Fuzzy Information and Decision Processes
M.M. Gupta and E. Sanchez (eds.)

ON SEVERAL REPRESENTATIONS OF AN UNCERTAIN BODY OF EVIDENCE

Didier DUBOIS
Département d'Automatique
Centre d'Etudes et de Recherches de Toulouse
2 Avenue Edouard Belin
31055 TOULOUSE CEDEX
FRANCE

Henri PRADE
Langages et Systèmes Informatiques
Université Paul Sabatier
118 route de Narbonne
31062 TOULOUSE CEDEX
FRANCE

ABSTRACT

This paper deals with various kinds of Sugeno's fuzzy measures which, with the exception of probability measures, have been introduced recently in the literature : g_λ –fuzzy measures, Shafer's belief and plausibility functions, Zadeh's possibility measures, necessity measures among others. First are recalled the existing axiomatic theories dealing with large classes of fuzzy measures : Shafer's belief theory and the triangular norm-based approach. From both approaches emerge three remarkable families of fuzzy measures : the probability, possibility and necessity measures. Shafer's belief and plausibility functions can be represented via a so-called basic probability assignment (which is nothing but a random set) ; triangular norm-based fuzzy measures can be expressed in terms of a density ; the relations existing between these two representations, basic assignment and density, are investigated for the various families of introduced fuzzy measures. The remainder of the paper is more particularly devoted to the study of the consistency relationship existing between possibilistic and probabilistic information . A distinction is made between physical and epistemic possibility. It is shown how to derive an epistemic possibility distribution from statistical evidence (i.e. a given histogram) in a rather natural way. The results are in agreement with the loose connections which exist, according to common sense, between the probable, the possible and the credible.

KEY WORDS
probability ; possibility ; necessity ; belief ; plausibility ; random set ; fuzzy set ; triangular norm ; epistemic possibility ; Dempster's rule.

1 - INTRODUCTION

Probability theory was the only well-established existing approach to uncertainty until recently. The appearance of alternative approaches such as Shafer's belief theory or Zadeh's possibility theory has introduced new points of view on uncertainty. Considering an event, we may try to evaluate its probability, its feasibility, its possibility of occurrence or how much it seems credible for instance. All those evaluations, performed in the framework of the above-mentioned theories, are based on some sets of numbers, generally normalized in some sense, which model our state of knowledge and which have to be combined in accordance with the characteristic axioms of those theories. As far as those models are actually in agreement with what is usually meant by "probable", "possible" or "credible", the various evaluations as well as the basic sets of numbers which are used cannot be completely unrelated, even if strictly speaking they represent independent informations. Indeed according to common sense there exists some loose connections between the probable, the possible and the credible. This paper is an attempt to explore these connections.

First, we recall the existing axiomatic theories dealing with large classes of Sugeno'fuzzy measures, namely Shafer's belief theory and the triangular norm-based approach recently proposed by Dubois and Prade. From both approaches emerge three remarkable families of uncertainty measures : the probability, possibility and necessity measures. Then, we focus our attention on the relations existing between the basic probability assignment, on which are based Shafer's belief and plausibility functions, and densities which enable the convenient expression of any uncertainty measures issued from the triangular norm-based approach. The following section is devoted to the study of the consistency relations existing between possibilistic and probabilistic information. In this discussion, a distinction is made between physical possibility which concerns the feasibility of events and epistemic possibility which is relative to their occurrence. It is shown how to build an (epistemic) possibility distribution from a probability density in a rather natural way. Concluding remarks emphasize the results which have been obtained, their interest and stress open questions.

2 - TWO AXIOMATIC APPROACHES TO THE MODELLING OF UNCERTAINTY

In 1972, Sugeno [20] introduced the concept of a fuzzy measure in order to depart from the too rigid framework of probability theory. A fuzzy measure is a set function whose characteristic property is only monotonicity rather than additivity. More precisely, assuming that the universe X is finite for sake of simplicity (in this paper, X will be always supposed to be finite except explicit statement of the contrary), a fuzzy measure is a set function g from an algebra $\mathcal{A}$ (e.g. the set $\mathcal{P}(X)$ of subsets of X) defined on X to the real interval [0,1], such that

$$\begin{cases} \text{i) } g(\emptyset) = 0 \ ; \quad \text{ii) } g(X) = 1 \\ \text{iii) } \forall A \in \mathcal{A}, \forall B \in \mathcal{A}, \text{ if } A \subset B, \text{ then } g(A) \leqslant g(B) \end{cases} \quad (1)$$

The following inequalities hold

$$\forall A \in \mathcal{A}, \forall B \in \mathcal{A}, g(A \cap B) \leqslant \min(g(A), g(B)) \quad (2)$$

$$\forall A \in \mathcal{A}, \forall B \in \mathcal{A}, g(A \cup B) \geqslant \max(g(A), g(B)) \quad (3)$$

The characteristic axioms and properties of two worth-considering classes of fuzzy measures are now recalled.

2.1 - Triangular norm-based approach to fuzzy measures.

Only basic results are given here ; for further details the reader is referred to Prade [15], Dubois [4] and more particularly to Dubois and Prade [8].

Axioms (1) are very general and may be particularized in various ways in order to obtain noticeable classes of fuzzy measures. The following axiom seems rather natural :

$$\forall A \in \mathcal{A}, \forall B \in \mathcal{A}, \text{ if } A \cap B = \emptyset, \text{ then } g(A \cup B) = g(A) * g(B) \quad (4)$$

where * is some operator under which [0,1] is closed. (4) expresses that the grade of uncertainty of disjoint events A and B only depends upon the grade of uncertainty of A and the grade of uncertainty of B.

The algebraic structure of $\mathcal{A}$ induces compatibility constraints on * which lead to choose * among the triangular conorms. A triangular conorm (see Schweizer and Sklar [17])is a two place real-valued function whose domain is the unit square [0,1] x [0,1], and which satisfies the following conditions

$$\begin{cases} \text{i) } 1 * 1 = 1 \ ; \ 0 * a = a * 0 = a \text{ (boundary conditions)} \\ \text{ii) } a * b \leqslant c * d \text{ whenever } a \leqslant c \text{ and } b \leqslant d \text{ (monotonicity)} \\ \text{iii) } a * b = b * a \text{ (symmetry)} \\ \text{iv) } a * (b * c) = (a * b) * c \text{ (associativity)} \end{cases} \quad (5)$$

The main triangular conorms are the maximum operator (max (a,b)), the probabilistic sum (a+b-a.b), the bounded sum, min (1,a+b), and the so-called T_W^* operator (defined by the boundary conditions and by $T_W^*(a,b) = 1$, $\forall (a,b) \in (0,1]^2$). The following inequalities hold between them

$$\forall (a,b) \in [0,1]^2, \quad \max(a,b) \leqslant a+b-a.b \leqslant \min(1,a+b) \leqslant T^*_W(a,b) \quad (6)$$

Moreover, every conorm * is such that

$$\forall (a,b) \in [0,1]^2, \ \max(a,b) \leqslant a*b \leqslant T^*_W(a,b) \quad (7)$$

An interesting consequence of (4) is

$$\forall A \in \mathcal{A}, \ g(A) * g(\bar{A}) = 1 \quad (8)$$

A conorm-based set function g (i.e. satisfying (4) is uniquely defined by the knowledge of the conorm * and the values of g over the set of singletons of X. Indeed, let $X = \{x_1, \ldots, x_n\}$, $A = \{x_{i_1}, \ldots, x_{i_p}\}$ and $g_i = g(\{x_i\})$, then

$$g(A) = g\Big(\bigcup_{j=1,p} \{x_{i_j}\}\Big) = g_{i_1} * \ldots * g_{i_p} \quad (9)$$

The normalization condition g(X)=1 entails

$$g_1 * \ldots * g_n = 1 \quad (10)$$

The set of numbers $\{g_i\}_{i=1,n}$ will be called the "density" attached to the set function g. Obviously, (9) is very interesting from a computational point of view.

Choosing the bounded sum for *, (4) yields

$$\text{if } A \cap B = \emptyset, \text{ then } g(A \cup B) = \min(1, g(A) + g(B)) \quad (11)$$

(9) and (10) give respectively

$$g(A) = \min\Big(1, \sum_{x \in A} g(\{x\})\Big) \quad (12)$$

and

$$\sum_{i=1}^{n} g_i \geqslant 1 \qquad (13)$$

letting this latter expression be an equality, we clearly recover probability measures.

The probabilistic sum does not lead to a well-known family of set functions. However, in this case (8) is equivalent to

$$\forall A \in \mathcal{A},\ \max(g(A), g(\overline{A})) = 1 \qquad (14)$$

Similarly, letting $a * b = \max(a,b)$, (4) yields

$$\forall A \in \mathcal{A},\ \forall B \in \mathcal{A},\ g(A \cup B) = \max(g(A), g(B)) \qquad (15)$$

(It can be shown that (15) is effectively equivalent to (4) with $* = \max$, i.e. the condition $A \cap B = \emptyset$ is not required in this case). We recognize Zadeh's possibility measures [23]. Possibility measures satisfy (14) also and (10) reads

$$\max_{i=1,n} g_i = 1 \qquad (16)$$

and (9) yields, using the notation Π for possibility measures

$$\forall A \in \mathcal{A},\ \Pi(A) = \max_{x \in A} \Pi(\{x\}) \qquad (17)$$

The density is usually called possibility distribution and (17) is trivially extended when $X = \mathbb{R}$ by

$$\Pi(A) = \sup_{x \in A} \pi(x) \qquad (18)$$

where the possibility distribution π is a mapping from $\mathbb{R}$ to $[0,1]$ such that $\sup_{x \in \mathbb{R}} \pi(x) = 1$.

Note that the way of estimating the possibility of an event, which is only based on the most favorable case (it strongly departs from probability calculus where all the favorable cases are cumulated) is in agreement with the idea of possibility which means feasibility or "happenability". Besides, (14) expresses that at least one of two contradictory well-defined events must be possible, which is rather natural.

Sugeno's g_λ -fuzzy measures [20] whose characteristic axiom is

$\forall A \in \mathcal{A}, \forall B \in \mathcal{A}$, if $A \cap B = \emptyset$, then

$$g_\lambda(A \cup B) = g_\lambda(A) + g_\lambda(B) + \lambda . g_\lambda(A) . g_\lambda(B) \qquad (19)$$

with $\lambda > -1$, correspond to the parametered conorm

$$a*b = \min(1, a+b+\lambda a.b) = \min(1, \varphi^{-1}(\varphi(a) + \varphi(b))) \qquad (20)$$

where $\varphi(t) = \operatorname{Ln}(1+\lambda t)$

provided that the normalization condition

$$\prod_{i=1,n} (1+\lambda g_i) = 1 + \lambda \qquad (21)$$

holds. As pointed out by Wierzchoń [21], $\hat{\varphi}_o g_\lambda$ is nothing but a probability measure, with $\hat{\varphi} = \frac{1}{\varphi(1)} . \varphi$.

If g is a fuzzy measure, the set function g' defined by

$$\forall A \in \mathcal{A},\ g'(A) = 1 - g(\overline{A}) \qquad (22)$$

is also a fuzzy measure ; (22) expresses the duality existing between g and g'. If g is a conorm-based set function satisfying (4), then the characteristic axiom of the dual measure g' is

$\forall A \in \mathcal{A}$, $\forall B \in \mathcal{A}$, if $A \cup B = X$,

then $g'(A \cap B) = g'(A) \perp g'(B) \qquad (23)$

where $\perp$ denotes the triangular norm dual of $*$, defined by

$a \perp b = 1 - (1-a) * (1-b)$.

A triangular norm satisfies (5) where the boundary conditions are replaced by

$$0 \perp 0 = 0 \ ;\ 1 \perp a = a \perp 1 = a \qquad (24)$$

The triangular norms dual of the following conorms : maximum, probabilistic sum, bounded sum, T_W^* are respectively the minimum operator $(a \perp b = \min(a,b))$, the product, the so called T_m operator $(T_m(a,b) = \max(0, a+b-1))$ and T_W (defined by (24) and $T_W(a,b) = 0$, $\forall (a,b) \in [0,1)^2$). We have

$\forall (a,b) \in [0,1]^2$,

$$T_W(a,b) \leqslant \max(0, a+b-1) \leqslant a.b \leqslant \min(a,b) \qquad (25)$$

and for any triangular norm $\perp$:

$$\forall (a,b) \in [0,1]^2,\ T_W(a,b) \leqslant a \perp b \leqslant \min(a,b) \qquad (26)$$

Properties of triangular norm-based set functions (i.e. satisfying (23)) can be easily deduced by duality from those of conorm-based set functions. We have now

$$\cdot\ \forall A \in \mathcal{A},\ g'(A) \perp g'(\overline{A}) = 0 \qquad (27)$$

$$g'(A) = g(\bigcap_{x \notin A} (X-\{x\})) = g'_{i_{p+1}} \perp \ldots \perp g'_{i_n} \quad (28)$$

where $A = \{x_{i1}, \ldots, x_{ip}\}$ and $g'_i =$

$g'(X-\{x_i\}) = 1 - g_i$ if $\{g_i\}_{i=1,n}$

is the density attached to the conorm-based set function g, dual of g'.
$g'(\emptyset) = 0$ entails

$$g'_1 \perp \ldots \perp g'_n = 0 \quad (29)$$

Choosing $\perp = T_m$ yields

if $A \cup B = X$, then
$$g'(A \cap B) = \max(0, g'(A) + g'(B) - 1) \quad (30)$$

the normalization condition is

$$\sum_{i=1}^{n} g'_i \leq n-1 \quad (31)$$

while (27) reads

$$g'(A) + g'(\bar{A}) \leq 1 \quad (32)$$

Probability measures obviously satisfy (30) and letting (31) be an equality yields their normalization condition.

Probability measures are pointwisely their own dual in the transformation (22).

Choosing $\perp$ = product does not yield something very well-known, in spite of its simplicity. Then, (27) is equivalent to

$$\forall A \in \mathcal{A},\ \min(g'(A), g'(\bar{A})) = 0 \quad (33)$$

Choosing $\perp = \min$, (23) is equivalent to

$$\forall A \in \mathcal{A},\ \forall B \in \mathcal{A},\ g'(A \cap B) = \min(g'(A), g'(B)) \quad (34)$$

Fuzzy measures satisfying (34) are dual of possibility measures, and as suggested by Dubois and Prade [7] can be named "necessity measures" since the necessity g'(A), denoted in the following by N(A), of an event A is the grade of impossibility of the opposite event, $1-\Pi(\bar{A})$ from (22) where Π is the dual of N. (29) reads

$$\min_{i=1,n} N(X-\{x_i\}) = 0 \quad (35)$$

and is equivalent to (16) since $N(X-\{x_i\}) = 1-\Pi(\{x_i\})$. Necessity measures also satisfy (33) ; (33) entails that at most one of two contradictory well-defined events may be somewhat necessary, which is rather natural. By duality, (17) yields

$$\forall A \in \mathcal{A},\ N(A) = \min_{x \notin A} N(X-\{x\}) \quad (36)$$

with $N(X-\{x\}) = 1 - \Pi(\{x\})$

and (18) corresponds to

$$N(A) = \inf_{x \notin A} (1 - \pi(x)) \quad (37)$$

Lastly, because of (22) and (33) we have the entailment

$$N(A) > 0 \Rightarrow N(\bar{A}) = 0 \Leftrightarrow \Pi(A) = 1 \quad (38)$$

which expresses that an event somewhat necessary must be completely possible. Then, we have

$$\forall A \in \mathcal{A},\ N(A) \leq \Pi(A) \quad (39)$$

The dual of a g_λ-fuzzy measure is a g_μ-fuzzy measure with $\mu = -\frac{\lambda}{1+\lambda}$ (Prade [14], Dubois Prade [7]). In other words, we have

$$\forall A \in \mathcal{A},\quad g_{-\frac{\lambda}{1+\lambda}}(A) = 1 - g_\lambda(\bar{A}) \quad (40)$$

2 . 2 - Shafer's theory of evidence.

Independently from the development of fuzzy set and possibility theory, Shafer [19] has proposed a theory of evidence where he has introduced the mathematical concept of belief function. See [12] and [22] for a discussion.

A belief function is a set function Bel from $\mathcal{P}(X)$ to [0,1] such that

$$\begin{cases} \text{i) } Bel(\emptyset) = 0 \ ;\ \text{ii) } Bel(X) = 1 \\ \text{iii) } \forall m \in \mathbb{N},\ \forall A_i \subset X,\ i=1,m, \\ Bel(\bigcup_{i=1}^{m} A_i) \geq \sum_{i=1}^{m} Bel(A_i) - \sum_{i<j} Bel(A_i \cap A_j) \\ + \ldots + (-1)^{m+1} Bel(\bigcap_{i=1}^{m} A_i) \end{cases} \quad (41)$$

(41) implies (1) and then a belief function is a particular case of Sugeno's fuzzy measures. For m = 2, (41) yields

$\forall A \subset X,\ \forall B \subset X,$

$$Bel(A \cup B) \geq Bel(A) + Bel(B) - Bel(A \cap B) \quad (42)$$

then, $\forall A \subset X,\ \forall B \subset X$

$$Bel(A \cap B) \geq \max(0, Bel(A) + Bel(B) - 1)$$

and $Bel(A \cap B) \leq \min(Bel(A), Bel(B))$ (43)

and

$$\forall A \subset X, \quad Bel(A) + Bel(\bar{A}) \leq 1 \qquad (44)$$

Shafer [19] has shown that any belief function is uniquely defined through the specification of a mapping m from $\mathcal{P}(X)$ to [0,1] called "basic probability assignment", satisfying

$$m(\emptyset) = 0 \ ; \quad \sum_{A \subset X} m(A) = 1 \qquad (45)$$

then, we have

$$\forall A \subset X, \quad Bel(A) = \sum_{B \subset A} m(B) \qquad (46)$$

Conversely, m is obtained from Bel by

$$\forall A \subset X, \quad m(A) = \sum_{B \subset A} (-1)^{|A-B|} Bel(B) \qquad (47)$$

where |A| denotes the cardinality of the set A.

A so-called commonality function Q can be defined [19] from m by

$$\forall A \subset X, \quad Q(A) = \sum_{\substack{B \subset X \\ A \subset B}} m(B) \qquad (48)$$

Note that $Q(\emptyset) = 1$. Bel can be expressed directly in terms of Q :

$$\forall A \subset X, \ Bel(A) = \sum_{B \subset \bar{A}} (-1)^{|B|} \ Q(B) \qquad (49)$$

Thus, Q provides another way of specifying a belief function.

A subset A of X such that $m(A) > 0$ is called a focal element of the belief function defined through m. Probability measures are particular cases of belief functions, their focal elements are singletons. "m(A) measures the belief that one commits exactly to A, not the total belief that one commits to A" (Shafer [19]). The set of focal elements may be thought as the set of the possible localizations of the truth, probabilistically weighted by m.

By assigning basic probabilities to subsets which are not singletons, the basic probabilities and hence the degrees of belief of smaller subsets (which correspond to more precise localizations of the eventual truth) are diminished because of the constraint (45). Note that (46) can be written

$$\forall A \subset X, \quad Bel(A) = \sum_{B \subset X} m(B). \ (\inf_{x \in B} \chi_A(x)) \qquad (50)$$

where χ_A is the characteristic function of the subset A. Thus, Bel(A) appears to be the expectation that the truth lies certainly in A, calculated from the set of probabilistic weights $\{m(B), B \subset X \text{ and } B \neq \emptyset\}$ attached to the possible localizations of the truth, since $\inf_{x \in B} \chi_A(x) = 1$ if $B \subset A$ and is zero otherwise.

If Bel is a belief function, the set function Pl defined by duality is called a plausibility function

$$\forall A \subset X, \quad Pl(A) = 1 - Bel(\bar{A}) \qquad (51)$$

Clearly, a plausibility function is a fuzzy measure in the sense of Sugeno.

Perhaps, credibility functions would be a better name for belief functions since "credible" sounds like "plausible". Moreover, in the evidential vocabulary, the more credible an event is, the less plausible is the opposite event and conversely.

We have

. $\forall A \subset X, \forall B \subset X,$

$$Pl(A \cup B) \leq Pl(A) + Pl(B) - Pl(A \cap B) \qquad (52)$$

. $\forall A \subset X, \forall B \subset X, \max(Pl(A), Pl(B)) \leq Pl(A \cup B)$

and $Pl(A \cup B) \leq \min(1, Pl(A) + Pl(B))$ (53)

. $\forall A \subset X, \ Pl(A) + Pl(\bar{A}) \geq 1$ (54)

and then $\forall A \subset X, \ Pl(A) \geq Bel(A)$ (55)

Pl is also uniquely characterized by means of the basic probability assignment m :

$$\forall A \subset X, \quad Pl(A) = \sum_{B \,:\, B \cap A \neq \emptyset} m(B) \qquad (56)$$

Note that $Pl(\{x\}) = Q(\{x\})$; (56) can be written

$$\forall A \subset X, \ Pl(A) = \sum_{B \subset X} m(B).(\sup_{x \in B} \chi_A(x)) \qquad (57)$$

Thus, Pl(A) appears to be the expectation that the truth lies possibly in A since $\sup_{x \in B} \chi_A(x) = 1$ if $A \cap B \neq \emptyset$ and is zero otherwise.

Shafer [19] has specially studied several particular cases of belief functions, among others the simple support functions and the consonant support functions.

A simple support function is defined by, given $Y \subset X$ ($Y \neq \emptyset$), $s \in [0,1]$:

$$\forall A \subset X,\ Bel(A) = \begin{cases} 0 \text{ if } A \not\supset Y \\ s \text{ if } A \supset Y \text{ and } A \neq X \\ 1 \text{ if } A = X \end{cases} \quad (58)$$

Its basic probability assignment is such that $m(Y) = s$; $m(\bar{Y}) = 1 - s$ and $m(A) = 0$ if $A \neq Y$ and $A \neq X$. The belief function defined by (58) corresponds to a situation where the possible localization of the truth is non-ambiguous and where the statement "the truth lies in Y" is corroborated to the degree s. When $s = 1$, our certainty that the truth lies in Y is total.

A consonant support function is a belief function whose focal elements are nested, i.e. its focal elements can be arranged in order so that each is contained in the following one : $A_1 \subset A_2 \subset \ldots \subset A_m$ and if $A \neq A_i$, then $m(A) = 0$. Simple support functions are particular cases of consonance. In case of consonance, all the possible localizations of the truth are not conflicting because they are nested. It can be easily shown that a consonant support function is nothing but a necessity measure and then satisfies (34) and conversely a necessity measure is always a consonant support function.

Note that consonant support functions were considered at length by the English economist Shackle [18] more than twenty years ago ; Shackle's theory was based on the "grade of potential surprize" attached to an event which exactly corresponds to the necessity of the opposite event. Then, possibility measures are particular cases of plausibility functions. Thus if Pl is attached to a basic probability assignment m whose focal elements are nested, we have

$$\forall A \subset X,\ A \neq \emptyset,\ Pl(A) = \max_{x \in A} \pi(x)$$

with $\pi(x) = Pl(\{x\})$. (59)

It can be shown [19] that the commonality function in case of consonance is given by

$$\forall A \subset X,\ A \neq \emptyset,\ Q(A) = \min_{x \in A} \pi(x) \quad (60)$$

Probability measures are the only belief functions which are also plausibility functions, but probability is not consonant ! Indeed the basic probability assignment underlying a probability measure completely departs from the one underlying a possibility measure : in the first case possible localizations of the truth are precise (they correspond to singletons) and are associated with frequencies, in the latter case, possible localizations of truth are more or less imprecise but consonant.

Banon [1] proved that for $\lambda \geq 0$, a g_λ -fuzzy measure is a belief function ; noticing that the mapping from $(-1,0]$ to $[0,+\infty)$ defined by $\lambda \longrightarrow \mu = \frac{-\lambda}{1+\lambda}$ is involutive and one-to-one and owing to formula (40), it can be easily deduced that g_λ - fuzzy measures for $-1 < \lambda \leq 0$ are plausibility functions.

Besides, it has been also shown (Prade [15], Dubois, Prade [8]) that a fuzzy measure which satisfies (23) for the triangular norm'product' is a belief function.

Thus, from both approaches --triangular norm based approach and belief function approach-- three remarkable families of fuzzy measures emerge : the probability, the possibility and the necessity measures.

3 - BASIC PROBABILITY ASSIGNMENT AND DENSITY

3.1.- Some examples

A belief or a plausibility function can be build from a basic probability assignment which is a set-function and thus contrasts with a density. Practically speaking, on a set X whose cardinality is n, we need $2^n - 2$ (taking into account (45)) values in order to define a belief or a plausibility measure from a basic probability assignment while n-1 (taking into account the normalization condition) values are sufficient to define a fuzzy measure which can be expressed in terms of a density.

Several examples of belief or plausibility functions, which are also fuzzy measures satisfying (4) or (23) and which can be defined in terms of a density consequently, have been given in the precedent section. Then, in such cases, the set of values $\{m(A), A \subset X\}$ can be estimated out of the set $\{g(\{x_i\}) = g_i, i=1,n\}$ only, and reciprocally owing to (47) and (51). Let us give some results

. for a probability measure

$$m(A) = \begin{cases} g_i & \text{if } A = \{x_i\} \\ 0 & \text{otherwise} \end{cases} \quad (61)$$

. for a g_λ -fuzzy measure (Banon [1]),

$$m(A) = \lambda^{|A|-1} \cdot \prod_{x_i \in A} g_i \quad (62)$$

where $\prod$ denotes the product.

. for a fuzzy measure satisfying (23) with the norm product (Dubois, Prade [8]).

$$m(A) = \prod_{x_i \in A} g_i \cdot \prod_{x_i \in \bar{A}} (1-g_i) \quad (63)$$

Let us examine the case of possibility and necessity measures more particularly. Their focal elements are nested ; then, we can arrange a nested sequence

$A_1 = \{x_1\}$, $A_2 = \{x_1, x_2\}$,...,

$A_i = \{x_1, x_2,\ldots,x_i\}$, ..., $A_n = X$

such that $\forall A \neq A_i$, $m(A) = 0$.

However, it is possible that for some i, $m(A_i) = 0$. With the notation $m_i = m(A_i)$, i=1,n and $\pi_i = \pi(x_i) = \Pi(\{x_i\})$ where Π is the possibility measure defined from m; applying (56) , we get

$\pi_1 = m_1 + m_2 + \ldots + m_n = 1$ (from (45) since $m(\emptyset)=0$)

$\pi_2 = m_2 + \ldots + m_n$

....

$\pi_n = m_n$

or more compactly

$$\forall\, i = 1,n \qquad \pi_i = \sum_{j=i}^{n} m_j \quad (64)$$

and conversely

$$\forall i = 1,n \quad , \quad m_i = \pi_i - \pi_{i+1} \quad (65)$$

with $\pi_{n+1} = 0$ and $\pi_1 = 1$.

Taking into account $m(\emptyset) = 0$, the normalization condition $\max_i \pi_i = 1$ is equivalent to this of m (i.e.(45)). For a simple support function defined by (58) and expressed in terms of π_i via (36), we have

$$\pi_i = \begin{cases} 1 & \text{if } x_i \in Y \\ 1-s & \text{if } x_i \notin Y \end{cases} \quad (66)$$

A complete investigation of the links between the triangular norm approach and Shafer's theory is still to be carried out.

3.2. - About Dempter's rule and the condition $m(\emptyset) = 0$ [16]

A possibility distribution may be generally viewed as a fuzzy set (Zadeh [23]) ; as noticed by Goodman [9], [10] and by Nguyen [13], a basic probability assignment defines a random set. Goodman [9], [10] proves that a fuzzy set is equivalent to a class of random sets ; when X is finite, this equivalence corresponds to a linear system of n+1 linear equations, where μ is the membership function of the fuzzy set, ($\mu(x_i)$ is the plausibility of $\{x_i\}$, see (56)) :

$$\begin{cases} \mu(x_i) = \sum_{A \subset X,\ A \supset \{x_i\}} m(A) & i=1,n \\ 1 = \sum_{A \subset X} m(A) \end{cases} \quad (67)$$

It is a linear system with coefficients 0 or 1, in 2^n unknowns ; here, $m(\emptyset)$ is not apriori supposed to be zero. In case of consonance, this linear system (67) reduces to (64) with $\pi_i = \mu(x_i)$, and then m can be uniquely determined from μ .

As noticed by Goodman [10], Dempster's rule of combination of evidence [2], [19], corresponds to the intersection of two statistically independent random sets of the same base space. Given two basic probability assignments m_1 and m_2, Dempster's rule enables to build a basic probability assignment m corresponding to their intersection [19], defined by $m(\emptyset) = 0$ and

$$m(A) = \frac{\sum_{i,j:\ A_i \cap B_j = A} m_1(A_i).m_2(B_j)}{1 - \sum_{i,j:\ A_i \cap B_j = \emptyset} m_1(A_i).m_2(B_j)} \quad (68)$$

It is easy to check that $\sum_{A \subset X} m(A)=1$.
Thus, Dempster's rule corresponds to the intersection of two satistically independent random sets of the same base space, after renormalization.

As pointed out by Dempster [2] and Shafer [19], the combination is more easily expressed in terms of commonality functions since we have

$$\forall A \subset X,\ Q(A) = \frac{Q_1(A).\ Q_2(A)}{K} \quad (69)$$

where K denotes the denominator of fraction (68).

Zadeh [24] has questioned the validity of the normalization of (68). It is worth noticing that the problem of normalization vanishes as soon as we admit that a basic probability assigned to the empty set ∅ may be non zero and we define the "intersection" of m_1 and m_2 by

$$m^*(A) = \sum_{\substack{i,j \\ A_i \cap B_j = A}} m_1(A_i).m_2(B_j) \quad (70)$$

It is easy to check we have $\sum_{A \subset X} m^*(A) = 1$, even if $m_1(\emptyset) \neq 0$ or $m_2(\emptyset) \neq 0$.
In the remainder of this section 3.2, the normalization condition of a basic probability assignment m will be only given by $\sum_{A \subset X} m(A) = 1$ and the requirement $m(\emptyset) = 0$ is dropped. Keeping definition (48) unchanged, (69) becomes

$$\forall A \subset X,\ Q(A) = Q_1(A).\ Q_2(A) \quad (71)$$

Taking into account (50), Bel(A) appears to be the expectation that the truth lies certainly in A, computed from the set of probabilistic weights $\{m(B),\ B \subset X$ and $B \neq \emptyset\}$ attached to the possible localizations of the truth.

According to this interpretation in terms of expectation, when $m(\emptyset)$ may be non-zero, (46) must be modified into

$$\forall A \subset X,\ Bel(A) = \sum_{\substack{B \subset A \\ B \neq \emptyset}} m(B) \quad (72)$$

(56) remains unchanged, but (51) becomes

$$\forall A \subset X,\ Pl(A) + Bel(\bar{A}) = 1 - m(\emptyset) \quad (73)$$

However, we keep $\forall A \subset X$, $Pl(A) \geqslant Bel(A)$. As soon as $m(\emptyset) \neq 0$, Bel and Pl are no longer normalized :

$$Bel(X) < 1 \ ; \ Pl(X) < 1 \quad (74)$$

We still have $Bel(\emptyset)$ and $Pl(\emptyset)$ equal to zero.
If the empty subset of X has a non-zero probabilistic weight, it means that a possible localization of the truth is outside of X, then it may seem natural that the credibility and plausibility that the truth lies in X are strictly less than 1 ; then, (73) means that 'A non plausible' does not entail that $\bar{A}$ is completely credible.

Let us consider the particular case of consonance . Due to (64), $m(\emptyset) = 0$ if and only if π is normalized (since $\pi(x_1)$ $= \max_{x \in X} \pi(x) = 1 - m(\emptyset)$. $m(\emptyset)$ is the complement to 1 of the height of the fuzzy set whose membership function is the possibility distribution π .

From (73) and (59), then we get

$$\forall A \subset X,\ A \neq \emptyset,\ Bel(A) = \min_{x \in \bar{A}} [(\max_{x \in X} \pi(x)) - \pi(x)] \quad (75)$$

Let m_1 and m_2 be two consonant basic probability assignments whose focal elements are respectively the increasing nested sequences $A_1 \subset \ldots \subset A_m$ and $B_1 \subset \ldots \subset B_\ell$ (if $m_1(\emptyset) \neq 0$, we add $A_0 = \emptyset$; if $m_2(\emptyset) \neq 0$, we add $B_0 = \emptyset$). The combination of m_1 and m_2 by Dempster's rule does usually yield a basic probability assignment m which is not consonant since m is non-zero only on the subsets $A_i \cap B_j$, $i=1,m$, $j=1,\ell$ which are not nested generally.

However, we have from (71)

$$\forall x \in X,\ Q(\{x\}) = Q_1(\{x\}).\ Q_2(\{x\}) \quad (76)$$

i.e. since $Q_i(\{x\}) = Pl_i(\{x\})$ $= \pi_i(x)$, $i=1,2$, (76) expresses that $x \longmapsto Q(\{x\})$ is the membership function of a fuzzy set which is the intersection (performed using product) of the fuzzy sets whose membership functions are respectively π_1 and π_2.

Consonance is lost and moreover $\{Q(\{x\}),\ x \in X\}$ does not represent all the information. Indeed the number of unknowns is generally greater than the number of equations in the system

(67) :

$$\begin{cases} Q(\{x\}) = \sum_{\substack{A \subset X \\ A \supset \{x\}}} m(A) \\ 1 = \sum_{A \subset X} m(A). \end{cases}$$

If m_1 and m_2 are two simple support functions (a particular case of consonance) in the sense of (58) with parameter s_1 and s_2 and reference set Y_1 and Y_2 respectively, m^* (which is equal to m if $Y_1 \cap Y_2 \neq \emptyset$) is defined by

$$\begin{aligned} m^*(Y_1 \cap Y_2) &= s_1 . s_2 \\ m^*(Y_1) &= s_1(1-s_2) \\ m^*(Y_2) &= s_2(1-s_1) \\ m^*(X) &= (1-s_1)(1-s_2) \\ m^*(A) &= 0 \text{ if } A \neq Y_1 \cap Y_2, Y_1, Y_2, X \end{aligned} \tag{77}$$

when $Y_1 \neq Y_2$, $Y_2 \not\subset Y_1$, $Y_1 \not\subset Y_2$.

Note that this result is no longer consonant since $Y_1 \cap Y_2$, Y_1, Y_2, X are nested if and only if Y_1 and Y_2 are nested which is not assumed in (77).

However m^* can still be equivalently represented by means of a fuzzy set whose membership μ is defined by (67) because in this case, the linear system (67) has the same number of equations and of unknowns.

We have

$$\begin{aligned} \mu(x) &= 1 & &\text{if } x \in Y_1 \cap Y_2 \neq \emptyset \\ \mu(x) &= 1-s_2 & &\text{if } x \in Y_1 \cap \overline{Y}_2 \\ \mu(x) &= 1-s_1 & &\text{if } x \in \overline{Y}_1 \cap Y_2 \\ \mu(x) &= (1-s_1)(1-s_2) & &\text{if } x \in \overline{Y}_1 \cap \overline{Y}_2 \end{aligned} \tag{78}$$

which is consistent with the intersection of fuzzy sets defined by the product of membership grades (since π_1 and π_2, the possibility distributions which are attached respectively to m_1 and m_2 and which may be regarded as fuzzy sets are defined by (66)). μ is not normalized if $Y_1 \cap Y_2 = \emptyset$.

The two intended purposes of this section were

1) to question the normalization condition $m(\emptyset) = 0$ of a basic probability assignment m and to study its implications for Dempster's rule of combination especially.
2) to show in what way Dempster's rule of combination is related to fuzzy set intersection based on product. In that perspective a natural question is raised : what combination rule would be related to fuzzy set intersection based on min operation (for instance) ? Normalization of fuzzy sets and normalization of basic probality assignments appear to be clearly related.

3.3- Approximating a belief or a plausibility function

Sometimes, it may be interesting to approximate a belief function Bel or a plausibility function Pl, whose basic probability assignment m cannot be reduced to an equivalent density, by a fuzzy measure which can be expressed in terms of a density. It can be carried out in several ways (Prade [14]) :

. by means of a probability measure. By "equidistributing" the values of the basic probability assignment, it is easy to build the probability density

$$\forall x \in X,\ p(x) = \sum_{B \supset \{x\}} \frac{1}{|B|} . m(B) \tag{79}$$

The probability measure defined by $Prob(A) = \sum_{x \in A} p(x)$ satisfies the following inequalities :

$$\forall A \subset X,\ Bel(A) \leqslant Prob(A) \leqslant Pl(A) \tag{80}$$

which can be checked by noticing that

$$Prob(A) = \sum_{B \subset A} m(B) + \sum_{x \in A} \sum_{\substack{B \supset \{x\} \\ B \not\subset A}} \frac{1}{|B|} \cdot m(B)$$

$$= \sum_{B \subset X} \frac{|B \cap A|}{|B|} \cdot m(B) \tag{81}$$

$$\text{and } Pl(A) = Prob(A) + \sum_{B \cap A \neq \emptyset} \frac{|B \cap \overline{A}|}{|B|} . m(B) \tag{82}$$

(80) expresses in what way Prob approximates Bel and Pl. When m is a probability density (i.e. $m(A) = 0$ as soon as A is not a singleton), we have $\forall x \in X$, $p(x) = m(\{x\})$.

. by means of a possibility and a necessity measure
We have the inequality, $\forall A \subset X$,

$$Pl(A) = \sum_{B \subset X} (\sup_{x \in A} \chi_B(x).m(B)) \geqslant \sup_{x \in A} (\sum_{B \subset X} \chi_B(x).m(B)) = \sup_{x \in A} Pl(\{x\}) \quad (83)$$

The right part of (83) is normalized if and only if $\exists x \in X$, $Pl(\{x\}) = 1$, i.e.

$$\bigcap_{\substack{B \subset X \\ m(B) \neq 0}} B \neq \emptyset \quad (84)$$

$\{Pl(\{x\}), x \in X\}$ can be then considered as the possibility distribution of a possibility measure Pos satisfying $\forall A \subset X$, $Pl(A) \geqslant Pos(A)$; if the basic probability assignment m is consonant, then the plausibility measure and the possibility measure are equal. The condition (84) expresses that there exists some consistency among all the possible localizations of the truth, since they have a common non-empty part. Note that the basic probability assignment m underlying a probability measure does not satisfy (84) generally, since then m takes non-zero values only on singletons ; thus, a probability measure cannot be "approximated" by a possibility measure, using (83).

Dually, we have, taking into account

$$Bel(A) = \sum_{\bar{B} \supset \bar{A}} m(B) = \sum_{B \subset X} (\inf_{x \in \bar{A}} \chi_{\bar{B}}(x)).m(B) \quad (85)$$

The inequality

$$\forall A \subset X, Bel(A) \leqslant \inf_{x \in \bar{A}} (\sum_{B \subset X} \chi_{\bar{B}}(x).m(B))$$

$$= \inf_{x \in \bar{A}} (1 - Pl(\{x\})) \quad (86)$$

if (84) holds, in the right part of (86) we recognize the necessity Nes(A) of A where Nes is the necessity measure attached to the possibility distribution $\{Pl(\{x\}), x \in X\}$.

Then we have, if and only if (84) is satisfied :

$$\forall A \subset X, 0 \leqslant Bel(A) \leqslant Nes(A) \leqslant Pos(A) \leqslant Pl(A) \leqslant 1 \quad (87)$$

Thus, as soon as (84) holds, if $Bel(A) > 0$, then $Pl(A) = 1$. (cf. (38)).

The inequality (87) expresses the consistency of the possibility and the necessity measures with the plausibility and the belief functions.

Besides, it is easy to check that the probability measure which derives via (81) from the basic probability assignment of Bel and Pl does not satisfy, for all $A \subset X$, $Nes(A) \leqslant Prob(A) \leqslant Pos(A)$, where Pos and Nes are defined by (83) and (86). It shows that the two ways of approximating a belief or a plausibility function presented here correspond to different points of view.

4 - CONSISTENT PROBABILITY AND POSSIBILITY

4.1. Preliminary discussion

An event A may be considered from several points of view : what is its possibility, its probability or its necessity for instance ? Then, we are faced with the question of the relationship between the probable and the possible (or between the probable and the necessary) since intuitively, it seems such a link exists, even if it is vague.

Let us start with an example used by Zadeh [23] in his introductory paper on possibility theory. Consider the statement "Hans eats u eggs for breakfast" where u takes values in the set $\{1, 2, 3, 4, \ldots\}$. The following table, taken from [23], gives the possibility $Pos_u(\{x\})$ (corresponding here to the "degree of ease" with which Hans is able to eat x eggs for breakfast) and the probability $Prob_u(\{x\})$ of the same event, for various values x of the variable u.

x	1	2	3	4	5	6	7	8
$Pos_u(\{x\})$	1	1	1	1	0.8	0.6	0.4	0.2
$Prob_u(\{x\})$	0.1	0.8	0.1	0	0	0	0	0

According to Zadeh [23] the degree of consistency γ of the probability density $p_i = Prob_u(\{x_i\})$, i=1,n with the possibility distribution $\pi_i = Pos_u(\{x_i\})$, i=1, n is defined by

$$\gamma = \sum_i \pi_i . p_i \quad (88)$$

This consistency is all the best that γ is closer to 1. $\gamma = 1$ if and only if $\pi_i = 1$ as soon as $p_i > 0$. According to

Zadeh [23] "it is an approximate formalization of the heuristic observation that a lessening of the possibility of an event tends to lessen its probability - but not vice-versa".

In fact, we have to make a distinction between two different uses of the word "possible". The first one corresponds to "what can be done", the second one to "what may happen". This dichotomy is reminding of the difference between de re modality and de dicto modality (see Hacking [11]). Thus, we have to distinguish between, on the one hand "It is possible for Hans to eat four eggs for breakfast" and on the other hand "It is possible that Hans eats three eggs for breakfast". The first example deals with Hans' ability and the second one with what may actually happen. It must be clear that what may happen can be done - but not vice-versa. Thus, in our example, Hans is definitely able to eat four eggs for breakfast, although it never happened (at least if the probability p_4 is strictly zero). The values of possibilities which are given here correspond to what Hans is able to do, but not to what he may do ; these values are a complementary information to the frequency histogram of the number of eggs Hans ate breakfast which tells us about Hans' actual behavior. (88) estimates the consistency between Hans' behavior and Hans' ability.

Beside the possibility degrees representing Hans' ability, it may be interesting to estimate the possibility that a given event happens, for instance that Hans eats x eggs for breakfast ; that latter possibility must be less or equal to the degree of ease with which the event can happen, since what happens can be done. If we consider only {0-1} - valued possibilities, in our example, the only possible events which may actually happen taking only into account the available frequency information, are that Hans eats one, two or three eggs for breakfast, while from four eggs the possibility will be zero. The problem we are facing here is the "possibilistic" interpretation of histograms concurrently to the usual probabilistic interpretation. Such an interpretation would enable to answer the question "How many eggs does Hans eat for breakfast ?" which calls for a possibility distribution (which may be viewed as a fuzzy number) rather than for a precise number with its attached probability.

4.2. Interpretation of Histograms

Shackle [18] interpreted the possibility of an event as the absence of surprise felt when it occurs. An event which often occurs is not very surprising and then it seems very possible that it happens ; on the other hand, events which are not very possible do not often occur and are surprising. Then, by an inductive reasoning, we are conducted to suppose that if an event seldom occurs, it must be less possible than events which often occur. The "possibilistic" interpretation of histograms may be carried out in several ways, at first glance, at least.

Let $\{h(x) \in R^+ , x \in X\}$ be an histogram. It is supposed that the events which are collected occurred in the same circumstances and that their amount is sufficient in order to equalize frequencies and probabilities.

Then, a quite simple way to deduce a possibility distribution is to normalize the height of the histogram :

$$\forall x \in X, \quad \pi_h(x) = \frac{h(x)}{\sup_{x \in X} h(x)} \quad ; \qquad (89)$$

thus, the most frequent event receives a degree of possibility equal to 1 (Dubois, Prade [5], [7]). By equalizing frequencies and probabilities, we get

$$\forall x \in X, \quad p_h(x) = \frac{h(x)}{\sum_{x \in X} h(x)} \qquad (90)$$

In the lack of any information other than the histogram, the possibility of occurrence of an elementary event is thus supposed to be directly linked with its number of occurrences. From (89) and (90), we get

$$\forall x \in X, \quad \pi_h(x) \geqslant p_h(x) \qquad (91)$$

This inequality agrees with the intuitive idea according to which the more probable an event is, the more possible we must consider its occurrence. However, (91) concerns elementary events only ; the extension of such an inequality to any events (i.e. to events which are not necessarily represented by singletons) depends on the shape of the histogram. Indeed, if Pos_h and $Prob_h$ respectively denote the possibility measure and the probability measure built from π_h and p_h, the following inequality does not hold for any subset A of X

$$Pos_h(A) = \sup_{x \in A} \pi_h(x) \geqslant Prob_h(A) = \sum_{x \in A} p_h(x) \qquad (92)$$

where h is any positive real-valued mapping.

If $X = \mathbb{R}^+$ and $p_h(x) = \frac{h(x)}{\int_X h(x)dx}$, then (92) holds for any $A \subset X$, for instance,

with $h(x) = \begin{cases} 1 \text{ if } x \in I \subset X \\ 0 \text{ otherwise} \end{cases}$

with $h(x) = \max(0, a(1-\frac{x}{b}))$ $(a > 0$ and $b > 0)$ or with $h(x) = a\, e^{-bx}$ $(a > 0,$ $b > 0)$, but not with $h(x) = \max(0, a-\sqrt{x})$ $(a > 0)$.

If we want to have the inequality (92) for any $A \subset X$, we have to look for possibility distributions which are consistent with the density (90) in the sense of (92). The following result (Dubois [3]) gives a sufficient condition in order that a possibility distribution π be consistent with a probability density p. Consider the subsets of X defined by

$$\forall \alpha \in [0, \sup p],\ C(\alpha) = \{x \in X,\ p(x) \leq \alpha\} \quad (93)$$

They are nested, $\forall \beta > \alpha$, $C(\beta) \supset C(\alpha)$.

Then, if $\forall x \in X$, $\pi(x) \geq \mathrm{Prob}(C(p(x)))$

$$= \sum_{t \in C(p(x))} p(t) \quad (94)$$

π is consistent with p.

Indeed, $\forall A \subset X$,

$$\mathrm{Pos}(A) = \sup_{x \in A} \pi(x) \geq \sup_{x \in A} \mathrm{Prob}(C(p(x)))$$

$$= \mathrm{Prob}(C(\alpha^*))$$

where $\alpha^* = \sup_{x \in A} p(x)$. We have $A \subset C(\alpha^*)$

and then $\forall A \subset X$, $\mathrm{Pos}(A) \geq \mathrm{Prob}(A)$.

Moreover, we check that $\sup_{x \in X} \mathrm{prob}(C(p(x)))$

$$= \mathrm{Prob}(C(\sup p)) = 1 \quad \text{Q.E.D.}$$

But the condition (94) is not necessary ; however, it becomes necessary if we demand π and p to have "similar" shapes, i.e. more precisely, $\forall x \in X$, $\forall x' \in X$,

$$\pi(x) \geq \pi(x') \iff p(x) \geq p(x') \quad (95)$$

Indeed, we have then

$\forall x \in X$, $C(p(x)) = \{x' \in X, p(x') \leq p(x)\}$

$= \{x' \in X,\ \pi(x') \leq \pi(x)\}$;

then, $\pi(x) = \mathrm{Pos}(C(p(x)))$ and because of (92) $\mathrm{Pos}(C(p(x))) \geq \mathrm{Prob}(C(p(x)))$, which yields (94).

Among the possibility distributions which satisfy (94), hence are consistent with a probability density p in the sense of (92), there exists one whose interpretation in terms of necessity is remarkable. It is the topic of the remainder of this section.

First, consider the example of a game of heads or tails with a biased coin. We have $X = \{x_1 = \text{heads},\ x_2 = \text{tails}\}$ and $1 \geq p_1 \geq \frac{1}{2} \geq p_2 = 1 - p_1 \geq 0$ where $p_i = \mathrm{Prob}(\{x_i\})$, $i=1,2$. If heads are the most frequent outcomes, we may say that there is some necessity to get heads or that there is some impossibility to get tails. An estimation, which may seem to be natural, of this necessity n_1 is the excess of probability $p_1 - p_2$ in favor of heads ; the possibility π_2 to get tails will be then equal to $1-n_1$ (the necessity of an event corresponds to the impossibility of the opposite event) ; lastly, we should have $\pi_1 = 1$ and $n_2 = 0$ where π_1 is the possibility corresponding to heads and n_2 the necessity corresponding to tails (since the excess of probability is not in favor of x_2 and in any cases, we must have (see (38)) $n_1 > 0 \Longrightarrow \pi_1 = 1$). If $p_1 = p_2 = \frac{1}{2}$, we get $n_1 = n_2 = 0$ and $\pi_1 = \pi_2 = 1$, i.e. if the coin is fair, heads or tails are equally possible and there is no necessity that one comes out rather than the other, which is natural since probabilities are "equidistributed". Necessity appears as soon as we are out of the domain of pure randomness. If $p_1=1$ and $p_2=0$, we have $n_1=1$, $n_2=0$ and $\pi_1=1$, $\pi_2=0$ which agree with the fact the only possible outcomes are heads.

Let us generalize this approach. The elements of X are supposed to be ranked in decreasing order with respect to their probability of occurrence : $p_1 \geq p_2 \geq \dots \geq p_n$ where $p_i = \mathrm{Prob}(\{x_i\})$.

Let $A_i = \{x_1, \dots x_i\}$. Let $N(A_i)$ denote the total excess of probability of the elements of A_i with respect to the element having the greatest probability outside of A_i (i.e. x_{i+1} if $i \leq n-1$) ;

we have

$$\begin{cases} N(A_i) \triangleq \sum_{j=1}^{i} (p_j - p_{i+1}) \leq \text{Prob}(A_i) \\ \qquad\qquad i=1,n-1 \\ N(A_n) = N(X) = 1 \end{cases} \quad (96)$$

More generally, $N(A) \triangleq \max_{\substack{i \\ A_i \subset A}} N(A_i) \leq \text{Prob}(A)$

(96bis)

Particularly, $N(A) = 0$ if $x_1 \notin A$ (96ter)

Let us check that the set function defined by (96) is a necessity measure. $N(A \cap B) = \max_{\substack{i \\ A_i \subset A \cap B}} N(A_i)$; because the A_i are nested, if $\max_{\substack{i \\ A_i \subset A}} N(A_i) \triangleq N(A_{i^*})$ and $\max_{\substack{i \\ A_i \subset B}} N(A_i) \triangleq N(A_{j^*})$ with for instance $i^* \geq j^*$, then $A_{i^*} \subset A \cap B$ and A_{i^*} is the greatest A_i contained in $A \cap B$; therefore $N(A \cap B) = \min(N(A), N(B))$; if $\nexists i^*$ or $\nexists j^*$, then $N(A) = 0$ or $N(B) = 0$ and then $N(A \cap B) = 0$.

Let us build the possibility distribution π which enables to express N via formula (37). We must have $\pi_i = \pi(x_i) = 1 - N(X - \{x_i\}) = 1 - N(A_{i-1})$. Then

$$\begin{cases} \pi_1 = 1 \\ \pi_{i+1} = 1 - \sum_{j=1}^{i} (p_j - p_{i+1}), \end{cases}$$

$i=1,n-1$

Taking into account the constraint $\sum_{i=1}^{n} p_i = 1$, we get

$$\begin{cases} \pi_1 = p_1 + p_2 + \ldots + p_n \\ \pi_2 = \qquad 2p_2 + \ldots + p_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \vdots \\ \pi_n = \qquad\qquad n.p_n \end{cases}$$

and more compactly

$$\pi_i = i.p_i + \sum_{j=i+1}^{n} p_j \qquad i=1,n$$

and reciprocally (97)

$$\forall i=1,n \quad , \quad p_i = \frac{1}{i}.\pi_i - \sum_{j=i+1}^{n} \frac{1}{j.(j-1)} \pi_j$$

(98)

The π_i's are normalized if and only if the p_i's are. It is easy to check that this possibility distribution satisfies (94). Anyway from (96) we have

$$\forall A \subset X \ , \quad \Pi(A) \geq \text{Prob}(A) \geq N(A) \qquad (99)$$

where Π is the possibility measure, dual of N, based on $\{\pi_i\}_{i=1,n}$. The inequalities (99) express the consistency between the possibility, the probability and the necessity measures : the more necessary (ineluctable) an event is, the more probable it must be and the more possible its occurrence must be.

Thus, from the intuitive idea of the concept of necessity, it has been possible in a natural way to associate a possibility distribution with a probability density ; then, this possibility distribution leads to possibility and necessity measures which are consistent with the probability measure. Conversely, in section 3.3 it has been shown how to derive a probability density and thus a probability measure by equidistributing a basic probability assignment (see formula (79)) ; then, the probability measure is bounded by the belief and the plausibility functions (whose particular cases are necessity and possibility measures) issued from the basic probability assignment, similarly to (99). Let us prove that when the belief and plausibility functions reduce to necessity and possibility measures, the probability density defined by (79) satisfies (98) also, i.e. possibility, probability and necessity are in the same relation in both approaches.

The expression of (79) is

$$p_i = \sum_{A \supset \{x_i\}} \frac{1}{|A|} . m(A) \qquad i=1,n$$

Since m is the basic probability assignment of a consonant belief function, the only subsets of X on which m is possibly non-zero are nested and can be ordered in the following way : $A_i = \{x_1 \ldots, x_i\}$, i=1,n (see section 3.1.).

Then, taking into account (79), we get

$$p_i = \sum_{j=i}^{n} \frac{1}{j} \cdot m_j = \sum_{j=i}^{n} \frac{1}{j} (\pi_j - \pi_{j+1}) \quad (100)$$

It is easy to recognize (98) in this expression.

Then from (100) we deduce

$$m_i = i.(p_i - p_{i+1}), \ i=1,n \quad (101)$$

with $p_{n+1} = 0$. Besides, we have

$$\pi_i - \pi_{i+1} = i.(p_i - p_{i+1}) \quad (102)$$

which entails

$$\pi_i = \pi_{i+1} \Longleftrightarrow p_i = p_{i+1} \quad (103)$$

and

$$\pi_i = 0 \Longleftrightarrow p_i = 0 \quad (104)$$

Lastly, (97) can be easily written without supposing the p_i's are ordered

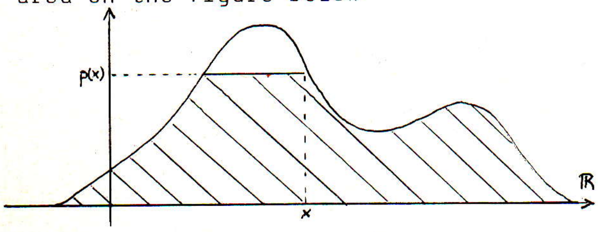
$$\pi_i = \sum_{j=1}^{n} \min(p_i, p_j) \quad i=1,n \quad (105)$$

Contrastedly with (89) and (90), here, each π_i does not depend only on p_i, but also on the other p_j's. The possibility of occurrence of an elementary event depends on its own probability and on the difference between this probability and the probabilities of the other elementary events.

When $X = \mathbb{R}$, (105) is generalized by

$$\forall x \in \mathbb{R} \quad \pi(x) = \int_X \min(p(x), p(t)) \, dt \quad (106)$$

Thus $\pi(x)$ corresponds to the shaded area on the figure below



Owing to formula (105), the table given in the example discussed in 4 .1. can now be completed :

amount of eggs	1	2	3	4	5
physical possibility	1	1	1	1	0.8
possibility of occurence	0.3	1	0.3	0	0
probability	0.1	0.8	0.1	0	0

amount of eggs	6	7	8
physical possibility	0.6	0.4	0.2
possibility of occurrence	0	0	0
probability	0	0	0

Observe that the possibility of occurrence is bounded by the probability and by the physical possibility.

5. - CONCLUDING REMARKS

Among uncertainty measures, probability, possibility and necessity measures play a remarkable role even if other fuzzy measures may be worth considering. Their respective axiomatics emphasize the differences existing between them.

The dual concepts of possibility and necessity may refer to several slightly different meanings, especially from the one hand the physical possibility and necessity which relate to the restrictions on possible events due to material constraints and on the other hand the epistemic possibility and necessity which relate to the potential surprise caused by the occurrence of events. Even if in this latter case, it seems possible to derive the (epistemic) possibility distribution from a probability density (via a formula such as (106)), it does not mean that the possibility of an event can be expressed only in terms of its probability (Indeed, the only existing relation between the possibility, the necessity and the probability of an event is the inequalities (99)).

This situation contrasts with g_λ -fuzzy measures which are equivalent to probability measures through a regular transformation (see 2.1.). Moreover, formula (106) and (89) enables to derive possibility distributions from statistical evidence, which may be important for fuzzy set membership function estimation.

The relation existing between Shafer's belief theory and the triangular norm approach to fuzzy measures still need to be clarified. However, relations between probability assignments and densities have been explicited for all introduced families of fuzzy measures which belong to both frameworks. Moreover, consistency inequalities have been encountered each time we exploit a given representation of evidence from several points of view. Lastly, links between combination of evidence via Dempster's rule and intersection of fuzzy sets have been stressed.

REFERENCES

[1] Banon, G. (1978) Distinction entre plusieurs sous-ensembles de mesures floues. Proc. of the International Colloquium on Fuzzy Set Theory and Applications, Marseille, Sept. 20-22, 1978.

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

[3] Dubois, D. (1980) Sur les liens entre les notions de probabilité et de possibilité (Quelques remarques) CNRS Round Table "Quelques Applications Concrètes Utilisant les Derniers Perfectionnements de la Théorie du Flou", Lyon, June 23-25, 1980.

[4] Dubois, D. (1980) Triangular norms for fuzzy sets. Proc. 2nd Int. Seminar on Fuzzy Set Theory, Johannes Kepler Univ. Linz, Austria Sept. 15-20, 1980 (E.P. Klement, ed.) pp. 39-68.

[5] Dubois, D. Prade, H. (1978) Fuzzy Algebra, Analysis, Logics. (Part E) TR-EE 78-13, Purdue University, Indiana.

[6] Dubois, D., Prade, H. (1979) Outline of fuzzy set theory. in Advances in Fuzzy set theory and Applications (M.M. Gupta, R.K. Ragade, R.R. Yager, eds.) North Holland, pp. 27-48.

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

[8] Dubois, D., Prade, H. (1982) A class of fuzzy measures based on triangular norms - A general framework for the combination of uncertain information. Int. J. of General Systems Vol. 8, n° 1. pp.43-61.

[9] Goodman, I.R. (1980) Identification of fuzzy sets with a class of canonically induced random sets. Proc. 19th IEEE Conference on Decision and Control, Albuquerque, New Mexico, Dec. 10-12, 1980, pp. 353-357.

[10] Goodman, I.R. (1981) Fuzzy sets as equivalence classes of random sets. in Recent Developments in Fuzzy Set and Possibility Theory (R.R. Yager, ed) Pergamon Press, New-York.

[11] Hacking, I. (1975) All kinds of possibility. Philosophical Review, vol. 84 pp. 321-347.

[12] Kampé de Fériet, J. (1980) Une interprétation des mesures de plausibilité et de credibilité au sens de G. Shafer et de la fonction d'appartenance définissant un ensemble flou de L. Zadeh. Pub. IRMA Lille, vol. 2, Fasc. 6, No. 2, pp. II-01, II-02.

[13] Nguyen, H.T. (1978) On random sets and belief functions. J. Math. Anal. & Appl. vol. 65, pp. 531-542.

[14] Prade, H. (1979) Nomenclature of fuzzy measures. Proc. 1st Int. Seminar on Fuzzy Set Theory, Johannes Kepler Univ., Linz, Austria, Sept. 24-29, 1979 (E.P. Klement, ed) pp. 7-25. Summary in BUSEFAL n°1, Winter 1979-80, L.S.I., Univ. P. Sabatier, Toulouse, pp. 26-28.

[15] Prade, H. (1980) Une approche des "mesures floues" basée sur les normes triangulaires. CNRS Round table "Quelques Applications Concrètes Utilisant les Derniers Perfectionnements de la Théorie du Flou", Lyon, June 23-25,1980.

[16] Prade, H. (1981) On the link between Dempster's rule of combination of evidence and fuzzy set intersection. BUSEFAL, n° 8, Fall 1981, L.S.I., Univ. P. Sabatier, Toulouse, pp. 60-64.

[17] Schweizer, B., Sklar, A. (1963) Associative functions and abstract semi-groups. Publ. Math. Debrecen Vol. 10 pp. 69-81.

[18] Shackle, G.L.S. (1961) Decision, Order and Time in Human Affairs. Cambridge University Press, (U.K.).

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

[20] Sugeno, M. (1974) Theory of Fuzzy Integrals and its Applications. Ph. D. Thesis, Tokyo Institute of Technology.

[21] Wierzchon, S.T. (1981) An algorithm for identification of fuzzy measure. IOPM "ORGMASZ" 01-877 Warszawa, Poland.

[22] Williams, P.M. (1978) On a new theory of epistemic probability. British J. for the Philosophy of Science Vol.29, pp. 375-387.

[23] Zadeh, L.A. (1978) Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems Vol. 1, N° 1, pp. 3-28.

[24] Zadeh, L.A. (1979) On the validity of Dempster's rule of combination of evidence. Memo UCB/ERL N° 79/24, Univ. of California, Berkeley.