Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT

Fuzzy Sets and Systems 142 (2004) 143–161

FUZZY sets and systems

www.elsevier.com/locate/fss

On the use of aggregation operations in information fusion processes

Didier Dubois*, Henri Prade

I.R.I.T.-CNRS Université Paul Sabatier, 118 route de Narbonne, Toulouse 31062, Cedex 4, France

Abstract

This position paper discusses the role of the existing body of fuzzy set aggregation operations in various kinds of problems where the process of fusion of items coming from several sources is central. Several kinds of membership functions can be useful according to the nature of the information to be merged: numerical vs. ordinal inputs, preferences vs. uncertain data, observations vs. constraints. In each case, some aggregation operations look more plausible or feasible than others. The aim of this discussion is to suggest directions for putting at work the results of recent mathematical investigations in the structure of aggregation operations.

Keywords: Aggregation operations; Information fusion; Multi-criteria decision; Database merging

1. Introduction

The last 10 years have witnessed a considerable interest in information fusion processes, related to the improved sophistication of information and communication sciences and their tremendous social impact. In parallel, a mathematical study of aggregation operations has been systematically carried out under the joint umbrellas of fuzzy set theory and non-classical decision theory. Traditional aggregation operations such as the weighted average, which plays a key-role in probability and classical decision theories are now acknowledged as particular cases of more general families of aggregation operations, such as Choquet integrals [44]. These traditional aggregation functions have also been articulated with logical connectives stemming from many-valued logics, and interpreted as fuzzy set unions or intersections. The latter have been generalized in the theory of triangular norms [47]. Other aggregation operations have been laid bare like symmetric sums [60], and more recently uninorms [38], null-norms [11], etc. Overall, an impressive organized

* Tel.: +33-61-556-331; fax: +33-61-556-239.
E-mail address: dubois@irit.fr (D.Dubois).

doi:10.1016/j.fss.2003.10.038

collection of aggregation operations is available, most of the time within a well-understood mathematical framework [12].

However, the question of how such mathematical tools can be applied to fusion processes has not been studied in a systematic way although some particular areas already take advantage of this advanced state-of-the-art, such as multicriteria decision making, or pattern recognition. The role of this position paper is to contribute to bridging the gap between aggregation operations and their potential users. Some information fusion paradigms are outlined, and in each of them we try to suggest which kinds of aggregation operations look best adapted. Modelling assumptions are also discussed, as the choice of a formal framework may restrict the possible choices of appropriate aggregation modes. This discussion is also greatly influenced by the results of a European Working Group on fusion processes edited by Bloch and Hunter [9].

2. The roadmap of information fusion

Information fusion consists of merging, or exploiting conjointly, several sources of information so as to answer questions of interest and make proper decisions. More precisely, the aims of information fusion can be one or several of the following:

- to improve the available knowledge about the current state of the world;
- to update the current information on a case of interest;
- to lay bare a consensual opinion, the global point of view of a group;
- to derive or improve generic knowledge by means of data.

It may be interesting to merge various kinds of information items, according to the context: sensor data, databases, expert opinions, human reports, beliefs, preferences, desires, generic knowledge such as default rules or regulations, plausible conclusions of inferences, and so on. Such items of knowledge may take various forms depending on how they are modelled formally: numbers, intervals, fuzzy sets, probability distributions, ordering relations, terms of natural language, and also sets of logical formulas in a classical or non-classical logics, especially “if–then” rules.

For the sake of clarity, it makes sense to distinguish between several kinds of information items according to their intended meaning [9, p. 1010]:

Observations: They reflect how the world is (or was) believed to be in a particular situation of interest. They may consist of numerical data, evidence, reported facts, beliefs, measurements, etc. Some observations can be called objective when supplied by a sensor [1]. Other can be called subjective if reported by a person who describes a situation, for instance the opinion of an expert on the value of the failure rate of some component in a physical system [13]. Objective observations can be tainted with uncertainty when sensors are not reliable. So they may take the form of uncertainty distributions, instead of point values. Subjective observations may take the form of assertions formulated in linguistic terms to which degrees of belief can be attached.

Knowledge: As opposed to observations that pertain to particular situations, knowledge means information that describes how the world is *generally*. In other words, it refers to some population, some class of situations, well defined or not. For instance statistical data modelled by means of a probability function pertain to a well-defined population. In contrast, the class of situations referred to by common-sense knowledge is often unclear. It is known that generally birds fly, but not clear which set of birds is precisely alluded to in such a statement. Statistical knowledge can be termed objective as it is induced from a collection of objective observations made on a representative sample of situations. Objective knowledge can take the form of probability distributions, mathematical models of laws of nature, etc. Common-sense knowledge is subjective and is often expressed by rules having exceptions, fuzzy rules and so on. It is often cast in a logical setting. It is also more qualitative than objective knowledge and may take the form of plausibility orderings on a given set of possible events or propositions.

Preference: Preference information consists of subjective descriptions of an individual or a group of people's desires about how THEY WOULD LIKE the world to be. It contrasts with subjective observations where a person is only a witness and produces a testimony. Individual or collective preferences are often either modelled quantitatively by means of utility functions, or qualitatively by means of ordering relations (see [10]). More recently, qualitative decision theory has promoted the notion of utility functions with values in finite ordered scales [32]. It is possible to adopt a setting for preference modelling that is both relational and numerical, using for instance fuzzy relations [37]. Of course, one may have unreliable information about preference. Preferences can be incomplete (partially revealed preferences), imprecise or uncertain, and often conflicting.

Regulations: They are pieces of generic information describing how the world SHOULD BE GENERALLY, according to some law. It tells what is forbidden, what is compulsory, what is permitted (explicitly or implicitly). Regulations pertain to a class of situations, not to a single situation. They may contradict individual or collective preferences. There is a great interest to-date in the formal modelling of regulations, especially in connection with computer security problems, and the handling of restricted access to websites (see [14] for instance). Regulations are often modelled by rules and expressed in some deontic logic. The merging of regulations is motivated by the detection of inconsistencies.

Given the variety of information items, it is not surprising that the choice of a fusion mode is not unique and depends on:

- the nature of the items to be merged: knowledge fusion and preference fusion may require different specific families of aggregation functions;
- the representation framework: qualitative information fusion cannot use the same mathematical tools as quantitative information fusion.

Yet we claim that to some extent it makes sense to reinterpret most information items in a unique formal setting by means of a *profile*, that is, a function μ from a set S (of possible worlds, of states, of alternatives) to a partially ordered set L acting as a scale. This is what can be called, in the wide sense, and abstractly speaking, the membership function of a fuzzy set, even if the original definition [67] is much more restrictive. Our view is also more general than the lattice-valued fuzzy sets of Goguen [41]. This claim is not an attempt to reduce the variety of representation frameworks by taking a dogmatic stand. It only accounts for the fact that fuzzy set theory is a formal framework that looks versatile enough to model, sometimes in an approximate way, the kinds of information items reported above. This framework can be either qualitative or quantitative. Besides, a generalized membership function can be interpreted in terms of a preference profile or as an uncertainty distribution, or yet have a deontic flavour. We do not claim that all items of information take this form. We only suggest that they can be interpreted, at least approximately, in this form, as shown by the following examples.

For instance, if the information comes in the form of a set K of logical formulas in some language, S is the set of interpretations of the language and the characteristic function of the set of models of K represents a (two-valued) profile ($L=\{0,1\}$) which precisely captures the meaning of K. If K is expressed in possibilistic logic [22], the induced profile will precisely be the membership function of a fuzzy set of models expressing a possibility distribution on interpretations [68]. More generally, Zadeh [69] has suggested a systematic translation of linguistic statements into possibility distributions. Besides, probability distributions can be viewed as special kinds of profiles, and families of probability measures as well as belief functions can be approximated (even if sometimes roughly) by means of numerical possibility distributions [26,27]. Lastly, preference relations on S can be viewed as ordinal profiles. The simplest theory of comparative uncertainty (based on ordering relations on events) is comparative possibility theory, first proposed by Lewis [50]. A comparative possibility relation on the power set of a finite set is completely characterized by its ordinal profile on S [29].

Taking the fairly general, even if slightly restrictive, view that any item of information can be represented, after a suitable semantic transformation, into a profile for some set S and scale L, it is clear that fusion processes can be modelled by suitable profile aggregation operations. This point makes it clear that the setting of aggregation operations, provided it is sufficiently extended to the aggregation of qualitative and ordinal profiles (not only those taking values in the unit interval), is a natural one for discussing many information fusion problems at hand. It does not mean that information fusion should be implemented on the profiles underlying the information items. For instance, we do not rule out the syntactic fusion of belief bases in classical logic. We only insist that the meaning of such fusion operations is better understood if expressed as a combination of profiles, taking advantage of the classification of aggregation operations.

3. On unipolar vs. bipolar scales

The set L acting as co-domain of a membership function μ is at least partially ordered by some irreflexive and transitive relation $>$. L generally contains distinguished elements, especially a top, denoted 1 and a bottom denoted 0, respectively, such that $\forall\alpha\in L$, $\alpha\neq 0$, 1, it holds $0>\alpha>1$. Sometimes, a specific role is played by a "mid-point" v. In order to define it, it is assumed that L possesses an order-reversing map n which is a decreasing involutive bijection from L to itself such that $n(0)=1$. Then a mid-point v is such that $n(v)=v$. On a totally ordered set, when it exists, it is unique.

How these distinguished elements are interpreted in the various application settings is crucial for a proper understanding of the role of the various existing aggregation operations. They can have a positive, a negative or a neutral flavour. For instance, the top value is supposed to rate the best possible situation, and the bottom has a neutral flavour. Then L is said to be a positive scale. On the contrary L is called a negative scale if the bottom rates the worst possible situation and the top has a neutral meaning. In those cases, mid-points, if any, play no specific role. Moreover, mapping L to itself via an order-reversing map n turns a negative scale into a positive one, and conversely. In contrast, a bipolar scale is such that the top value rates the best possible situation, and the bottom rates the worst possible one. The mid-point then serves as the neutral landmark separating positive grades from negative ones. The real line is often used as a prototype of bipolar scale. When L is simply a partially ordered set, distinguishing between these notions is more difficult.

In the theory of fuzzy sets, [0, 1]-valued membership grades clearly belong to a bipolar scale, since 1 means full membership and 0 full non-membership. Then the value 0.5 represents the crossover point which is a balance between membership and non-membership (sometimes debatably interpreted as uncertainty about membership). The role of these three anchor-points is very clear in the properties of indices of fuzziness (e.g. [53]), since extreme fuzziness is reached by the fuzzy set where all elements have membership 0.5, while making a fuzzy set less fuzzy comes down to contrast intensification (reflected by the ambiguity ordering of Mukaidono [52]).

Now using membership functions in the various contexts described in Section 2 does not necessarily goes along with the bipolar view of membership grades.

In possibility theory [25,29,68], information pertaining to generic knowledge is modelled by possibility distributions π_x attached to an ill-known quantity (an uncontrollable state variable) and is valued on a negative scale. This is the view proposed by Shackle [58]:

$$\pi_x(s) = 0 \Leftrightarrow \text{“}x = s\text{”} \quad \text{never occurs (impossibility)},$$

$$\pi_x(s) = 1 \Leftrightarrow \text{“}x = s\text{”} \quad \text{is a normal, typical, unsurprising situation (neutral value)}.$$

For instance, consider the case of set-valued statistics modelled by a basic probability assignment m and a set $\{A_1, \ldots, A_n\}$ of focal elements such that $\sum_{i=1,\ldots,n} m(A_i) = 1$, where $m(A_i)$ is the frequency of observing exactly A_i (= the probability of knowing “$x \in A_i$” only). This is what is assumed in Shafer’s [59] evidence theory. The above conventions totally fit with the induced possibility distribution

$$\pi_x(s) = \sum_{s \in A} m(A).$$

These conventions also enable beliefs on the current situation to be described, assuming $\pi_{x\mathrm{now}}(s) = \pi_x(s)$, where x^{now} is the current state of the world, supposedly a typical case of the underlying population. More generally, possibility degrees of events $\Pi(E) = \sup_{x \in E} \pi_x(s)$ are also valued on a negative scale.

Possibility distributions valued on a negative scale can also be used in preference modelling for the description of soft constraints [17]. Then the variable x is (usually, but not necessarily) a decision variable, the value of which can be decided upon. Then π_x has the following meaning:

$$\pi_x(s) = 0 \Leftrightarrow \text{the choice } x = s \text{ is not feasible, is totally rejected},$$

$$\pi_x(s) = 1 \Leftrightarrow \text{the choice } x = s \text{ is fully feasible, not objected to}.$$

Then π_x is a kind of utility function expressing only negative utility values, but no positive preference.

Lastly, such possibility distributions can also account for the semantics of deontic statements expressing implicit permission with similar conventions. Then π_x has the following meaning:

$$\pi_x(s) = 0 \Leftrightarrow \text{the choice } x = s \text{ is forbidden},$$

$$\pi_x(s) = 1 \Leftrightarrow \text{the choice } x = s \text{ is not explicitly forbidden}.$$

Positive value scales in possibility theory have been seldom used. However, note that degrees of necessity being dual of degrees of possibility ($N(A) = 1 - \Pi(A^c)$, where A^c is the complement of A)

belong to a positive scale which is the possibility scale, reversed. Indeed $N(A)=1$ expresses full belief, while $N(A)=0$ expresses a complete lack of belief in A, but not full disbelief. However, the consideration of degrees of guaranteed possibility [21,33,66] has led to the study of value profiles having a positive flavour.

When pertaining to observations, a guaranteed possibility distribution δ_x describes cases that have been actually observed and are more or less relevant for the description of a system. It is a mapping from S to L with the following conventions:

$$\delta_x(s) = 0 \Leftrightarrow x = s \quad \text{has not been observed,}$$

$$\delta_x(s) = 1 \Leftrightarrow x = s \quad \text{has been observed and is totally relevant.}$$

$\delta_x(s)$ can be interpreted as a degree of evidential support, as suggested by Weisbrod [66]. Typically, δ_x represents a database, a memory of cases, etc., while in this setting π_x rather captures the semantics of integrity constraints [33]. The coherence condition between guaranteed and usual possibility distributions is that $\delta_x(s) \leqslant \pi_x(s)$ for all values s [62]; for instance $\delta_x(s) > 0$ implies $\pi_x(s) = 1$. The pair (δ_x, π_x) forms an interval-valued fuzzy set where the upper and lower membership functions have different meanings.

When modelling preference, it is natural to represent positive profiles that represent desires, gains and the like, under the following conventions [5]:

$$\delta_x(s) = 1 \Leftrightarrow \text{the choice } x = s \text{ is a desired, fully satisfactory state,}$$

$$\delta_x(s) = 0 \Leftrightarrow \text{the choice } x = s \text{ offers no guarantee of satisfaction.}$$

The latter case means that $x = s$ is either rejected or indifferent. When modelling deontic values, explicit permission can be taken into account, namely [21]

$$\delta_x(s) = 1 \Leftrightarrow \text{the choice } x = s \text{ is explictly permitted (e.g. it has precedents),}$$

$$\delta_x(s) = 0 \Leftrightarrow \text{the choice } x = s \text{ is not explicitly permitted (but maybe legal anyway).}$$

The case of numerical bipolar scales for preference modelling is frequent in utility theory when modelling the negative impact of losses and the positive impact of gains simultaneously on the real line. It is much less common with an ordinal scale of preference (see [45]). In the case of uncertainty modelling, probability measures are clearly valued on a bipolar uncertainty scale since $P(A) = 0$ means that event A is impossible and $P(A) = 1$ means that event A is sure. The neutral value is 0.5, but the normalisation constraint of probability distributions severely restricts its role of neutral value to binary universes.

Note that if a value in a bipolar scale L is ill-known, and represented by an interval, L can be viewed as a negative scale for the upper bound of this interval, and as a positive scale for the lower bound of this interval.

Conversely, it is interesting to put together a positive and a negative scale and build a bipolar scale. For instance the set $L\Pi = \{(a,b), \max(a,b) = 1, a, b \in [0,1]\}$ can be viewed as a bipolar scale ordered by the following relation: [19], [40]

$$(a,b) \leqslant (c,d) \text{ iff either } a = c = 1 \text{ and } b \leqslant d,$$

or $b = d = 1$ and $a \geqslant c$,

or $a = 1$ and $d = 1$.

The bottom of this scale is (1,0), its top is (0, 1) and its neutral point is (1, 1). The canonical example of such a scale is the set of pairs $(\Pi(A^c), \Pi(A))$ of degrees of possibility for events A and their complements. When $(\Pi(A^c), \Pi(A)) \leqslant (\Pi(B^c), \Pi(B))$, it means that B is at least as likely (certain or plausible) as A, (1, 0) means "impossible", (0, 1) means "sure" and (1, 1) means "unknown". Giang and Shenoy [40] use it as a bipolar preference scale where the midpoint (1, 1) means "indifferent".

4. Aggregation operations

An aggregation operator is a family of functions $\{f^n, n \in N\}$, called aggregation operations, where f^n attaches to each n-tuple $(\alpha_1, \ldots, \alpha_n)$ of values from L another value $f^n(\alpha_1, \ldots, \alpha_n)$ in L [12]. Minimal requirements on aggregation operations are that they are monotonically increasing in the wide sense with respect to each argument, that the tuple containing the bottom value of L yields the bottom of L ($f^n(0, \ldots, 0) = 0$), and likewise for the top value of L ($f^n(1, \ldots, 1) = 1$). The monotonicity property can be strengthened in two steps:

- global strict increasingness: if $\forall i$, $\alpha_i > \beta_i$ then $f^n(\alpha_1, \ldots, \alpha_n) > f^n(\beta_1, \ldots, \beta_n)$. This property holds for the minimum and the maximum operations.
- Strict increasingness in each place: if $\forall i$, $\alpha_i \geqslant \beta_i$ and $\exists i$, $\alpha_i > \beta_i$ then $f^n(\alpha_1, \ldots, \alpha_n) > f^n(\beta_1, \ldots, \beta_n)$. It rules out the minimum and the maximum operations.

These properties express various forms of compatibility with the natural partial ordering of vectors of values in L, induced by the ordering in L (also called Pareto-ordering in economics). It is reasonable to consider global strict increasingness as a minimal requirement for practical applications.

There are at least four classes of aggregation operations one may consider as being currently investigated.

- Operations generalizing the notion of conjunction. They are basically the minimum and all those functions f bounded from above by the minimum operation.
- Operations generalizing the notion of disjunction. They are basically the maximum and all those functions f bounded from below by the maximum operation.
- Averaging operations: all those functions lying between the maximum and the minimum. Some of them coincide with the Boolean conjunctions (like the geometric mean on the unit interval), or the Boolean disjunctions on $\{0, 1\}$.
- Operations not affected by the midpoint of L: $f^{n+1}(\alpha_1, \ldots, \alpha_n, v) = f^n(\alpha_1, \ldots, \alpha_n)$. They are not so much studied, but uninorms [38] are a typical family thereof.

Protopypes of generalized conjunctions and disjunctions on the unit interval are triangular norms and co-norms (see [47], for an extensive treatise). They are associative, commutative and monotonically increasing in the wide sense, with suitable identity and zeros taken as the top and the bottom of the scale. Basic conjunctions of that kind are the minimum operation, the product, and the linear operation $\max(a+b-1,0)$. Strict t-norms, isomorphic to the product, are strictly increasing. Basic disjunctions are the maximum operation, the "probabilistic sum" $a+b-a\cdot b$, and the bounded sum $\min(1,a+b)$. Virtually all other practically interesting (e.g. continuous) are isomorphic transforms of these ones or locally coincide with one of these transforms. A companion family of generalized conjunctions are copulas [61] which are instrumental in the combination of marginal probability distributions and capture stochastic dependence. Two-place copulas need not be associative but lie between the minimum and the linear triangular norms. Triangular-norms which are copulas are characterized by a Lipschitz condition.

Averaging operations are best exemplified by the arithmetic mean, and its isomorphic transforms. Virtually no averaging operation is associative, but other weaker properties such as bisymmetry are usually requested. The latter expresses the ability to aggregate entries of a matrix either starting by aggregating elements in rows, and then the partial results, or starting with aggregating elements in columns. It is known for a long time [39] that the only associative averaging operations are of the form median(a,b,α) for $0<\alpha<1$, a poorly expressive family of operations.

Uninorms are associative operations on the unit interval, monotonically increasing in the wide sense, whose identity v is such that $0<v<1$. It plays the role of a neutral element on the scale [0, 1], and can typically be chosen as 0.5. Uninorms are often obtained by "gluing" together a triangular norm (on $[0,v]$) and a triangular conorm (on $[v,1]$) [38]. Some uninorms are associative on (0, 1) such as Dombi's operation $ab/(ab+(1-a)(1-b))$. Such uninorms are transforms of the addition on the real line by monotonic functions [47, p. 224]. They are adapted to working on bipolar scales. This operation is also a symmetric sum, like the arithmetic mean, which means it is invariant via a De Morgan-like transformation (taking $1-\cdot$ as the negation).

Considering the case of qualitative value scales (such as finite chains), the range of operations is much restricted. For instance, the basic triangular norms on finite chains cannot be strictly increasing, and only counterparts of the linear t-norm and the minimum remain. Other well-behaved symmetric aggregation operations on finite ordinal scales seem to be constant on significant subsets of their domains (e.g. [36]) which make these aggregations not so attractive in practice. For instance the associative operation median(a,b,α) for $\alpha\in L\backslash\{0,1\}$, is constant whenever $\min(a,b)\leqslant\alpha\leqslant\max(a,b)$. Qualitative aggregation operations suffer from a lack of discrimination power expressed by the fact that they cannot be strictly increasing, but only globally strictly increasing (like the minimum and the maximum operations) when a unique finite scale is used for all arguments and the range of the aggregation operator (see [30]). Namely n-tuples $(\alpha_1,\ldots,\alpha_n)$ of values from L are partitioned into no more equivalence classes than the number of elements in the scale L. Discrimination power can be improved by pairwise comparisons of tuples $(\alpha_1,\ldots,\alpha_n)$ and $(\beta_1,\ldots,\beta_n)$, neglecting arguments i with equal values $\alpha_i=\beta_i$ (for instance the discrimin partial ordering which refines the minimum operation, see [18]); or using lexicographic techniques (the leximin ordering, for instance). Similar critiques can be addressed to the pessimistic and optimistic criteria of Dubois et al. [32], which use weighted versions of minimum and maximum operations. The problem of their refinement by lexicographic schemes is has recently been considered by Fargier and Sabbadin [35].

If the finite scale is bipolar, it is even more difficult to define aggregation operations having good mathematical properties (see for instance [43]). When they have good properties, aggregation operations have little discrimination power. For instance, the qualitative preference functional considered by Giang and Shenoy [40] for decision under uncertainty takes values on the bipolar scale $L\Pi$ introduced at the end of Section 3, considered as a utility scale. This preference functional maps acts, viewed as n-tuples $f=((\alpha_1,\beta_1),\dots,(\alpha_n,\beta_n))$ of values of $L\Pi$, to $L\Pi$ itself using possibility weights $(\pi_1,\dots,\pi_n)$ such that $\max_{i=1,n}\pi_i=1$. The utility of an act f, is computed as the pair

$$u(f)=\left(\max_{i=1,\dots,n}\min(\pi_i,\alpha_i),\ \max_{i=1,\dots,n}\min(\pi_i,\beta_i)\right)\in L\Pi.$$

This form results from simple very natural axioms on possibilistic lotteries, which are counterparts to the Von Neumann and Morgenstern axioms [64]: complete preorder of acts, increasingness in the wide sense according to the ordering in $L\Pi$, substitutability of indifferent lotteries, and the assumption that any consequence of an act is valued on $L\Pi$. Yet, this criterion has similar drawbacks as median (a,b,α): Whenever there two states i and j are such that $\alpha_i=1$ and $\beta_j=1$ (respectively, a bad or neutral, and a good or neutral state) and these states have maximal possibility $\pi_i=\pi_j=1$, then $u(f)=(1,1)$ results, expressing indifference.

In the purely ordinal case, it is generally very difficult to perform the aggregation of ordering relations, a problem extensively studied in social choice [57]. Starting from several complete pre-ordering relations $R_1,\dots,R_n$, it is impossible, under mild assumptions, to perform an aggregation yielding another complete preordering R that does not coincide with any of the original relations, so long as the global ordering between two objects (xRy) only depends on the local ordering relations between these two objects (xR_iy). Counterparts to conjunctive and disjunctive aggregations can be defined as follows:

conjunctive: $xR_\wedge y$ iff $\forall i,\ xR_iy$,

disjunctive: $xR_\vee y$ iff $\exists i,\ xR_iy$.

However, it is clear that $R_\wedge$ is certainly not complete and will be hardly discriminant. On the other hand, $R_\vee$ may fail to be transitive. Numerical data are hard to get from human observers, who rather deliver qualitative reports. So, the more qualitative we go, the less arbitrary representations we use, but the more difficult it is to perform useful and expressive aggregations.

5. Various information fusion problems

The various information items surveyed in Section 1 can be, at least approximately, modelled by profiles on a suitable set S of alternatives. However, not all aggregation operators make sense in any context. Here, we briefly discuss three kinds of information fusion problems: the merging of uncertain observations, the merging of preference and the merging of logical databases. For the merging of regulations, see [14].

5.1. Merging uncertain observations

The problem of merging uncertain observations can be summarized as follows: given a set of sources of information, generally heterogeneous ones, supplying data about the value of some parameter or the description of a situation, find the most plausible values of the observed parameter or the most credible description of the situation, trying to discard erroneous pieces of information, while remaining faithful to the available information. By a source, we mean a lot of different things:

it can be a human (an expert), a sensor, or a database, hence a potential heterogeneity of the pieces of information.

This is a problem met in various areas, for instance:

(i) in robotics, data coming from various sensors must be merged in order to establish the positioning of autonomous robots [1];
(ii) in image processing, several remotely-sensed images of some area must be merged with maps of the same area; the same problem occurs with medical images that must be exploited conjointly with anatomic pictures (e.g. [8]);
(iii) in risk analysis, information items pertaining to the reliability of physical components are supplied by several experts and must be fused [13];
(iv) the fusion of databases, or of replies to queries addressed to several databases is also of current interest [2].

The uncertain observation merging problem is a matter of finding the truth in a reliable way: what is required is to make the best of the available information by discarding the wrong data when possible, and keeping the right information. Logical combinations are then natural candidates as fusion operators. Conjunctive combinations apply when all the sources are reliable, while disjunctive combinations deal with the case of unreliable sources hidden in a group of other reliable ones. Obviously, prioritized logical combinations may be considered, in particular when the sources are not equally reliable.

Averaging operations in uncertain observation aggregation are justified when the set of sources can be viewed as a single random source producing different inputs. In that case, indeed, the set of data to be fused can be interpreted as standard statistics. For instance, several successive independent measurements from a single sensor can be viewed as the result of a random experiment. Then the discrepancies between the sources can be explained in terms of random variability and fusion comes close to a filtering problem. However, in the case of unique measurements issued from heterogeneous sensors, or in the case of expert knowledge or databases, it is not clear that averaging combination modes make sense.

Uncertain observations are modelled in various formal settings, especially probability theory, possibility theory, random sets [31]. For instance the opinion of an expert on the value of a parameter, or the result of a measurement is sometimes modelled by means of a probability distribution which then stands as a profile [13]. Alternatively it can be modelled by an interval or a fuzzy interval whose membership function is viewed as a possibility distribution, especially when the information is linguistic. An alternative probabilistic approach consists in considering possibilistic profiles as likelihood functions $P(u_0|u)$ representing the probability of precisely observing $X_{\text{obs}} = u_0$ when actually $X = u$. The likelihood function $P(u_0|\cdot)$ is actually very similar to a possibility distribution. In the case of a random set model, the uncertain observation takes the form of a set of pairs $\{(A_i, m_i),\ i = 1,\ldots,k\}$. The profile of the information is defined as $Pl(u) = \Sigma_{i:\ u\in A_i} m_i$ only summarizes the original data, and subsumes both probability and possibility distributions [49].

The choice an aggregation operator is sometimes constrained by the mathematical framework. If the merging of probability degrees of events must result in a probability measure, then only a convex mixture is generally possible [13]. No conjunctive, nor disjunctive modes are allowed. Enlarging the mathematical framework may allow for such combination modes. For instance, the product of two probability measures on a finite set yields a belief function (a two-additive measure) and corresponds to a disjunction operation.

The choice of an aggregation operator also needs assumptions on the sources (Dubois and Prade [24,28,31,34]).

(i) When all sources are assumed to be reliable it is natural to perform a conjunctive fusion of the profiles. Generally the presence of partial conflicts between the profiles yields a subnormal resulting profile (such that $\pi(s) < 1$, $\forall s$). Applying the principle that whatever remains as possible in the end, however implausible it may look, is the truth, a normalization step takes place. There remains some degree of freedom for the choice of the conjunction operation to apply. When sources are assumed to be independent, a reinforcement effect between the profiles is justified. Typically, the pointwise product of profiles is performed (followed by a renormalization). The archetype of such a methodology is Dempster rule of combination in evidence theory [59]. Interestingly, it is seldom used when the profile is a probability distribution (because the conflict between several probability distributions is always rather high, viewed from a conjunctive combination point of view). When sources cannot be considered as independent, it is more cautious to use idempotent operations. The archetype of such a methodology is the minimum rule in possibility theory. Interestingly, no idempotent conjunctive fusion rule is known in the theory of evidence, except in particular cases, nor in probability theory. Note that renormalized versions of a conjunctive aggregation may fail to preserve associativity (except for the product, see [24]).

(ii) In the case of strongly conflicting heterogeneous sources, normalized conjunctive operations on profiles become discontinuous, hence numerically unstable [24]. Besides, performing an averaging operation between profiles is not always recommended, because it may suggest results that sources individually reject. For instance, the mean value of the arithmetic mean of two very distinct probability distributions may have zero probability. It seems more reasonable to use disjunctive operations on conflicting observations. It presupposes that at least one of these observations is correct, while other ones may be wrong. The result preserves the normalization of the operands but may become very imprecise. Avoiding such imprecision presupposes not to combine all observations. For instance, the conflict may be explained by the fact that observations pertain to different objects. One may also consider prioritizing the sources in terms of reliability, discarding the less reliable sources in conflict with more reliable ones [28]. Another way of reducing the resulting imprecision due to the use of a disjunctive operation is to use a quantified merging operation, like an ordered weighted minimum [34]. It is based on an assumption regarding how many sources are reliable among the ones providing observations. A more sophisticated approach is to exploit adaptive combination rules where the assumed number of reliable sources results from a conflict analysis between observations to be merged [28]. Techniques borrowed from logical approaches to inconsistency handling can be adapted, such as looking for maximal consistent subsets of uncertain observations on which conjunctive operations apply, and performing a disjunctive aggregation of the partial results [31].

(iii) Averaging operations are justified in the case of numerous identical independent sources which can be viewed as a single one governed by some unique stochastic process. This is typically what is done when synthesizing the (precise but scattered) results of random experiments. Statistics are based on averaging operations on data. It can be viewed as a very particular case of uncertain observation merging.

(iv) The so-called "naive" Bayesian approach to information fusion consists in computing the probability $P(u|u_1,\ldots,u_k)$ that $X = u$ when $X_{\text{obs}} = (u_1,\ldots,u_k)$ from k sources, from the knowledge of the likelihoods $P(u_i|u)$, and an independence assumption, namely

$$P(u_1,\ldots,u_k|u)=\prod_{i=1,\ldots,n}P(u_i|u),$$

using Bayes theorem. This computation presupposes some a priori knowledge on the value of X under the form of a probability measure P which is conjunctively combined with the likelihood functions. When this a priori knowledge is lacking, only $P(u_1,\ldots,u_k|u)$ is available and can be viewed as an estimation of the possibility that $X=u$ computed in accordance with the possibility theory setting for independent reliable sources. Hence, up to the existence of probabilistic a priori knowledge, the Bayesian technique for fusing uncertain observations is consistent with our setting of profile merging. It could be extended using copulas. Nevertheless the profile merging setting does not exclude a priori knowledge. For instance, counterparts of Bayes theorem exist in possibility theory and evidence theory [29].

5.2. Aggregation of preference profiles

The problem of preference aggregation consists in deriving a global preference profile achieving a consensus between the preference profiles supplied by the various sources. This new preference profile may be different from all the input ones provided that it remains close to all of them. Note that contrary to the fusion of uncertain observations there is no notion of "true state" involved. Generally, averaging operators are natural candidates for preference merging except if preference profiles are valued on a negative scale and express flexible constraints. More usually, they are valued on numerical positive or bipolar scales. Several types of preference merging problems exist:

(1) Multiagent fusion: It consists in finding a consensus among individuals expressing their preferences on a set of candidate choices. The aim is often to determine the average opinion likely to avoid extreme positions. This problem has been widely studied in social choice theory [57].

(2) Multicriteria decision making: It consists in rating and ranking individual decisions from several points of view. Multiple goals are involved and the problem is to find trade-offs between them [46].

(3) Decision under uncertainty: It consists in rating and ranking individual decisions whose consequences depends on the current state of the world which is partially unknown [55,64].

These problems are mathematically very similar (see, e.g. [16,20,23]): agents, criteria, and states play the same role of sources supplying preference profiles across potential decisions (candidates, objects, uncertain acts). Importance coefficients for agents or criteria play the same role in aggregation schemes as plausibility weights for states. While in uncertain observation fusion the normal situation is that all sources faithfully report correct information, and the abnormal situation is the case of conflicting reports, conflicts are the normal situation in preference merging. It leads to searching for consensus preference profiles. When this is impossible, incomparability phenomena results [63].

In the setting of preference merging, it is very natural to require idempotent aggregation operations. If all individuals have the same preference profiles, this preference profile should be the global one. If an act has the same consequence regardless of the state of the world, its rating should reflect the utility of this consequence only. Given the natural assumption of increasingness in the wide sense for aggregation operations, it is clear that they can only be averaging ones.

Despite their mathematical analogy, the methodology adopted for expressing and solving preference merging problems depends on the context. In multiagent fusion the tradition of social choice is to adopt purely ordinal approaches, leading to impossibility theorems (like Arrow's), that scholars try to bypass by finding suitably relaxed settings [57]. The reason for this approach is that the key-issue seems to be the difficulty of comparing preference scales of different individuals. They are generally supposed to be non-commensurate. Voting schemes considered to be rational in such a setting are often very drastic, and one is often led to resort to lexicographic choice rules, or to leave many incomparabilities unsolved.

In decision under uncertainty, the traditional approach is numerical and relies on expected utility (after Von-Neumann and Morgenstern [64] and Savage [55]), that is basically a weighted arithmetic mean. More recent approaches advocate the use of Choquet integrals [56]. Numerical approaches are natural because this field comes from economics where consequences of acts are expressed in terms of money, and statistics are sometimes available for the quantification of uncertainty. The commensurability problems that are patent in the multiagent problems are far less acute in decision under uncertainty due to the assumption of a unique set of consequences independently of the considered state; moreover decision under uncertainty often relies on the notion of certainty equivalent to an uncertain lottery. It enables to assume that utilities and probabilities can be compared on the same scale.

More recently, the emergence of qualitative utility theory has been observed in artificial intelligence [15]. This trend has tried to give up the numerical settings, because in robotics, or in recommender systems for instance, it is not so easy to formulate goals, or uncertainty in purely numerical terms. When preserving commensurability between uncertainty and utility on finite chains, Sugeno integral has been proposed as a counterpart to expected utility [32]. Besides, giving up all commensurability assumptions and adopting a purely relational setting leads to the same difficulties as ordinal social choice theories [16].

The case of multicriteria aggregation has led to many approaches (see [10]), according to the assumptions made about the information that can be supplied by a decision maker regarding the preference scales for criteria and the weighting scale for criteria importance. The most conservative approach consists in assuming that preference scales for each criteria are not commensurate and that no weighting pattern for criteria importance is available. Then the aggregation of profiles consists only in ranking vectors or ratings for decisions. It results in a partial ordering with very little discrimination power. The most data demanding traditional approach is multiattribute utility theory [46] where every preference scale can be made commensurate with every other one, and preference can be quantified. This approach essentially results in the extensive use of the weighted arithmetic mean. It fully fits the profile merging setting of this paper. More recently, the necessity of taking into account redundancies or synergies between criteria has led to the proposal of Choquet integral as a versatile tool for preference aggregation [42]. A balanced attitude between a full-fledged numerical modelling and the vector-maximization school is to adapt the multiagent preference merging paradigm to multicriteria decision making. However this school of thought (see [63]) admits of the existence of weights attached to criteria and adopts decision rules in the spirit of the Condorcet pairwise majority scheme: a decision d_1 is globally better than another d_2 if the sum of the weights of criteria that prefer d_1 to d_2 is greater than the sum of the weights of criteria that prefer d_2 to d_1. Generally, this kind of procedure generates intransitive preference relations. This decision analysis school also leaves room to incomparability when the preference profiles of two decisions are severely conflicting, and proposes methods to exploit the resulting ill-behaved preference relation.

Bellman and Zadeh's approach to decision-making in a fuzzy environment [3] strongly departs from the tradition of compensatory operations in preference merging, by proposing a fusion of preference profiles using a minimum operation. A counterpart to this proposal was made quite early in decision under uncertainty by Wald [65] and is known as the maximin criterion. This technique for combining preference profiles is severely non compensatory. It comes down to requiring that all criteria should be satisfied in order to select a decision. This request may sound too demanding. However, we have advocated the idea that this approach makes sense when combining flexible constraints, none of which should be violated completely [17]. The preference scale is then negative since the top of L corresponds to a feasible decision, and the bottom to an unfeasible one. In some practical problems (like avoiding delayed jobs in scheduling, for instance) a balanced decision achieving all goals to some extent may look more satisfactory than a decision that fully meets almost all goals and totally violates some of them. It is clear that many averaging schemes (like the arithmetic mean) may yield optimal solutions of the latter kind, while optimization under the minimum rule tends to balance the local satisfaction levels. So, the minimum rule and, to a large extent, other conjunctive aggregation operations are adapted to reasoning with flexible or valued constraints in artificial intelligence [7]. It is well known that negative utilities and positive ones are not combined likewise by decision makers. The latter are more careful with losses than with gains. A maxmin approach to combining negative utilities (which membership functions of fuzzy constraints are) seems to be plausible, while averaging operations look natural for combining positive preferences.

5.3. Fusion of logical databases

This section is concerned with the combination of n sets of propositional formulas representing deductive databases. Generally if several databases supply information on a given topic, one might expect that putting them together in order to improve the overall knowledge leads to an inconsistent database. The question is then how to make the best of the available knowledge by coping with such inconsistency. This problem has been considered in Artificial Intelligence for some years [2]. To cope with inconsistency there are basically three attitudes:

(1) Extract a suitable consistent subset of formulas, and do away with the other pieces of information.
(2) Keep the inconsistent database, but use some non-trivial inconsistency-tolerant approach to inference for question–answering.
(3) Merge the databases while preserving consistency.

At first glance, database fusion looks very different from the profile merging approach. Yet some existing approaches come down to profile merging [51,54]. One of the most common methodology is "distance-based merging" (see [48] for a general framework).

The first step is to derive a distance from a database. Let K be a logical database, containing propositional formulas described by means of propositional variables in a set V. Let S be the set of interpretations of the language (maximal conjunctions of literals). Let $[K]$ be the set of models of K (a subset of interpretations where all formulas of K are true). A natural distance between interpretations is the Hamming distance $H(w,w^*)$ evaluating the number of literals with different signs in w and w^*. The distance between w and the database K is $d(w,K)=\min_{w^*\in[K]} H(w,w^*)$. If w is a model of K, $d(w,K)=0$. The greater $d(w,K)$, the stronger the violation of K by w.

On such a basis, the process of merging two databases K_1 and K_2 which use the same language proceeds as follows [48]:

(1) Compute $d(w,K_1)$ and $d(w,K_2)$.
(2) Compute $\delta(w)=d(w,K_1)\oplus d(w,K_2)$ using an aggregation operation $\oplus$.
(3) Retrieve a merged belief base K_{12} as one having as models:

$$[K_{12}] = \{w \text{ minimizing } \delta(w)\}.$$

Hence the merging process is clearly based on the aggregation of profiles. Clearly, if K_1 and K_2 are not expressed in the same language, there is a preliminary problem of translation between sets of propositional variables. This issue is always present in practical fusion problems, be they numerical or logical as it is not obvious whether sources refer to the same frame of discernment or not.

The range of distances is the set of integers instead of the unit interval. Typical choices for operation $\oplus$ are minimum, maximum, the sum, the weighted sum, the leximax ranking. A drawback of the method is that, for n sources, the merging has to be done in one step, because the scheme is not associative even if the operation $\oplus$ is. Indeed, if K_{12} is merged with K_3, $d(w,K_{12})$ has to be combined with $d(w,K_3)$, but $d(w,K_{12})$ generally differs from $\delta(w)$.

The framework of possibilistic logic enables this difficulty to be solved and the merging operation to be performed at the syntactic level [6]. Given a belief base K, and $a\in(0,1)$, define $\pi(w)=a^{d(w,K)}$. Then the fusion of the possibility distributions by operation $\otimes=$ max, min, product, weighted product, leximin corresponds to the fusion of the distances by operation $\oplus=$ min, max, sum, weighted sum, leximax, respectively.

The syntactic counterpart of the possibility distribution π is a possibilistic knowledge base, that is a set of pairs $\Sigma=\{(\phi_j,\alpha_j),\ j=1,\dots,n\}$ where α_j is a priority level understood as a lower bound of the degree of necessity of $[\phi_j]$. The possibility distribution on interpretations induced by Σ is

$$\pi(w)=\min\{1-\alpha_j,\ w\notin[\phi_j]\},$$

that is, for each interpretation we check for the violated formula with a maximal priority.

Consider possibilistic databases Σ_i provided by k sources. Each possibilistic database Σ_i is associated with a possibility distribution π_i which is its semantic counterpart. The syntactic encoding of the above-mentioned fusion modes proceeds as follows. Given a semantic combination rule f, we look for a syntactic combination C such that:

$$f(\pi_1,\dots,\pi_k)=\pi_{C(\Sigma_1,\dots,\Sigma_k)},$$

where $\Sigma_C=C(\Sigma_1,\dots,\Sigma_k)$ is the result of syntactically merging $\Sigma_1,\dots,\Sigma_k$.

For instance the result $\Sigma_{\min}$ obtained with $f=$ minimum is simply the union of the Σ_is. The result $\Sigma_{\max}$ obtained with $f=$ maximum is, for $k=2$:

$$C(\Sigma_1,\Sigma_2)=\Sigma_{\max}=\{(\phi_i\vee\psi_j,\min(\alpha_i,\beta_j))\,|\,(\phi_i,\alpha_i)\in\Sigma_1 \text{ and } (\psi_j,\beta_j)\in\Sigma_2\}.$$

It is the intersection of the fuzzy set of possibilistic consequences of Σ_1 and the fuzzy set of those of Σ_2. More generally, the aggregation of π_1 and π_2 using a monotonic operation $\otimes$ such that $1 \otimes 1 = 1$ corresponds to the following syntactic fusion of Σ_1 and Σ_2:

$$\{(\phi_i \vee \psi_j, 1-(1-\alpha_i)\otimes(1-\beta_j)) \mid (\phi_i,\alpha_i)\in\Sigma_1,\ (\psi_j,\beta_j)\in\Sigma_2\}$$
$$\cup\{(\phi_i, 1-(1-\alpha_i)\otimes 1)) \mid (\phi_i,\alpha_i)\in\Sigma_1\}\cup\{(\psi_i, 1-1\otimes(1-\beta_i))|(\psi_i,\alpha_i)\in\Sigma_2\}.$$

For instance, if $\otimes =$ arithmetic mean, then $1-(1-\alpha_i)\otimes(1-\beta_j)=(\alpha_i+\beta_j)/2$ and $1-(1-\alpha_i)\otimes 1=\alpha_i/2$.

The distance-based fusion can be embedded in this setting which preserves the associativity of semantic merging operations. To this end it is necessary to encode the distance $d(w,K)$ into a possibilistic knowledge base (see [4]). The result of the fusion is a possibilistic knowledge base whose most prioritary layer is the classical database searched for. This kind of fusion tools can also be useful for merging regulations or logically represented preferences [5].

6. Conclusion

This discussion paper has tried to point out major potential application fields for aggregation operations in the scope of information sciences and engineering. The main messages of this paper are as follows. Aggregation operators can be useful in a variety of information fusion problems. Many information fusion tasks, even in the case of syntactic fusion of logical databases, come down to the aggregation of profiles. The choice of an aggregation operation crucially depends on the type of value scale to be used: bipolar or not, quantitative or qualitative. Qualitative value scales do not allow for as many aggregation modes as numerical value scales. Qualitative bipolar value scales are especially constrained. There is a trade-off to be studied between expressiveness of the aggregation process (which is maximal for numerical frameworks) and its meaningfulness (in some problems the information is only available in ordinal form). The fuzzy set community has come up with an impressive body of formal results on aggregation operations mainly using the unit interval as a reference value scale. The ambition of this paper is merely to suggest some hints towards their practical use, and to point out the need for more research on aggregation modes under qualitative settings.

References

[1] M.A. Abidi, R.C. Gonzalez (Eds.), Data Fusion in Robotics and Machine Intelligence, Academic Press, New York, 1992.
[2] C. Baral, S. Kraus, J. Minker, V.S. Subrahmanian, Combining knowledge bases consisting in first order theories, Comput. Intell. 8 (1) (1992) 45–71.
[3] R. Bellman, L.A. Zadeh, Decision making in a fuzzy environment, Management Sci. 17 (1970) B141–B164.
[4] S. Benferhat, D. Dubois, S. Kaci, H. Prade, Possibilistic merging and distance-based fusion of propositional information, Ann. Math. Artif. Intell. 34 (2002) 217–252.
[5] S. Benferhat, D. Dubois, S. Kaci, H. Prade, Bipolar representation and fusion of preferences in the possibilistic logic framework, Proc. Eighth Internat. Conf., Principles of Knowledge Representation and Reasoning (KR2002), Toulouse, France, Morgan Kaufmann Publishers, San Francisco, CA, 2002, pp. 421–432.

[6] 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, Germany, 1997, pp. 141–161.
[7] S. Bistarelli, H. Fargier, U. Montanari, F. Rossi, T. Schiex, G. Verfaillie, Semiring-based CSPs and valued CSPs: frameworks, properties, and comparison, Constraints 4 (1999) 199–240.
[8] I. Bloch, Fusion of image information under imprecision and uncertainty: numerical methods, in: G. Della Riccia, H. Lenz, R. Kruse (Eds.), Data Fusion and Perception, CISM Courses and Lectures, Vol. 431, Springer, Berlin, 2001, pp. 135–168.
[9] I. Bloch, A. Hunter (Eds.), Special Issue on Data and Knowledge Fusion, Int. J. Intell. Systems 16 (10 & 11) (2001) 1107–1320.
[10] D. Bouyssou, T. Marchant, M. Pirlot, P. Perny, A. Tsoukias, P. Vincke P, Evaluation Models: A Critical Perspective, Kluwer Academic Publisher, Boston, 2000.
[11] T. Calvo, B. De Baets, J. Fodor, The functional equations of Alsina and Frank for uninorms and null-norms, Fuzzy Sets and Systems 120 (2001) 15–24.
[12] T. Calvo, G. Mayor, R. Mesiar (Eds.), Aggregation operators: new trends and applications, Studies in Fuzziness and Soft Computing, Vol. 97, Physica-Verlag, Heidelberg, 2002.
[13] R.M. Cooke, Experts in Uncertainty, Oxford University Press, Oxford, UK, 1991.
[14] F. Cuppens, L. Cholvy, C. Saurel, J. Carrère, Merging regulations: analysis of a practical example, Internat. J. Intell. Systems 16 (11) (2001) 1223–1244.
[15] J. Doyle, R. Thomason, Background to qualitative decision theory, The AI Magazine 20 (2, Summer 1999) (1999) 55–68.
[16] D. Dubois, H. Fargier, P. Perny, On the limitations of ordinal approaches to decision-making, Proc. Eighth Internat. Conf., Principles of Knowledge Representation and Reasoning (KR2002), Toulouse, France, Morgan Kaufmann Publishers, San Francisco, CA, 2002, pp. 133–144.
[17] D. Dubois, H. Fargier, H. Prade, Possibility theory in constraint satisfaction problems: handling priority, preference and uncertainty, Appl. Intell. 6 (1996) 287–309.
[18] D. Dubois, H. Fargier, H. Prade, Refinements of the maximin approach to decision-making in fuzzy environment, Fuzzy Sets and Systems 81 (1996) 103–122.
[19] D. Dubois, J.C. Fodor, H. Prade, M. Roubens, Aggregation of decomposable measures with application to utility theory, Theory and Decision 41 (1996) 59–95.
[20] D. Dubois, M. Grabisch, F. Modave, H. Prade, Relating decision under uncertainty and multicriteria decision making models, Internat. J. Intell. Systems 15 (10) (2000) 967–979.
[21] D. Dubois, P. Hajek, H. Prade, Knowledge-driven versus data-driven logics, J. Logic Language Inform. 9 (2000) 65–89.
[22] D. Dubois, J. Lang, H. Prade, Possibilistic logic, in: D.M. Gabbay, C.J. Hogger, J.A. Robinson, D. Nute (Eds.), Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 3, Oxford University Press, Oxford, 1994, pp. 439–513.
[23] D. Dubois, J.-L. Marichal, H. Prade, M. Roubens, R. Sabbadin, The use of the discrete Sugeno integral in decision-making: a survey, Internat. J. Uncertainty, Fuzziness Knowledge-based Systems 9 (2001) 539–561.
[24] D. Dubois, H. Prade, Representation and combination of uncertainty with belief functions and possibility measures, Comput. Intell. 4 (1988) 244–264.
[25] D. Dubois, H. Prade, Possibility Theory: An Approach to the Computerized Processing of Uncertainty, Plenum Press, New York, 1988.
[26] D. Dubois, H. Prade, Consonant approximations of belief functions, Internat. J. Approx. Reason. 4 (5/6) (1990) 419–449.
[27] D. Dubois, H. Prade, When upper probabilities are possibility measures, Fuzzy Sets and Systems 49 (1992) 65–74.
[28] D. Dubois, H. Prade, Possibility theory and data fusion in poorly informed environments, Control Eng. Practice 2 (5) (1994) 811–823.
[29] D. Dubois, H. Prade, Possibility theory: qualitative and quantitative aspects, in: P. Smets (Ed.), Handbook on Defeasible Reasoning and Uncertainty Management Systems—Vol. 1: Quantified Representation of Uncertainty and Imprecision, Kluwer, Dordrecht, The Netherlands, 1998, pp. 169–226.

[30] D. Dubois, H. Prade, Refining aggregation operations in finite ordinal scales, Proc. Second Internat. Conf. in Fuzzy Logic and Technology (EUSFLAT 01), Leicester, UK, 2001, pp. 175–178.
[31] D. Dubois, H. Prade, Possibility theory in information fusion, in: G. Della Riccia, H. Lenz, R. Kruse (Eds.), Data Fusion and Perception, CISM Courses and Lectures, Vol. 431, Springer, Berlin, 2001, pp. 53–76.
[32] D. Dubois, H. Prade, R. Sabbadin, Decision-theoretic foundations of possibility theory, European J. Oper. Res. 128 (2001) 459–478.
[33] D. Dubois, H. Prade, P. Smets, Not impossible vs guaranted possible in fusion and revision, Proc. Sixth European Conf. (ESCQARU 2001), Toulouse, France, Springer, Berlin, 2001, pp. 522–531.
[34] D. Dubois, H. Prade, R.R. Yager, Merging fuzzy information, in: J.C. Bezdek, D. Dubois, H. Prade (Eds.), Fuzzy Sets in Approximate Reasoning and Information Systems, The Handbooks of Fuzzy Sets Series, Kluwer Academic Publishers, Dordrecht, 1999, pp. 335–401.
[35] H. Fargier, R. Sabbadin, Qualitative decision under uncertainity: back to expected utility, in: Proc. International Joint Conf. on Artificial Intelligence (IJCAII'03), Acapulco, Mexico, 2003, pp. 303–308.
[36] J. Fodor, Smooth associative operations on finite ordinal scales, IEEE Trans. Fuzzy Systems 8 (2000) 791–795.
[37] J. Fodor, M. Roubens, Fuzzy Preference Modelling and Multicriteria Decision Support, Kluwer Academic Publishers, Dordrecht, 1994.
[38] J. Fodor, R. Yager, A. Rybalov, Structure of uninorms, Internat. J. Uncertainty, Fuzziness and Knowledge-based Systems 5 (1997) 411–427.
[39] L.W. Fung, K.S. Fu, An axiomatic approach to rational decision-making in a fuzzy environment, in: L.A. Zadeh, K.S. Fu, K. Tanaka, M. Shimura (Eds.), Fuzzy Sets and Their Application to Cognitive and Decision Processes, Academic Press, New York, 1975, pp. 227–256.
[40] P.H. Giang, P.P. Shenoy, A comparison of axiomatic approaches to qualitative decision-making using possibility theory, Proc. 17th Internat. Conf. Uncertainty in Artificial Intelligence, Morgan Kaufmann, San Francisco, CA, 2001, pp. 162–170.
[41] J.A. Goguen, L-fuzzy sets, J. Math. Anal. Appl. 8 (1967) 145–174.
[42] M. Grabisch, The application of fuzzy integrals in multicriteria decision-making, European J. Oper. Res. 89 (1996) 445–456.
[43] M. Grabisch, B. De Baets, J. Fodor, On symmetric pseudo-additions and pseudo-multiplications: is it possible to build rings on $[-1,+1]$? Proc. Nineth Internat. Conf. Information Processing and Management of Uncertainty in Knowledge based Systems (IPMU2002), Annecy, France, 2002, pp. 1349–1355.
[44] M. Grabisch, T. Murofushi, M. Sugeno (Eds.), Fuzzy Measures and Integrals, Physica-Verlag, Heidelberg, Germany, 1999.
[45] M. Grabish, The symmetric Sugeno integral, Fuzzy Sets and Systems, 139 (3) (2003) 473–490.
[46] R. Keeney, H. Raiffa, Decision With Multiple Objectives: Preferences and Value Trade-offs, Wiley, New York, 1976.
[47] E.P. Klement, R. Mesiar, E. Pap, Triangular Norms, Kluwer Academic Publishers, Boston, 2000.
[48] S. Konieczny, J. Lang, P. Marquis, Distance-based merging: a general framework and some complexity results, in: Proc. Eighth Internat. Conf. on Principles of Knowledge Representation and Reasoning (KR2002), Toulouse-France, Morgan Kaufmann Publishers, San Francisco, CA, 2002, pp. 97–108.
[49] J. Gebhardt, R. Kruse, Parallel combination of information sources, in: D. Dubois, H. Prade (Eds.), Belief Change, Handbook of Defeasible Reasoning and Uncertainty Management Systems, Vol. 3, Kluwer Academic Publishers, Dordrecht, 1998, pp. 393–440.
[50] D. Lewis, Counterfactuals, Basil Blackwell, London, 1973.
[51] P. Liberatore, M. Schaerf, Arbitration (or how to merge knowledge bases), IEEE Trans Data and Knowledge Eng. 10 (1998) 76–90.
[52] M. Mukaidono, On some properties of fuzzy logic, Systems-Computers-Control 3 (3) (1975) 36–43.
[53] N. Pal, J.C. Bezdek, Quantifying different facets of fuzzy uncertainty, in: D. Dubois, H. Prade (Eds.), Fundamentals of Fuzzy Sets, Kluwer Academic Publishers, Dordrecht, 2000.
[54] P.Z. Revesz, On the semantics of arbitration, Internat. J. Algebra Comput. 7 (1997) 133–160.
[55] L.J. Savage, The Foundations of Statistics, Dover, New York, 1972.
[56] D. Schmeidler, Subjective probability and expected utility without additivity, Econometrica 57 (1989) 571–587.

[57] A.K. Sen, Social choice theory, in: K. Arrow, M.D. Intrilligator (Eds.), Handbook of Mathematical Economics, Elsevier, Amsterdam, 1986, pp. 1173–1181 (Chapter 22).
[58] G.L.S. Shackle, Decision Order & Time in Human Affairs, 2nd Edition, 1969, Cambridge University Press, Cambridge, UK, 1961.
[59] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, NJ, 1976.
[60] W. Silvert, Symmetric summation: A class of operations on fuzzy sets, IEEE Trans. Systems, Man Cybernet. 9 (10) (1979) 657–659.
[61] B. Schweizer, A. Sklar, Probabilistic Metric Spaces, North-Holland, Amsterdam, 1983.
[62] L. Ughetto, D. Dubois, H. Prade, Implicative and conjunctive fuzzy rules—A tool for reasoning from knowledge and examples, Proc. 15th US Nat. Conf. in Artificial Intelligence(AAAI-99), Orlando, 1999, pp. 214–219.
[63] P. Vincke, Multicriteria Decision-Aid, Wiley, New York, 1992.
[64] J. Von Neumann, O. Morgenstern, Theory of Games and Economic Behaviour, Princeton University Press, Princeton, NJ, 1944.
[65] A. Wald, Statistical Decision Functions, Wiley, New York, 1950.
[66] J. Weisbrod, A new approach to fuzzy reasoning, Soft Comput. 2 (1998) 89–99.
[67] L.A. Zadeh, Fuzzy sets, Inform. Control 8 (1965) 338–353.
[68] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1 (1978) 3–28.
[69] L.A. Zadeh, A theory of approximate reasoning, in: J.E. Hayes, D. Michie, L.I. Mikulich (Eds.), Machine Intelligence, Vol. 9, Elsevier, New York, 1979, pp. 149–194.