Additivity and monotonicity of measures of information defined in the setting of Shafer's evidence theory

Didier Dubois Henri Prade

Université Paul Sabatier
Laboratoire Langages et Systèmes Informatiques
118, route de Narbonne
31062 TOULOUSE CEDEX

1 - Introduction

In a recent paper [6], the authors survey different measures of information which have been proposed in the setting of Shafer's evidence theory [10]. More particularly three measures are considered :

$$. \ HC(m) = -\sum_{A \subseteq X} m(A) \ . \ \ln[Cr(A)] \quad (1)$$

$$. \ HI(m) = +\sum_{A \subseteq X} m(A) \ . \ \log_2[|A|] \quad (2)$$

$$. \ HD(m) = -\sum_{A \subseteq X} m(A) \ . \ \ln[Pl(A)] \quad (3)$$

where ln denotes the Napierian logarithm,
where m denotes a basic probability assignment, i.e. a set function from $\mathcal{P}(X)$ (the set of subsets of a finite universe X) to [0,1] such that

$$m(\emptyset) = 0 \ ; \ \sum_{A \subseteq X} m(A) = 1 \quad (4)$$

where |A| denotes the cardinality of the ordinary subset A and where Cr and Pl are the credibility and plausibility functions defined from m by

$$Cr(A) = \sum_{B \subseteq A} m(B) \; ; \; Pl(A) = \sum_{A \cap B \neq \emptyset} m(B) \quad . \tag{5}$$

A subset A such that $m(A) > 0$ is called a focal element.

The measure HC was introduced by Höhle [9] and as explained in [6] can provide an estimate of the amount of confusion between the focal elements of m (HC is zero only when there is one focal element and is maximum when the weighting expressed by m is equally shared among a maximum number of focal elements which are not included in each other).

The measure HI was introduced by Higashi and Klir [8] in the setting of Zadeh's possibility measures [15] (a particular case of plausibility functions) and extended to Shafer's framework in Dubois, Prade [4]. HI provides an estimate of the imprecision of the focal elements since HI is zero only if all the focal elements are singletons (which corresponds to a maximum of precision and turns out to be the case where m reduces to a regular probability density) and HI is maximum if is such that $m(X) = 1$ (the universe is the only focal element, which corresponds to a situation of total ignorance). HI can be viewed as a weighted Hartley measure [8].

The measure HD was introduced by Yager [13] and provides an estimate of the disjointness of focal elements, since HD is zero only when the set of focal elements has an intersection which is non-empty and HD is maximum if m reduces to a probability density (the focal elements are singletons and thus are all disjoint). HD is zero in particular when the focal elements are nested (i.e. they can be linearly ordered with respect to set inclusion), i.e. when the plausibility function is defined from m is nothing but a possibility measure [3], [5].

In this short note we study additivity and monotonicity properties which are not presented in [6] (except for HI in the particular case of possibility measures), as well as the multiplicative behavior of a measure of specificity previously introduced by Yager. We need to have extended definitions for the concepts of projection and Cartesian product in order to be able to present additivity properties and the extension of the concept of set inclusion in order to introduce monotonicity properties. We first consider additivity issues with the necessary perequisites ; then we give a background on inclusion in the framework of Shafer's evidence theory and discuss monotonicity properties.

2 - Projection, Cartesian product in Shafer's evidence theory

Let m be a basic probability assignment on XxY. Its projections on X and Y are respectively defined by

$$\forall A \subseteq X,\ m_X(A) = \sum_{A=\mathrm{proj}(S;X)} m(S)\ ;\ \forall B \subseteq Y,\ m_Y(B) = \sum_{B=\mathrm{proj}(S;Y)} m(S) \qquad (6)$$

where proj(S;X) and proj(S;Y) denote the projections of S on X and Y respectively. Thus the focal elements of m_X (resp. m_Y) are just the projections on X (resp. on Y) of those of m. This projection coincides with the projection of a possibility distribution as defined by Zadeh [15], when m is equivalent to a possibility distribution. See Shafer [11].

Given two basic probability assignments m_X and m_Y on X and Y respectively, their Cartesian product $m = m_X x m_Y$ is defined by (see Shafer [11])

$$\forall A \subseteq X,\ \forall B \subseteq Y,\ m(AxB) = m_X(A)\ .\ m_Y(B) \qquad (7)$$

where the focal elements of m are only obtained as Cartesian products of focal elements of m_X and m_Y. (7) can be viewed as a direct application of Dempster's rule of combination [1] [10] when we consider the combination of the cylindrical extensions of m_X and m_Y on XxY. However note that (7) does <u>not</u> coïncide with Zadeh's definition of Cartesian product [2] (where min is used instead of the product) when m is equivalent to a possibility distribution, since (7) entails $Pl(\{(x,y)\}) = Pl_X(\{x\})\ .\ Pl_Y(\{y\})$.

3 - Additivity properties

We need the following

Lemma 1 : Let f be a set function such that

$\forall A, \forall B, f(AxB) = f_1(A)\ .\ f_2(B)$

Then, the following equality holds

$$\sum_{A,B} m_1(A).m_2(B)\ .\ \ln(f(AxB)) = \sum_A m_1(A).\ln(f_1(A)) + \sum_B m_2(B).\ln(f_2(B))$$

where m_1 and m_2 are two basic probability assignments on X and Y respectively.

Proof :

$$\sum_{A,B} m_1(A).m_2(B).\ln(f(AxB)) = \sum_{A,B} m_1(A).m_2(B)[\ln(f_1(A))+\ln(f_2(B))]$$

$$= \sum_{A,B} m_1(A).m_2(B).\ln(f_1(A)) + \sum_{A,B} m_1(A).m_2(B).\ln(f_2(B))$$

$$= \sum_{A} m_1(A).\ln(f_1(A)) .[\sum_{B} m_2(B)]+ \sum_{B} m_2(B).\ln(f_2(B)) .[\sum_{A} m_1(A)]$$

$$= \sum_{A} m_1(A).\ln(f_1(A)) + \sum_{B} m_2(B).\ln(f_2(B))$$ (due to the normalization of m_1 and m_2).

Q.E.D.

We have

. $|AxB| = |A| . |B|$

. $Cr(AxB) = Cr_X(A).Cr_Y(B)$

. $Pl(AxB) = Pl_X(A).Pl_Y(B)$ (see Shafer [11])

where Cr and Pl are defined by (5) from m (obtained as the Cartesian product of m_X and m_Y defined by (7)) and where Cr_X and Pl_X (resp. Cr_Y and Pl_Y) are defined by (5) from m_X (resp. m_Y). Then using Lemma 1, the following additivity properties can be easily established

$$HC(m_X x m_Y) = HC(m_X) + HC(m_Y) \quad (8)$$

$$HI(m_X x m_Y) = HI(m_X) + HI(m_Y) \quad (9)$$

$$HD(m_X x m_Y) = HD(m_X) + HD(m_Y) \quad (10)$$

Remark : Yager [12] introduced a so-called specificity measure for possibility distributions, which has been extended to any kind of basic probability assignments [13], [4]. This index is defined by

$$Sp(m) = \sum_{A} \frac{m(A)}{|A|} \quad (11)$$

Sp provides an estimate of the precision of the focal elements since Sp is maximum (i.e. is equal to 1) if all the focal elements are singletons and is minimum (then equal to $\frac{1}{|X|}$) in case of total ignorance. Note that HI and Sp have opposite behaviors, the former estimates the imprecision and the latter the precision. It can be easily checked that Sp satisfies the multiplicative property :

$$Sp(m_X x m_Y) = Sp(m_X) . Sp(m_Y) \quad (12)$$

Indeed :

$$\sum_{A,B} \frac{m(A\times B)}{|A\times B|} = \sum_{A,B} \frac{m_X(A).m_Y(B)}{|A|.|B|} = \sum_{A} \frac{m_X(A)}{|A|} \left(\sum_{B} \frac{m_Y(B)}{|B|}\right).$$

4 - A sub-additivity property of HI :

Besides we have the following sub-additivity property for HI :

$$HI(m) \leq HI(m_X) + HI(m_Y) \tag{13}$$

where m_X and m_Y are obtained as the projections of m on X and Y respectively.

Proof :

$$HI(m_X) = \sum_{A\subseteq X} m_X(A).\log_2[|A|] = \sum_{A\subseteq X} \Big(\sum_{A=proj(S;X)} m(S)\Big).\log_2[|A|]$$

$$= \sum_{A\subseteq X} \sum_{A=proj(S;X)} m(S).\log_2[|proj(S;X)|]$$

$$= \sum_{S\subseteq X\times Y} m(S).\log_2[|proj(S;X)|].$$

Therefore

$$HI(m_X) + HI(m_Y) = \sum_{S\subseteq X\times Y} m(S).\log_2[|proj(S;X)|.|proj(S;Y)|]$$

$$\geq \sum_{S\subseteq X\times Y} m(S).\log_2[|S|] = HI(m)$$

Q.E.D.

Note that when $m = m_X \times m_Y$, (13) holds with equality.

5 - Super-multiplicative behavior of Yager's specificity in case of possibility distributions

The specificity measure Sp defined by (11), in general does not satisfy the super-multiplicativity property

$$Sp(m) \geq Sp(m_X) \, . \, Sp(m_Y) \tag{14}$$

which would be analogous to (13), m being at least precise as the Cartesian product of its projections $m_X \times m_Y$.

Counter-example of (14) :

$X = \{x_1, x_2, x_3\}$, $Y = \{y_1, y_2, y_3\}$. m is pictured on Fig. 1 and defined by

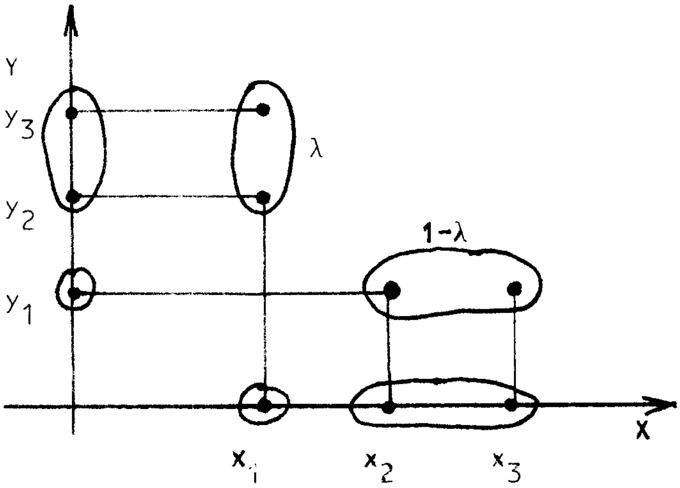

Figure 1

$$m(\{(x_1,y_2),(x_1,y_3)\}) = \lambda$$
$$m(\{(x_2,y_1),(x_3,y_1)\}) = 1-\lambda$$

Then

$$m_X(\{x_1\}) = \lambda \ ; \ m_X(\{(x_2,x_3)\}) = 1-\lambda$$
$$m_Y(\{y_1\}) = 1-\lambda \ ; \ m_Y(\{(y_2,y_3)\}) = \lambda$$

We have

$$Sp(m_X) = \lambda + \frac{(1-\lambda)}{2} = \frac{1+\lambda}{2}$$
$$Sp(m_Y) = \frac{\lambda}{2} + 1-\lambda = \frac{(2-\lambda)}{2}$$
$$Sp(m) = \frac{\lambda}{2} + \frac{1-\lambda}{2} = \frac{1}{2}$$

For $\lambda = \frac{1}{2}$ we have $Sp(m_X).Sp(m_Y) = \frac{9}{16} > \frac{1}{2} = Sp(m)$, which contradicts (14). It can be checked that, $\forall \lambda \in [0,1]$, $Sp(m_X).Sp(m_Y) = \frac{1}{4}(1+\lambda)(2-\lambda) \geq \frac{1}{2} = Sp(m)$. Note that on this example $HI(m) = HI(m_X x m_Y)$, i.e. a limiting case of (13).

However the inequality (14) still holds if m reduces to a possibility distribution.

<u>Proof of (14) in case of a possibility distribution</u>

In case of a possibility distribution the focal elements are nested. Let $F_1 \subseteq \ldots \subseteq F_r$ be the r focal elements of m defined on XxY. Let $s_i = \frac{1}{|proj(F_i;X)|}$ and $t_i = \frac{1}{|proj(F_i;Y)|}$. We have $s_i.t_i \leq \frac{1}{|F_i|}$. Then in order to establish (14) in this particular case, it is sufficient to prove, denoting $\lambda_i \stackrel{\Delta}{=} m(F_i)$,

$$(\sum_{i=1}^{r} \lambda_i.s_i) \ . \ (\sum_{i=1}^{r} \lambda_i.t_i) \leq \sum_{i=1}^{r} \lambda_i.s_i.t_i \qquad (15)$$

since $\sum_{i=1}^{r} \lambda_i \ . \ \frac{1}{|F_i|} = Sp(m)$ is an upper bound of the right part of the inequality. Due to the fact that the projections of the focal elements are also nested, we have $s_1 \geq \ldots \geq s_r > 0$ and $t_1 \geq \ldots \geq t_r > 0$. With $\lambda'_i = \frac{\lambda_i.t_i}{\sum_{i=1}^{r} \lambda_i.t_i}$ the inequality (15) can be rewritten

$$\sum_{i=1}^{r} \lambda_i . s_i \leq \sum_{i=1}^{r} \lambda'_i . s_i \qquad (16)$$

where $\sum_{i=1}^{r} \lambda'_i = 1$. Note that $\lambda'_1 \geq \lambda_1$ (since $\sum_{i=1}^{r} \lambda_i . t_i \leq \sum_{i=1}^{r} \lambda_i . t_1 = t_1$) and $\lambda'_r \leq \lambda_r$ (since $\sum_{i=1}^{r} \lambda_i . t_i \geq \sum_{i=1}^{r} \lambda_i . t_r = t_r$). More generally it can be checked that if

$\lambda'_i \geq \lambda_i$, then $\forall\, j < i$, $\lambda'_j \geq \lambda_j$ and if

$\lambda'_i \leq \lambda_i$, then $\forall j > i$, $\lambda'_j \leq \lambda_j$ since the t_i's are decreasingly ordered. Thus, $\exists\, k$, $\forall\, i \in [\![1,k]\!]$, $\lambda'_i \geq \lambda_i$ and $\forall\, i \in [\![k+1,r]\!]$, $\lambda'_i \leq \lambda_i$. Since the s_i's are decreasingly ordered, the weighted mean of the s_i's computed with the λ'_i's is then greater than or equal to the one computed with the λ_i's. Therefore (16) holds.

Q.E.D.

Remark :

When m_X and m_Y reduce to possibility distributions π_X and π_Y, there is another way, apart from (7), for defining their Cartesian product, namely by directly combining the two possibility distributions with the min operation ; let $m_X \otimes m_Y$ denote the Cartesian product of m_X and m_Y, defined as the basic probability assignment equivalent to $\min(\pi_X, \pi_Y)$. Then HI still satisfies the additivity property [8] [6]

$$HI(m_X \otimes m_Y) = HI(m_X) + HI(m_Y) \qquad (17)$$

Note that due to (9) and (17) the results of the two combinations $m_X x m_Y$ (using Dempster's rule) and $m_X \otimes m_Y$ (using min operation on possibility distributions) have exactly the same amount of imprecision in terms of HI, but $m_X \otimes m_Y$ reduces to a possibility distribution while $m_X x m_Y$ does not in general.

As shown by the following counter-example, we have not an analogous property for Sp, namely

$Sp(m_X \otimes m_Y) \neq Sp(m_X).Sp(m_Y)$.

$X = \{x_1, x_2, x_3\}$; $\pi_X(x_1) = 1$, $\pi_X(x_2) = 0.8$, $\pi_X(x_3) = 0.5$

$Y = \{y_1, y_2\}$; $\pi_Y(y_1) = 1$, $\pi_Y(y_2) = 0.6$

Let $\pi_{X,Y} = \min(\pi_X, \pi_Y)$. We have

$\pi_{X,Y}(x_1,y_1) = 1$; $\pi_{X,Y}(x_2,y_1) = 0.8$; $\pi_{X,Y}(x_1,y_2) = \pi_{X,Y}(x_2,y_2) = 0.6$;

$\pi_{X,Y}(x_3,y_1) = \pi_{X,Y}(x_3,y_2) = 0.5$.

In [4] it is shown that the expression (11) of Sp can be rewritten, when m reduces to a possibility distribution π,

$$Sp(m_\pi) = \sum_{i=1}^{n} \frac{\pi(x_i) - \pi(x_{i+1})}{i} \quad (18)$$

where the n elements of X are decreasingly ordered according to π and where $\pi(x_{n+1}) = 0$ by convention. Then in our example we have

$$Sp(m_{\pi_X}) = \frac{1-0.8}{1} + \frac{0.8-0.5}{2} + \frac{0.5-0}{3} \simeq 0.516$$

$$Sp(m_{\pi_Y}) = \frac{1-0.6}{1} + \frac{0.6-0}{2} = 0.7$$

$$Sp(m_{\pi_{X,Y}}) = \frac{1-0.8}{1} + \frac{0.8-0.6}{2} + \frac{0.6-0.6}{3} + \frac{0.6-0.5}{4} + \frac{0.5-0.5}{5} + \frac{0.5-0}{6}$$

$$\simeq 0.411 \neq Sp(m_{\pi_X}) \,.\, Sp(m_{\pi_Y}) \simeq 0.361$$

6 - HC and HD are not sub-additive

The measures of information HC and HD do not satisfy a sub-additivity property similar to (13).

Counter-example :

Let us consider the basic probability assignment m and its projections m_X and m_Y pictured on Fig. 2, where $m(S_1) = \lambda$, $m(S_2) = 1-\lambda$, and $S_1 \cap S_2 = \emptyset$.

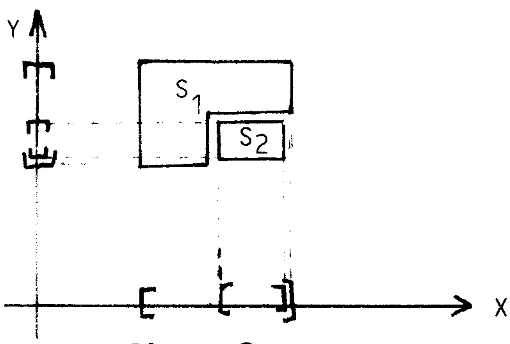

Figure 2

Then $m_X(proj(S_1;X)) = \lambda$ and $m_X(proj(S_2;X)) = 1-\lambda$ with $proj(S_2;X) \subseteq proj(S_1;X)$; $m_Y(proj(S_1,Y)) = \lambda$ and $m_Y(proj(S_2;Y)) = 1-\lambda$ with $proj(S_2;Y) \subseteq proj(S_1;Y)$. Then we have the following results

	Cr(S)	Pl(S)	$Cr_X(proj(S;X))$	$Cr_Y(proj(S;Y))$	$Pl_X(proj(S;X))$	$Pl_Y(proj(S;Y))$
S_1	λ	λ	1	1	1	1
S_2	$1-\lambda$	$1-\lambda$	$1-\lambda$	$1-\lambda$	1	1

Then $HC(m) = -(\lambda \ln \lambda + (1-\lambda).\ln (1-\lambda)) = HD(m)$

$HC(m_X) = -(1-\lambda).\ln (1-\lambda) = HC(m_Y)$

The sub-additivity property will be violated if

$$- HC(m) = + \lambda \,.\, \ln \lambda + (1-\lambda).\ln(1-\lambda) < + 2(1-\lambda).\ln(1-\lambda) = - HC(m_X) - HC(m_Y)$$

i.e. if $\lambda . \ln \lambda < (1-\lambda).\ln(1-\lambda)$

i.e. if $\lambda < 1-\lambda$, which is feasible.

Clearly we have

$$\forall \lambda, \ HD(m) > HD(m_X) + HD(m_Y) = 0$$

Q.E.D.

However if m reduces to an ordinary probability density on XxY, HC = HD is just Shannon entropy and in this case the sub-additivity property holds.

7 - Inequalities between the measures of information of m, m_X and m_Y.

Since $|S| \geq \max(|proj(S;X)|,|proj(S;Y)|)$, it can be easily checked that

$$HI(m) \geq \max(HI(m_X),HI(m_Y)) \quad (19)$$

$$Sp(m) \leq \min(Sp(m_X),Sp(m_Y)) \quad (20)$$

For establishing inequalities which are similar to (19) for HC and HD, we need the following results.

$$\forall S \subseteq XxY, \ Cr(S) \leq \min(Cr_X(proj(S;X)),Cr_Y(proj(S;Y))) \quad (21)$$

$$\forall S \subseteq XxY, \ Pl(S) \leq \min(Pl_X(proj(S;X)),Pl_Y(proj(S;Y))) \quad (22)$$

where Cr, Cr_X and Cr_Y are respectively defined from m and its projections m_X and m_Y.

Proof :

$$Cr_X(proj(S;X)) = \sum_{A \subseteq proj(S;X)} m_X(A)$$

$$= \sum_{A \subseteq proj(S;X)} \ \sum_{A=proj(T;X)} m(T)$$

$$= \sum_{proj(T;X) \subseteq proj(S;X)} m(T) \geq \sum_{T \subseteq S} m(T) = Cr(S)$$

Similarly it can be shown that

$$Pl_X(proj(S;X)) \geq Pl(S)$$

Q.E.D.

Then it can be checked that

$$HC(m) \geq \max(HC(m_X),HC(m_Y)) \quad (23)$$

$$HD(m) \geq \max(HD(m_X),HD(m_Y)) \quad (24)$$

where m_X and m_Y are the projections of m.

Proof :

$$HC(m_X) = - \sum_{A \subseteq X} m_X(A).\ln[Cr_X(A)]$$

$$= - \sum_{A \subseteq X} \Big(\sum_{A=proj(S;X)} m(S)\Big).\ln[Cr_X(A)]$$

$$= - \sum_{A \subseteq X} \Big(\sum_{A=proj(S;X)} m(S).\ln[Cr_X(proj(S;X))]\Big)$$

$$\leq - \sum_{S \subseteq X \times Y} m(S).\ln[Cr(S)] = HC(m)$$

A similar proof holds for HD. **Q.E.D.**

8 - Generalized inclusion in the framework of Shafer's evidence theory

Several extensions of the idea of set-inclusion can be contemplated in the framework of Shafer's evidence theory. See Dubois, Prade [7] for a complete discussion. In this paper, we shall only consider the following strong extension of the idea of inclusion. Given two basic probability assignments m and m' on X, by definition

$m \subseteq m' \Leftrightarrow$

i) $\forall A \in F(m), \exists B \in F(m'), B \supseteq A$

ii) $\forall B \in F(m'), \exists A \in F(m), A \subseteq B$

iii) $\exists w : 2^X \times 2^X \to [0,1]$, with $w(A,B) = 0 \ \forall A \notin F(m)$, or $B \notin F(m')$

such that $\forall A \subseteq X$, $m(A) = \sum_{B: A \subseteq B} w(A,B)$

$$m'(B) = \sum_{A: A \subseteq B} w(A,B)$$

where $F(m)$ and $F(m')$ are the sets of focal elements of m and m' respectively. This definition turns to be equivalent to the one proposed by Yager in [14] ; see [7]. When m and m' reduce to possibility distributions, this generalized inclusion turns to be the usual fuzzy set inclusion (pointwisely defined as an inequality between the membership functions).

9 - Monotonicity properties

We have the following results

$m \subseteq m' \Rightarrow HI(m) \leq HI(m')$ (25)

$m \subseteq m' \Rightarrow Sp(m) \geq Sp(m')$ (26)

$m \subseteq m' \Rightarrow HD(m) \geq HD(m')$ (27)

Proof :

First, let us establish the following Lemma_2, where f is a positive set-function such that $A \subseteq B \Rightarrow f(A) \leq f(B)$,

$$m \subseteq m' \Rightarrow \sum_{A \subseteq X} m(A).f(A) \leq \sum_{A \subseteq X} m'(A).f(A).$$

Indeed

$$\sum_{A \subseteq X} m(A).f(A) = \sum_{A \subseteq X} \left(\sum_{B: A \subseteq B} w(A,B) \right).f(A)$$

$$\leq \sum_{A \subseteq X} \sum_{B: A \subseteq B} w(A,B).f(B)$$

$$= \sum_{B \subseteq X} \sum_{A: A \subseteq B} w(A,B).f(B) = \sum_{B \subseteq X} m'(B).f(B)$$

Q.E.D.

Clearly if $A \subseteq B$ then $|A| \leq |B|$ and $1/|A| \geq 1/|B|$. Therefore (25) holds and since we also have, for a function g such that $A \subseteq B \Rightarrow g(A) \geq g(B)$,

$m \subseteq m' \Rightarrow \sum_{A \subseteq X} m(A).g(A) \geq \sum_{A \subseteq X} m'(B).g(B)$, (26) holds to.

Besides in [7] it is shown that

$m \subseteq m' \Rightarrow \forall A \subseteq X$, $Cr(A) \geq Cr'(A)$ and $Pl(A) \leq Pl'(A)$ where Cr' and Pl' are defined from m'. Then we have

if $m \subseteq m'$, then $-\ln[Pl(A)] \geq -\ln[Pl'(A)]$

Therefore if $A \subseteq B$, then $-\ln[Pl(A)] \geq -\ln[Pl'(B)]$. Finally using a proof similar to the one of Lemma 2, we obtain $HD(m) \geq HD(m')$. **Q.E.D.**

The monotonicity properties (25)-(26) are satisfying, since the idea of imprecision and of specificity are very related to inclusion. It generalizes properties which were already known in case of possibility distributions. (27) is also intuitively satisfying, since the disjointness of focal elements cannot decrease when the basic probability assignment becomes smaller in the sense of the generalized inclusion.

When $m \subseteq m'$, there is no inequality between HC(m) and HC(m') as shown by the two following examples pictured on Fig. 3.a and Fig. 3.b :

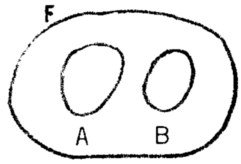

Figure 3.a

a) m'(F) = 1

m(A)+m(B) = 1 ; $A \subseteq F$; $B \subseteq F$.

Clearly $m \subseteq m'$ (since w(A,F) = m(A) and w(B,F) = m(B)) ;

HC(m') = 0

and HC(m) = -[m(A).ln(m(A)) + m(B)ln(m(B))] > 0

b) m'(F) + m'(G) = 1

m(A) = 1, $A \subseteq F \cap G$

Clearly $m \subseteq m'$ (since w(A,F) = m'(F) and w(A,G) = m'(G)),

HC(m) = 0

and HC(m') = -[m(F).ln(m(F))+m(G)ln(m(G))] > 0 .

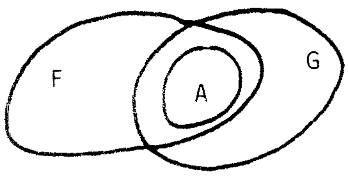

Figure 3.b

The non-monotonicity of HC with respect to inclusion is not very surprising since the amount of confusion between focal elements estimated by HC, is not clearly related to the extended inclusion we introduce between basic probability assignments.

10 - Concluding remarks

In this paper we investigate the additivity and monotonicity properties of different measures of information which have been recently introduced in the framework of Shafer's evidence theory. The existence of these properties shows that it is possible to extend information theory in a nice way beyond its probabilistic setting. Moreover, it is noticeable that different measures of information are necessary in order to characterize the information which can be represented in Shafer's evidence theory.

References

[1] Dempster, A.P. (1967) Upper and lower probabilities induced by a multivalued mapping. Ann. Math. Statistics, 38, 325-339.

[2] Dubois, D., Prade, H. (1980) Fuzzy Sets & Systems : Theory and Applications. Academic Press, New York.

[3] Dubois, D., Prade, H. (1982) On several representations of an uncertain body of evidence. In : Fuzzy Information and Decision Processes (M.M. Gupta, E. Sanchez, Eds.), North-Holland, Amsterdam, 167-181.

[4] Dubois, D., Prade, H. (1985) A note on measures of specificity for fuzzy sets. Int. J. General Systems, 10, 279-283. Preprint in BUSEFAL n° 19 (L.S.I., Univ. P. Sabatier, Toulouse), 1984, 83-89.

[5] Dubois, D., Prade, H. (1985) Théorie des Possibilités. Applications à la Représentation des Connaissances en Informatique. Masson, Paris.

[6] Dubois, D., Prade, H. (1985) Mesures d'information pour les distributions de possibilité et les fonctions de croyance. Actes Journées "Analyse de problèmes décisionnels dans un environnement incertain et imprécis", Reims, France, July 11-13, 1985. Published by "GR 22, CNRS, Univ. P. et M. Curie, Paris", pp. 67-80.

[7] Dubois, D., Prade, H. (1985) A set-theoretic view of belief functions : Logical operations and approximations by fuzzy sets. To appear in Int. J. General Systems.

[8] Higashi, M., Klir, G.J. (1983) Measures of uncertainty and information based on possibility distributions. Int. J. General Systems, 9, 43-58.

[9] Höhle, U. (1982) Entropy with respect to plausibility measures. Proc. 12th IEEE Int. Symp. Multiple-Valued Logic, Paris, 167-169.

[10] Shafer, G. (1976) A Mathematical Theory of Evidence. Princeton University Press.

[11] Shafer, G. (1985) Belief functions and possibility measures. In : The Analysis of Fuzzy Information (J.C. Bezdek, Ed.), CRC Press, Boca Raton, Fl., Vol. 1, to appear.

[12] Yager, R.R. (1982) Measuring tranquility and anxiety in decision-making : an application of fuzzy sets. Int. J. General Systems, 8, 139-146.

[13] Yager, R.R. (1983) Entropy and specificity in a mathematical theory of evidence. Int. J. of General Systems, 9, 249-260.

[14] Yager, R.R. (1985) The entailment principle for Dempster-Shafer granules. Tech. Rep. MII-512, Iona College, New Rochelle, New York (22 p.).

[15] Zadeh, L.A. (1978) Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets & Systems, 1, (1), 3-28.