $\nu(\max(\nu(\alpha), \nu(\beta)) = \min(\alpha, \beta)$.

In the DASB settings the most usual conjunctions and disjunctions are respectively continuous triangular norms (t-norms, for short) and co-norms [65]. They are fuzzy set connectives that extend the intersection and the union of sets. On top of min and max, examples of associative conjunctive operators include, *product* ($\alpha \cdot \beta$) and *linear product* ($\max(0, \alpha + \beta - 1)$) and examples of disjunctive operators, are the *probabilistic sum* ($\alpha + \beta - \alpha \cdot \beta$) and the *bounded sum* ($\min(1, \alpha + \beta)$), where α and β stand for possibility degrees for state w according to two agents. These additional operations do not make sense on classificatory or ordinal scales.

While conjunctions min and product belong to the DENUM setting, some operations are not definable. For instance, there is no order-reversing map on the set of integers, hence we cannot define a non-idempotent disjunction associated to the probabilistic sum of terms of the form α^i, α^j, that is a non-idempotent dual to the sum $i + j$ of integers, while $\min(i, j)$ and $\max(i, j)$ make sense.

Non-idempotent conjunctions make sense if we can assume independence between coherent sources, which justifies a property such as $\pi_1 * \cdots * \pi_n < \pi_i$ for the conjunction of information items, expressing a reinforcement effect toward impossibility. On the contrary, if the dependence between sources is unknown, the use of idempotent connectives (min and max) looks appropriate [49]. It could lead to an additional postulate that would make sense on numerical plausibility scales introducing an assumption about source independence that is absent from the basic postulates.

Remark 1. The scale classification at the beginning of this subsection stands in contrast with the one usual in measurement theory [76], distinguishing between ordinal scales (DORD), ratio scales (the positive real line up to a positive multiplicative factor), and interval scales (the real line up to a positive affine transformation). The two latter scales are generally not used for representing degrees of uncertainty (more usually for degrees of utility). Nevertheless, let us mention Ma and Liu [77] who use representations of epistemic states as a mapping $\rho : W \to \mathbb{Z}$, the set of relative integers (including $\pm\infty$) that is insensitive to integer translations (for any constant k, ρ and $\rho + k$ represent the same information item). The degree $\rho(w)$ represents a degree of relative plausibility of w with the understanding that only the difference $\rho(w) - \rho(w')$ reflects the extent to which w is more plausible than w', while $\rho(w)$ gives no clue on the absolute plausibility strength. In such a setting, the natural fusion rule is the pointwise addition $\rho + \rho'$, that preserves the equivalence of relative plausibility functions up to an additive constant. However, we cannot use max nor min that fail to preserve it. One can construct

a Spohn ordinal conditional function from a relative plausibility function ρ as: $\kappa_\rho(w) = \max_{w' \in W} \rho(w') - \rho(w)$. It can be checked that $\kappa_{\rho+\rho'} = \kappa_\rho + \kappa_{\rho'} - \min_{w' \in W}(\kappa_\rho(w') + \kappa_{\rho'}(w'))$, the combination rule for kappa functions [71].

6.2. Rational fusion rules on plausibility scales

We denote by $\pi_i(w)$ the degree of plausibility of alternative w according to source i, and consider the problem of merging n possibility distributions π_i. The consistency degree between two possibility distributions is $Cns_*(\pi_i, \pi_j) = \max_w \pi_i(w) * \pi_j(w)$, where $*$ is a conjunctive operation. It ranges from 1 when there is a common w that is fully possible, to 0 when the supports do not overlap. It depends on the choice of the conjunction. If the conjunction is associative, its extension $Cns_*(\pi_1, \ldots, \pi_n)$ to n sources is obvious and is 1 (resp., > 0) if all cores (resp. supports) of the π_i's overlap.

The basic properties make sense as they stand, using the core, the support and the plausibility ordering induced by π_i, and interpreting minimal commitment as a maximisation of possibility degrees. One may nevertheless provide a more specialised version of some postulates:

1. Unanimity
   - *Possibility preservation* $\pi_1 * \ldots * \pi_n \leq f(\pi_1, \ldots, \pi_n)$ for some conjunction operation $*$ such that $a > 0$, $b > 0$ implies $a * b > 0$.
   - *Impossibility preservation* $\mathcal{S}(f(\pi_1, \ldots, \pi_n)) \subseteq \bigcup_{i=1}^n \mathcal{S}(\pi_i)$.
2. *Information monotonicity.* If $Cns_*(\pi_1, \ldots, \pi_n) \neq 0$ and $\pi_i \leq \pi_i', \forall i = 1, \ldots, n$ then $f(\pi_1, \ldots, \pi_n) \leq f(\pi_1', \ldots, \pi_n')$.
3. *Consistency enforcement.* The result must not be strongly inconsistent, i.e., $\mathcal{S}(f(\pi_1, \ldots, \pi_n)) \neq \emptyset$.
4. *Optimism.* If $Cns_*(\pi_1, \ldots, \pi_n) = 1$ then $\pi_1 * \ldots * \pi_n \geq f(\pi_1, \ldots, \pi_n)$, and $\max(\pi_1, \ldots, \pi_n) \geq f(\pi_1, \ldots, \pi_n)$ otherwise.
5. *Insensitivity to Vacuous Information.* $f(\pi_1, \ldots, \pi_{n-1}, \pi^?) = f(\pi_1, \ldots, \pi_{n-1})$, where $\pi^?(w) = 1\ \forall w \in W$.
6. *Commutativity.*
7. *Minimal commitment.* The obtained possibility distribution should be the least specific among those that satisfy the above postulates.

Let us examine the properties of conjunctive combination rules. It is fairly easy to prove using conjunction min:

Proposition 8. *If* $Cns_{\min}(\pi_1, \ldots, \pi_n) > 0$, *the minimum rule* $\pi_{\min} = \min(\pi_1, \ldots, \pi_n)$ *satisfies all basic properties of information fusion.*

However, the minimum rule does not satisfy Consistency Enforcement in case supports of the possibility distributions do not overlap (the result is the empty set). It is then not fair either, as the result in case of strong inconsistency between sources, leaves no trace of them. Replacing the minimum by an Archimedean continuous t-norm ($t(\alpha, \alpha) < \alpha$ if $0 < \alpha < 1$) such as product yields a connective that satisfies the same properties, but minimal commitment is in some sense supplemented by an independence assumption justifying a reinforcement effect making a state less possible than the minimal possibility degree granted by sources to this state. Note that the linear product t-norm is excluded from further consideration as it violates possibility preservation. In fact, it can be useful if the possibility of lying untruthful sources exists [85].

To get rid of the difficulty with Consistency Enforcement, one assumption that is often implicitly made is the "No Impossible World" assumption, here $\forall w \in W, \forall i = 1, \ldots, n : \pi_i(w) > 0$, which then lead some authors to accept conjunctive rules as the only rational merging rules. The strong form of the Consistency Enforcement postulate says that if there is some common ground between sources with a non-zero possibility degree, it should include the truth, also a strong form of optimism. Trying to satisfy this property justifies a renormalised conjunctive operation of the form, $\forall w \in W$:

$$\hat{*}(\pi_1, \ldots, \pi_n)(w) = Cns_*(\pi_1, \ldots, \pi_n) \to_* (\pi_1(w) * \ldots * \pi_n(w))$$

where $\alpha \to_* \beta = \sup\{\gamma : \alpha * \gamma \leq \beta\}$ is the residual operator of $*$ [65], which yields 1 if $\alpha \leq \beta$. Two instances of this rule, respectively in the QUALFI or DORD settings and in the numerical ones were already proposed in [45]:

$$\hat{\min}(\pi_1, \ldots, \pi_n)(w) = \begin{cases} \min_{i=1}^n \pi_i(w) \text{ if less than } Cns_{\min}(\pi_1, \ldots, \pi_n) \\ 1 \text{ otherwise;} \end{cases} \quad (5)$$

$$\hat{prod}(\pi_1, \ldots, \pi_n)(w) = \frac{\pi_1(w) \cdot \cdots \cdot \pi_n(w)}{\max_{v \in W} \pi_1(v) \cdot \cdots \cdot \pi_n(v)} \quad (6)$$

The rule (5) violates Impossibility Preservation (Unanimity 1b) in the situation of disjoint supports, since everything becomes possible in this case. The second part of the optimism axiom may be violated as well due to the renormalisation factor which may make a state w fully plausible even if no source is proposing it with full possibility. Moreover, the rule (5) loses the associativity of the minimum rule, while its variant (6) with product instead of minimum preserves associativity [45]. However, the latter rule makes sense only as long as $\bigcap_{i=1}^n \mathcal{S}(\pi_i) \neq 0$. If this condition is not met, the rule again violates Consistency Enforcement, because it then is not defined mathematically. Note that this combination rule is exactly the one recently proposed by Laverny and Lang [71] on positive integers:

$$\kappa \oplus \kappa' = \kappa + \kappa' - \min_{w' \in W}(\kappa(w') + \kappa'(w')),$$

in order to combine ranking functions of Spohn, once the latter is mapped back to the unit interval (DENUM setting) via a suitable transformation. Ma and Liu [77] propose postulates to justify this combination rule in terms of addition of relative plausibility functions ρ recalled in Remark 1. These postulates partially reflect some basic properties advocated here, but there is no full-fledged information comparison ordering between relative plausibility functions that are defined in terms of integer-valued functions invariant with respect to a translation (but for comparing the cores). Moreover, the use of addition or product presupposes a reinforcement effect that makes sense only if sources can be considered independent, and that can only be expressed in a numerical setting.

Then, in order to recover all our basic postulates in the possibilistic setting, one is led to use disjunctive combination rules and extend the MCS rule (4) from sets to fuzzy sets. We can do it in at least two ways:

$$\text{MCS-1}(\pi_1, \ldots, \pi_n) = \max_{I \in MCS(\{\mathcal{C}(\pi_1), \ldots, \mathcal{C}(\pi_n)\})} *_{i \in I} \pi_i \quad (7)$$

$$\text{MCS-0}(\pi_1, \ldots, \pi_n) = \max_{I \in MCS(\{\mathcal{S}(\pi_1), \ldots, \mathcal{S}(\pi_n)\})} *_{i \in I} \pi_i \quad (8)$$

In fact each of MCS-1, MCS-0 selects maximal consistent subsets in a specific way, checking consistency on cores or supports respectively. Rule MCS-0 may yield a subnormalised but not empty result. Once, this principle chosen, the same reasoning holds as in the crisp case, and we obtain expected properties for these two fusion rules:

Proposition 9. *For* $* = \min$, *the extended-MCS rules (7) and (8) satisfy all basic fusion properties.*

Proof. It is clear that these fusion rules coincide with the minimum rule if $Cns_{\min}(\pi_1, \ldots, \pi_n) = 1$. Moreover, if all input distributions have disjoint supports then both rules yield the maximum of

Table 2
Possibilistic fusion rules on 3-level plausibility scales.

Rules	$C_1 \cap C_2 \neq \emptyset$	$C_1 \cap C_2 = \emptyset, S_1 \cap S_2 \neq \emptyset$	$S_1 \cap S_2 = \emptyset$
min	$(C_1 \cap C_2, S_1 \cap S_2)$	$(\emptyset, S_1 \cap S_2)$	$(\emptyset, \emptyset)$
$\hat{\text{min}}$	$(C_1 \cap C_2, S_1 \cap S_2)$	$(S_1 \cap S_2, S_1 \cap S_2)$	(W, W)
MCS-1	$(C_1 \cap C_2, S_1 \cap S_2)$	$(C_1 \cup C_2, S_1 \cup S_2)$	$(C_1 \cup C_2, S_1 \cup S_2)$
MCS-0	$(C_1 \cap C_2, S_1 \cap S_2)$	$(\emptyset, S_1 \cap S_2)$	$(C_1 \cup C_2, S_1 \cup S_2)$
MCS-cuts	$(C_1 \cap C_2, S_1 \cap S_2)$	$(C_1 \cup C_2, S_1 \cap S_2)$	$(C_1 \cup C_2, S_1 \cup S_2)$

all possibility distributions. So

$$\min(\pi_1, \ldots, \pi_n) \leq \text{MCS-}i(\pi_1, \ldots, \pi_n) \leq \max(\pi_1, \ldots, \pi_n), i = 0, 1$$

and Unanimity postulate clearly holds (these rules are idempotent moreover). Informational Monotony, Consistency Enforcement (strong for MCS-1), Insensitivity to Vacuous Information and Commutativity are obvious. These rules satisfy Fairness, since when $Cns_{\min}(\pi_1, \ldots, \pi_n) < 1$, $\forall i, \exists w \in \mathcal{C}(\pi_i)$, MCS-1$(\pi_1, \ldots, \pi_n)(w) = 1$ and $\forall i, \exists w \in \mathcal{S}(\pi_i)$, MCS-0$(\pi_1, \ldots, \pi_n)(w) > 0$. Optimism holds for MCS-1 in the sense that for maximal subgroups I of strongly consistent sources a conjunctive rule is applied and the use of the minimum operation corresponds to Minimal Commitment inside this group. Similarly, optimism also holds for MCS-0 considering consistent sources. Minimal Commitment also prevents us from choosing between the maximal strongly consistent subsets. □

MCS-1 is much demanding on mutual consistency of sources and already yields plain disjunction if cores of π_i are disjoint. MCS-0 is less demanding and more optimistic: it yields $\min(\pi_1, \ldots, \pi_n)$ if all supports overlap, but one may renormalise the resulting distribution. The same results hold if min is replaced by product, with the same proviso as before regarding minimal commitment. Note that there are more such rational fusion rules as we could use α-cuts $\{\mathcal{A}_\alpha(\pi_i) = \{w : \pi_i(w) \geq \alpha\}$ to compute maximal consistent subsets:

$$\text{MCS-}\alpha(\pi_1, \ldots, \pi_n) = \max_{I \in MCS(\{\mathcal{A}_\alpha(\pi_1), \ldots, \mathcal{A}_\alpha(\pi_n)\})} *_{i \in I} \pi_i.$$

All fusion rules MCS-α, $\alpha \in L\backslash\{0\}$ can be defined, and become more demanding as α increases.

Another fusion rule for possibility distributions called *MCS-cuts* applies the classical MCS rule to all cuts of the input possibility distributions has been recently proposed [30]. Its result T is defined by the family of sets T_α, $\alpha \in L\backslash\{1\}$ obtained by applying (4) to $\{\mathcal{A}_\alpha(\pi_1), \ldots, \mathcal{A}_\alpha(\pi_n)\}$.

$$T_\alpha = \bigcup_{I \in MCS(\mathcal{A}_\alpha(\pi_1), \ldots, \mathcal{A}_\alpha(\pi_n))} \bigcap_{i \in I} \mathcal{A}_\alpha(\pi_i) \tag{9}$$

It obviously satisfies all basic postulates but the sets T_α, $\alpha \in L\backslash\{1\}$ are in general no longer nested, which calls for an enlarged representation framework. Note that one may even have that $T_{\alpha_i} \cap T_{\alpha_j} = \emptyset$ for $i \neq j$. In the numerical settings DABS and DENUM, it yields a belief function, with mass function such that $m_T(T_{\alpha_i}) = \alpha_i - \alpha_{i+1}, i = 0, \ldots, m-1$.

To illustrate these fusion rules we can consider the simplest qualitative plausibility scale $L = \{1 > \alpha > 0\}$. This is when an information item is completely defined by non-empty supports and cores: $T = (C_T \subseteq S_T)$ defining an ordered partition $\mathcal{P}_T = (C_T, S_T \setminus C_T, W \setminus S_T)$. The results of applying the above rules for two sources yielding $(C_i, S_i), i = 1, 2$ are pictured on Table 2:

It can be checked that

- all fusion rules coincide if cores intersect;
- the min rule violates Consistency Enforcement and Fairness when $S_1 \cap S_2 = \emptyset$;
- in the same situation, the normalised $\hat{\text{min}}$ rule violates Impossibility Preservation, Optimism and Fairness;
- MCS-1 coincides with the max rule when cores are disjoint and is less optimistic than MCS-0;
- MCS-0 mends the min rule as it restores (weak) Consistency Enforcement and Fairness when the latter fails them;
- the MCS-cuts rule does not yield nested subsets, and especially we may have (middle column) $(C_1 \cup C_2) \cap (S_1 \cap S_2) = \emptyset$ if it turns out that $C_1 \cap S_2 = S_1 \cap C_2 = \emptyset$.

Remark 2. In [87], the following properties were proposed w.r.t a fusion operator $\circ_\pi$ applied to a set of possibility distributions π_i:

O1: π-Commutativity and π-associativity.
O2: π-Idempotence.
O3: π-Monotony. Given two sets $\{\pi_1, \ldots, \pi_n\}$ and $\{\pi'_1, \ldots, \pi'_n\}$ of possibility distributions such that $\forall i, \forall w, \pi_i(w) \geq \pi'_i(w)$, then $\circ_\pi(\pi_1, \ldots, \pi_n) \geq \circ_\pi(\pi'_1, \ldots, \pi'_n)$.
O4: π-impossible case and π-complete ignorance case.
Let $\pi_\emptyset$ ($\pi_\emptyset(w) = 0$, $\forall w \in W$) and π_W ($\pi_W(w) = 1$, $\forall w \in W$) be possibility distributions that stand for *complete contradiction* and *vacuous information* respectively, and then

1. $\circ_\pi(\pi_1, \ldots, \pi_n, \pi_\emptyset) = \circ_\pi(\pi_1, \ldots, \pi_n)$
2. $\circ_\pi(\pi_1, \ldots, \pi_n, \pi_W) = \circ_\pi(\pi_1, \ldots, \pi_n)$.

Our axiomatic framework does not consider associativity (O1) nor idempotence (O2) as imperative requirements, nor does it assume the possibility of inconsistent inputs (first part of O4). We also restrict axiom O3 (our property 2) to strongly consistent inputs. Up to these points, the above axioms do fit ours (O4-2 is our property 6: insensitivity to vacuous information).

There are two axioms related to O4, adapted from the probabilistic setting [24]:

(a) *zero preservation:* If $\exists w \in W, \forall i, \pi_i(w) = 0$, then $\circ_\pi(\pi_1, \ldots, \pi_n)(w) = 0$.

(b) *maximal plausibility:* If $\exists w \in W, \forall i, \pi_i(w) > 0$, then $\circ_\pi(\pi_1, \ldots, \pi_n)(w) > 0$.

These properties correspond to our postulate 1.

Remark 3. Quota merging rules can apply to possibility distributions [52]. In this case one can consider the maximal number k_* of strongly consistent sources (cores intersecting), and the maximal number $k^* > k_*$ of weakly consistent sources (supports intersecting), performing the conjunction of pieces of information from the former group, and the renormalised conjunction of pieces of information from the latter group. These two resulting information items can be further merged, in a prioritised way, giving more weight to the set of strongly consistent sources.

Remark 4. Chopra et al [22] also propose seven postulates for information fusion in the DENUM setting. We briefly restore them in our setting. Their axiom $\Delta 0$ corresponds to an optimistic form of impossibility preservation (even though no state is considered impossible in [22]), which would read $\pi(w) \leq \max_{i=1}^n \pi_i(w)$ in our setting. Axiom $\Delta 1$ is a truth-functionality axiom, akin to Independence of Irrelevant Alternatives, saying basically that $\pi(w)$ is a function of the values $\pi_i(w)$ only. Remember that this postulate does not contradict our setting only if $\pi_i(w) > 0, \forall i, \forall w \in W$. They also require commutativity ($\Delta 3$), ordinal unanimity ($\Delta 4$) and a form of Fairness ($\Delta 5$) similar to the one in the ordinal setting of Section 5.2. Axiom ($\Delta 6$) is a variant of the Optimism axiom cast in terms of a form of strict Pareto dominance Finally, there is also one postulate ($\Delta 2$) requiring that the range of values used in the fusion process is made of consecutive integers (a requirement which is only justified by the DENUM representation setting).

This axiomatic system is again quite in agreement with our basic postulates, even though there is no use of informational comparison between information items, nor any minimal commitment assumption, nor any postulate regarding the uselessness of uninformed sources. The lack of such considerations and the fact that

they compare their setting with the one of Arrow as well and discuss strategy-proofness, suggest their framework is not solely devoted to the search of truth in a body of information items, but like authors of [69], they consider information fusion as encompassing both preference and plausibility merging.

6.3. Relation with distance-based knowledge merging

Possibility distributions over a set consisting of interpretations of a Boolean language can be encoded as possibilistic knowledge bases [36]. So, aggregation operations that apply to possibility distributions can be encoded at the syntactic level into operations that merge totally ordered, or stratified, knowledge bases in possibilistic logic. The reader is referred to the literature for details on this topic, especially the works of Benferhat and colleagues [7,8,11,12,62] as well as Liu and colleagues [75,75,89–91,117].

It is nevertheless interesting to point out the link between the distance-based merging approaches that implement fusion rules in propositional logic, and the possibilistic fusion setting.

In the distance-based merging approach [66], given a propositional knowledge base K_i, a distance between interpretations and K_i is defined as

$$d(w, K_i) = \min_{w' \models K_i} d(w, w'),$$

where d is a distance between interpretations of the language, for instance the Hamming distance counting the number of instantiated variables that differ in w and w'. Viewing the set of models of K_i as the core of an information item, and using a Hamming distance, it is natural to interpret the integer $d(w, K_i)$ as the degree of disbelief of w given the belief base K_i, in the style of Spohn ranking functions. It is then easy to construct a possibility distribution expressing the relative plausibility of w as $\pi_i(w) = \epsilon^{d(w,K_i)}$, for some positive $\epsilon < 1$ (which brings the problem inside the DENUM setting). As explained in [7], it is then easy to reinterpret distance-based propositional knowledge-based merging rules in terms of aggregation operations in possibility theory:

- The maximum rule proposed in [66,68] corresponds to the normalised minimum rule (5) in possibility theory. Indeed, the interpretations w such that the function $\max_{i=1}^{n} d(w, K_i)$ is minimal correspond to the maximal elements of $\min_{i=1}^{n} \pi_i(w)$ and form the core of the resulting possibility distribution, which comes down to a normalisation. The result of the propositional fusion of $K_1, \ldots, K_n$ is then always a consistent classical propositional base K such that $[K] = \{w : \pi(w) = Cns_{\min}(\pi_1, \ldots, \pi_n)\}$. Note that this fusion rule is not associative, even if max is associative. The max rule is one of those proposed by Chopra et al. [22] in the DENUM setting.
- The sum rule proposed in [66,68] corresponds to the normalised product rule (6) in possibility theory. Indeed, the interpretations w such that $\sum_{i=1}^{n} d(w, K_i)$ is minimal correspond to the maximal elements of $*_{i=1}^{n} \pi_i(w)$ and form the core of the resulting possibility distribution, which comes down to a normalisation. The result of the propositional fusion of $K_1, \ldots, K_n$ is then always a consistent classical propositional base K such that $[K] = \{w : \pi(w) = Cns_{*}(\pi_1, \ldots, \pi_n)\}$. Note that a form of this fusion rule was already at work in the old expert system MYCIN (see the corresponding discussion in [45]). The sum rule is also one of those proposed in [22] and appears as well in [71,77].
- The GMax fusion rule proposed in [66,68] corresponds to the leximin rule [34] that refines the minimum operation applied to possibility distributions [7]. This operation results in a ranking of interpretations that is finer-grained than the scale of the inputs. One of the merging rules (called $\Delta_{\min 1}$) proposed in [22] encodes the GMax fusion rule by means of integers.
- Quota rules for merging knowledge bases are found in [54,91] and can be explained in the setting of possibility theory as well. In [75], an adaptive algorithm for merging stratified knowledge-bases (SKBs) is proposed. The algorithm first selects among maximal consistent subsets of sources based on assessing how (partially) consistent the information pieces in the subset are, in the spirit of possibilistic merging. However, this paper does not mention any information ordering between SKBs, refers neither to Information Monotonicity nor Optimism to advocate the proposed fusion operators.
- In the same vein, Hunter and Liu [59], proposed an adaptive algorithm for merging possibility distributions. Although information ordering, as defined in this paper was not used in [59], the assessment of information quality for ranking information items was indeed applied in order to generate partially maximal consistent subsets.

Note that, in the distance-based merging context, $\forall w \in W, \pi_i(w) > 0$ so that conjunctive merging rules always satisfy Consistency Enforcement and Fairness, hence all basic postulates of information fusion. This condition is not taken for granted in the usual possibilistic setting, which explains again the absence of MCS rules in the frameworks of distance-based merging and all settings based on the integer scale for grading disbelief.

These approaches explicitly use the connection between stratified knowledge-bases and plausibility ranking of possible worlds to define merging operations. This is also true for the approach in [117], where the set of possible worlds is ranked, based on the set of stratified knowledge bases, according to a Condorcet-like method after voting systems, prior to reconstructing a stratified base. This method satisfies most of the basic postulates as well but for those involving the information-ordering, a notion not used in that paper.

7. A principled view of information fusion in Dempster–Shafer theory

In evidence theory [97,116], further developed as the Transferable Belief Model [102], a piece of information T is modelled by a *basic belief assignment (bba)* m_T which is a mapping from 2^W to [0, 1] such that $\sum_{E \subseteq W} m_T(E) = 1$, and $m_T(\emptyset) = 0$. This piece of information must be interpreted as an unreliable testimony, whereby $m_T(E)$ is the probability that E is the correct information given by the source. It means that with probability $1 - m_T(E)$, the piece of information E is irrelevant. A set E with positive mass is called a *focal set*. The set of focal sets is denoted by $\mathcal{F}_m$. Given a bba, two conjugate evaluations can be defined, regarding the confidence it gives in propositions described by sets $A \subseteq W$:

- The degree of belief in a proposition A is the probability that A can be logically inferred from the agent's body of evidence: $Bel(A) = \sum_{E_i \subseteq A} m_T(E_i)$;
- The degree of plausibility of A is the probability that A is logically consistent with the agent's body of evidence: $Pl(A) = \sum_{E_i \cap A \neq \emptyset} m_T(E_i) = 1 - Bel(A^c)$, where A^c is the complement of A.

The uninformative item is captured by the *vacuous belief function*: $m^{\top}$ such that $m^{\top}(W) = 1$. It is well-known [97] that a bba subsumes probability distributions (when focal sets are only singletons), and possibility distributions (when focal sets are nested). In the first case, the belief and plausibility functions reduce to a probability measure; in the latter case they are necessity and possibility measures respectively.

7.1. Core, support, plausibility ordering

Various relevant notions can be defined as follows for a bba m_T. They are instrumental for instantiating the basic properties in evidence theory:

- **Support**: if $\mathcal{F}_T = \{A_1, \ldots, A_n\}$, then $\mathcal{S}(m_T) = \bigcup_{i=1}^{n} A_i$.[10]
- **Core**: $\mathcal{C}(m_T) = \bigcap_{i=1}^{n} A_i$.
- **Plausibility ordering**: It can be constructed in two ways:
  - using the *contour function* [97]: $\pi_T(w) = \sum_{A \subseteq W, w \in A} m_T(A) = Pl(\{w\})$; this is a natural option for comparing possible worlds ($w_1 \succeq_T w_2$ if and only if $\pi_T(w_1) \geq \pi_T(w_2)$).
  - directly using a kind of dominance between bba's [37]: $w_1 \succeq_T^{dom} w_2$ if and only if for any $A \subseteq W \setminus \{w_1, w_2\}$, $m_T(A \cup \{w_1\}) \geq m_T(A \cup \{w_2\})$.

It is clear that $w \in \mathcal{S}(m_T)$ if and only if $Pl(\{w\}) > 0$ and $w \in \mathcal{C}(m_T)$ if and only if $Pl(\{w\}) = 1$. Note that the core may be empty. When the focal sets are nested, the *contour* function π_T of the bba m_T coincides with a possibility distribution, and $Pl(A) = \max_{w \in A} \pi_T(w)$ is a possibility measure. Moreover, bba dominance $\succeq_T^{dom}$ is a partial ordering coherent with the contour function as indicated by the easy-to-check result:

Proposition 10. *: $\succeq_T^{dom}$ is a reflexive and transitive relation. Moreover, $w_1 \succeq_T^{dom} w_2$ implies $\pi_T(w_1) \geq \pi_T(w_2)$, and likewise $w_1 \succ_T^{dom} w_2$ implies $\pi_T(w_1) > \pi_T(w_2)$*

A bba is usually assumed to be self-consistent, i.e., as said earlier, $m(\emptyset) = 0$. A stronger form of self-consistency requests a non-empty core, that is, a normalised contour function. However, in the following we shall still call *bba* a mass function for which $m(\emptyset) \neq 0$. Such bba will be called weakly consistent if $m(\emptyset) \neq 1$, and strongly inconsistent otherwise.

7.2. Mutual consistency

We now examine issues related to the notion of conflict between belief functions.

The degree of inconsistency (or conflict) of two bbas m_1 and m_2 is often measured by the mass bearing on the empty set as the result of the conjunction of m_1 and m_2 viewed as independent random sets:

$$Inc(m_1, m_2) = \sum_{A \cap B = \emptyset} m_1(A) \cdot m_2(B).$$

It is the counterpart of $1 - Cns_*(\pi_1, \pi_2)$ in possibility theory, using conjunction product. However, this index is questionable. It has been pointed out in [74] that $m_{1,2}(\emptyset)$ is not a measure of discrepancy between bba's, since two identical bba's may have a non-zero degree of conflict [11]. Moreover, assuming independence is an extra assumption that may or may not be used in the fusion process.

Alternatively we can adopt definitions that do not rely on independence: 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 $Inc(m_1, m_2) < 1$)
- *Strongly (or logically [29]) mutually consistent* if $\forall E \in \mathcal{F}, \forall E' \in \mathcal{F}' : E \cap E' \neq \emptyset$ (note that it does imply that $Inc(m_1, m_2) = 0$).

It is easy to see that these conditions can only be partially stated in terms of contour functions. Indeed, m and m' are:

- weakly mutually consistent if and only if $\max_{w \in W} \min(\pi_m(w), \pi_{m'}(w)) \neq 0$
- strongly mutually consistent whenever $\max_{w \in W} \min(\pi_m(w), \pi_{m'}(w)) = 1$.

The converse of the last statement does not hold, since $\max_{w \in W} \min(\pi_m(w), \pi_{m'}(w)) = 1$ presupposes the use of belief functions with non-empty cores.

Note that the above definitions do not involve the mass functions, and are all-or-nothing concepts, which may not be fully satisfactory. Strong mutual consistency is nevertheless very natural as it means that each pair of focal sets coming from each belief function is consistent. It is weaker than the strong consistency of their contour functions. However, weak mutual consistency can be judged too weak. An alternative stronger definition of consistency may be as follows [29,40]:

Definition 2. Two mass functions m_1 and m_2 are said to be mutually probabilistically consistent if there exists a joint mass function $x(B, C), B \in \mathcal{F}_1, C \in \mathcal{F}_2$ such that

$$\sum_{B \in \mathcal{F}_1} x(B, C) = m_2(C), \forall C \in \mathcal{F}_2; \tag{10}$$

$$\sum_{C \in \mathcal{F}_2} x(B, C) = m_1(B), \forall B \in \mathcal{F}_1; \tag{11}$$

$$x(B, C) = 0 \text{ whenever } B \cap C = \emptyset. \tag{12}$$

$$\sum_{B \in \mathcal{F}_1, C \in \mathcal{F}_2} x(B, C) = 1. \tag{13}$$

The two first sets of equalities say that m_1 and m_2 are marginals of $x(\cdot,\cdot)$; the third one forbids to allocate positive mass to pairs of conflicting focal sets (then $x(\cdot,\cdot)$ is strongly self-consistent).

Given a belief function, it is well-known that the set of probabilities $\mathcal{M}(m) = \{P : P(A) \geq Bel(A), \forall A\}$ is non-empty and convex. It is called the credal set induced by m. The above definition of consistency is equivalent to saying that the two credal sets $\mathcal{M}(m_1)$ and $\mathcal{M}(m_2)$ have a non-empty intersection [20], which explains the name of this type of consistency. A degree of probabilistic consistency between m_1 and m_2 can then be defined as

$$PC(m_1, m_2) = \sup_x \sum_{B \in \mathcal{F}_1, C \in \mathcal{F}_2, B \cap C \neq \emptyset} x(B, C)$$

where x is any joint mass assignment with marginals m_1 and m_2 (not necessarily obeying (16)). This consistency index has better behaviour than $1 - Inc(m_1, m_2)$. First, it does not presuppose independence between m_1 and m_2. Moreover, it obeys intuitively satisfactory properties that the other index violates:

Proposition 11. *Index PC obeys the following properties:*

- *For any normalised mass function m, $PC(m, m) = 1$.*
- *$PC(m_1, m_2) = 1$ if and only if m_1 and m_2 are mutually probabilistically consistent.*

Proof. Indeed for the case when $m_1 = m_2 = m$, it is enough to choose $x(A, A) = m(A), \forall A \subseteq W$ to get $PC(m, m) = \sum_{A \subseteq W} m(A) = 1$. The other point is obvious as probabilistic consistency precisely consists in the possibility to assign joint masses in agreement to the given marginals only to non-conflicting sets. □

Note that even if m_1 and m_2 are weakly mutually consistent, they may be probabilistically inconsistent ($PC(m_1, m_2) = 0$), in contrast with $Inc(m_1, m_2)$ that will always be less than 1). Besides, strong mutual consistency ensures $PC(m_1, m_2) = 1$.

The drawback of our conflict index is that it needs to run a linear programming algorithm to compute it in practice. Recently,

[10] Often called the core in the literature, which is at odds with, e.g., probability theory where this set is also called the support.

[11] A distance function such as $d(m_1, m_2) = max_{A \subseteq W}(|m_1(A) - m_2(A)|)$ seems more appropriate for this kind of discrepancy.

Destercke and Burger [28] have proposed an axiomatic setting for measures of conflict, that has some connection with our postulates. However, they propose degrees of confict that use contour functions and are easier to compute.

7.3. Information ordering

In the literature, different information orderings in evidence theory have been proposed for comparing the information contents of bba's. We only consider the most basic ones (see e.g. [29,41])

Definition 3 (c-ordering). $m_1 \sqsubseteq_c m_2$ if $\forall w \in W,\ \pi_{m_1}(w) \leq \pi_{m_2}(w)$.

This is a straightforward extension of the specificity ordering of possibility theory. It very much lacks discrimination as it only relies on the plausibility of singletons.

Definition 4 (pl-ordering). [41] $m_1 \sqsubseteq_{pl} m_2$, if $\forall A \subseteq W,\ pl_1(A) \leq pl_2(A)$.

This kind of information ordering is much in line with the understanding of belief functions as a special family of probability functions since $m_1 \sqsubseteq_{pl} m_2$ is equivalent to the inclusion between the corresponding credal sets: $\mathcal{M}(m_1) \subseteq_{pl} \mathcal{M}(m_2)$.

The next definition requires the notion of a commonality function q induced by a bba m_T [97]: $q(A) = \sum_{E_i \supseteq A} m_T(E_i)$. Clearly degrees of commonality tend to be all the greater as masses are allocated to larger focal sets.

Definition 5 (q-ordering). [41] $m_1 \sqsubseteq_q m_2$ if $\forall A \subseteq W,\ q_1(A) \leq q_2(A)$.

It means that, by and large, m_2 assign masses to larger subsets than m_1, which sounds like m_2 being the last informative of the two bba's.

Definition 6 (Specialisation). [41,64] Let m_1 and m_2 be two bbas over W, m_1 is a specialisation of m_2, denoted by $m_1 \sqsubseteq_s m_2$ if and only if there exists a non-negative matrix joint bba with general term $x(A, B)$, $A \in \mathcal{F}_1, B \in \mathcal{F}_2$ such that

$$\sum_{A \in \mathcal{F}_1} x(A, B) = m_2(B), \forall B \in \mathcal{F}_2; \tag{14}$$

$$\sum_{B \in \mathcal{F}_2} x(A, B) = m_1(A), \forall A \in \mathcal{F}_1; \tag{15}$$

$$x(A, B) = 0 \text{ whenever } A \not\subseteq B. \tag{16}$$

$$\sum_{B \in \mathcal{F}_1, C \in \mathcal{F}_2} x(B, C) = 1. \tag{17}$$

Note that a necessary condition for specialisation (or s-ordering) is that $\forall A \in \mathcal{F}_1, \exists B \in \mathcal{F}_2, A \subseteq B$ and $\forall B \in \mathcal{F}_2, \exists A \in \mathcal{F}_1, A \subseteq B$. Moreover, this definition is in full agreement with the definition of probabilistic consistency between bba's (notice the strong similarity between their respective definitions).

Relationships among these orderings are as follows [41]: s-ordering implies pl-ordering and q-ordering, but the converse is false. Both are more demanding than the c-ordering. Hence the s-ordering is the strongest ordering of the four, and c-ordering the weakest. Moreover, they all coincide with the specificity ordering of possibility theory when the focal sets are nested. However, the pl-ordering and the q-ordering are not comparable and can be at odds with one another: it is easy to find bba's m_1 and m_2 such that $m_1 \sqsubset_{pl} m_2$ and $m_2 \sqsubset_q m_1$, in which case the contour functions are equal [29], which may baffle the intuition. These arguments plea for the use of specialisation as the most natural information ordering in the setting of evidence theory.

7.4. Instantiation of basic fusion postulates in evidence theory

There are several possible ways of instantiating the basic postulates of information fusion for belief functions, due to several notions of mutual consistency, information ordering, plausibility ordering, and the like.

In the following we give one possible instantiation which may be appealing because it is as little constrained as possible. To this end we use specialisation, dominance ordering, and strong mutual consistency. For the sake of simplicity we write the postulates for two bba's m_1 and m_2, denoting by m_{12} the result of the fusion. In order to build a combination rule we also need a dependence structure between sources that is expressed by the joint bba, namely we must know the joint bba is $x(\cdot, \cdot)$ whose marginals are m_1 and m_2. On this basis, two straightforward extensions of the set-theoretic intersection and union can then be defined for random sets:

$$\forall E \subset W, m_{1\cap 2}(E) = \sum_{E=F\cap G} x(F, G). \tag{18}$$

$$\forall E \subset W, m_{1\cup 2}(E) = \sum_{E=F\cup G} x(F, G). \tag{19}$$

One assumption often made is independence between pieces of information provided by sources, so that $x(F, G) = m_1(F) \cdot m_2(G)$. Then these combination rules are the ones already known in the literature [42,100], except that, here, the joint bba is $x(\cdot, \cdot)$ instead of the product of bba's that requests an independence assumption between the two random sets. In the latter case, we denote $m_{1\cap 2}$ by $m_1 \odot m_2$, and $m_{1\cup 2}$ by $m_1 \oplus m_2$.

This dependence or independence assumption does not affect the postulates below. Namely, fusion postulates do not have to impose a particular dependence structure between sources.

1. **Unanimity**
   - *Possibility preservation*: If $w \in \mathcal{C}(m_1) \cap \mathcal{C}(m_2)$ then, $w \in \mathcal{C}(m_{12})$.
   - *Impossibility preservation*: If $w \notin \mathcal{S}(m_1) \cup \mathcal{S}(m_2)$, then $w \notin \mathcal{S}(m_{12})$.
   - An additional requirement can be *local ordinal unanimity* with respect to dominance ordering: for two any states w and w': if $w \succeq_1^{dom} w'$ and $w \succeq_2^{dom} w'$ then $w \succeq_{12}^{dom} w'$.
2. *Weak Information monotonicity*: If m_1 and m_2 are strongly consistent, and $m_1 \sqsubseteq_s m_1'$, $m_2 \sqsubseteq_s m_2'$, then $m_{12} \sqsubseteq_s m_{12}'$.
3. *Consistency enforcement*:
   - $\sum_{E \subseteq S} m_{12}(E) = 1$ (strong version).
   - $\Sigma_{E \subseteq S} m_{12}(E) > 0$ (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 $m_{12} \sqsubseteq_s m_{1\cup 2}$.
5. *Fairness*
   - *No Favourite*: If m_1 and m_2 are not strongly mutually consistent, then $m_{12} \not\sqsubseteq_s m_i, i = 1, 2$
   - *No Dismissal*: Each m_i should be weakly consistent with m_{12}.
6. *Insensitivity to vacuous information*: If $m_1(W) = 1$ then $m_{12} = m_2$.
7. *Symmetry*: $m_{12} = m_{21}$.
8. *Minimal commitment*: m_{12} should be minimally specific for specialisation.

Some comments are in order. For Consistency Enforcement, the choice between strong and weak version depends on whether the result of the combination rule should be normalised or not. The optimism postulate again expresses that we favour conjunctive rules of combination when sources strongly agree and that we should assume at least one source is correct if they disagree. The latter involves the disjunctive combination rule $m_{1\cup 2}$. In the

Fairness axiom, we have provided the two sides of it explicitly, because they are no longer redundant, since we do not use the same form of mutual consistency on each side. Finally, minimal commitment could be expressed in the sense of any suitable information ordering relation. The choice of a dependence structure between the sources expressed via the joint bba $x(\cdot,\cdot)$ reduces the scope of the minimal commitment axiom and the role of the postulates to the choice of a proper set-theoretic combination for focal sets.

7.5. Examples of fusion rules

Let us consider a few known combination rules in the theory of belief functions, to be evaluated in the light of the above postulates. Many combination rules have been proposed in evidence theory for merging information, apart from the well-known Dempster's rule of combination. For the sake of simplicity, we only focus on the main combination rules, although many variants have been proposed in the literature (see [101] for an extensive review).

$$m_{De}(C) = \begin{cases} \frac{\sum_{A,B:A\cap B=C} m_1(A)\cdot m_2(B)}{1-\sum_{A,B:A\cap B=\emptyset} m_1(A)\cdot m_2(B)}, \forall C\neq\emptyset, \\ 0 \text{ if } C=\emptyset. \end{cases}$$

(Dempster's rule) (20)

$$m_{Sm}(C) = \sum_{A,B\subseteq W, A\cap B=C} m_1(A)\cdot m_2(B) = m_1\odot m_2(C)$$

(Smets' rule) [101] (21)

$$m_{Ya}(C) = \begin{cases} \sum_{A,B:A\cap B=C} m_1(A)\cdot m_2(B) \text{ if } C\neq W,\emptyset \\ m_1(W)\cdot m_2(W) + \sum_{A\cap B=\emptyset} m_1(A)\cdot m_2(B) \text{ if } C=W, \\ 0 \text{ if } C=\emptyset. \end{cases}$$

(Yager's rule) [114] (22)

$$m_{DP}(C) = \begin{cases} \sum_{A,B:A\cap B=C} m_1(A)\cdot m_2(B) \\ \quad + \sum_{A,B:A\cup B=C, A\cap B=\emptyset} m_1(A)\cdot m_2(B), \forall C\neq\emptyset, \\ 0 \text{ if } C=\emptyset. \end{cases} \quad (23)$$

The most celebrated fusion rule is an associative operation called Dempster rule of combination [27,97]. All four fusion rules presuppose independence between sources, as an additional assumption, which enforces the choice of $x(\cdot,\cdot) = m_1(\cdot)\cdot m_2(\cdot)$. The main difference between Dempster's rule and the three other rules respectively proposed in [102] (see also [100]) [113], and [45] concern the way the mass $(m_1\odot m_2)(\emptyset) = \sum_{A,B:A\cap B=\emptyset} m_1(A)\cdot m_2(B)$ is reallocated. In Dempster's rule, the renormalisation by division enforces consistency preservation, when the two bba's are weakly consistent (otherwise the operation is not defined). Smets's rule simply keeps this mass on $\emptyset$, whilst Yager's rule assigns it to W. The DP rule keeps the mass $m_1(A)\cdot m_2(B)$ on $A\cap B$ whenever this intersection is not empty, and gives it to $A\cup B$ otherwise.

All four fusion rules coincide with each other if the two bba's are strongly consistent. Then all postulates are satisfied. When $\sum_{A\cap B=\emptyset} m_1(A)\cdot m_2(B) = 1$, m_{De} is not defined due to a total conflict between the sources, which violates the Consistency Enforcement postulate, like for the normalised conjunctive rules in possibility theory. When the two bba's are weakly mutually consistent, the result is consistent since $m_{De}(\emptyset) = 0$. For instance, suppose that m_1 and m_2 are weakly mutually consistent in a minimal way, namely if there is a single pair $A\in\mathcal{F}_1, B\in\mathcal{F}_2$, such that $A\cap B\neq\emptyset$, which is a situation of severe conflict. Then $m_{De}(A\cap B) = 1$, so $m_{De}\not\sqsubseteq m_1\oplus m_2$ as soon as there are $C\in\mathcal{F}_1, D\in\mathcal{F}_2$, such that $A\cap B\not\subseteq C\cup D$, since the mass $m_{De}(A\cap B)$ cannot flow to all focal sets of $m_1\oplus m_2$. The afore-mentioned situation shows that Dempster's rule of combination is over-optimistic in case of weak consistency; it may fail to satisfy the second optimism condition, due to renormalisation (it would satisfy it if we replace it by the weaker condition $\mathcal{S}(m_{12})\subseteq\mathcal{S}(m_1)\cup\mathcal{S}(m_2)$).

Smets [100] has advocated a non-normal version (21) of Dempster combination rule, which comes down to a mere intersection of independent random sets. In Smets rule, the mass assigned to the empty set $m_S(\emptyset)$, may be different from 0, and can even be 1. Thus, 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 because it is not compatible with the failure of the Consistency Enforcement postulate. Optimism is recovered at the expense of internal consistency of the result, which is subnormalised in case of weakly consistent inputs. So, while Dempster rule is overoptimistic, Smets rule gives up internal consistency.

Yager [113] proposed another form of renormalisation that ensures a strongly internally consistent result, while making the combination applicable to any pair of bba's, by reassigning the mass $m_S(\emptyset)$ to the whole set W. It does not respect optimism and moreover, impossibility preservation is clearly violated. The failure of Optimism can be observed if the two bba's are not strongly consistent, that is when $m_{Ya}(W) > m_1(W)\cdot m_2(W)$. Then the condition $m_{Ya}\sqsubseteq_s m_1\oplus m_2$ may fail because $m_{Ya}(W)$ has become too large. In fact this rule is far too cautious in the presence of conflicts.

Regarding local ordinal unanimity we can prove the following

Proposition 12. *Dempster, Smets and Yager combination rules obey local ordinal unanimity with respect to dominance ordering.*

Proof. We prove for Smets rule first. Suppose $w_1\succeq_1^{dom} w_2$ and $w_1\succeq_2^{dom} w_2$, namely that for any $A\subseteq W\setminus\{w_1,w_2\}$, $m_i(A\cup\{w_1\})\geq m_i(A\cup\{w_2\})$ for $i=1,2$. Then we must prove that for any $C\subseteq W\setminus\{w_1,w_2\}$,

$\sum_{C\cup\{w_1\}=F\cap G} m_1(F)\cdot m_2(G)\geq\sum_{C\cup\{w_2\}=F\cap G} m_1(F)\cdot m_2(G)$.

Note that if $C\cup\{w_1\} = F\cap G$ then $F = F'\cup C\cup\{w_1\}$ and $F = G'\cup C\cup\{w_1\}$, with $F', G'\subseteq W\setminus\{w_1,w_2\}$ and likewise for w_2. The latter inequality reads:

$\sum_{F',G'\subseteq W\setminus\{w_1,w_2\}} m_1(F'\cup C\cup\{w_1\})\cdot m_2(G'\cup C\cup\{w_1\})$
$\geq\sum_{F',G'\subseteq W\setminus\{w_1,w_2\}} m_1(F'\cup C\cup\{w_2\})\cdot m_2(G'\cup C\cup\{w_2\})$
which holds since

$m_1(F'\cup C\cup\{w_1\})\cdot m_2(G'\cup C\cup\{w_1\})\geq m_1(F'\cup C\cup\{w_2\})\cdot m_2(G'\cup C\cup\{w_2\})$ from the working assumption.

As to Dempster rule of combination, note that the mass function obtained differs from the one in Smets rule only by a multiplicative coefficient.

Yager rule respects Local Ordinal Unanimity for dominance, since the obtained bba differs from Smets rule result only by adding a constant to the mass of W. □

The combination rule (23) directly extends the basic fusion rule (3) for two sets, from Section 4.1, and was proposed by Dubois and Prade [45]. Again, it coincides with Dempster's rule of combination and Smets' conjunctive rule, and Yager's as well if and only if the two mass functions are strongly consistent, that is, $\forall F$, $m_1(F) > 0$, $\forall G$, $m_2(G) > 0$, $F\cap G\neq\emptyset$.

Table 3
Properties of combination rules in evidence theory.

rule/Prop	Una	Mono	Vacuous	Cons	Opti	Fair	Min-Com
Dempster	*Yes*[1]	*Yes*	*Yes*	*Strong*[a]	*No*	*Yes*[a]	*No*
Smets	*Yes*[b]	*Yes*	*Yes*[b]	*No*	*Yes*	*No*	*Yes*[b]
Yager	*No*	*Yes*	*Yes*[b]	*Strong*	*No*	*Yes*	*Yes*
DP	*Yes*	*Yes*	*Yes*	*Strong*	*Yes*	*Yes*	*Yes*

[a] Only when defined.
[b] Trivially in case of strong inconsistency.

Proposition 13. *The DP fusion rule (23) satisfies all fusion postulates stated for belief functions.*

Proof It goes just as for the basic fusion rule for two sets (3). We only clarify the possibly not obvious issues. Impossibility preservation is ensured by the fact that masses $m_1(F) \cdot m_2(G)$ are never allocated outside $F \cup G$. As a consequence, $m_{12} \sqsubseteq_s m_{1\cup 2}$ (just reallocate to $F \cup G$ the masses $m_1(F) \cdot m_2(G)$ allocated to $F \cap G$ when $F \cap G \neq \emptyset$. For fairness, consider the No Favourite property: if $F \cap G = \emptyset$ the mass allocated to $F \cup G$ can be reallocated neither to F nor to G, and in general there is no focal set of m_i that will contain $F \cup G$, or if it does, the mass $m_{DP}(F \cup G)$ may well be larger than the sum of the masses $m_i(C)$ for $F \cup G \subseteq C$. The No Dismissal is obvious. Minimal commitment holds because, once the dependence structure is fixed, it is applied for each pair of focal sets, one for m_1, one for m_2 like for the rule (3). □

Note that Dempster rule and Smets rule are associative, while the other aggregation methods are not. However, Yager's rule is quasi-associative (apply Smets rule to the n inputs, then renormalise the result by transferring the mass on the empty set to the whole set W) and Dubois and Prade combination rule can be readily extended to $n > 2$ independent sources, using the MCS rule (4) on all n-tuples of focal sets: $\forall E \neq \emptyset$,

$$m_{DP-n}(E) = \sum_{F_1,\ldots,F_n : E = f_n^{MCS}(F_1,\ldots,F_n)} m_1(F_1) \cdot \ldots \cdot m_n(F_n) \tag{24}$$

The above results are summarised by Table 3 (all above rules are symmetric).

As already said, the fusion rules considered in this section can be generalised, replacing the product of bba's $m_1(F) \cdot m_2(G)$ by a suitably chosen joint mass function $x(E, F)$ whose marginals are m_1 and m_2. If we give up choosing a dependence structure, 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 probabilistically mutually consistent. However, there may be several minimally specific fusion rules, if we leave the choice of $x(E, F)$ open [29].

8. Information fusion in probability theory

A probability distribution can be used to serve either of two purposes: modelling of random phenomena (it is then said to be objective), and modelling belief (it is then said to be personal or subjective). In the latter case, a probability distribution is attached to the observer rather than to the observed phenomenon. This is what we assume here, since each information item we consider is attached to a source.

A subjective probability distribution is arguably the full-fledged opposite to set-valued representation of incomplete knowledge: it cannot represent incomplete knowledge understood as partial ignorance faithfully. On the contrary, probability distributions are tailored to aleatory uncertainty. This point has been abundantly documented in the literature. Three main points can be stressed [51]:

1. The uniform probability distribution is ambiguous, it cannot tell fair dice from unknown ones, but for the fact that the uniform distribution representing the former is objective, while the one representing the latter is subjective.
2. Probabilistic representations of ignorance are questionably scale or language-dependent: for instance a uniform distribution on a linear scale does not remain uniform after a non-linear monotonic transformation (e.g. logarithmic). This is acceptable if the distribution represents frequentist information, less so for ignorance.
3. In the presence of ambiguity or incomplete information, people do not make decision in agreement with expected utility based on a unique subjective probability distribution on states of nature, as revealed by Ellsberg paradox [53].

Actually a probability distribution should be understood as a weighted collection of conflicting singletons, in contrast with belief functions or possibility distributions that account for a collection of more or less reliable disjunctive sets representing incomplete information. Contrary to the basic fusion modes of disjunctive sets that come down to a conjunction or a disjunction, there are no such connectives available in probability theory (conjunction and disjunction are not closed operations for singletons). Only the setting of evidence theory can shed light on what could be a conjunction or a disjunction of probability distributions (e.g. the union of random singletons is a mass function bearing on focal sets with one and two elements).

Nevertheless, beginning in the 1960s, there is a large literature on the fusion of subjective probability measures, split in two schools. When parallel fusion is taken for granted (see for instance [24]), the result is often required to be a weighted average of original probabilities. Under the Bayesian approach [57,84], the fusion process is in fact a combined revision/fusion process due to the presence of a prior probability. Moreover, its application requires a lot of information to be available (for instance, prior probabilities).

As explained quite early by Walley [109], the framework of imprecise probability is much more convenient than the one of single subjective probability in order to discuss the problem of merging beliefs, since the latter are inherently imprecise. In fact, the representation of belief by imprecise probability is but a slight variation of the standard exchangeable betting behaviour setting [110]. Nowadays, imprecise probabilities have been extensively developed and found applications in many real-world scenarios [4].

8.1. Representing beliefs by convex sets of probabilities

In this section, it is assumed that information items take the form of convex sets of finitely additive probabilities on W, denoted by $\mathcal{M}$. This representation is justified by Walley as a rational approach to representing coherent beliefs, after some previous works by Smith [103] and Williams [112], extending De Finetti's framework for subjective probabilities [55]. We briefly provide an account of this view.

Like in the personalist approaches to probability [55], belief is measured by the propensity of an agent to buy or gamble with uncertain outcomes. Walley's theory presupposes that the real line models a well-defined currency system, where positive values represent gains and negative values represent losses. A gamble is a bounded mapping $X : W \to \mathbb{R}$, interpreted as a random function that delivers $X(w)$ currency units if the actual world turns out to be w. A set of gambles desirable for an agent is denoted by $\mathcal{D}$. It satisfies the following rationality properties [110]:

1. If $X \in \mathcal{D}$ then $\forall \lambda > 0, \lambda X \in \mathcal{D}$.
2. If $X \in \mathcal{D}$ and $Y \in \mathcal{D}$, then, $X + Y \in \mathcal{D}$.
3. If $X \geq 0$, $X \neq 0$ then $X \in \mathcal{D}$ (Accepting partial gain)
4. If $X \leq 0$, $X \neq 0$ then $X \notin \mathcal{D}$ (Refusing partial loss)

A set of gambles takes the form of a positive cone in $\mathbb{R}^W$, excluding the origin 0. There is a range of prices an agent accepts to

pay for getting gamble X. The maximal price, $\underline{E}(X)$, based on a set $\mathcal{D}$ of desirable gambles, is defined as

$$\underline{E}(X) = \sup\{x \in \mathbb{R} : X - x \in \mathcal{D}\},$$

where $[X - x](w) = X(w) - x$, also called a lower prevision. $\underline{E}(X)$ is superadditive ($\underline{E}(X+Y) \geq \underline{E}(X) + \underline{E}(Y)$), homogeneous ($\underline{E}(\lambda X) = \lambda \underline{E}(X), \forall \lambda > 0$), and such that $\underline{E}(X) \geq \inf\{X(w) : w \in W\}$. $\underline{E}(X)$ is called a lower prevision and can be interpreted as the lower bound of a family of expected values of X with respect to a set of probabilities called a *credal set*:

$$\mathcal{M}(\underline{E}) = \{P : E_P(X) \geq \underline{E}(X)\},$$

where $E_P(X)$ is the expected value of X w.r.t. probability P.

If an agent provides a set of allegedly maximal buying prices $\rho(X_1), \ldots, \rho(X_n)$ for gambles $X_1, \ldots, X_n$, this assignment is said to *ensure a sure loss* if the associated credal set $\mathcal{M}_\rho = \{P : E_P(X) \geq \rho(X_i), \forall i = 1, \ldots, n\}$ is empty. If not empty, the assignment is said to be *coherent* if $\forall i = 1, \ldots, n$, the lower expectation $\underline{E}(X_i) = \inf\{E_P(X) : P \in \mathcal{M}_\rho\}$ precisely equals $\rho(X_i)$, i.e. the proposed prices cannot be increased without changing the credal set.

Conversely, any convex subset $\mathcal{M}$ is characterised by the lower expectations on gambles $\underline{E}_{\mathcal{M}}(X) = \inf\{E_P(X) : P \in \mathcal{M}\}$ inducing the set of almost desirable gambles $\mathcal{D} = \{X : \underline{E}_{\mathcal{M}}(X) \geq 0\}$. Note that the upper expectation $\overline{E}(X) = -\underline{E}(-X)$ is the least price at which the agent accepts to sell X. In the classical exchangeable bet approach [55], the buying and selling prices are the same and $\overline{E}(X) = \underline{E}(X)$ is the expectation with respect to a single probability distribution.

Under the subjectivist interpretation due to Walley [110], epistemic states are characterised by a credal set. Of course, credal sets can also represent imprecise information about an otherwise precise random phenomenon. In any case, information supplied by a source can take the form of a credal set, which is more general than all other frameworks encountered so far. Credal sets $\mathcal{M}$ induce upper and lower probabilities on events, of the form

$$\underline{P}(A) = \inf\{P(A) : P \in \mathcal{M}\}; \quad \overline{P}(A) = \sup\{P(A) : P \in \mathcal{M}\} = 1 - \underline{P}(A^c)$$

Degrees of belief attached to propositions take the form of lower probabilities. However, the credal set $\mathcal{M}$ can generally not be recovered from the knowledge of the lower probabilities on events it induces, as in general $\mathcal{M} \subset \{P : P(A) \geq \underline{P}(A)\}$. Nevertheless there are noticeable set-functions that are special cases of coherent lower or upper probabilities [111]:

- 2-monotone (convex) capacities g such that $g(A \cup B) + g(A \cap B) \geq g(A) + g(B)$. Then $\mathcal{M}(g) = \{P : P(A) \geq g(A)\} \neq \emptyset$. It holds that $\underline{P}(A) = \inf\{P(A) : P \in \mathcal{M}(g)\} = g(A)$.
- Belief functions, that are ∞-monotone: $\mathcal{M}(m) = \{P : P(A) \geq Bel(A)\} \neq \emptyset$. It holds that $\underline{P}(A) = \inf\{P(A) : P \in \mathcal{M}(m)\} = Bel(A)$.
- Necessity measures: $\mathcal{M}(\pi) = \{P : P(A) \geq N(A)\} \neq \emptyset$. It holds that $\underline{P}(A) = \inf\{P(A) : P \in \mathcal{M}(\pi)\} = N(A)$.
- Sets: $\mathcal{M}(T) = \{P : P(T) = 1\}$.
- Single probability measures: $\mathcal{M} = \{P\}$ is then a singleton.

The information ordering between credal sets $\mathcal{M}$ and $\mathcal{M}'$ is defined by the inclusion ordering $\mathcal{M} \subseteq \mathcal{M}'$, whereby $\mathcal{M}$ is at least as informative as $\mathcal{M}'$. Two credal sets are mutually inconsistent if their intersection is empty. Hence, basic concepts of informational comparison and mutual consistency for credal sets are the same as for mere set-valued representations of information items. The plausibility ordering induced by $\mathcal{M}$ can be defined by means of the upper probability of singletons, in agreement with the possibility theory and belief function settings.

8.2. Walley's merging axioms

In an unpublished research note [109], Walley gave a detailed account of postulates for information fusion taking the form of credal sets representing agent opinions, emphasising four general requirements

a) the aggregated opinion should be coherent;
b) if all agents of the group desire a certain gamble, so should reflect the aggregated result;
c) an aggregation result should at least partially reflect the opinions of each agent (what Walley calls a *reconciliation*);
d) an aggregation result should reflect any level of indeterminacy shared by all agents.

These concerns prove to be in agreement with our approach to information fusion, and Walley's unpublished report can clearly be seen as a pioneering reflexion on information fusion. To implement these four concerns, fifteen axioms were precisely stated mathematically, in terms of credal sets, sets of desirable gambles, and lower previsions, respectively. Below, we recall these axioms (numbered after Walley and using his terminology) and compare them with our common properties listed above, emphasising the credal set point of view that is more in agreement with our set-based setting.

Suppose n sources supplying n credal sets $\mathcal{M}_i$ (with associated desirable sets of gambles $\mathcal{D}_i$, and lower previsions $\underline{E}_i(X)$), and let the merged result be $f(\mathcal{M}_1, \ldots, \mathcal{M}_n)$.

Criterion 1: (Coherence). $f(\mathcal{M}_1, \ldots, \mathcal{M}_n)$ should be a credal set (a convex set of probability measures).
This criterion is captured by our Property 3 (Consistency enforcement).

Criterion 2: (Unanimity). $f(\mathcal{M}_1, \ldots, \mathcal{M}_n) \subseteq H(\bigcup_{i=1}^n \mathcal{M}_i)$ where H denotes the convex hull.
This criterion is a variant of our property 1b (Impossibility preservation). The use of a convex hull is motivated by the need to get a credal set as a result, which $\bigcup_{i=1}^n \mathcal{M}_i$ is not. It can be justified in terms of desirable gambles as the property is provably equivalent to $\bigcap_{i=1}^n \mathcal{D}_i \subseteq \mathcal{D}$ where $\mathcal{D}$ is the set of desirable gambles associated to $f(\mathcal{M}_1, \ldots, \mathcal{M}_n)$ (requirement (b) above). In terms of lower previsions, it is $\underline{E}(X) \geq \min_{i=1}^n \underline{E}_i(X)$, which also provides $\min_{i=1}^n \underline{P}_i(A)$ as a lower bound to the resulting lower probabilities.

Criterion 3: (Reconciliation). If $\bigcap_{i=1}^n \mathcal{M}_i = \emptyset$, then $f(\mathcal{M}_1, \ldots, \mathcal{M}_n) \cap \mathcal{M}_i \neq \emptyset, \forall i = 1, \ldots, n$.
It corresponds exactly to our Fairness property 5. In terms of desirable gambles, Walley proves that this condition comes down to never let a gamble X be accepted by the group of sources if one of them finds its opposite $-X$ strictly desirable. In terms of lower previsions, it is $\underline{E}(X) \leq \min_{i=1}^n \overline{E}_i(X)$, which also provides $\min_{i=1}^n \overline{P}_i(A)$ as an upper bound to the resulting lower probabilities. Clearly, this property ensures that if the opinions of the sources are described by different single probability functions, the result can then never be an additive probability measure, as it should be a credal set containing all input distinct probabilities.

Criterion 4: (Indeterminacy) $\bigcap_{i=1}^n \mathcal{M}_i \subseteq f(\mathcal{M}_1, \ldots, \mathcal{M}_n)$
This is clearly our Possibility preservation property 1a; it implies that the resulting information cannot be more precise than what the group considers jointly to be possible. In terms of desirable gambles it says that the collective desirable gamble set is contained in the "additive closure" of the $\mathcal{D}_i$, that is, $\{\sum_{i=1}^n X_i : X_i \in \mathcal{D}_i, \forall i = 1, \ldots, n\}$. The translation of this postulate in terms of lower previsions is less palatable.
The two stronger forms below deal with situations when agents have inconsistent beliefs.

Criterion 5: (Strong indeterminacy): It consists in replacing $\sum_{i=1}^n X_i : X_i \in \mathcal{D}_i, \forall i = 1, \ldots, n$ by $\bigcup_{i=1}^n \mathcal{D}_i$ in the previous criterion: if no agent of the group finds X desirable, then the group should not find X desirable. It is far less obvious to

express in terms of credal sets, except that ultimately, it seems to enforce $\bigcap_{i=1}^n \mathcal{M}_i \subset f(\mathcal{M}_1, \ldots, \mathcal{M}_n)$, which is hardly acceptable (it is not clear why we forbid equality in the case of consistency between the credal sets). It also reads $\underline{E}(X) \geq \max_{i=1}^n \underline{E}_i(X)$.

Criterion 6: (Strong Pareto): This criterion is again hard to express with credal sets. In terms of gambles it says that a gamble X can stand as the result of merging if it is desirable for at least one agent, and is not strictly undesirable for other agents. Despite its intuitive appeal in this form, it violates coherence. This is because it reads $\underline{E}(X) \geq \min(\max_{i=1}^n \underline{E}_i(X), \min_{i=1}^n \overline{E}_i(X))$. That it may lead to inconsistency is patent if we notice that the set-functions $g_1(A) = \max_{i=1}^n \underline{P}_i(A)$ and $g_2(A) = \min_{i=1}^n \overline{P}_i(A)$ can be any capacity, that is, the credal sets induced by g_1 and g_2 can be empty.

Criterion 7: (Strong reconciliation): If for some subset I of integers, $\bigcap_{i\in I}^n \mathcal{M}_i \neq \emptyset$, then $f(\mathcal{M}_1, \ldots, \mathcal{M}_n) \bigcap_{i\in I}^n \mathcal{M}_i \neq \emptyset$.
This is clearly the optimistic form of the Fairness property, already mentioned in Section 3.2, according to which the outcome of merging should be determined by the merged opinions of consistent sources. One can consider it is Fairness plus the Optimism axiom applied to subset of sources.

Criterion 8: (Conjunction): If $\bigcap_{i\in I}^n \mathcal{M}_i \neq \emptyset$, then $f(\mathcal{M}_1, \ldots, \mathcal{M}_n) \subseteq \bigcap_{i=1}^n \mathcal{M}_i$.
This criterion is covered by the Optimism postulate. In terms of desirable gambles it writes $\sum_{i=1}^n X_i \in \mathcal{D}$ whenever $X_i \in \mathcal{D}_i, \forall i = 1, \ldots, n$.

Criterion 9: (Total reconciliation): $f(\mathcal{M}_1, \ldots, \mathcal{M}_n) \supseteq \bigcup_{i=1}^n \mathcal{M}_i$
In terms of desirable gambles, it says that a gamble can be accepted by the set of sources only if it is accepted by all of them. In terms of lower previsions, $\underline{E}(X) \leq \min_{i=1}^n \underline{E}_i(X)$. This is a very conservative property at odds with Optimism, and that is questionable as it allows for non-informative merged results (e.g. $\underline{P}(A) = 0$) from informative inputs (e.g., $\underline{P}_i(A) > 0, \forall i = 1, \ldots, n$).

Criterion 10: (Symmetry): All agents play the same role if no particular additional information is available that allows to prioritise some sources over the others. This is our Commutativity property.

Criterion 11: (Complete ignorance): An agent providing no information can be ignored during merging. This criterion is exactly our Insensitivity to Vacuous Information postulate.

Criterion 12: (Relative ignorance): An agent with beliefs that can be subsumed by the beliefs of each of the remaining agents can be ignored in the result of merging. This is a strengthening of Insensitivity to Vacuous Information, which makes sense if the aggregation is idempotent. It is insensitivity to redundant information, which could read for two sources and any representation framework: if $T_1 \sqsubseteq T_2$, then $f(T_1, T_2) = T_1$. However, it becomes questionable in the case of independent sources, when reinforcement effects are allowed.

Criterion 13: (Aggregation of aggregation): Merging can be done step-by-step, merging the partial result with the information of the next source. This is a weak version of associativity, but still very demanding. Along with the three first criteria, it enforces a disjunctive combination $H(\bigcup_{i=1}^n \mathcal{M}_i)$ as a result. Then the result of merging does not depend on whether the sources are mutually consistent or not.

Criterion 14: (Continuity): This criterion says that aggregated opinions should not be sensitive to small changes in inputs. This criterion is akin to the Non-Sensitivity property. But we have seen that this criterion is not compatible with the proper handling of inconsistent sources, that require a form of merging different from consistent ones. This criterion is easy to fulfil with consistent sources only.

Criterion 15: (Updating): This is a request for commuting revision and merging: The revision of the merged information should be the result of merging revised original information items. This criterion is out of scope of our study that restricts to the postulates of information fusion. However, in a more general perspective, it may be natural to consider fusion and revision to occur at the same time.

It is clear that there is a significant overlap between our postulates in the general case, and Walley criteria for merging convex sets of probabilities. In fact Walley's axiomatic setting covers our properties 1,3, 4, 5, 6, 7 explicitly. Walley mentions no counterpart to the Information monotonicity property 2, not does he use minimal commitment in the axioms. Moreover, the expression of the criteria is exactly the same as the corresponding properties for merging set-valued information items viewed as constraints in our paper. As a consequence, basic rules characterised by his criteria are the same as the ones we obtain for merging sets: the intersection of consistent sets of states is replaced by the intersection of credal sets, applying criteria 2 (our Impossibility Preservation) and 8 (Optimism); he nevertheless obtains the union of sets as an obvious consequence of criterion 4 (our Possibility Preservation) and the debatable criterion 9. Actually, our Information Monotonicity property is a consequence of the other ones for sets.

Note that Walley does not consider all his criteria on a par. He only considers 1–4 as essential (our properties 1, 3, 5), and criteria 5, 6, 9, 13, 14 as dubious (they correspond indeed to properties we either discarded or did not mention).

Besides, minimal commitment is actually a major building block of Walley's theory, yielding the so-called natural extension. It always computes the least informative credal set in agreement with available information (maximal credal set compatible with constraints [110]). This is equivalent to choosing a set of desirable gambles listed by an agent, using the 4 requirements (a-d) listed at the beginning of this subsection, and that coherent sets of desirable gambles should satisfy. The setting of imprecise probability theory has this feature, shared by possibility and belief function theory, namely its concern to represent information at the exact degree of inderminacy it contains. This is in opposition to single probability distributions that cannot account for incompleteness of information. In our general approach to information fusion, we found it instrumental to require minimal commitment as one of the basic postulates, even if it seems to exclude a purely Bayesian approach to uncertainty.

Moral and Sagrado [83] also consider postulates for merging credal sets partially inspired by Walley's. Interestingly, they again endorse Possibility and Impossibility Preservation, Insensitivity to Vacuous Information, Commutativity, Optimism (requiring the conjunction of credal sets as the proper fusion rule if they intersect), and a form of Information Monotonicity. They suggest to simplify the input family of credal sets, by deleting the redundant ones, and replacing them by intersections of maximal consistent subsets of the remaining ones. Then they propose a rather complex merging process based on convex combinations.

8.3. Basic fusion rules in imprecise probability theory

Walley [109] comes up with a number of merging rules that result from the conditions he posed, and that according to the above analysis fit our framework. For simplicity, we only mention the following ones, that are counterparts to fusion rules already encountered in other settings:

- If the information items are mutually consistent, the conjunction rule obtains: $f_n^{\bigcap}(\mathcal{M}_1, \ldots, \mathcal{M}_n) = \bigcap_{i\in I}^n \mathcal{M}_i$.

- If the information items are pairwise mutually inconsistent, the convex disjunction rule obtains: $f_n^{\cup}(\mathcal{M}_1, , \mathcal{M}_n) = H(\bigcup_{i=1}^n \mathcal{M}_i)$, also called *unanimity* rule by Walley.
- In the general case, Walley proposes the imprecise probability version of the MCS rule (4): $f_n^{MCS}(\mathcal{M}_1, \ldots, \mathcal{M}_n) = H(\bigcup_{I \in MCS(\{1,\ldots,n\})} \bigcap_{i \in I}^n \mathcal{M}_i \neq \emptyset)$.

Noticeably, the three combination rules are the same as in the case of set-based information items, replacing states by probability measures. Like in the case of merging sets of states, the Walley-unanimity (disjunctive) rule is questionable when the credal sets proposed by sources are globally consistent (it violates Optimism), and the conjunction rule is not coherent when sources are inconsistent. On the other hand, the MCS rule satisfies all of our postulates, as well as those of Walley, but for the dubious criteria 5, 6, 9, 13, 14. It satisfies the updating criterion (Walley says), and reduces to the conjunction rule if sources are globally inconsistent and to the unanimity rule if they are pairwise mutually inconsistent.

8.4. Probabilistic fusion methods

The unanimity (disjunction) fusion is especially interesting when merging single probabilities, which, as soon as they are distinct are pairwise conflicting. Namely, if $\mathcal{M}_i = \{P_i, i = 1, \ldots, n\}$, then $f_n^{\cup}(\mathcal{M}_1, \ldots, \mathcal{M}_n) = H(\{P_1 \ldots P_n\})$. As pointed out by Walley, this fusion rule is at odds with usual approach to probabilistic fusion that results in a single probability measure that differs from all other input probability measures, hence violating Fairness. This is in particular true for one of the two main approaches to probabilistic fusion described as follows:

- The weighted average rule, promoted by Cooke [24], of the form $f_n^{av}(P_1, \ldots, P_n) = \sum_{i=1}^n \alpha_i P_i$ where $\sum_{i=1}^n \alpha_i = 1$. This rule proves to be the only one ensuring that $f(P_1(A), \ldots, P_n(A))$ is a probability measure, assuming $f(1, \ldots, 1) = 1$ and $f(0, \ldots, 0) = 0$, and being stable under marginalisation (Mc Conway [81], Lehrer and Wagner [108])[12].
- The Bayesian fusion rule, which assumes pieces of information take the form of observations $w_1, \ldots, w_n$, along with conditional probability functions (likelihoods) $P(w_i|w)$ of observing w_i if the real state is w. Then a prior probability P on W is supposed to be available. It enables Bayes rule to be applied, yielding $P(w|w_1, \ldots, w_n)$, assuming in the simplest case that $P(w_1, \ldots, w_n|w)$ is obtained as $\prod_{i=1}^n P(w_i|w)$ (this is called naive Bayesian fusion).

There are several difficulties when trying to apply our framework to the above probabilistic combination rules. The weighted average rule is generally not symmetric. Considering the mere arithmetic average so as to stay in agreement with our assumptions, the issue is whether we consider a probability measure P as a singleton $\{P\}$, i.e. as a very special credal set, as Walley [109] did, or if we see P as just a kind of distribution pointing out some states are more probable than other ones.

In the first view, as already pointed out, the result of the merging by arithmetic mean will be none of the input probabilities, hence violating Impossibility Preservation. This point of view could be challenged if we admit that sources can have different beliefs (disjoint credal sets) about the state of the world while still being in relative agreement about what this state is. In other words, for the subjective Bayesian, the notion of mutual consistency between credal sets adopted by Walley will sound too strong: the fusion problem is not so much to check the overlap between $\{P_1\}$ and $\{P_2\}$ viewed as credal sets (which sounds barren), as the one of measuring consistency of P_1 and P_2, viewed as restricting subsets of possible states.

And indeed, if we adopt the latter view with a strict Bayesian approach, we can encode any set-valued information item T as a uniform probability over its support following Laplace principle of insufficient reason. Then we are bound to use the arithmetic mean of characteristic functions of sets or fuzzy sets (properly renormalised in the guise of probability distributions) as the main fusion tool to be evaluated in the face of our postulates. It is obvious that several postulates are then verified: Possibility and Impossibility Preservation (with respect to possible states, not probabilities as in the credal set view), Consistency Enforcement, Fairness, Commutativity. However, Insensitivity to Vacuous Information is clearly not respected, if vacuous means uniform over W (but again no Bayesian probabilistic information is vacuous...). Moreover, we are at a loss defining mutual consistency and information ordering for probability functions, which may prevent us from writing the other postulates. Mutual consistency may just refer to non-overlapping supports of distributions (a very weak form in probabilistic terms). For information ordering, one way out is to compare the entropies of the probability distributions. But then, Optimism would be violated by the arithmetic mean, since the entropy of the average of two overlapping distributions may be greater than the ones of each input distribution, while Optimism requires that the resulting informativeness should increase from merging consistent inputs.

Interestingly, the comparison between entropies of two probability distributions yields an ordering that refines the specificity ordering of possibility distributions that can be obtained by a suitable transformation (called the Lorentz curve) of these probability distributions (see [35] for results in the finite setting and an overview of this literature). In other words, if we consider probability distributions as one (very constrained) way of representing imprecise information (as is the case with subjective probabilities), looking for postulates of merging in the style of our approach seems to lead us back to the set-theoretic setting, for mutual consistency and information ordering.

This is even more patent in the case of Bayesian fusion, where input information takes the form of likelihood functions that are well-known to be a kind of possibility distributions [23,39]. When the prior information takes the form of a uniform distribution over W, what naive Bayesian fusion achieves is a variant of the normalised product fusion rule (6) of possibility distributions under vacuous prior information (the normalisation factor is such as to recover a probability measure). The Bayesian fusion actually combines revision (of the prior probability) and fusion (of the likelihood functions); see [52] for more details and the counterpart to Bayesian fusion in possibility theory, using possibilistic priors.

Note that as probabilities are special belief functions, the fusion rules of evidence theory considered in the previous section still apply to combine independent probability functions P_1 and P_2, if we allow for the result to lie in a more general family of set-functions:

- The conjunctive rule of Smets leads to a submormalised distribution such that $P(w) = P_1(w) \cdot P_2(w)$. It can be renormalised in a standard way (this is Dempster rule) or by assigning the mass $\sum_{w_1 \neq w_2} P_1(w_1) \cdot P_2(w_2)$ to W, which is no longer a probability distribution. This mass is generally quite high.
- The DP fusion rule coincides with the disjunctive rule when applied to probabilities and yields a mass function bearing on singletons and doubletons:

$$\begin{aligned} m_{DP}(\{w\}) &= P_1(w) \cdot P_2(w) = m_{\cup}(\{w\}) \\ m_{DP}(\{w_1, w_2\}) &= P_1(w_1) \cdot P_2(w_2) + P_1(w_2) \cdot P_2(w_1) \\ &= m_{\cup}(\{w_1, w_2\}) \end{aligned}$$

[12] This result has a counterpart in possibility theory in terms of weighted maximum of possibility measures [47].

The corresponding belief function is defined by $Bel_{DP}(A) = P_1(A) \cdot P_2(A)$ since, in general, $Bel_{\cup}(A) = Bel_1(A) \cdot Bel_2(A)$ for the disjunctive rule [42]. This approach is almost never used since it is always assumed that merging probabilities must yield probabilities. Yet the DP-rule, which is a special case of MCS, seems to be reasonable as it accounts for the discrepancies between sources in a style similar to Walley's unanimity rule. The latter computes the credal set $\mathcal{M} = \{\lambda P_1 + (1-\lambda)P_2 : \lambda \in [0,1]\}$ and thus yields a more informative result than the DP rule, since the lower prevision it provides is $\underline{P}(A) = \min(P_1(A), P_2(A)) > Bel_{DP}(A)$, but it is not a belief function.

8.5. Ontic vs. epistemic views of credal sets

At this point it is useful to reconsider imprecise probabilities in the scope of the opposition between frequentist and subjectivist views of imprecise probability. Indeed, this debate sheds light on the meaning of credal set fusion as envisaged by Walley [109]. If we adopt the frequentist standpoint, a probability measure P represents a model of a random phenomenon, hence it can be viewed as a possible world, and a credal set represents imprecise information about the unknown probability. Hence we can argue that we are back to the case of set-based information (Section 4.1), replacing deterministic states of the world by stochastic ones. This is the sensitivity analysis interpretation of credal sets: a credal set is just a standard epistemic set, containing the actual probability distribution. No surprise then that Walley's postulates fit our general setting, and were precisely pioneering it. And Walley's unanimity rule looks convincing as a way to reconcile credal sets representing imprecise probabilistic models of reality obtained from several sources (in particular, the convex hull of alternative precise probabilities).

However, adopting the subjectivist standpoint under Walley's approach, a credal set $\mathcal{M}$ does not represent imprecise knowledge about an ill-known subjective probability, it represents the agent's beliefs in a singular event, just as in the previous section of this paper. Walley considers [110] that there is no such thing as an ill-known subjective probability: beliefs in imprecise probability theory are directly modelled by lower previsions and in particular, lower probabilities (this is the so-called *direct interpretation* of lower prevision [110], section 2.10). Hence, under this view a credal set is not a model of an ill-known probability: it is the lower prevision function that represent the agent's epistemic state. Then $\mathcal{M}$ is an ontic set [25,82], a precise mathematical representation of the lower prevision function, which in turn may represent imprecise knowledge of a determinic state $w \in W$. Under this view, Walley's setting for fusion becomes somewhat questionable. The support of $\mathcal{M}$ should be the set of states having positive upper prevision $\mathcal{S}(\mathcal{M}) = \{w : \overline{P}(w) > 0\}$ as in possibility and evidence theories, not $\mathcal{M}$ itself (which is only compatible with an objectivist view of probabilities). Interestingly, even if $\mathcal{M}_1 \cap \mathcal{M}_2 = \emptyset$, the supports of $\mathcal{S}(\mathcal{M}_1)$ and $\mathcal{S}(\mathcal{M}_2)$ may overlap (obvious, considering the case singletons). So, while as epistemic representations of a stochastic reality, $\mathcal{M}_1$ and $\mathcal{M}_2$ totally conflict, they may still be partially mutually consistent as ontic representations of agent's beliefs about a precise state in W. It suggests that directly expressing merging rules in terms of lower or upper previsions may lead to postulates different from those using credal sets. Typically, Possibility Preservation, in the form of Walley criterion 4, whose expression in terms of lower previsions is too complex; one may argue it should be written as $\mathcal{S}(\mathcal{M}_1) \cap \mathcal{S}(\mathcal{M}_2) \subseteq f(\mathcal{S}(\mathcal{M}_1), \mathcal{S}(\mathcal{M}_2)))$. The fact that the interpretation of credal sets may affect the definition of some basic notions pertaining to them is not surprising. The same holds for the notions of independence in imprecise probability theory that differ if credal sets are imprecise descriptions of probability functions or precise characterisations of lower previsions [110].

These considerations suggest more work is needed to put information fusion of lower previsions in the proper subjectivist perspective, and especially to find some that are in agreement with fusion rules in possibility and evidence theories.

9. Conclusion

In this paper, we have provided a general framework for analysing fusion operators proposed in different settings, in a unified way. We have investigated various representation formats including classical sets, knowledge bases, total pre-orders, possibility theory, imprecise probabilities, and evidence theory, considering a few representative sampling of operators. It is clear that the analysis may be carried out even more systematically, as well as applied to other settings, whether numerical or ordinal, e.g., ranking functions [104] (that we nevertheless also discuss briefly).

Our aim was to propose a unified view of information fusion, as the task of finding what is true out of a collection of testimonies, by studying the conflicts between them, bypassing the boundaries of various disciplines and formalisms. We could lay bare 8 universal postulates that seem to fit this purpose, and we showed they are at work in most known fusion rules in the literature from various unrelated areas. Some properties not among the basic ones can still be useful, in the sense that a fusion rule will be criticised if it does not satisfy them, like associativity (unlikely in the presence of conflicting information), or insensitivity to small changes (e.g. Dempster's rule is oversensitive to small changes of input values). Some other properties can be useful in some situations but not possessed by many fusion rules (e.g., idempotency). Our postulates have no claim to be relevant to other types of aggregation tasks such as voting, or preference merging, or multicriteria decision-making. We focalise here on imprecise or uncertain testimonies in the form of elementary pieces of information.

A natural line for further work is to adapt the basic properties to prioritised merging. It would lead us to an enlarged setting where belief revision and fusion could be considered in the same range of formal framework, since belief revision can be viewed as a form of prioritised fusion, or yet constrained fusion, and Bayesian fusion methods are in the same vein. Some relevant preliminary results can be found in [26,32,60,67]. The study of combination rules for dependent sources is another open issue.

Acknowledgements

Part of the research was carried out while Didier Dubois was appointed as a World Class Visiting Professor at Queen's University Belfast. This support is gratefully acknowledged. This work has also been partially supported by EPSRC project EP/J012149/1.

References

[1] M.A. Abidi, R.C. Gonzalez, Data Fusion and Machine Intelligence, Academic Press, 1992.
[2] C.E. Alchourrón, P. Gärdenfors, D. Makinson, On the logic of theory change: Partial meet functions for contraction and revision, Symb. Logic 50 (1985) 510–530.
[3] K.J. Arrow, A.K. Sen, K. Suzumura (Eds.), Handbook of Social Choice and Welfare, vol. 1, North-Holland, 2002.
[4] T. Augustin, F.P. Coolen, G. de Cooman, M.C.M. Troffaes (Eds.), Introduction to Imprecise Probability, Wiley, New York, N.Y., 2014.
[5] G. Beliakov, A. Pradera, T. Calvo, Aggregation Functions: A Guide for Practitioners, Springer, Berlin, 2008.
[6] N.D. Belnap, J.M. Dunn, G. Epstein, A useful four-valued logic, Modern Uses of Multiple-Valued Logic, D. Reidel Publishing Company, 1977, pp. 8–37.
[7] S. Benferhat, D. Dubois, S. Kaci, H. Prade, Possibilistic merging and distance-based fusion of propositional information, Ann. Math. Artif. Intell. 34 (1-3) (2002) 217–252.

[8] S. Benferhat, D. Dubois, H. Prade, From semantic to syntactic approaches to information combination in possibilistic logic, in: B. Bouchon-Meunier (Ed.), Aggregation and Fusion of Imperfect Information, Physica Verlag, Heidelberg, 1998, pp. 141–161.

[9] S. Benferhat, D. Dubois, H. Prade, Possibilistic and standard probabilistic semantics of conditional knowledge bases, J. Logic Comput. 9 (1999) 873–895.

[10] S. Benferhat, D. Dubois, H. Prade, M. Williams, A framework for revising belief bases using possibilistic counterparts of Jeffrey's rule, Fundam. Inf. 11 (2009) 1–18.

[11] S. Benferhat, S. Kaci, Logical representation and fusion of prioritized information based on guaranteed possibility measures: Application to the distance-based merging of classical bases, Artif. Intell. 148 (1–2) (2003) 291–333.

[12] S. Benferhat, C. Sossai, Reasoning with multiple-source information in a possibilistic logic framework, Inf. Fus. 7 (1) (2006) 80–96.

[13] I. Bloch, Fusion of image information under imprecision and uncertainty: numerical methods, in: G. Della Riccia, H.-J. Lenz, R. Kruse (Eds.), Data Fusion and Perception, CISM Courses and Lectures, Vol. 431, 2001, pp. 135–168.

[14] I. Bloch, Fuzzy sets for image processing and understanding, Fuzzy Sets Syst. 281 (2015) 280–291.

[15] I. Bloch, A. Hunter, Fusion: General concepts and characteristics, Int. J. Intell. Syst. 16 (10–11) (2001) 1107–1136. Special Issue on Data and Knowledge Fusion.

[16] G. Bonanno, Rational choice and AGM belief revision, Artif. Intell. 173 (12-13) (2009) 1194–1203.

[17] A. Bronselaer, D.V. Britsom, G. De Tré, A framework for multiset merging, Fuzzy Sets Syst. 191 (2012) 1–20.

[18] A. Bronselaer, D.V. Britsom, G. De Tré, Propagation of data fusion, IEEE Trans. Knowl. Data Eng. 27 (5) (2015) 1330–1342.

[19] A. Bronselaer, D.V. Britsom, G. De Tré, Pointwise multi-valued fusion, Inf. Fus. 25 (2015) 121–133.

[20] A. Chateauneuf, Combination of compatible belief functions and relation of specificity, Advances in the Dempster-Shafer Theory of Evidence, John Wiley & Sons, Inc., New York, NY, USA, 1994, pp. 97–114.

[21] L. Cholvy, S. Moral, Merging databases: problems and examples, Int. J. Intell. Syst. 16(10) (2001) 1193–1221.

[22] S. Chopra, A.K. Ghose, T.A. Meyer, Social choice theory, belief merging, and strategy-proofness, Inf. Fus. 7 (1) (2006) 61–79.

[23] G. Coletti, R. Scozzafava, Conditional probability, fuzzy sets, and possibility: a unifying view, Fuzzy Sets Syst. 144 (1) (2004) 227–249.

[24] R. Cooke, Experts in Uncertainty, Department of Mathematics, Delft University of Technology, Delft, Holland, 1990.

[25] I. Couso, D. Dubois, Statistical reasoning with set-valued information: Ontic vs. epistemic views, Int. J. Approx. Reason. 55 (7) (2014) 1502–1518.

[26] J. Delgrande, D. Dubois, J. Lang, Iterated revision as prioritized merging, in: Proceedings of KR'06, 2006, pp. 210–220.

[27] A.P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Ann. Stat. 28 (1967) 325–339.

[28] S. Destercke, T. Burger, Toward an axiomatic definition of conflict between belief functions, IEEE T. Cybern. 43 (2) (2013) 585–596.

[29] S. Destercke, D. Dubois, Idempotent conjunctive combination of belief functions: Extending the minimum rule of possibility theory, Inf. Sci. 181 (18) (2011) 3925–3945.

[30] S. Destercke, D. Dubois, E. Chojnacki, Possibilistic information fusion using maximal coherent subsets, IEEE Trans. Fuzzy Syst. 17 (1) (2009) 79–92.

[31] D. Dubois, Belief structures, possibility theory and decomposable confidence measures on finite sets, Comput. Artif. Intell. (Bratislava) 5 (5) (1986) 403–416.

[32] D. Dubois, Information fusion and revision in qualitative and quantitative settings - steps towards a unified framework, in: W. Liu (Ed.), Proceedings of the 11th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty - , ECSQARU 2011, Belfast, UK, June 29-July 1, 2011, volume 6717 of Lecture Notes in Computer Science, Springer, 2011, pp. 1–18.

[33] D. Dubois, H. Fargier, P. Perny, Qualitative decision theory with preference relations and comparative uncertainty: An axiomatic approach, Artif. Intell. 148 (1-2) (2003) 219–260.

[34] D. Dubois, H. Fargier, H. Prade, Beyond min aggregation in multicriteria decision: (ordered) weighted min, discri-min, leximin, in: R.R. Yager, J. Kacprzyk (Eds.), The Ordered Weighted Averaging Operators, Kluwer Academic Publishers, Norwell, MA, USA, 1997, pp. 181–192.

[35] D. Dubois, E. Hüllermeier, Comparing probability measures using possibility theory: A notion of relative peakedness, Int. J. Approx. Reason. 45 (2) (2007) 364–385.

[36] D. Dubois, J. Lang, H. Prade, Possibilistic logic, in: Handbook of Logic in Artificial Intelligence and Logic Programming, 3, Oxford University Press, Oxford, UK, 1994, pp. 439–513.

[37] D. Dubois, W. Liu, J. Ma, H. Prade, A principled discussion of information combination rules in different representation settings, in: Proceedings 3rd International Conference on Soft Computing and Pattern Recognition (SoCPaR'11), Dalian, IEEE, 2011, pp. 446–451.

[38] D. Dubois, W. Liu, J. Ma, H. Prade, et al., Toward a general framework for information fusion, in: V. Torra, et al. (Eds.), Proceedings Modeling Decisions for Artificial Intelligence – MDAI 2013, Barcelona, Spain, volume 8234 of LNCS, Springer, 2013, pp. 37–48.

[39] D. Dubois, S. Moral, H. Prade, A semantics for possibility theory based on likelihoods, J. Math. Anal. Appl. 205 (1997) 359–380.

[40] D. Dubois, H. Prade, On the unicity of Dempster rule of combination, Int. J. Intell. Syst. 1 (2) (1986) 133–142.

[41] D. Dubois, H. Prade, A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets, Int. J. Gen. Syst. 12 (1986b) 193–226.

[42] D. Dubois, H. Prade, The principle of minimum specificity as a basis for evidential reasoning, in: B. Bouchon, R. Yager (Eds.), Uncertainty in Knowledge-Based Systems, Springer Verlag, 1987, pp. 75–84.

[43] D. Dubois, H. Prade, Une approche ensembliste de la combinaison d'informations imprécises ou incertaines, Revue d'Intell. Artif. 1 (4) (1987) 23–42.

[44] D. Dubois, H. Prade, Possibility Theory: An Approach to Computerized Processing of Uncertainty, Plenum Press, 1988.

[45] D. Dubois, H. Prade, Representation and combination of uncertainty with belief functions and possibility measures, Comput. Intell. 4 (1988) 244–264.

[46] D. Dubois, H. Prade, On incomplete conjunctive information, Comput. Math. Appl. 15 (1988) 797–810.

[47] D. Dubois, H. Prade, Aggregation of possibility measures, in: J. Kacprzyk, M. Fedrizzi (Eds.), Multiperson Decision Making using Fuzzy Sets and Possibility Theory, Kluwer, Dordrecht, 1990, pp. 55–63.

[48] D. Dubois, H. Prade, When upper probabilities are possibility measures, Fuzzy Sets Syst. 49 (1992) 65–74.

[49] D. Dubois, H. Prade, Possibility theory and data fusion in poorly informed environments, Control Eng. Pract. 2 (1994) 811–823.

[50] D. Dubois, H. Prade, Possibility theory: Qualitative and quantitative aspects, in: D. Gabbay, P. Smets (Eds.), Handbook of Defeasible Reasoning and Uncertainty Management Systems, 1, Kluwer, 1998, pp. 169–226.

[51] D. Dubois, H. Prade, Formal representations of uncertainty, in: D. Bouyssou, D. Dubois, M. Pirlot, H. Prade (Eds.), Decision-Making - Concepts and Methods, Wiley, 2009, pp. 85–156. chapter 3.

[52] D. Dubois, H. Prade, R.R. Yager, Merging fuzzy information, in: J. Bezdek, D. Dubois, H. Prade (Eds.), Fuzzy Sets in Approximate Reasoning and Information Systems, The Handbooks of Fuzzy Sets Series, Kluwer, Boston, Mass., 1999, pp. 335–401.

[53] D. Ellsberg, Risk, ambiguity and the Savage axioms, Q. J. Econ. 75 (1961) 643–669.

[54] P. Everaere, S. Konieczny, P. Marquis, Disjunctive merging: Quota and gmin merging operators, Artif. Intell. 174 (12-13) (2010) 824–849.

[55] B. De Finetti, Theory of Probability, New York: Wiley & Sons., 1974.

[56] P. Gärdenfors, Knowledge in Flux, MIT Press, 1988.

[57] C. Genest, J.V. Zidek, Combining probability distributions: A critique and an annoted bibliography, Stat. Sci. 1 (1) (1986) 114–135.

[58] A. Grove, Two modellings for theory change, J. Philos. Logic 17 (1988) 157–180.

[59] A. Hunter, W. Liu, A context-dependent algorithm for merging uncertain information in possibility theory, IEEE Trans. Syst. Man Cybern. Part A 38 (6) (2008) 1385–1397.

[60] A. Hunter, W. Liu, Knowledge base stratification and merging based on degree of support, Proceedings of the 10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'09), LNAI, Springer, 2009, pp. 383–395.

[61] T. Inagaki, Interdependence between Safety-Control Policy and Multiple-Sensor Schemes Via Dempster-Shafer Theory, IEEE Trans. Reliab. 40 (2) (1991) 182–188.

[62] S. Kaci, S. Benferhat, D. Dubois, H. Prade, A principled analysis of merging operations in possibilistic logic, in: Proceedings of the 16th Conference in Uncertainty in Artificial Intelligence UAI,'00, Stanford, 2000, pp. 24–31.

[63] H. Katsuno, A.O. Mendelzon, Propositional knowledge base revision and minimal change, Artif. Intell. 52 (1991) 263–294.

[64] F. Klawonn, P. Smets, The dynamic of belief in the transferable belief model and specialization-generalization matrices, Proceedings of UAI, Morgan Kaufmann, San Mateo, CA, 1992, pp. 130–137.

[65] E.P. Klement, R. Mesiar, E. Pap, Triangular Norms, Kluwer Academic, Dordrecht, 2000.

[66] S. Konieczny, J. Lang, P. Marquis, Distance based merging: A general framework and some complexity results, in: Proceedings of the 8th International Conference on Knowledge representation and Reasoning (KR'2002), Toulouse, 2002, pp. 97–108.

[67] S. Konieczny, R.P. Pérez, Merging information under constraints: A logical framework, J. Logic Comput. 12 (2002) 773–808.

[68] S. Konieczny, R. Pino-Pérez, On the logic of merging, in: Proceedings of KR'98, 1998, pp. 488–498.

[69] S. Konieczny, R. Pino-Pérez, Logic based merging, J. Philos. Logic 40 (2) (2011) 239–270.

[70] J. Lang, P. Liberatore, P. Marquis, Propositional independence - formula-variable independence and forgetting, J. Artif. Intell. Res. 18 (2003) 391–443.

[71] N. Laverny, J. Lang, From knowledge-based programs to graded belief based programs, part i: On-line reasoning, Synthese 147 (2) (2005) 277–321.

[72] D. Lewis, Counterfactuals and comparative possibility, J. Philos. Logic 2 (4) (1973) 418–446.

[73] P. Liberatore, M. Schaerf, Arbitration (or how to merge knowledge bases, IEEE Trans. Knowl. Data Eng. 10 (1) (1998) 76–90.

[74] W. Liu, Analyzing the degree of conflict among belief functions, Artif. Intell. 170 (2006) 909–924.

[75] W. Liu, G. Qi, D.A. Bell, Adaptive merging of prioritized knowledge bases, Fundam. Inf. 73 (3) (2006) 389–407.

[76] R.D. Luce, D.H. Krantz, P. Suppes, A. Tversky, Foundations of Measurement, Academic Press, New York, 1990.
[77] J. Ma, W. Liu, A framework for managing uncertain inputs: an axiomization of rewarding, Int. J. Approx. Reason. 52 (7) (2011) 917–934.
[78] R.P. Mahler, Statistical Multisource-Multitarget Information Fusion, Artech House, Boston, 2007.
[79] P. Maynard-Reid-II, D.J. Lehmann, Representing and aggregating conflicting beliefs, CoRR (2002). Cs.AI/0203013.
[80] P. Maynard-Reid-II, Y. Shoham, Belief fusion: Aggregating pedigreed belief states, J. Logic, Lang. Inf. 10 (2001) 183–209.
[81] K.J. McConway, Marginalization and linear opinion pools, J. Am. Stat. Assoc. 76 (374) (1981) 410–414.
[82] S. Moral, Comments on "statistical reasoning with set-valued information: Ontic vs. epistemic view" by Inés Couso and Didier Dubois, Int. J. Approx. Reason. 55 (7) (2014) 1578–1579.
[83] S. Moral, J. Sagrado, Aggregation of imprecise probabilities, in: B. Bouchon-Meunier (Ed.), Aggregation and Fusion of Imperfect Information, Physica-Verlag, Heidelberg, 1997, pp. 162–188.
[84] P.A. Morris, Combining expert judgments : a Bayesian approach, Manag. Sci. 23 (1977) 679–693.
[85] D. Mundici, The logic of Ulam games with lies, in: C. Bicchieri, M. Della Chiara (Eds.), Knowledge, Belief and Strategic Interaction, Cambridge Univ. Press, 1992, pp. 275–284.
[86] R.R. Murphy, Dempster-Shafer theory for sensor fusion in autonomous mobile robots, IEEE Trans. Robot. Autom. 14 (2) (1998) 197–206.
[87] M. Oussalah, Study of some algebrical properties of adaptive combination rules, Fuzzy Sets Syst. 114 (3) (2000) 391–409.
[88] J. Pearl, System Z: a natural ordering of defaults with tractable applications to nonmonotonic reasoning, in: Proceedings of 3d Conference on Theoretical Aspects of Reasoning about Knowledge, TARK'90, Pacific Grove, California, 1990, pp. 121–135.
[89] G. Qi, W. Liu, D.A. Bell, A comparison of merging operators in possibilistic logic, Proceedings of the 4th International Conference, Knowledge Science, Engineering and Management, KSEM, Belfast, Northern Ireland, UK, September 1-3, 2010, volume 6291 of Lecture Notes in Computer Science, Springer, 2010, pp. 39–50.
[90] G. Qi, W. Liu, D.A. Bell, Measuring conflict and agreement between two prioritized knowledge bases in possibilistic logic, Fuzzy Sets Syst. 161 (14) (2010) 1906–1925.
[91] G. Qi, W. Liu, D.A. Bell, Quota-based merging operators for stratified knowledge bases, Proceedings of the of 1st International Conference on Knowledge Science, Engineering and Management (KSEM06), LNAI, Springer-Verlag, 2006, pp. 341–353.
[92] N. Rescher, R. Manor, On inference from inconsistent premises, Theory Dec. 1 (1970) 179–219.
[93] P.Z. Revesz, On the semantics of theory change: Arbitration between old and new information, Proceedings of the Twelfth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, Washington, DC, 1993, pp. 71–82.
[94] S.A. Sandri, D. Dubois, H.W. Kalfsbeek, Elicitation, assessment and pooling of expert judgements using possibility theory, IEEE Trans. Fuzzy Syst. 3 (1995) 313–335. (Corrections, 479).
[95] S. Schockaert, H. Prade, An inconsistency-tolerant approach to information merging based on proposition relaxation, in: Proceedings of the 24th AAAI Conference on Artificial Intelligence, 2010, pp. 363–368.
[96] G.L.S. Shackle, Decision Order and Time In Human Affairs, Cambridge University Press, U.K, 1961.
[97] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.
[98] G. Shafer, Non-additive probabilities in the work of Bernoulli and Lambert, Arch. Hist. Exact Sci. 19 (1978) 309–370.
[99] P. Smets, Belief functions, in: P. Smets, A. Mamdani, D. Dubois, H. Prade (Eds.), Non-Standard Logics for Automated Reasoning, 1988, pp. 253–286.
[100] P. Smets, The combination of evidence in the transferable belief model, IEEE Trans. Pattern Anal. Mach. Intell. 12 (5) (1990) 447–458.
[101] P. Smets, Analyzing the combination of conflicting belief functions, Inf. Fusion 8 (4) (2007) 387–412.
[102] P. Smets, R. Kennes, The transferable belief model, Artif. Intell. 66 (2) (1994) 191–234.
[103] C.A.B. Smith, Consistency in statistical inference and decision, J. R. Stat. Assoc., B 23 (1961) 1–25.
[104] W. Spohn, Ordinal conditional functions: A dynamic theory of epistemic states, in: W.L. Harper, B. Skyrms (Eds.), Causation Decis., Belief Chang. Stat., Kluwer Academic Publishers, Dordrecht, The Netherlands, 1988, pp. 105–134.
[105] W. Spohn, A general non-probabilistic theory of inductive reasoning, in: R.D. Shachter, T.S. Levitt, L.N. Kanal, J.F. Lemmer (Eds.), Uncertainty in Artificial Intelligence, Vol. 4, Elsevier Science, 1990, pp. 149–158.
[106] W. Spohn, The Laws of Belief: Ranking Theory and its Philosophical Applications, Oxford University Press, UK, 2012.
[107] A. Tapia-Rosero, A. Bronselaer, R. De Mol, G. De Tré, Fusion of preferences from different perspectives in a decision-making context, Inf. Fus. 29 (2016) 120–131.
[108] C. Wagner, K. Lehrer, Rational Consensus in Science and Society, D. Reidel, 1981.
[109] P. Walley, The Elicitation and Aggregation of Beliefs, University of Warwick, 1982 Technical Report.
[110] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, 1991.
[111] P. Walley, Measures of uncertainty in expert systems, Artif. Intell. 83 (1996) 1–58.
[112] P.M. Williams, Notes on conditional previsions, Int. J. Approx. Reasoning 44 (3) (2007) 366–383. Original version available in 1975.
[113] R. Yager, On the Dempster-Shafer framework and new combination rules, Inf. Sci. 41 (1987) 93–138.
[114] R. Yager, On the specificity of a possibility distribution, Fuzzy Sets Syst. 50 (1992) 279–292.
[115] R.R. Yager, Entropy and specificity in a mathematical theory of evidence, in: L. Liu, R.R. Yager (Eds.), Classic Works of the Dempster-Shafer Theory of Belief Functions, Studies in Fuzziness and Soft Computing, Springer, 2008, pp. 291–310.
[116] R.R. Yager, L.P. Liu (Eds.), Classic Works of the Dempster-Shafer Theory of Belief Functions, Springer, 2008.
[117] A. Yue, W. Liu, A. Hunter, Approaches to constructing a stratified merged knowledge base, Proceedings of the Ninth European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'07), Springer, 2007, pp. 54–65.
[118] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets Syst. 1 (1978) 3–28.