



**HAL**  
open science

## Multiple-sources informations fusion - a practical inconsistency-tolerant approach

Didier Dubois, H el ene Fargier, Henri Prade

► **To cite this version:**

Didier Dubois, H el ene Fargier, Henri Prade. Multiple-sources informations fusion - a practical inconsistency-tolerant approach. 8th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU 2000), Jul 2000, Madrid, Spain. hal-03405306

**HAL Id: hal-03405306**

**<https://hal.science/hal-03405306>**

Submitted on 27 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destin ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.

# Multiple-sources information fusion - a practical inconsistency-tolerant approach

Didier Dubois - H el ene Fargier - Henri Prade

Institut de Recherche en Informatique de Toulouse  
Universit  Paul Sabatier – 118 route de Narbonne  
31062 Toulouse Cedex 4 – France  
Email: {dubois, fargier, prade}@irit.fr

## Abstract

A multiple-sources information system should be able to provide the user with the most possibly accurate answers about the values of the attributes describing the objects of interest. In this paper, the attribute values, which are supposed to range on linearly ordered domains, may be imprecisely given by the sources, which moreover may not be fully reliable. After a brief background on the general principles underlying the fusion of imprecise information in the framework of possibility theory, the case of single-valued attributes is first considered, before dealing with the multiple-valued case. The handling of known relations linking attributes, in a fusion process, is also briefly discussed.

**Keywords:** fusion; possibility theory; inconsistency, maximal consistent subsets.

## 1. Introduction

Information fusion plays an important role in advanced information systems. Indeed, the use of several sources of information often allows to deliver answers to the users, which are more precise than the ones that each source could provide individually. However, the different sources of information are not always equally reliable, and the pieces of information they give may be partially inconsistent.

There exists different theoretical frameworks for modelling the combination of uncertain information (e.g., Abidi and Gonzales, 1992; Luo and Kay, 1995; Gebhardt and Kruse, 1998). Beside probabilistic approaches which have been

extensively studied and used in multiple-sensors fusion, or in expert opinion aggregation (e.g., Cooke, 1991), logical settings have been considered more recently, in particular for databases fusion (e.g., Cholvy, 1998).

The possibility theory framework is used in this paper since it is rather well-suited for representing "poor" information when uncertainty can only be expressed in a qualitative manner; see (Dubois and Prade, 1994, 1995) for a general introduction. This paper proposes a discussion of the practical use of a simplified version of a possibilistic approach, which however can handle inconsistencies and take into account the different levels of reliability of the sources. The problems raised by the exploitation in the combination process of expert knowledge, given under the form of relations linking the variables whose values are provided by the sources, is also discussed, as well as the case of multiple-valued attributes.

## 2. Possibilistic information fusion

Let us consider some single-valued attribute represented by a parameter  $x$  ranging on a referential  $U$ , for a given object.  $U$  is supposed to be linearly ordered; e.g.,  $U$  is a subset of the real line. Let  $A_i$  represent the piece of information issued from source  $i$  to be understood as  $x \in A_i$ . As it is often the case in practice,  $A_i$  is assumed to be an interval, or more generally a fuzzy interval, which restricts the possible values of  $x$ . In practice, precise

values provided by sources may be only indicative; if it is the case, it is better to replace them by some genuine interval acknowledging some possible error. In the approach developed in this paper, the fuzzy interval  $A_i$  will be also supposed to have a trapezoidal membership function with core  $N_i$  and support  $S_i$ , which will be written  $A_i = (N_i, S_i)$ .

When fusing information coming from two sources, two situations may be encountered:

- either the pieces of information are sufficiently consistent and it can be assumed that the two sources are reliable, which allows for the cross-checking and the refinement of information;
- or the pieces of information are not sufficiently consistent and we are forced to admit that at least one of the sources is wrong.

The hypothesis that the two sources are reliable leads to use a conjunctive combination. Let  $\mu_{A_i}(u)$

be the degree of possibility that  $x = u$  knowing that a source asserts that  $x \in A_i$ . There are two main types of conjunctive combination:

- *Idempotent combination*: the result is

$$\mu_{A_1 \cap A_2}(u) = \min(\mu_{A_1}(u), \mu_{A_2}(u)) \text{ for all } u \in U.$$

This combination can be straightforwardly extended to  $n$  sources and does not assume any dependency between the sources (information can be provided in a redundant and repeated way). However the number of sources which are agreeing on some possible values of  $x$  is not taken into account in this case.

- *Product-based combination*: the result is

$$\mu_{A_1 \cap A_2}(u) = \mu_{A_1}(u) \cdot \mu_{A_2}(u) \text{ for all } u \in U.$$

This combination, which, as the previous one, can be justified in the setting of a probabilistic approach, requires independent sources, which is not always the case in practice.

**N.B.** Nilpotent combination operators such that  $\mu_{A_1 \cap A_2}(u) = \max(0, \mu_{A_1}(u) + \mu_{A_2}(u) - 1)$  seem to be of a more restricted use since values which are

judged to have a low possibility (without being impossible) by both sources are finally considered as impossible after the combination step.

The height of  $A_1 \cap A_2$ ,  $ht(A_1 \cap A_2) = \sup_u \min(\mu_{A_1}(u), \mu_{A_2}(u))$  expresses the confidence we can have in the hypothesis that both sources are reliable. Particularly, if  $ht(A_1 \cap A_2) = 0$ , this hypothesis cannot hold. However if  $ht(A_1 \cap A_2) = 1$  this hypothesis is only plausible, and the conjunctive combination remains optimistic since it changes this plausibility into certainty.

A more cautious hypothesis is that at least one of the source is right. In this case the pieces of information are combined disjunctively:

$$\mu_{A_1 \cup A_2}(u) = \max(\mu_{A_1}(u), \mu_{A_2}(u)) \text{ for all } u \in U.$$

This assumes that the values rejected by both sources can be rejected and only these ones. For a general presentation of fusion operations, see (Dubois and Prade, 1994, 1995).

### 3. The proposed approach to fusion

#### 3.1 Principle

In a previous work (Dubois and Prade, 1994), adaptive combination operators have been proposed, which are intermediary between disjunctive and conjunctive modes, and whose expression depends on the degree of consistency  $ht(A_1 \cap A_2)$  between the two sources. These combination operators may be computationally heavy to handle and do not preserve the shape of the fuzzy intervals. It is why we propose here a simplified approach (for trapezoidal membership functions), which enables us to have a combination process which is less costly. Since we can never exclude that one of the sources is wrong, the combination will perform the union of the supports  $S_1$  and  $S_2$  of  $A_1$  and  $A_2$ . Besides, if their cores  $N_1$  and  $N_2$  are overlapping, their intersection will be performed, in order, as far as possible, to single out a small subset of highly

plausible values for the parameter under study. Otherwise, the pieces of information will be combined disjunctively. Moreover, since the referential is ordered, we can further assume that the trapezoidal shape of the membership functions should be preserved, which leads to replacing  $S_1 \cup S_2$  and  $N_1 \cup N_2$  by their convex hulls, denoted by  $[S_1 \cup S_2]$ ,  $[N_1 \cup N_2]$ . Let  $\oplus$  denote the combination of pieces of information  $A_1$  and  $A_2$ , then we have:

$$A_1 \oplus A_2 = ([N_1 \cap N_2], [S_1 \cup S_2]) \text{ if } N_1 \cap N_2 \neq \emptyset \\ = ([N_1 \cup N_2], [S_1 \cup S_2]) \text{ otherwise.}$$

If  $N_1 = [a_1, b_1]$ ,  $N_2 = [a_2, b_2]$  and  $a_1 \leq a_2$ , then:

- $N_1 \cap N_2 \neq \emptyset$  if and only if  $b_1 \geq a_2$ ;
- $[N_1 \cap N_2] = [a_2, \min(b_1, b_2)]$  otherwise.
- $[N_1 \cup N_2] = [a_1, \max(b_1, b_2)]$ .

Let us point out that if  $N_1 \cap N_2 \cap N_3 \neq \emptyset$ ,  $(A_1 \oplus A_2) \oplus A_3 = A_1 \oplus (A_2 \oplus A_3)$ , i.e., the combination is associative in this case and can be extended to  $n$  sources. Unfortunately, associativity no longer holds with inconsistent sources. However, the combination can be adapted to the case of  $n$  sources, as explained now.

Let  $A_1, \dots, A_n$ , be  $n$  pieces of fuzzy information. Let  $K$  be a maximal subset (in the sense of inclusion) made of consistent sources, then we can compute:

$$A_1 \oplus A_2 \oplus \dots \oplus A_n = (\bigwedge_{i \in K} N_i, [\bigwedge_{i=1, n} S_i])$$

If there are several such maximal subsets  $K^*_j$  of sources then we shall take

$$[\bigwedge_j \bigwedge_{i \in K^*_j} N_i]$$

in place of  $\bigwedge_{i \in K} N_i$ .

### 3.2 Justifying the fusion procedure

Viewing the core  $N_j$  of each piece of information as the set of models of a logical formula, the collection of these pieces of information can be regarded as a propositional knowledge base. If the intersection of these cores is empty, the corresponding knowledge base is

inconsistent. In presence of an inconsistent knowledge base, one may either look for a consistent subbase by eliminating formulas, or extends the notion of logical deduction in a way which is inconsistency-tolerant. The first approach supposes that we have reasons for choosing the formulas to be eliminated (it might be here sources for which we would be sure of their lack of reliability). We assume that this is not the case and we are led to adopt the second attitude. A consequence from an inconsistent knowledge base will be considered as sure if this is a consequence from all maximal consistent subbases (Rescher and Manor, 1970; Benferhat *et al.*, 1997). Building a maximal consistent subbase amounts to keep the greatest number of formulas, and thus to eliminate as little information as possible. If we do not know what formula (or source) to eliminate, we would have no reason of choosing between the maximal consistent subbases.

From a semantical point of view, this amounts to consider that the models of an inconsistent knowledge base are only made of the union of the models of all the maximal consistent subbases. This corresponds exactly, in terms of sources, to the combination of cores described above, namely  $\bigwedge_j \bigwedge_{i \in K^*_j} N_i$  where  $\bigwedge_{i \in K^*_j} N_i$  is the set of models of  $K^*_j$ . The idea is here of applying a logical approach to fusion problems. It should be pointed out that performing  $\bigwedge_{i \in K^*_j} N_i$  relies on an hypothesis of optimistic combination as already said. It is impossible that all the sources in a super-set of  $K^*_j$  are altogether reliable. The chosen combination avoids to have to decide about a sub-group of sources which are not wrong, and assumes the greatest possible number of reliable sources, for improving information.

### 3.3 Algorithm

The  $K^*_j$  's can be obtained by means of a sorting procedure on the cores taking advantage of the ordered structure of  $U$ . We have to look for the maximal consistent subsets of  $\{N_1, \dots, N_n\}$ . Let  $N_i = [a_i, b_i]$ . We have the following result:

The  $m$  maximal consistent subsets of  $\{N_1, \dots, N_n\}$  are  $K_1, \dots, K_m$ , such that  $K_j = A_j - B_j$  where  $A_j = \{N_k, a_k \leq c_{s(j)}\}$  and  $B_j = \{N_k, b_k < c_{s(j)}\}$ , where the  $c_{s(j)}$  are obtained using the following procedure:

i) Rank in increasing order the  $2n$  numbers  $\{a_i, i = 1, n\} \cup \{b_i, i = 1, n\}$ ; let  $c_1 \leq \dots \leq c_{2n}$  denote the result.

Moreover, let us denote

type(i) = a if  $c_i$  is of the form  $a_k$  and  
 = b if  $c_i$  is of the form  $b_k$ .

ii) and let  $s(1), \dots, s(m)$  be the  $m$  positions in the sequence of  $c_i$  such that  $\text{type}(s(j)) = b$  and  $\text{type}(s(j) - 1) = a$ .

Indeed any intersection of intervals is of the form  $[a_i, b_j]$  (with maybe  $i = j$ ). The intersections of maximal consistent subsets of  $\{N_1, \dots, N_n\}$  have a minimal size in the sense of inclusion. Thus, let  $[a_i, b_j]$  be such minimal interval. If we could find  $b_k$  such that  $a_i < b_k < b_j$  or  $a_k$  such that  $a_i < a_k < b_j$  this means that  $N_k$  is included in any maximal consistent subset including  $N_i$  and  $N_j$ , and that the interval  $[a_i, b_j]$  is not minimal, which leads to a contradiction. Then any position  $j$  in the sequence of  $c_i$  such that  $\text{type}(j) = b$  and  $\text{type}(j - 1) = a$ , determines a maximal consistent subset and any maximal consistet subset is characterized by such a position in the sequence. The maximal consistent subset of intersections  $[a_i, b_j]$  is defined by the subset of the  $N_k$  's which contain  $[a_i, b_j]$ , then  $a_k \leq a_i$  (which is equivalent to  $a_k \leq b_j$  since  $[a_i, b_j]$  is minimal) and  $b_k \geq b_j$ . Thus all the  $N_k$  such that  $a_k$  are ranked before  $b_j$  except those whose  $b_k$  are also before  $b_j$ , i.e.,  $K_j = A_j - B_j$ .

This leads to the following algorithm starting from the sequence  $c_1 \leq \dots \leq c_{2n}$ .

Algorithm:  
 List =  $\emptyset$

$K = \emptyset$   
 for  $i = 1, 2n - 1$   
 If  $\text{type}(i) = a$  add to  $K$   $N_k$  such that  $c_i = a_k$   
 if also  $\text{type}(i + 1) = b$  add  $K$  to List  
 If not delete from  $K$   $N_k$  such that  $c_i = b_k$   
 end for  
 List will contain all the maximal consistent subsets of  $\{N_1, \dots, N_n\}$ .

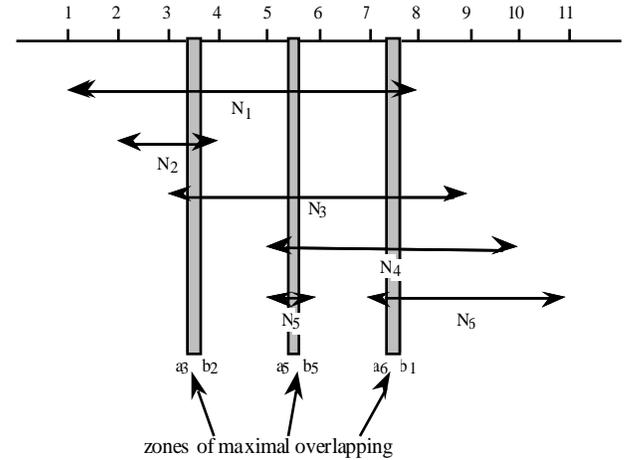
NB : if  $a_i = b_i$ , the ordering should be done as if  $a_i < b_i$ . If  $a_i = b_j$  the ordering should be done as if  $a_i < b_j$  for taking into account the intersection between  $N_i$  and  $N_j$  which then reduces to a point.

### 3.4 Example

$N_1 = [1, 8]$ ;  $N_2 = [2, 4]$ ;  $N_3 = [3, 9]$ ,  $N_4 = [5, 10]$ ,  
 $N_5 = [5, 6]$ ;  $N_6 = [7, 11]$ .

We have the ordering:  $a_1, a_2, a_3, b_2, a_4, a_5, b_5, a_6, b_1, b_3, b_4, b_6$ .

The algorithm first finds  $K = \{N_1, N_2, N_3\}$ . Then finding  $b_2$ ,  $K$  is kept in List.  $N_2$  is deleted from  $K$ , and  $N_4$  and  $N_5$  are added. Then the algorithm finds  $b_5$  and adds  $\{N_1, N_4, N_3, N_5\}$  to List. It deletes  $N_5$  from  $K$  and then adds  $N_6$ , then finds  $b_1$ . Then  $\{N_1, N_3, N_4, N_6\}$  is added to List. We might stop there, but the algorithm still makes  $K$  empty.



obtained interval :  $[[3,4] \cup [5,6] \cup [7,8]] = [3,8]$

NB: Let  $k = \max\{|K|, \bigcap_{i \in K} N_i \neq \emptyset\}$  be the maximal number of consistent sources; the maximality in the sense of inclusion can be replaced by a cardinality criterion which only keeps the consistent subsets of cardinality  $k$ . Nevertheless, although this enables us

to get a more precise result, it presupposes an independence assumption between the sources since it is based on a counting. If  $[\bigwedge_{i=1,n} S_i]$  is found to be too large, one can apply the above method to the supports instead of the cores and take  $[\bigwedge_j \bigvee_{i \in K^*j} S_i]$  as support for the result, where  $K^*j$  denotes a maximal consistent subset of  $\{S_1, \dots, S_n\}$ .

#### 4. Sources with unequal reliability levels

The information about the reliability of a source can be expressed under two different forms:

- either one can estimate the probability that a source is reliable (because one can have access to the proportion of times when it provided wrong information in the past, for instance),
- or we have a more qualitative information only, such as the considered source has 'a small reliability', or is 'very reliable', or is more reliable than another one. This seems to be often the case in practice.

In any case, it is assumed that the set of sources can be partitioned into subgroups of sources of equal reliability which can be linearly ordered from the more reliable to the less reliable one. This reliability ordering can depend on the considered attribute, since some sources may be more reliable for some topics than for others.

The proposed fusion procedure consists in

- i) applying the approach of Section 3 inside each subgroup where the sources play the same role;
- ii) combining the partial results by starting with the most reliable group using the following principle. Let  $i$  and  $j$  be two groups of sources such that  $i$  is more reliable than  $j$ . Then if the information given by  $i$  is consistent with the information given by  $j$  a conjunctive combination is performed. Otherwise only the information given by source  $i$  is kept. When the information to be combined is represented under the form of fuzzy intervals  $A_i$  and  $A_j$ , a

weighted combination is performed, where the importance of the most reliable source is 1 and the importance of the other source is the degree of consistency between the two sources ( $\text{height}(A_i \cap A_j)$ ).

Let us assume for instance that we have two groups of sources: "confidential" sources and "public" sources. The former are supposed to be more reliable than the latter. After the fusion step inside each group we get two fuzzy intervals represented by the two pairs  $(N_1, S_1)$  and  $(N_2, S_2)$  for the "confidential" and the "public" sources respectively. Then, after fusion we shall get the following result  $(N, S)$ :

$$\begin{aligned} N &= N_1 \cap N_2 \text{ if } N_1 \cap N_2 \neq \emptyset \\ &= N_1 \text{ otherwise} \\ S &= [S_1 \cup S_2]. \end{aligned}$$

The disjunctive fusion of supports is justified in order not to neglect any information, since it may happen that a less reliable source has the right information.

**Remark:** In case of two sources of unequal reliability, we might think of applying an approach based on conditioning in the spirit of Bayesian updating. This would lead to consider the least reliable sources as providing an a priori information to be revised by the information coming from the most reliable sources. The above proposal agrees with the conditioning view. But this latter approach does not provide any result if the sources are altogether inconsistent.

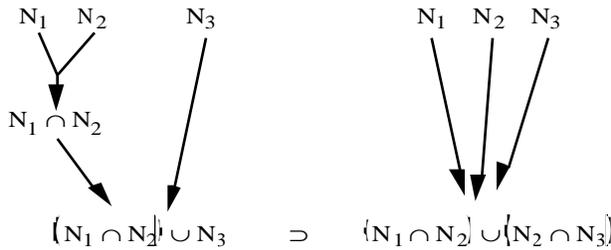
In some situations, it may happen that multiple pieces of information are given by the same source (possibly at different time instants). Three approaches may be considered for this problem of multiple values inside a source

- i) if information may become obsolete one may consider that the most recent information is the right one. This view albeit natural is not always suitable;

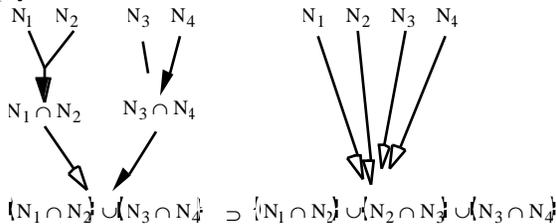
- ii) regard each piece of information as provided by a particular source;
- iii) perform partial fusion inside each source, before merging the unique, thus obtained, pieces of information (one by source).

The two latter approaches are not equivalent for the cores. In general, the results obtained by each of the two methods are not comparable: any of the two may give more precise results than the other in some cases.

Examples : Source 1 gives  $N_1$  and  $N_2$  such that  $N_1 \cap N_2 \neq \emptyset$ , source 2 gives  $N_3$  such that  $N_2 \cap N_3 \neq \emptyset$ , and  $N_3 \cap N_1 = \emptyset$ . Then we get



If source 1 gives  $N_1$  and  $N_2$  such that  $N_1 \cap N_2 \neq \emptyset$ , source 2 gives  $N_3$  and  $N_4$  such that  $N_3 \cap N_4 \neq \emptyset$ ; assume  $N_2 \cap N_3 \neq \emptyset$ , the other intersections being empty, then we can obtain :



it is enough to combine the two previous cases for obtaining results which are not comparable. Let 3 sources

source 1 :  $N_1$   $N_2$   
 source 2 :  $N_3$   $N_4$   
 source 3 :  $N_5$

Hypothesis :  $N_1 \cap N_2 \neq \emptyset$ ,  $N_2 \cap N_5 \neq \emptyset$ ,  $N_3 \cap N_4 \neq \emptyset$ ,  $N_2 \cap N_3 \neq \emptyset$  (all other intersections empty).

Then the inside-source combination followed by a combination of the partial results gives  $R_1 = (N_1 \cap$

$N_2) \cup (N_3 \cap N_4) \cup N_5$ . The direct combination gives  $R_2 = (N_1 \cap N_2) \cup (N_2 \cap N_5) \cup (N_2 \cap N_3) \cup (N_3 \cap N_4)$ . Clearly  $R_1 \not\subset R_2$  since  $N_5 \not\subset R_2$ , and  $R_2 \not\subset R_1$  since  $N_2 \cap N_3 \not\subset R_1$ .

Consequently the choice of a preliminary fusion step for the values given by the same source before the global fusion step should be motivated by the nature of the pieces of information coming from the same source, when they make a whole, proper to the source. By contrast, if the pieces of information are felt as heterogeneous, without any explicit link between them, it would be better to consider each piece of information as a particular source.

### 5. Interaction of expertise and fusion

Let us assume that we have at our disposal some (fully reliable) expertise linking the parameters which can be estimated by the sources, through relations or equations. If an equation applies to pieces of information given by source  $S_1$  for an object, it should also applies to the corresponding pieces of information given by source  $S_2$  which describes the same object.

The available expertise should enable us to improve the description of the considered object by modifying the result of the fusion. The question is to know if the expertise should be applied first inside each source and then we perform the fusion of the pieces of information refined thanks to the expertise, or if we only apply the expertise to the result of the fusion performed on the initial data. In case of conjunctive combination (as performed on the cores in the approach of Section 3), the two methods give the same result. This is no longer true with disjunctive fusion (which is applied to the supports), as shown by the following example:

*Example* :

Data provided by the sources:

Source A :  $X = [10, 12]$ ,  $Y = [2, 3]$ ,  $Z = [0, 15]$

Source B :  $X = [0, 11]$ ,  $Y = [3, 4]$ ,  $Z = [1, 6]$

Source C :  $X = [3, 7], Y = 2, Z = [1, 5]$

Result of the fusion:

$X = [0, 12], Y = [2,4], Z = [0, 15]$

Available expertise:  $Z = X - Y$

Result obtained by applying the expertise to the result of the fusion :

$X = [2, 12] Y = [2, 4], Z = [0, 10]$

But the correct method modifies the information first by applying the expertise for each source; it gives for the same initial data :

Source A :  $X = [10, 12], Y = [2,3], Z = [7,10]$

Source B :  $X = [4, 10], Y = [3, 4], Z = [1, 6]$

Source C :  $X = [3, 7], Y = 2, Z = [1, 5]$

and the result of the final fusion is :

$X = [3, 12] Y = [2, 4], Z = [1, 10]$

The first treatment, computationally simpler, does not propagate the constraints as much as possible and thus lead to evaluations which have a support which may be too large.

## 6. Fusing conjunctive information

### 6.1 Multiple-valued attributes

An attribute is said to be *multiple-valued* if it can take several values simultaneously. For instance, the attribute "age of the child(ren) of a person" may include several values if the person has several children. The representation of this type of data (called 'conjunctive data' ) when it is pervaded with uncertainty, requires the use of possibility distributions on the power set  $2^U$ , the set of subparts of referential  $U$ . Then a multiple-valued attribute  $x$  will be associated with the possibility distribution  $\pi_x$  such that

$$\forall E \subseteq U, \pi_x(E) = \text{Possibility}(x = E).$$

Clearly, the representation of imprecision for multiple-valued attributes is exponential, in general. Since the referential  $U$  is supposed to be linearly ordered, we might assume that  $\pi_x(E) > 0$  holds only if  $E$  is an interval, but this is not entirely satisfactory. Disjoint unions of intervals cannot be excluded.

Practically, we have to use a simplified, thus approximate, representation proposed by Dubois and Prade (1988) and Yager (1987, 1988). Let us first consider the case of an all-or-nothing distribution ( $\pi_x(E) \in \{0,1\}$ ) and let

$$E(x) = \{E, \pi_x(E) = 1\}$$

the set of conjunctive values possible for  $x$ . An approximate representation of  $E(x)$  can be made by means of two nested subsets  $(E_*, E^*)$  with  $E_* \subseteq E^*$  such that

$$E_* = \bigwedge \{E \in E(x)\}; E^* = \bigvee \{E \in E(x)\}.$$

$E_*$  is the set of individual values *certainly* taken by  $x$ . The complement of  $E^*$  is the set of individual values completely excluded for  $x$ . We may have  $E_* = \emptyset$  (one does not know any value for sure); one may also have  $E^* = U$  (none value is known for impossible).

In the general case where  $\pi_x$  can take intermediary degrees of possibility,  $\pi_x$  can be approximated by a pair of nested fuzzy sets such that

$$\forall u \in U, \mu_{E_*}(u) = \inf_{u \in E} 1 - \pi_x(E),$$

$$\mu_{E^*}(u) = \sup_{u \in E} \pi_x(E).$$

It is easy to see that  $\forall E \in E(x), E_* \subseteq E \subseteq E^*$  but the converse is wrong: one may have  $E_* \subseteq E \subseteq E^*$  and  $\pi_x(E) = 0$ . It can be checked that (Dubois and Prade, 1988) we have

$$\mu_{E_*}(u) > 0 \Rightarrow \mu_{E^*}(u) = 1$$

i.e., in terms of core and support  $S(E_*) \subseteq N(E^*)$ ; individual values which are somewhat certain are among the completely possible ones.  $(E_*, E^*)$  is thus an "enclosing" approximation of  $E(x)$ , which is more imprecise (values which are not possible may be regarded as possible). However the gain in simplicity for the representation is spectacular since we have only to handle pairs of fuzzy sets, which can be compared with one fuzzy set for single-valued attributes.

### 6.2 Fusing multiple-valued attribute information

Regarding information fusion, principles remain the same as for single-valued attributes but they lead to different algorithms on the approximate representation. The fusion of multiple-valued attributes does not seem to have received much attention in the scientific literature, although it leads to fusion methods which are dual from the previous ones. Indeed a conjunctive combination of conjunctive intervals is a set union, and a disjunctive combination of conjunctive intervals is a set intersection. For instance, if source  $i$  provides information  $(E_{*i}, E^*_{*i})$  for  $i = 1$  or  $2$ , and if these pieces of information are consistent, it is easy to see that the conjunctive combination gives

$$(E_{*1}, E^*_{*1}) \cap (E_{*2}, E^*_{*2}) = (E_{*1} \cup E_{*2}, E^*_{*1} \cap E^*_{*2})$$

We can see that there are more values which are certain and less values which are possible, This shows that the imprecision has diminished. Indeed

$$\{E, E_{*1} \cup E_{*2} \subseteq E \subseteq E^*_{*1} \cap E^*_{*2}\} = \{E, E_{*1} \subseteq E \subseteq E^*_{*1}\} \cap \{E, E_{*2} \subseteq E \subseteq E^*_{*2}\}$$

$$(\text{since } E \subseteq E'' \text{ and } E' \subseteq E'' \Leftrightarrow E \cup E' \subseteq E'')$$

$$E \subseteq E' \text{ and } E \subseteq E'' \Leftrightarrow E \subseteq E' \cap E'').$$

Thus the above result is indeed the approximation of  $E_1(x) \cap E_2(x)$ . In the same way, the disjunctive combination  $E_1(x) \cup E_2(x)$ , to be used if  $E_1(x) \cap E_2(x) = \emptyset$  leads to compute

$$(E_{*1}, E^*_{*1}) \cup (E_{*2}, E^*_{*2}) = (E_{*1} \cap E_{*2}, E^*_{*1} \cup E^*_{*2})$$

This can be similarly justified. The precision has decreased since there are more possible values and less certain ones.

Note that the case where  $E_1(x) \cap E_2(x) = \emptyset$  can be tested on the approximate representation by means of the following sufficient condition for inconsistency

$(E_{*1}, E^*_{*1})$  is inconsistent with  $(E_{*2}, E^*_{*2})$  if and only if  $E_{*1} \cup E_{*2} \not\subseteq E^*_{*1} \cap E^*_{*2}$  (an individual value which is certain becomes impossible) which entails (but not the converse)  $E_1(x) \cap E_2(x) = \emptyset$ .

The disjunctive combination may become useless if  $E_{*1} \cap E_{*2} = \emptyset$  and  $E^*_{*1} \cup E^*_{*2} = U$ , since everything then becomes possible. The advantage of the approximate representation of the imprecision of multiple-valued attribute information is that it leads to partial reuse of fusion algorithms from the single-valued case. Let us assume that we take for  $E^*_{*i}$  and  $E_{*i}$  a representation by pairs (core, support), we will then have an imprecise conjunctive information under the form of four nested subsets:  $N(E_{*i}) \subseteq S(E_{*i}) \subseteq N(E^*_{*i}) \subseteq S(E^*_{*i})$  due to the property recalled in 6.1.

The information combination in the multiple-valued case is made in a similar way as in the single-valued case. The inconsistency test in the general case (with intermediary levels of possibility) becomes :

$$(E_{*1}, E^*_{*1}) \text{ is inconsistent with } (E_{*2}, E^*_{*2}) \text{ if}$$

and only if

$$N(E_{*1}) \cup N(E_{*2}) \not\subseteq N(E^*_{*1}) \cap N(E^*_{*2})$$

or

$$S(E_{*1}) \cup S(E_{*2}) \not\subseteq S(E^*_{*1}) \cap S(E^*_{*2})$$

If  $N(E_{*i}) = [a_i, b_i]$ ,  $N(E^*_{*i}) = [c_i, d_i]$ , we have to check separately  $N(E_{*i}) \not\subseteq N(E^*_{*1}) \cap N(E^*_{*2})$  for  $i = 1, 2$ .

The delicate point is the search for maximal consistent subsets. It seems that this problem has never been considered in the literature as far as we know. For the moment, it is unclear if efficient methods can be found, although the problem can be stated rigorously.

## 7. Conclusion

This paper has sketched a simple approach to information fusion, which allows for a treatment of inconsistencies between sources. Some questions which are not often addressed in the literature, such as the handling of multiple-valued attributes, or of relations between (single-valued) attributes. A

comparaison with other combination modes existing in the possibility theory is still to be made. Moreover, the proposed approach is mainly based on the manipulation of the cores and the supports of the distributions; a generalization to any level cuts would be of interest.

## References

- M.A. Abidi, R.C. Gonzales, (eds.) *Data Fusion in Robotics and Machine Intelligence*. Academic Press, New York, 1992
- S. Benferhat, D. Dubois, H. Prade, Some syntactic approaches to the handling of inconsistent knowledge bases — Part 1: The flat case. *Studia Logica*, 58, 17-45, 1997.
- L. Cholvy, Reasoning about merged information. In: *Belief Change*, (D. Dubois, H. Prade, eds.), Vol.3 in Handbook of Defeasible Reasoning and Uncertainty Management Systems (DM. Gabbay, P. Smets, eds.), Kluwer Acad. Publ., Dordrecht, 233-263, 1998.
- R.M. Cooke, *Experts in Uncertainty*, Oxford University Press, Oxford, 1991.
- D. Dubois, H. Prade, Incomplete conjunctive information. *Comput. Math. Appli.*, 15 (10), 797-810, 1988.
- D. Dubois, H. Prade, Possibility theory and data fusion in poorly informed environments. *Control Eng. Practice*, 2, 811-823, 1994.
- D. Dubois, H. Prade, La fusion d'informations imprécises. *Traitement du Signal*, 11(6), 447-458, 1995.
- J. Gebhardt, R. Kruse, Parallel combination of information sources, In: *Belief Change*, (D. Dubois, H. Prade, eds.), Vol.3 in Handbook of Defeasible Reasoning and Uncertainty Management Systems (DM. Gabbay, P. Smets, eds.), Kluwer Acad. Publ., Dordrecht, 393-439, 1998.
- R.C. Luo, M.G. Kay, *Multisensor Integration and Fusion for Intelligent Machines and Systems*, Ablex Publ. Corp. Norwood, N.J., 1995.
- N. Rescher, R. Manor, On inference from inconsistent premises. *Theory and Decision*, 1, 179-219, 1970.
- R. R.Yager, Toward a theory of conjunctive variables, *Int. J. of General Systems*, 13, 203-227, 1987.
- R.R.Yager, Set based representations of conjunctive and disjunctive knowledge, *Information Sciences*, 41, 1-22, 1988.