Data Fusion System Evaluation in Possibility Theory

Gilles Zunino
Ecole des Hautes Etudes Industrielles
et Laboratoire I^{3}D (USTL)
13 rue de Toul
59 046 Lille Cedex FRANCE
☎ (+33) 3.28.38.48.58
Fax (+33) 3.28.38.48.04
Gilles.Zunino@hei.fr

A.M. Jolly – D. Jolly
Laboratoire d'Automatique I^{3}D
Université des Sciences et Technologies de Lille
Bâtiment P2
59 655 Villeneuve d'Ascq Cedex FRANCE
☎ (+33) 3.20.43.42.45
Fax (+33) 3.20.43.48.23
{Anne-Marie.Desodt,Daniel.Jolly}@univ-lille1.fr

Abstract : Data fusion aim is to combine various sources of information in order to refine them. Fusion can occur in several frameworks (probabilities, belief fonctions, possibility theory . . .) wherein a large number of combination scheme have already been studied. This paper addresses information quality evaluation after a fusion process in the framework of possibility theory. The meaning of "information improvement" is tied very closely to the application. After an attempt to bring out some criterions of quality in information, we propose a new way to evaluate information gain in data fusion process. An example of application in shown.

Keywords : Data fusion, possibility theory, information quality, specificity.

1 INTRODUCTION

In the framework of possibility theory a large number of fusion rules exists. They can be roughly classified in two great families : symetrical and dissymetrical. With symetrical rules [Bloch, 1996], all informations to be combined play the same role. A dissymmetric data fusion rule [Dubois and Prade, 1994] is a belief change which operates on the current knowledge when a new evidence is acquired. Currently, a large amount of work has been done to create new fusion rules but only a few authors [Kewley, 1993] tryied to evaluate the changes in term of information quality.

Before introducing how a data fusion system can be evaluated, some prerequisites are needed. First, it will be assumed that at least one source is correct, even if it impossible to say which one. This states that the data fusion system is not able to give a valid result from invalid inputs. Second, the "closed world" will be assumed. This means that the universe of discourse Ω is supposed to contain all the possible events. As a consequence, all possibility distribution of Ω will be normalised.

Evaluating a symetrical data fusion system is a hard task because combination can be done in three different modes : conjunction, disjunction and compromise. Conjunction must be used when sources agree : it can be seen as the intersection between informations. In this case, the more specific the resulting distribution is, the more the fusion process seems to achieved its goal.

The disjunctive mode is used when most of the sources disagree : it can be seen as the union of the sources. Intuitively, if it is impossible to bring out a consensus, all the sources are potentially right. If there is a large number of different sources which all disagree, the result tends to the total ignorance. Evaluating information in this context can be done by computing how much the total ignorance and the result of the union differs.

Compromise is a trade-off situation : sources do not really agree but they do not really disagree. The result reflects the "average decision" i.e. every information given by an input can be found in the result, maybe with a different possibility degree. Criterions for evaluating such an operator depend on the application.

Some symetrical operators (said "adaptative operators" [Deveughele and Dubuisson, 1996]) are able to combine continuously from disjunction to conjunction. If sources are concordant, the operator will apply a conjunctive combination and if the sources disagree, an union is made resulting in a disjunctive behavior. They do not have a simple, well established mode of operation : it depends on the conflict between sources. Since it evolves from disjunction to conjunction, when sources agree, the measure of quality should reduce to whatever used for evaluating conjuntive operator and when sources disagree, it should reduce to whatever used to evaluating disjunctive operator.

Dissymetrical fusion refines a current knowledge with a new evidence. Evaluating such a system cannot be done with algorithms described in this paper which addresses only symetrical fusion process.

First, some well-known existing measure on possibility distributions (cardinal, specificity, ...) are recalled. Next, based on those measures, an evaluation process for the three combination modes is proposed. A new algorithm for evaluating adaptative operators is described. The evaluation process presented is then applied to a simple example.

2 SOME EXISTING MEASURES ON POSSIBILITY DISTRIBUTIONS

In the framework of possibility theory several measures exists. For instance cardinal and specificity are commonly used. In order to evaluate the gain in a combination an intuitive idea is to compute measures on both inputs and output and compare them using a criterion. An evaluation can be : compute the specificity of each inputs and make sure the result is more specific than all the inputs.

In the following sections cardinal, specificity, non specificity, entropy, fuzziness degree and distances between two distributions and their properties are recalled.

2.1 CARDINAL

The cardinal of a possibility distribution π [De Luca and Termini, 1972] tries to evaluate the number of items in π. When π is discrete, the left hand side expression of equation (1) is used. The right hand side of equation (1) suits continuous distributions :

$$|\pi| = \sum_{\Omega} \pi(\omega) d\omega \qquad\qquad |\pi| = \int_{\Omega} \pi(\omega) d\omega \tag{1}$$

Cardinality can be seen as a measure of imprecision because it is minimal when the underlying possibility distribution is reduced to a singleton and maximal when π denotes the total ignorance.

2.2 SPECIFICITY

The specificity evaluates how much a possibility distribution is far from the most specific distribution. The most specific distribution take the value 1 at only one point in the universe Ω and 0 everywhere else. The specificity is computed using :

$$S_P(\pi) = \int_0^1 \frac{1}{|\pi_\alpha|}.d\alpha \quad \text{with} \quad \begin{cases} \pi_\alpha(\omega) & = 1 \text{ if } \pi(\omega) \geq \alpha \\ & = 0 \text{ otherwise} \end{cases} \tag{2}$$

It is straightforward to show that $S_P(\pi_A) = 1$ if and only if there is ω in Ω such that $A = \{\omega\}$. For two sets A and B with associated possibility distributions π_A and π_B S_p holds $A \subseteq B \Rightarrow Sp(\pi_A) \geq Sp(\pi_B)$.

2.3 NON SPECIFICITY

Non specificity (also called “U-Uncertainty”) tries to evaluate how much the information supported by this possibility distribution may contains alternatives. It’s a direct evaluation of imprecision since the more an information is imprecise, the less it is specific. This is computed using :

$$H(\pi) = \int_0^1 \log_2 |\pi_\alpha|.d\alpha \quad \text{with} \quad \begin{cases} \pi_\alpha(\omega) & = 1 \text{ if } \pi(\omega) \geq \alpha \\ & = 0 \text{ otherwise} \end{cases} \tag{3}$$

H is minimum (and equals to 0) if and only if there is ω in Ω such that $A = \{\omega\}$. For two sets A and B with associated possibility distributions π_A and π_B S_p we have $A \subseteq B \Rightarrow H(\pi_A) \leq H(\pi_B)$.

2.4 SHANNON'S ENTROPY

The SHANNON's entropy is heavily used in probability theory. If the probability degree $p(x)$ is replaced with a possibility degree $\pi(\omega)$, the entropy becomes :

$$H_S(\pi) = -\int_\Omega \pi(\omega)\log_2 \pi(\omega) d\omega \tag{4}$$

When the distribution is specific and certain, H_S is equal to 0. Replacing probabilities with possibilities is no the only way to extend SHANNON's entropy to the possibilistic framework. All the classical ways to transform probabilities to possibilities can be used [Dubois and Prade, 1988].

2.5 FUZZY INDICATION

This indication tries to quantify how much a set is fuzzy. A large number of such a criterion have been introduced by several authors. Among them, [Kaufmann, 1977] uses a distance d between the possibility distibutions π and π_{ref} derived from π. The reference distribution π_{ref} is often the nearest crisp subset of the fuzzy set associated with π. It can be computed in two different ways : the optimistic case and the pessimistic case as shown in equation (5). The real γ is used a scale factor since f should maps into $[0,1]$:

$$f(\pi) = \gamma d(\pi, \pi_{ref}) \quad \text{with} \quad \begin{cases} \pi_{opt}(\omega) & = 1 \text{ if } \omega \text{ in Support}(\pi) \\ & = 0 \text{ otherwise} \end{cases} \quad \begin{cases} \pi_{pess}(\omega) & = 1 \text{ if } \omega \text{ in Kernel}(\pi) \\ & = 0 \text{ otherwise} \end{cases} \tag{5}$$

The optimistic case is built using the support of the distribution and the pessimistic case is built using the core of the distribution : the cardinal of π_{opt} is always larger or equals to the cardinal of π_{pess}.

If d is the HAMMING's distance and $\gamma = \frac{1}{\max_\Omega d(\pi,\pi_{ref})}$, f reduces to the KAUFFMAN's indice which takes the value 0 as soon as the set is a crisp set and grows to the maximum when the fuzzy set is such that $\forall\omega \in \Omega, \pi(\omega) = 0.5$.

When the distribution is specific and certain or unspecific and uncerain, this indication gives 0. If the possibility distribution is not normalised (we excluded this scenario), the result is 0 too. KAUFFMAN's indication is unable to distinguish between normalized and a non normalized distributions.

2.6 DISTANCES

It is also possible to evaluate how much two distributions agree with a distance. Such a criterion is called "indication of conflict" between two distributions. This can be useful to evaluate how much the output of a data fusion process is far from a reference distribution (for instance an "*a priori*" information). The reference distribution can also be a linear combination of π_{pess} and π_{opt} (defined in equation (5)) like $\pi_{ref} = \lambda\pi_{opt} + (1-\lambda)\pi_{pess}$ where $\lambda \in [0,1]$. Let d be a distance from Ω^2 to the real interval $[0,1]$ and n a monotonically non-decreasing function of d verifying $n(0) = 1$ and $n(\infty) = 0$. The criterion I is given by :

$$I(\pi_1, \pi_{ref}) = n(d(\pi, \pi_{ref})) \qquad \begin{cases} n_1(d) = [1+d]^{-1} \\ n_2(d) = e^{-d} \\ n_3(d) = 1 - \frac{d}{d_0} \quad \text{if} \quad \exists\, d_0 \,/\, 0 < \sup_{(\pi_1,\pi_2)\in\Omega^2} d(\pi_1,\pi_2) \le d_0 < \infty \end{cases} \tag{6}$$

The right hand side of the previous equation (6) recall some classical choices for function n. When d_0 exists, the n_3 is frequently recommended because of its linearity. If d is a normalised distance, $d_0 = 1$.

The choice of d has a great influence on the final results. A lot of well known distances like the HAUSDORFF one have been studied in [Zunino et al., 1997]. Here, we will use the distance of BHATTACHARYYA d_{Bat} [Kailath, 1967] defined in the continuous case by (when the integrals converge) :

$$\begin{aligned} P_A(x) &= \frac{\pi_1(x)}{\int_\Omega \pi_1(x)dx} & R(P_A, P_B) &= \int_\Omega \sqrt{P_A(x)P_B(x)}dx \\ P_B(x) &= \frac{\pi_2(x)}{\int_\Omega \pi_2(x)dx} & d_{Bat}(\pi_1,\pi_2) &= \sqrt{1 - R(P_A, P_B)} \end{aligned} \tag{7}$$

All the indications above are centered on giving a measure on distributions. The following paragraph explain how they can used to evaluate a symetrical fusion process.

3 EVALUATION OF SYMETRICAL OPERATORS

It is very hard (and maybe quite impossible) to evaluate the quality of an information without depending on the context and this remark applies to evaluation of data fusion systems. Fusion algorithms are expensive, at least in term of computation time. Our evaluation can bee seen as an estimation the work done by the fusion operator i.e. the amount of modifications made from the input set to give the result. If this “modification rate” is low, the fusion operator is not really needed because added value introduced is negligible.

Evaluating the added value of a fusion operator cannot be dissocied form the operation mode : conjunction, disjunction or compromise. Some operators like adaptative one do not have a fixed behavior, making the evaluation more difficult and computing the added value introduced by such operator will be disccussed after the three basic modes.

3.1 CONJUNCTION

When fusion is done in conjunctive mode, the aim is to dismiss as much alternatives as possible. Refining means keeping only information available in all sources and so decrease data uncertainty as stated in [Kewley, 1993]. Due to the mathematical operation used, this leads to an increase of the belief in the area common to all the distributions (i.e. the sets intersection). Let’s suppose the fusion process did not exist. From the input set, only a source π_A would be considered. If the fusion process exists, the result π_R and π_A can be used to derive the amount of added value during fusion. Choosing π_A is a critical task. Here, it will be assumed that among all the input distributions; π_A is the closest to π_R according to BHATTACHARYYA’s distance (equation (7)). This implies find π_A such that $d_{Bat}(\pi_A, \pi_R)$ is the maximum.

The fusion process achieved his goal if the resulting distribution is more specific (a large number of alternatives have been suppressed), if the cardinal and, non specificity and Entropy have decreased. In order to evaluate the added value during fusion, thoses atributes are computed on π_A and π_R and the results are substracted, normalized (see equation (8)) and then summed.

$$E_C = \frac{|\pi_A| - |\pi_R|}{\max(|\pi_A|, |\pi_R|)} + \frac{S_p(\pi_R) - Sp_p(\pi_A)}{\max(S_p(\pi_R), S_p(\pi_A))} + \frac{H(\pi_R) - H(\pi_A)}{\max(\|H(\pi_R)\|, \|H(\pi_A)\|)} + \frac{H_S(\pi_A) - H_S(\pi_R)}{\max(\|H_S(\pi_A)\|, \|H_S(\pi_R)\|)} \quad (8)$$

The more E_C is low, the less fusion process introduced added value. If $E = 0$, the fusion did not perform any work.

3.2 DISJUNCTION

Let ’s now suppose data are known to disagree. This implies the use of a cautious fusion process (i.e. fusion in disjunctive mode) which will ensures that each input is present in the output. In this case, the fusion process should collect all alternatives in a new possibility distribution and make this distribution as specific as possible. The last condition is needed to make sure a fusion process which returns a possibility distribution covering the whole universe will be flagged as bad. If the fusion process did not exist, the best choice would be to consider the nearest crisp set (given by its possibility distribution π_{NS}) of input’s union. As before in the conjunctive case, we will compute the added value using π_{NS}. The cardinal and non specificity should have decreased and the specificity should have increased. The evaluation E_D is given by :

$$E_D = \frac{|\pi_{NS}| - |\pi_R|}{\max(|\pi_{NS}|, |\pi_R|)} + \frac{S_p(\pi_R) - Sp_p(\pi_{NS})}{\max(S_p(\pi_R), S_p(\pi_{NS}))} + \frac{H(\pi_{NS}) - H(\pi_R)}{\max(\|H(\pi_{NS})\|, \|H(\pi_R)\|)} \quad (9)$$

As with E_C , the more E_D is low, the less fusion process introduced added value.

3.3 COMPROMISE

The compromise operator is used when data do not agree but not really disagree. It is indeed necessary to keep all the alternatives (just like in a disjunction process) but change the belief degree in order to give the result looks like the “average” distribution. As before, if the data fusion process did not exist, a correct choice would be to consider only the nearest distribution π_A from the intersection of all the inputs. The word done during the fusion can be evaluated using the distance from the result to π_A added to the variation of non specificity and entropy like stated in the following equation :

$$E_M = d_{Bat}(\pi_A, \pi_R) + \frac{H(\pi_A) - H(\pi_R)}{\max(\|H(\pi_A)\|, \|H(\pi_R)\|)} + \frac{H_S(\pi_R) - H_S(\pi_A)}{\max(\|H_S(\pi_R)\|, \|H_S(\pi_A)\|)} \quad (10)$$

As with E_C , the more E_M is low, the less fusion process introduced added value.

4 EVALUATION OF ADAPTATIVE OPERATORS

We are able to evaluate a fusion process operating in one of the three basic modes. An adaptive operator continuously evolves from conjunctive to disjunctive mode. It seems interesting to compute how much the operator behaves conjunctively, disjunctively and like a mean. Let's suppose the fusion was done using a conjunctive operator like $\min$, the result would be $\pi_{R_{\min}}$. If the fusion was made using a disjunctive operator like $\max$, the result would be $\pi_{R_{\max}}$ and if the fusion was made using the arithmetic mean, the result would be $\pi_{R_{\text{mean}}}$. The amount of the conjunctive behavior (noted I_{Conj}), the amount of disjunctive behavior (noted I_{Disj}) and the amount of compromise behavior (noted I_{Mean}) will be derived from the three distributions $\pi_{R_{\min}}$, $\pi_{R_{\max}}$ and $\pi_{R_{\text{mean}}}$ using the BHATTACHARYYA's distance (see equation (11)). The evaluation of the added value introduced by the adaptive operator E_A can be computed using :

$$\begin{cases} I_{Conj} = d_{Bat}(\pi_R, \pi_{R_{\min}}) \\ I_{Disj} = d_{Bat}(\pi_R, \pi_{R_{\max}}) \\ I_{Mean} = d_{Bat}(\pi_R, \pi_{R_{\text{mean}}}) \end{cases} \qquad E_A = I_{Conj}E_C + I_{Disj}E_D + I_{Mean}E_M \tag{11}$$

The π_A distribution needed is as before, the closest distribution, computed according to each section : conjunction, disjunction and compromise.

5 EXAMPLE

In this section the evaluation process presented is applied to a very simple fusion process. Let's consider a tank containing a liquid continually mixed. The tank is equipped with three sensors A, B, C watching the temperature respectively at the bottom, the middle and the top of a tank. The aim of the mixer is to keep the temperature equal everywhere into the tank. Periodicaly, one measure is acquired from each sensor and a fusion process is used to compute the tank's temperature. Intuitively, the most adapted fusion mode seems to be the mean. For illustration purpose, the fusion will be done in the three basic modes : conjunction, compromise and disjunction.

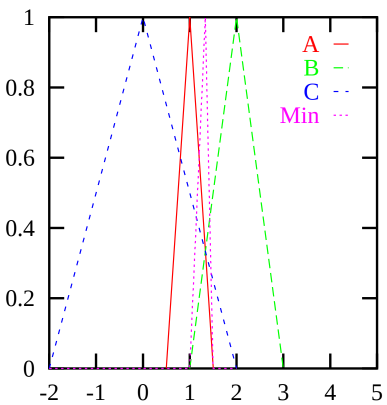

Figure 1: Conjunction.

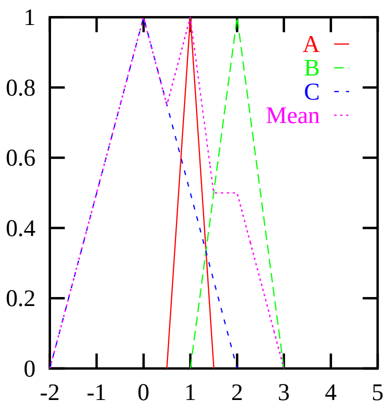

Figure 2: Compromise.

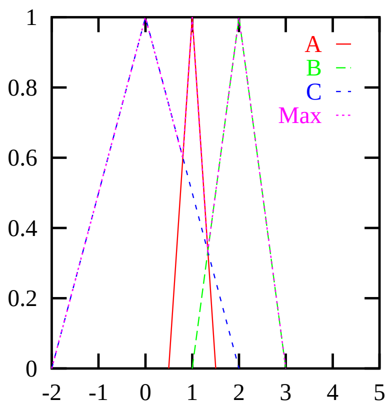

Figure 3: Disjunction.

Figure 4: Fusion in the three basic modes.

In this simulation (see figure (4)), the distributions are supposed to be triangular without a loss of generality. Each distribution is given by a center point and a gap applied on each side. The first one, A, is centered on 1 with a gap value of 0.5 so A extends from 0.5 to 1.5. The second one B is centered on 2 with a gap of 1 and the third C is centered on 0 with a gap of 2. In this example, the fusion has been done in the three basic modes : conjunctive (using $\min$ operator), disjunctive (using $\max$ operator) and compromise (using the mean operator).

Let's now have a look at the evaluation process (see figure (8)). It presents the result of the evaluation in the three basic mode. The curves have been made using the following algorithm : fix a gap value for A in the range $[0.5; 1.6]$ (for $\min$ evaluation), $[0.5; 1.6]$ (for mean evaluation) and $[0.5; 2.5]$ (for $\max$ evaluation) and leave other distributions unchanged. Next, compute the result of the fusion, evaluate it and put a point with coordinates (gap,evaluation) on the curve. The horizontal axis give the gap and the vertical axis the result of evaluation. On the $\min$ data fusion process evaluation, when the distribution has a gap of 1, the added value introduced by the operator is low. On the $\max$ data fusion process evaluation, when the distribution has a gap of 2, the added value introduced by the operator is low.

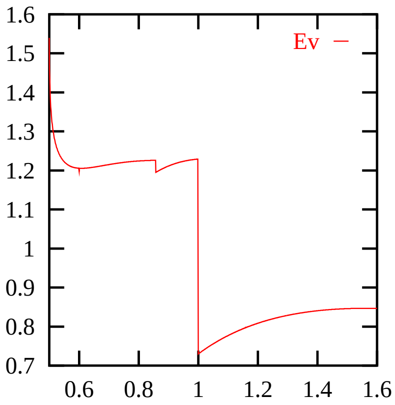

Figure 5: Evaluation for min.

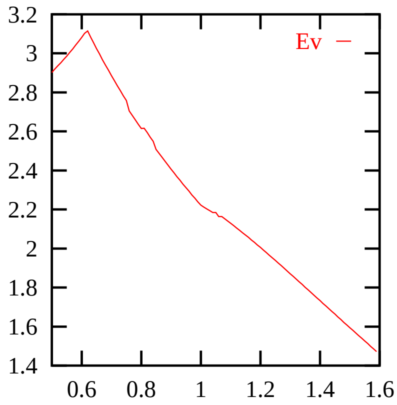

Figure 6: Evaluation for mean.

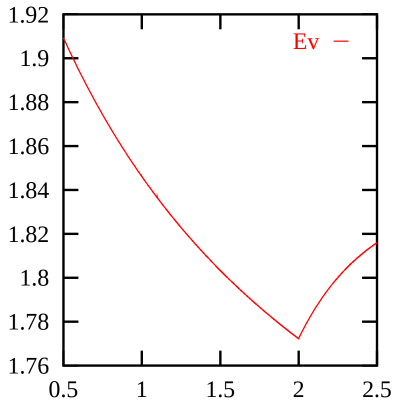

Figure 7: Evaluation for max.

Figure 8: Fusion in the three basic modes.

6 CONCLUSION

The notion of "added value" has been introduced for data fusion algorithms. It tries to quantify the amount of work done by the operator to give the result. Four cases have been distingued : the three basic data fusion modes : conjunctive, disjunctive, compromise and the adaptative operator. The adaptive operator evaluation scheme has been built on top of the three basic modes, using three new measures on an adaptive operator : a measure of conjunctive behavior, disjunctive behavior and compromise behavior. The "added value" measure is useful to banch the fusion process. If the gain is considered beeing too low, the fusion process may simply be bypassed in order to limit computation time.

Our work is now going to focuse on the evaluation of dissymetric data fusion operations and the evaluation of data fusion systems where result given by a fusion operator is used as an input to another fusion operator.

References

[Bloch, 1996] Bloch, I. (1996). Information combination operators for data fusion : A comparative review with classification. *IEEE SMC*, 26(1):447–458.

[De Luca and Termini, 1972] De Luca, A. and Termini, S. (1972). A definition of non-probabilistic entropy in the setting of fuzzy sets theory. In *Information and Control*, volume 20, pages 301–312.

[Deveughele and Dubuisson, 1996] Deveughele, S. and Dubuisson, B. (1996). The influence of a conflict index in the frame of the adaptative combinaison. In *CESA'96 IMACS Symp. Robotics and Cybernetics*, pages 98–103, Lille.

[Dubois and Prade, 1988] Dubois, D. and Prade, H. (1988). *Théorie des Possibilités*. Masson, 2 edition.

[Dubois and Prade, 1994] Dubois, D. and Prade, H. (1994). A survey of belief revision and updating rules in various uncertainty models. *Int. J. of Intelligent Systems*, 9:61–100.

[Kailath, 1967] Kailath, T. (1967). The BHATTACHARYYA distance measure in signal detection. *IEEE Trans. Commun. Tech.*, 1:607–637.

[Kaufmann, 1977] Kaufmann, A. (1977). *Introduction à la théorie des ensembles flous*, volume 4, chapter 92, pages 39–58. Masson.

[Kewley, 1993] Kewley, D. (1993). A model for evaluating data fusion systems. *IEEE – 1058-6393/93*, pages 273–277.

[Zunino et al., 1997] Zunino, G., Jolly-Desodt, A., and Jolly, D. (1997). Indication of conflict in data fusion. In *IFAC Symp. on Intelligent Components and Instruments for Control Applications SICICA'97*, pages 393–398, Annecy.