14

A Set-Theoretic View of Belief Functions

Logical Operations and Approximations by Fuzzy Sets*

Didier Dubois and Henri Prade

Abstract. A body of evidence in the sense of Shafer can be viewed as an extension of a probability measure, but as a generalized set as well. In this paper we adopt the second point of view and study the algebraic structure of bodies of evidence on a set, based on extended set union, intersection and complementation. Several notions of inclusion are exhibited and compared to each other. Inclusion is used to compare a body of evidence to the product of its projections. Lastly, approximations of a body of evidence under the form of fuzzy sets are derived, in order to squeeze plausibility values between two grades of possibility. Through all the paper, it is pointed out that a body of evidence can account for conjunctive as well as a disjunctive information, i.e. the focal elements can be viewed either as sets of actual values or as restrictions on the (unique) value of a variable.

Key words: Theory of evidence, Possibility measure, Fuzzy set, Knowledge representation

Introduction

The framework of plausibility and credibility (or belief) functions[24] or, equivalently that of the random sets[19] encompasses both probability theory and possibility theory[7, 38]. It is now acknowledged that fuzzy sets[35] viewed as possibility distributions, are, using Shafer's terminology, contour functions of consonant belief functions[4, 17] or in the terminology of random sets, one-point coverages of random sets[13, 22, 28]. In a recent paper[25] Shafer carefully examines the rules of calculation of fuzzy sets and possibility measures as opposed to their counterparts for belief functions. It turns out that the main difference lies in the use of Dempster rule for combining belief functions versus fuzzy set-intersection for combining possibility measures. Dempster rule applied to the combination of possibility measures does

* This paper is based on a presentation at the 1st IFSA Conference, held in Palma de Mallorca, Spain, July 1–6, 1985.

not yield a possibility measure while a fuzzy set-intersection does. This paper is a contribution to the debate between possibility measures and belief functions. First, combination rules for belief functions in the spirit of Dempster rule are described; they are counterparts of fuzzy set-theoretic union, complementation, products and projection. This set-theoretic view of belief functions points out the fundamental identity of both approaches to combining. Next, an extensive study of the concept of inclusion of bodies of evidence is carried out. Four definitions are proposed and compared. The existence of two antagonistic points of view on bodies of evidence is stressed and it brings some light to discriminate between definitions of inclusion. The following section is devoted to projections and products of belief functions, and the links between a body of evidence and the product of its projections. Lastly the problem of approximating belief functions by consonant bodies of evidence is considered, and best approximations, which squeeze a plausibility measure between two possibility measures, are calculated.

1 Shafer's Theory of Evidence Revisited

In this section, basic notions are introduced in a concise manner. It borrows from several already published works [4, 24, 25, 38] to which the reader is referred for proofs or detailed explanations. However some new issues are raised, especially the convexity of the set of belief functions and the difference between conjunctive and disjunctive items of information. This last point follows some early remarks by Zadeh[37] and a more elaborated discussion by Yager[31] in the framework of fuzzy sets and linguistic variables. Moreover, the allocation of a probability weight on the empty set is no longer forbidden.

1.1 Uncertainty Measures Induced by a Body of Evidence

According to Shafer[24], a body of evidence is modelled by a weighted set of logical statements, each referring to a subset A of a frame of discernment Ω. This frame of discernment corresponds to a point of view on a problem, and contains the possible values of some variable x. A body of evidence supplies information about the actual value of x (which is some element in Ω), with the following conventions, given here in a finite setting for simplicity. Let $\mathcal{F}$ be a family of subsets of Ω. A body of evidence is viewed as a pair $(\mathcal{F}, m)$ where m is a mapping from 2^Ω to the unit interval such that $m(A) > 0$ if and only if $A \in \mathcal{F}$. Any element of A of $\mathcal{F}$ is called a focal element, because part of the available information focuses on A. $m(A)$ is the relative weight of the statement "$x \in A$", and is viewed as the share of total belief committed to this statement exactly, and not to any other statement of the form "$x \in B \subset A$". m is called a basic assignment and satisfies the following requirement

$$\sum_{A \subseteq \Omega} m(A) = 1 \tag{1}$$

where 1 stands for the amount of total belief. The set of bodies of evidence on Ω is denoted as $\mathcal{B}(\Omega)$. In Shafer's book a basic assignment satisfies the additional condition

$$m(\varnothing) = 0 \tag{2}$$

which claims that no belief should be committed to the impossible event. (2) is a normalization condition which looks reasonable if the statement "$x \in \Omega$" is taken for granted. However in some instances one may be uncertain as to whether Ω is definitely exhaustive, or whether assigning a value to x is ever meaningful. For instance, if x is the age of cars belonging to some population where some individuals may have no cars (Zadeh[42]). Such situations can be conveniently handled by letting $m(\varnothing) > 0$. See also Dubois and Prade[4], Yager[29], Zadeh[41, 42] for further discussions. A body of evidence satisfying (2) is said to be *normal*.

Viewed as an allocation of probability over subsets of Ω, a body of evidence is also a random set[19]. However it can be equivalently represented by one of the following set-functions

$$\forall\ A \subseteq \Omega,\ \mathrm{Cr}(A) = \sum_{\varnothing \neq B \subseteq A} m(B), \tag{3}$$

$$\forall\ A \subseteq \Omega,\ \mathrm{Pl}(A) = \sum_{B \cap A \neq \varnothing} m(B), \tag{4}$$

$$\forall\ A \subseteq \Omega,\ Q(A) = \sum_{A \subseteq B} m(B). \tag{5}$$

Cr is called a belief function by Shafer[24], but we had rather call it a *credibility measure* since $\mathrm{Cr}(A)$ gathers the pieces of evidence which support A. Pl is called a plausibility measure since $\mathrm{Pl}(A)$ gathers the pieces of evidence which make the occurrence of A possible. Pl and Cr are related through the duality relation

$$\forall\ A,\ \mathrm{Pl}\,(A) + \mathrm{Cr}\left(\bar{A}\right) = 1 - m\,(\varnothing) \tag{6}$$

i.e. $\mathrm{Pl}(A)$ accounts for evidence which does not support the opposite event $\bar{A}$ nor events "outside Ω" (i.e. $\varnothing$). Q is called a *commonality function* by Shafer[24] and gathers pieces of evidence supported by event A. So far, its usefulness has been purely technical. Note that

$$\mathrm{Pl}\,(\varnothing) = \mathrm{Cr}\,(\varnothing) = 0; \qquad Q(\varnothing) = 1 \tag{7}$$

$$\mathrm{Pl}\,(\Omega) = \mathrm{Cr}\,(\Omega) = 1 - m\,(\varnothing)\,; \qquad Q\,(\Omega) = m\,(\Omega)\,. \tag{8}$$

When $m(\varnothing) = 0$, Shafer[24] has proved that Cr is order-n superadditive $\forall\ n \in \mathbb{N}$. This property still holds when $m(\varnothing) > 0$, for $\mathrm{Cr} + m(\varnothing)$, hence for Cr too. Then the basic assignment is still expressed in terms of the credibility measures as

$$\forall\ A, m\,(A) = \sum_{B \subseteq A} (-1)^{[A-B]}\left(\mathrm{Cr}(B) + m(\varnothing)\right) = \sum_{B \subseteq A} (-1)^{|A-B|}\mathrm{Cr}(B) \tag{9}$$

where $|A-B|$ is the cardinality of the set-difference $A-B$. See Shafer[24] for other inversion formulae (Pl in terms of Q, etc....). Pl and Cr are monotonic increasing with respect to set-inclusion, while Q is monotonic decreasing. As a consequence of (6), Pl is subadditive, which reads at order n:

$$\mathrm{Pl}\,(A_1 \cap A_2 \ldots \cap A_n) \leqq \sum_{\substack{I \subseteq \{1,\ldots,n\} \\ I \neq \varnothing}} (-1)^{|I|+1}\,\mathrm{Pl}\left(\bigcup_{i\in I} A_i\right). \tag{10}$$

The set of plausibility measures on Ω is isomorphic to $\mathcal{B}(\Omega)$ and has an interesting structure. Namely it is a convex set since the convex combination $\sum_{i=1}^{n} \alpha_i \cdot \mathrm{Pl}_i$ of subadditive functions Pl_i is subadditive too. The coefficients α_i are such that $\sum_{i=1}^{n} \alpha_i = 1,\ \alpha_i \geqq 0,\ \forall\, i$. The plausibility measure $\mathrm{Pl} = \sum_{i=1}^{n} \alpha_i \cdot \mathrm{Pl}_i$ is called a *mixture*. Let $(\mathcal{F}_i, m)$ be the body of evidence associated with Pl_i. Then, that associated with Pl is $(\mathcal{F}, m)$ such that

$$\mathcal{F} = \bigcup_{i=1,n} \mathcal{F}_i;\ \forall\, A \subseteq \Omega, m\,(A) = \sum_{i=1}^{n} \alpha_i m_i\,(A)\,.$$

The same remark holds for credibility measures. Let $\mathcal{B}^+(\Omega)$ be the set of normal bodies of evidence. Clearly $\mathcal{B}^+(\Omega)$ is a convex subset of $\mathcal{B}(\Omega)$.

1.2 Possibility, Necessity, Probability

Two extreme cases of plausibility measures can be obtained by adding constraints on the set of focal elements.

a) *$\mathcal{F}$ contains only singletons*, i.e. $\forall\, A \in \mathcal{F},\ \exists\, \omega \in A,\ A = \{\omega\}$. This occurs if and only if Cr = Pl and is a probability measure P. m is a probability assignment in the usual sense ($P(\{\omega\}) = m(\{\omega\})$).

The set functions Pl and Cr can be viewed as upper and lower probabilities (Dempster[1]) since any probability measure P generated from $(\mathcal{F}, m)$ by the following allocation procedure

i) $\forall\, A \in \mathcal{F}$ choose $\omega_A \in A$
ii) set $P(\{\omega\}) = \sum_{\omega_A = \omega} m(A), \forall\, \omega \in \Omega$

satisfies the following inequalities:

$$\forall\, A, \mathrm{Cr}(A) \leqq P(A) \leqq \mathrm{Pl}\,(A) \tag{11}$$

when $(\mathcal{F}, m)$ is normal.

Dempster[1] has proved that the set of probability measures satisfying (11) is convex and is the convex closure of the set of probability measures obtained by the procedure (i)–(ii).

b) *$\mathcal{F}$ contains only a nested sequence of subsets $E_1 \subseteq E_2 \ldots \subseteq E_p$.* It occurs if and only if $\forall\ A, B \subseteq \Omega$

$$\mathrm{Cr}\,(A \cap B) = \min\left(\mathrm{Cr}\,(A)\,, \mathrm{Cr}\,(B)\right) \tag{12}$$

$$\mathrm{Pl}\,(A \cup B) = \min\left(\mathrm{Pl}\,(A)\,, \mathrm{Pl}\,(B)\right). \tag{13}$$

Cr is called a consonant belief function by Shafer[24] and Pl a possibility measure by Zadeh[38]. A possibility measure is denoted Π, and the duality relationship (6) justifies the name of "necessity measure"[2] for consonant belief functions. They are also called certainty measures by Zadeh[40], and shall be denoted N in the following.

The set $\pi(\Omega)$ of possibility measures is not convex. Indeed if $\mathcal{F}$ and $\mathcal{F}'$ both define nested sequences, then generally $\mathcal{F} \cup \mathcal{F}'$ does not, so that $\alpha\Pi + (1 - \alpha)\Pi'$ is not always a possibility measure. A possibility measure Π such that $\forall\ A, \Pi(A) \in \{0, 1\}$ is called a *crisp* possibility measure. Any crisp possibility measure derives from a unique focal element which is a subset E of Ω, i.e. ($\mathcal{F} = \{E\}$).

These two extreme cases of bodies of evidence correspond to precise but scattered pieces of uncertain information (Case (a)) and imprecise but consonant pieces of information (Case (b)). The nature of the relevant uncertainty measure (possibility or probability) is dictated by the structure of the available body of evidence. Generally a body of evidence is neither consonant nor precise. A body of evidence $(\mathcal{F}, m)$ is said to be *consistent* if and only if $\bigcap_{A \in \mathcal{F}} A \neq \varnothing$. This condition is weaker than the consonant constraint of nested focal elements, but still expresses some agreement between the various statements which form the body of evidence.

The following result indicates that in some sense probability measures and possibility measures are the basic concepts in the theory of evidence:

Proposition 1. *Any plausibility measure other than a possibility or a probability measure is a convex combination of a probability measure and possibility measures which are not Dirac functions.*

Proof. For any subset A of Ω, denote Π_A the possibility measure such that $\{A\}$ is its set of focal elements. Let Pl be a plausibility measure. Then (4) also reads

$$\mathrm{Pl}\,(A) = \sum_{B \subseteq \Omega} m\,(B) \cdot \Pi_B\,(A)\,.$$

Now if B is a singleton, then Π_B is a Dirac function, so that the plausibility measure defined by

$$P\,(A) = \frac{\sum_{|B|=1} m\,(B) \cdot \Pi_B\,(A)}{\sum_{|B|=1} m\,(B)}$$

is a probability measure when it exists. Q.E.D.

As a consequence, if we identify the set of crisp possibility measures with 2^Ω, the set of subsets of Ω, through the bijection $A \mapsto \Pi_A$ such that $m(A) = 1$, the set $\mathcal{B}(\Omega)$ can be viewed as the convex hull of 2^Ω, while the set $\mathcal{P}(\Omega)$ of probability measures is the convex hull of the subset of singletons of Ω.

1.3 Possibility Measures as Fuzzy Sets

A possibility or a probability measure is entirely characterized by the set $\{\mathrm{Pl}(\{\omega\})|\omega \in \Omega\}$; $\mathrm{Pl}(\{\omega\})$ is the one-point coverage function, in terms of random sets[13] and is called a contour function by Shafer[24]. In the case of probability measures, $\mathrm{Pl}(\{\omega\}) = P(\{\omega\})$ and $\forall\ A,\ \mathrm{Pl}(A) = \sum_{\omega \in A} \mathrm{Pl}(\{\omega\})$. In case of a possibility measure

$$\Pi\left(A\right) = \max_{\omega \in A} \Pi\left(\{\omega\}\right);\ N(A) = \min_{\omega \in \bar{A}}\ 1 - \Pi\left(\{\omega\}\right). \tag{14}$$

When $\Pi(\Omega) = 1$, we have $\max_{\omega \in \Omega} \Pi(\{\omega\}) = 1$.

In the following, $\Omega = \{\omega_1, \ldots, \omega_n\}$ has n elements, $P(\{\omega_i\})$ is denoted p_i, and $\Pi(\{\omega_i\})$ is denoted π_i, for the sake of simplicity. When Π has values only in {0,1}, the function $\mu_F : \Omega \mapsto [0,1]$ defined by

$$\mu_F\left(\omega_i\right) = \pi_i \tag{15}$$

is the characteristic function of a set. In the general case it is the membership function of a fuzzy set[35] F.

Let $F_\alpha = \{\omega|\mu_F(\omega) \geqq \alpha\}$ be the *α-cut* of F. When Ω is finite the set $\{F_\alpha|\alpha \in [0,1]\}$ of α-cuts is finite, and it is proved[4] that it is the set of focal elements of the possibility measure such that $\mu_F(\omega) = \Pi(\{\omega\})$. More specifically assume $\pi_1 = 1 \geqq \pi_2 \geqq \cdots \geqq \pi_n \geqq \pi_{n+1} = 0$ and let $A_i = \{\omega_1, \ldots, \omega_i\}$. Then the basic assignment m is defined in terms of the π_i's by:[4]

$$\begin{cases} m(A) = 0 \text{ if } \ \nexists\ i : A = A_i \\ m\left(A_i\right) = \pi_i - \pi_{i+1} \end{cases}. \tag{16}$$

In the general case, $\mathrm{Pl}(\{\omega\})$ may still be interpreted as the membership grade of ω in a fuzzy set F. However the knowledge of $\{\mathrm{Pl}(\{\omega\})|\omega \in \Omega\}$ is not enough to recover the body of evidence $(\mathcal{F}, m)$. Moreover F is not always a normalized fuzzy set. Namely, even if $(\mathcal{F}, m)$ is normal,

$$\exists\ \omega : \mathrm{Pl}\left(\{\omega\}\right) = 1 \text{ if and only if } \ (\mathcal{F}, m) \text{ is consistent.}$$

Moreover when Pl is a probability measure, it rather corresponds to the idea of a fuzzy point,[16] since the grade of complete membership (1) is shared among the singletons in that case. The characterization of plausibility measures Pl such that $\forall\ \omega,\ \mathrm{Pl}(\{\omega\}) = \mu_F(\omega)$, given μ_F, is done by Goodman[13] in the setting of random sets.

1.4 Disjunctive versus Conjunctive Evidence

In the preceding paragraphs, a set is viewed as restricting the possible values of a variable x, and these values are supposedly mutually exclusive. Similarly fuzzy sets are viewed as fuzzy restrictions[36]. There is another view of sets, as containing values which are actually taken by x. This point of view is considered by Yager[31] in terms of linguistic variables and by Prade and Testemale[21] in the framework of fuzzy relational databases. In the first case variables are single-valued and the body of evidence is said to be *disjunctive*. In the second case, variables are multiple-valued, and the body of evidence is said to be *conjunctive*. The difference between conjunctive and disjunctive fuzzy sets has been pointed out by Zadeh[37].

Example 1. "John is *tall*" means that John's height is some number restricted by the fuzzy set "tall".

"John stayed in Paris from 1980 to 1984" means that {1980, 1981, 1982, 1983, 1984} is a set of years when John actually stayed in Paris.

In the case of conjunctive knowledge, $\forall\ B \subseteq A$, if "$x = A$" is true then "$x = B$" is also true, so that the entailment principle[39] works backwards (Yager[31]). As a consequence, the quantity $Q(A)$, i.e. the commonality number, defined by (5), is the actual grade of credibility of "$x = A$" in the case of a conjunctive body of evidence, instead of $\mathrm{Cr}(A)$, as pointed out by Zadeh. Notice that, for singletons the identity

$$Q\left(\{\omega\}\right) = \mathrm{Pl}\left(\{\omega\}\right) \tag{17}$$

holds, and moreover if $\mathcal{F}$ is consonant then, equivalently

$$\forall\ A, B, Q\left(A \cup B\right) = \min\left(Q\left(A\right), Q\left(B\right)\right). \tag{18}$$

In the consonant case, $\{Q(\{\omega\}) | \omega \in \Omega\}$ also characterizes the body of evidence and

$$\forall\ A, Q\left(A\right) = \min_{\omega \in A} Q\left(\{\omega\}\right). \tag{19}$$

In the conjunctive context, the membership function μ_F defined by (15) is no longer viewed as a possibility distribution, but what could be termed as a "necessity" or "certainty" distribution since $\mu_F(\omega_i) = Q(\{\omega_i\})$ is now the grade of certainty that ω_i is a value of x. The grade of possibility is then defined by

$$\phi\left(A\right) = 1 - Q\left(\bar{A}\right) = \max_{\omega \notin A} 1 - \mu_F\left(\omega\right). \tag{20}$$

Note that when F is such that there are at least two elements ω' and ω'' such that $\mu_F(\omega') = \mu_F(\omega'') = 0$, then $\forall\ \omega, \phi(\{\omega\}) = 1$. Indeed $\mu_F(\omega) = 0$ does not forbid ω as a *value* of x but only let this statement be contingent (total uncertainty). In other words "$x = A$" means that x takes *at least* all values in A. Lastly note that from (19) and (20)

$$\phi(A) = \Pi_{\bar{F}}\left(\bar{A}\right), Q(A) = 1 - \Pi_{\bar{F}}(A) = N_{\bar{F}}\left(\bar{A}\right) \tag{21}$$

where $\Pi_{\bar{F}}$ is the possibility measure where the underlying possibility distribution is the membership function of the complement $\bar{F}$ of F, i.e. $1 - \mu_F$.

The notion of conjunctive versus disjunctive types of information seems to be an important issue in knowledge representation, and is encountered in the next section, as a by-product.

2 Set-Theoretic Operations on Bodies of Evidence

Dempster[1] has introduced a rule of combination for two disjunctive normalized bodies of evidence $(\mathcal{F}_1, m_1)$, $(\mathcal{F}_2, m_2)$, consistently with Bayes rule of conditioning. It reads:

$$\forall\ A \subseteq \Omega, (m_1 \cap m_2)(A) = \sum_{B \cap C = A} m_1(B) \cdot m_2(C) \tag{22}$$

$$\forall\ A \subseteq \Omega, m(A) = \frac{(m_1 \cap m_2)(A)}{1 - (\{m_1 \cap m_2\})\varnothing}. \tag{23}$$

Equation (22) can be justified in statistical terms on the basis of the independence of the sources which provide $(\mathcal{F}_1, m_1)$ and $(\mathcal{F}_2, m_2)$. Equation (23) underlies a complete reliability of these sources, and is a normalization technique. The term $(m_1 \cap m_2)(\varnothing)$ reflects the amount of dissonance between the sources, and is eliminated. Equation (22) can be viewed as performing the intersection of independent random sets[4, 13].

Contrastedly, if Π_1 and Π_2 are two possibility measures, with possibility distributions $\pi_1 = \mu_{F_1}, \pi_2 = \mu_{F_2}$, a possibility measure Π_{12} can be obtained from the possibility distribution $\pi_{12} = \mu_{F_1 \cap F_2}$ where the fuzzy set-theoretic intersection is defined by a triangular norm[8, 23]:

$$\pi_{12} = \pi_1 * \pi_2. \tag{24}$$

The main candidates for $*$ are $a * b = \min(a, b)$; $a \cdot b$; $\max(0, a + b - 1)$.[2, 8] Assuming the complete reliability of the sources leads to normalize π_{12} into

$$\forall\ \omega, \pi(\omega) = \frac{\pi_1(\omega) * \pi_2(\omega)}{\max_{\omega \in \Omega} \pi_1(\omega) * \pi_2(\omega)}. \tag{25}$$

(24) and (25) are possibilistic counterparts of (22) and (23) respectively.

It was pointed[4] that if Π_1 and Π_2 are combined via (22) what is obtained is generally not a possibility measure. This is because when $\mathcal{F}_1$ and $\mathcal{F}_2$ are consonant, the set $\mathcal{F}_{1\cap 2} = \{A \cap B | A \in \mathcal{F}_1, B \in \mathcal{F}_2\}$ is generally not consonant. Besides, using (22) yields a not necessarily normalized plausibility function Pl_{12} where[4]:

$$\mathrm{Pl}_{12}\left(\{\omega\}\right) = \pi_1\left(\omega\right) \cdot \pi_2\left(\omega\right) \tag{26}$$

which is a particular instance of (24) where $*$ is the product. Hence Dempster rule is closely related to a fuzzy set intersection. But generally

$$\mathrm{Pl}_{12}\left(A\right) \geqq \Pi_{12}\left(A\right) = \max_{\omega \in A} \pi_1\left(\omega\right) \cdot \pi_2\left(\omega\right) \tag{27}$$

i.e. Π_{12} is more informative than Pl_{12}.

The set of combination operations for fuzzy sets is richer than for bodies of evidence since all connectives of propositional logic can be extended to the combination of fuzzy sets and this extension is not unique. Strangely enough counterparts of set-union, set-complementation, etc.. . .have not been considered for bodies of evidence, but in the mathematical literature of random sets[12, 14].

In the following, these connectives are defined at an elementary level[1], thus casting Dempster rule in a set-theoretic framework, and enriching the set of combination rules. This view of plausibility measures reflects the standpoint of logic and contrasts with the measure-theoretic view which Dempster had when he introduced his concept of upper and lower probabilities and expectations.

2.1 The Union of Bodies of Evidence

The union of two bodies of evidence $(\mathcal{F}_1, m_1)$ and $(\mathcal{F}_2, m_2)$ on Ω is defined, in the spirit of (22)–(23) by the basic assignment $m_1 \cup m_2$ such that

$$\forall\ A \subseteq \Omega, \left(m_1 \cup m_2\right)\left(A\right) = \sum_{B \cup C = A} m_1\left(B\right) \cdot m_2\left(C\right). \tag{28}$$

Note that (28) is (22) where $\cap$ is changed into $\cup$. While the intersection of two bodies of evidence only keeps the items of information asserted by both sources, the union does not reject anything. Especially if $m_1(\varnothing) = m_2(\varnothing) = 0$, it is easy to check that $(m_1 \cup m_2)(\varnothing) = 0$, i.e. the union does not generate any conflict, and the normalization step (23) is useless here. The resulting set of focal elements is indeed $\mathcal{F}_{1\cup 2} = \{A \cup B \mid A \in \mathcal{F}_1, A \in \mathcal{F}_2\}$.

The union of two bodies of evidence is more easily performed via the credibility measure since:

Proposition 2. *Let* $\mathrm{Cr}_1 \cup \mathrm{Cr}_2$ *be the credibility measure associated with* $m_1 \cup m_2$. *Then* $\forall\ A \subseteq \Omega, (\mathrm{Cr}_1 \cup \mathrm{Cr}_2)(A) = \mathrm{Cr}_1(A) \cdot \mathrm{Cr}_2(A)$.

[1] During the course of the investigation whose results are reported here, we became aware of similar attempts by Yager[32] and Oblow[20].

Proof.

$$(\mathrm{Cr}_1 \cup \mathrm{Cr}_2)(A) = \sum_{\varnothing \neq B \cup C \subseteq A} m_1(B) \cdot m_2(C)$$
$$= \sum_{\varnothing \neq B \subseteq A} m_1(B) \left(\sum_{\varnothing \neq C \subseteq A} m_2(C) \right). \quad \text{Q.E.D.}$$

Notice that in the case of intersection of bodies of evidence, the counterpart of Proposition 2 holds for the commonality numbers only, since, as noted by Shafer[24], (22) implies

$$(Q_1 \cap Q_2)(A) = Q_1(A) \cdot Q_2(A). \tag{29}$$

The notion of conjunctive and disjunctive knowledge can shed light on these properties, if we recall, following Yager[31], that in the presence of conjunctive information, "$x = A$ or $x = B$" translates into $x = A \cap B$ and "$x = A$ and $x = B$" translates into "$x = A \cup B$".

Example 2. John stayed in Paris from 1980 till 1982 and from 1982 till 1984 is equivalent to "John stayed in Paris from 1980 till 1984".

But if we happen to know from two sources that he stayed in Paris from 1980 till 1983 *or* from 1981 till 1984, then the only sure resulting item of information is that he stayed in Paris from 1981 till 1983.

Now, if we remember that the commonality numbers play, for a conjunctive body of evidence, the same role as the credibility degrees in a disjunctive body of evidence, it is clear that (29) is the mirror image of Proposition 2, and achieves an "or" of two conjunctive bodies of evidence.

As a consequence of (29) and Proposition 2, the union and intersection of bodies of evidence are commutative and associative. If we denote by Ω (resp.: $\varnothing$) the body of evidence such that $m(\Omega) = 1$ (resp.: $m(\varnothing) = 1$), that is, total ignorance (resp.: the null value "not applicable") for variable x in the disjunctive interpretation, we have

$$\forall\, m, m \cap \Omega = m; \qquad m \cup \varnothing = m. \tag{30}$$

Now, applying (22) and (28) on subsets A of Ω, i.e. $m(A) = 1$, we recover the usual set-intersection and union in 2^Ω. But these operations are not idempotent on $\mathcal{B}(\Omega)$. Indeed, Proposition 2 and (29) lead to

$$\forall\, \omega \in \Omega, (\mathrm{Pl}_1 \cap \mathrm{Pl}_2)(\{\omega\}) = \mathrm{Pl}_1(\{\omega\}) \cdot \mathrm{Pl}_2(\{\omega\}) \tag{31}$$
$$(\mathrm{Pl}_1 \cup \mathrm{Pl}_2)(\{\omega\}) = \mathrm{Pl}_1(\{\omega\}) + \mathrm{Pl}_2(\{\omega\}) - \mathrm{Pl}_1(\{\omega\}) \cdot \mathrm{Pl}_2(\{\omega\}). \tag{32}$$

Note that the intersection and the union in $\mathcal{B}(\Omega)$ are not stable on the subset of possibility measures. This is because if $\mathcal{F}_1$ and $\mathcal{F}_2$ are consonant, then generally neither $\mathcal{F}_{1\cap 2}$ nor $\mathcal{F}_{1\cup 2}$ are. But (32) as (31), correspond to well known fuzzy set-theoretic operations. The set of probability measures is not closed under the union operation, since $P_1 \cup P_2$ corresponds to a set of focal elements some of which are 2-element sets. Strictly speaking, the closure property does not hold for intersection since the intersection of two probability measures is no longer normalized (the intersection of singletons is generally empty!). The closure property is recovered through normalization (23) i.e. using Dempster rule as a whole.

Lastly the union of two consistent bodies of evidence is consistent while their intersection may no longer be so.

2.2 Complement of a Body of Evidence

The complement of a body of evidence $(\mathcal{F}, m)$ is $(\neg\mathcal{F}, \bar{m})$ defined by

$$\forall\ A \subseteq \Omega, \bar{m}(A) = m\left(\bar{A}\right) \tag{33}$$

so that $\neg\mathcal{F} = \{\bar{A} | A \in \mathcal{F}\}$. This complementation is formally involutive. Moreover the union and intersection satisfy De-Morgan laws since

$$\begin{aligned}\forall\ \overline{(m_1 \cup m_2)}(A) = (m_1 \cup m_2)\left(\bar{A}\right) &= \sum_{B \cup C = \bar{A}} m_1(B) \cdot m_2(C) \\ &= \sum_{\bar{B} \cap \bar{C} = A} \bar{m}_1\left(\bar{B}\right) \cdot \bar{m}_2\left(\bar{C}\right) = (\bar{m}_1 \cap \bar{m}_2)(A).\end{aligned}$$

It is easy to see that (33) reduces to usual set complementation when $m(A) = 1$. Moreover if $\mathcal{F}$ is consonant, then $\neg\mathcal{F}$ is also consonant, so that (33) also reduces to fuzzy set complementation when applied to a possibility measure, i.e. the set of possibility measures is closed under complementation. But the set of probability measures is not for $|\Omega| > 2$ since all focal elements in $\neg\mathcal{F}$ then contain $|\Omega| - 1$ elements.

$\mathcal{B}(\Omega)$ is *not* a Boolean algebra. Indeed union and intersection are not idempotent. Moreover the laws of contradiction and excluded middle are not valid, i.e.

$$\text{for } m \in \mathcal{B}(\Omega) - 2^{\Omega}, \qquad \text{generally} \quad m \cap \bar{m} \neq \varnothing; \qquad m \cup \bar{m} \neq \Omega.$$

Actually it can be checked that $(m \cap \bar{m})(\varnothing) > 0, (m \cup \bar{m})(\Omega) > 0$ which expresses that these laws *somewhat* hold. If $\mathcal{F} = \{A, B\}$ then $\neg\mathcal{F} = \{\bar{A}, \bar{B}\}$ and $(m \cap \bar{m})(A \cap \bar{B}) > 0, (m \cup \bar{m})(A \cup \bar{B}) > 0$, etc....

Hence $\mathcal{B}(\Omega)$ has the same algebraic structure as the set of fuzzy subsets of Ω, $[0,1]^{\Omega}$, under the product, probabilistic sum, and usual complementation of fuzzy sets, i.e.

$$\begin{aligned}\mu_{F \cap G}(\omega) &= \mu_F(\omega) \cdot \mu_G(\omega) \\ \mu_{F \cup G}(\omega) &= \mu_F(\omega) + \mu_G(\omega) - \mu_F(\omega)\,\mu_G(\omega) \\ \mu_{\bar{F}}(\omega) &= 1 - \mu_F(\omega).\end{aligned}$$

Moreover these fuzzy set-theoretic operations are consistent with set-theoretic operations in $\mathcal{B}(\Omega)$, under independence assumption, up to stability of $\pi(\Omega)$.

An interesting feature of complementation in $\mathcal{B}(\Omega)$ is that it turns a disjunctive body of evidence into a conjunctive one. To see it consider the simple case $\mathcal{F} = \{A\}$, and A restricts the possible values of x. Then $\bar{A}$ is a set of values which are forbidden for x. Let $\bar{x}$ be the variable which takes values which x does not take. It is clear that $\overline{A} \subseteq \overline{x}$ ($\overline{x}$ takes at least all values in $\overline{A}$) is equivalent to x is A (the value of x is restricted by A). In the general case, the same transformation occurs, and any focal element $\bar{A} \in \urcorner F$ is a set of values which x certainly does not take (with weight $m(A)$). This transformation in the nature of evidence provides some explanation of the following property.

Proposition 3. *Let $\bar{Q}$ be the commonality function associated with the complement $(\urcorner\mathcal{F}, \bar{m})$ of a disjunctive body of evidence $(\mathcal{F}, m)$. Then*

$$\forall\ A, \mathrm{Cr}\,(A) = \bar{Q}\left(\bar{A}\right) - m\,(\varnothing)\,.$$

Proof.

$$\sum_{\varnothing \neq B \subseteq A} m\,(B) = \sum_{\bar{A} \subseteq \bar{B} \neq \Omega} \bar{m}\left(\bar{B}\right) = \bar{Q}\left(\bar{A}\right) - \bar{m}\,(\Omega)\,. \qquad \text{Q.E.D.}$$

This result stresses that Cr and Q play the same role in each type of knowledge, disjunctive and conjunctive respectively.

An important remark is that, reciprocally, the complement of a conjunctive body of evidence is *not* a disjunctive body of evidence in the sense defined in this paper. To see it, consider the case of the conjunctive statement '$A \subseteq x$', then, defining $\bar{x}$ as above, all we know about $\bar{x}$ is that any subset of $\bar{A}$ is a possible conjunctive set of values for $\bar{x}$, so that the knowledge about $\bar{x}$ is a possibility distribution on $2^{\bar{A}}$, say π, such that $\forall\ B \subseteq \bar{A},\ \pi(B) = 1$ means B is a possible set of values for $\bar{x}$ (i.e. x possibly does not take any value in B). Hence π defines disjunctive knowledge *over* $2^{\bar{A}}$. The usual disjunctive information is recovered as a particular case, setting $\pi(B) = 1$ if and only if B is a singleton in $\bar{A}$ and 0 otherwise. This type of higher-order disjunctive information is not a mere game of the mind; it is often encountered in data-bases with multiple-valued attributes, when one wishes to represent the possible sets of tongues spoken by an individual, for instance (see Prade and Testemale[21]).

These remarks weaken the apparent strength of the involution property of the complementation operation in $\mathcal{B}(\Omega)$.

2.3 Inclusions

Concepts of inclusion can also be introduced on $\mathcal{B}(\Omega)$. Given a normal body of evidence $(\mathcal{F}, m)$, the interval $[\mathrm{Cr}(A),\ \mathrm{Pl}(A)]$ can be viewed as the range of the probability of A induced by the lack of precision of the focal elements (see *1.2.*). In other words, the body of evidence $\mathcal{F}$ defines a (convex) set of probability measures on Ω, say $\mathcal{C}(\mathcal{F})$.

A normal body of evidence $(\mathcal{F}, m)$ can be viewed as included in $(\mathcal{F}',\ m')$ as soon as $\mathcal{C}(\mathcal{F}) \subseteq \mathcal{C}(\mathcal{F}')$. In terms of the plausibility and credibility measures (Pl, Cr) and $(\mathrm{Pl}', \mathrm{Cr}')$, this is equivalent to:

$$\forall\ A \in \Omega, [\mathrm{Cr}\,(A)\,, \mathrm{Pl}\,(A)] \subseteq \left[\mathrm{Cr}'\,(A)\,, \mathrm{Pl}'\,(A)\right]. \tag{34}$$

We shall write $(\mathcal{F}, m) \subseteq (\mathcal{F}', m')$ when (34) holds true. (34) reduces to (11) when $\mathrm{Cr} = \mathrm{Pl} =$ a probability measure. Note that because $\mathrm{Cr}(A) = 1 - \mathrm{Pl}(\bar{A})$, any of the following inequalities is equivalent to (34):

$$\mathrm{Cr}\,(A) \geqq \mathrm{Cr}'\,(A)\,, \forall\ A \in \Omega, \tag{35}$$

$$\mathrm{Pl}\,(A) \leqq \mathrm{Pl}'\,(A)\,\forall\ A \in \Omega. \tag{36}$$

Disjunctive and Conjunctive Inclusions

The definition of inclusion can be extended from $\mathcal{B}^+(\Omega)$ to $\mathcal{B}(\Omega)$, taking (36) as the actual definition. Note that (35) is not equivalent to (36) for bodies of evidence which are not normal. Indeed, in the general case, (36) is equivalent to:

$$\mathrm{Cr}\,(A) + m\,(\varnothing) \geqq \mathrm{Cr}'\,(A) + m'\,(\varnothing)\,, \forall\ A \in \Omega$$

due to the definitions of Cr and Pl. (36) induces some relationships between the respective contents of $\mathcal{F}$ and $\mathcal{F}'$ such that $(\mathcal{F}, m) \subseteq (\mathcal{F}', m')$. In the following the *core* (resp.: *support*) of $(\mathcal{F}, m)$ is the intersection (resp.: union) of focal elements and denoted $C(\mathcal{F})$ (resp.: $S(\mathcal{F})$). The following necessary condition for inclusion relationship is noticeable:

Proposition 4. *If,* $(\mathcal{F}, m) \subseteq (\mathcal{F}', m')$ *then*

i) $S(\mathcal{F}) \subseteq S(\mathcal{F}');\ C(\mathcal{F}) \subseteq C(\mathcal{F}')$,
ii) $\forall\ A' \subseteq \mathcal{F}',\ \exists A \in \mathcal{F},\ A \subseteq A'$.

Proof. $\forall\ \omega,\ \mathrm{Pl}(\{\omega\}) = 1$ if and only if $\omega \in C(\mathcal{F})$. From (36) if $\omega \in C(\mathcal{F})$ then $\mathrm{Pl}(\{\omega\}) = 1 = \mathrm{Pl}'(\{\omega\})$; hence $\omega \in C(\mathcal{F}')$. Besides $\forall\ \omega,\ \mathrm{Pl}(\{\omega\}) > 0 \Leftrightarrow \omega \in S(\mathcal{F})$. From (36) if $\omega \in S(\mathcal{F})$ then $0 < \mathrm{Pl}(\{\omega\}) \leqq \mathrm{Pl}'(\{\omega\})$; hence $\omega \in S(\mathcal{F}')$. To prove (ii), let $A' \in \mathcal{F}'$ contain no focal element in $\mathcal{F}$. Then

$$\mathrm{Pl}\left(\bar{A}'\right) = 1 > 1 - m'\,(A') \geqq \mathrm{Pl}'\left(\bar{A}'\right)$$

which contradicts (36). Q.E.D.

Conditions on the relative structure of $(\mathcal{F}, m)$ and $(\mathcal{F}', m')$ which would be necessary and sufficient to ensure $(\mathcal{F}, m) \subseteq (\mathcal{F}', m')$ seem to be difficult to produce. Inclusion $\subseteq$ has natural properties such as transitivity, and mutual inclusion implies equality (since Cr determines m). Notice also that

$$(\mathcal{F}, m) \cap (\mathcal{F}', m') \subseteq (\mathcal{F}, m) \subseteq (\mathcal{F}, m) \cup (\mathcal{F}', m') . \tag{37}$$

For instance

$$\begin{aligned}\left(\mathrm{Pl} \cap \mathrm{Pl}'\right)(A) &= \sum_{B \cap B' \cap A \neq \varnothing} m(B) \cdot m'(B') \\ &= \sum_{B \cap A \neq \varnothing} m(B) \cdot \left(\sum_{B \cap B' \cap A \neq \varnothing} m'(B') \right) \leqq \mathrm{Pl}(A) .\end{aligned}$$

The other inclusion in (37) can be obtained in a similar way.

If $(\mathcal{F}, m)$ and $(\mathcal{F}', m')$ both generate possibility measures with possibility distributions $\pi = \mu_F$ and $\pi' = \mu_{F'}$, then

$$(\mathcal{F}, m) \subseteq (\mathcal{F}', m') \Leftrightarrow F \subseteq F' \quad \text{(i.e. } \mu_F \leqq {\mu_F}'). \tag{38}$$

That is, the inclusion of bodies of evidence is completely consistent with Zadeh's[35] inclusion of fuzzy sets, hence with the usual inclusion in 2^Ω. To see it just notice that if $(\mathcal{F}, m) \subseteq (\mathcal{F}', m')$ then, as a particular case of (36), $\mathrm{Pl}(\{\omega\}) = \mu_F(\omega) \leqq \mathrm{Pl}'(\{\omega\}) = \mu_{F'}(\omega)$. Conversely if $F \subseteq F'$ then $\mathrm{Pl}(A) = \max\{\mu_F(\omega) | \omega \in A\} \leqq \mathrm{Pl}'(A) = \max\{\mu_{F'}(\omega) \mid \omega \in A\}$.

More surprising, and a disquieting fact at first glance, is that the complementation introduced in *2.2.* is *not* order-reversing for $\subseteq$. To see it first notice that due to Proposition 3

$$\forall\ A, \mathrm{Pl}(A) \leqq \mathrm{Pl}'(A) \Leftrightarrow \forall\ A, \bar{Q}(A) \geqq \bar{Q}'(A) \tag{39}$$

where $\bar{Q}$ and $\bar{Q}'$ are the commonality functions of the complementary bodies of evidence $(\neg\mathcal{F}, \bar{m})$, $(\neg\mathcal{F}', \bar{m}')$ respectively. Moreover $(\mathcal{F}, m) \subseteq (\mathcal{F}', m')$ does not imply any inequality between Q and Q', as proved by the following:

Counter-example 1 Let

$$\Omega = \{a, b, c\}, 0 < k < \frac{1}{2}.$$
$$\mathcal{F} = \{\{a\}, \Omega\}; m(\{a\}) = 1 - k, m(\Omega) = k.$$
$$\mathcal{F}' = \{\{a, b\}, \{a, c\}\}; m'(\{a, b\}) = k, m'(\{a, c\}) = 1 - k.$$

Then the reader can check that $(\mathcal{F}, m) \subseteq (\mathcal{F}', m')$; especially $\forall\ A \neq \{b\}, \{c\}, \varnothing,\ \mathrm{Pl}'(A) = 1$ and $\mathrm{Pl}(\{b\}) = \mathrm{Pl}'(\{b\}) = k,\ \mathrm{Pl}(\{c\}) = k < \mathrm{Pl}'(\{c\}) = 1 - k$. But $Q(\{c\}) = k < Q'(\{c\}) = 1 - k,\ Q(\{b, c\}) = k > Q'(\{b, c\}) = 0$. Q.E.D.

This lack of order-reversingness should not hurt our intuition because (36) is meaningul only for disjunctive evidence, but $(\neg\mathcal{F}, \bar{m})$ is conjunctive and the grade of credibility of A deduced from $(\neg\mathcal{F}, \bar{m})$ is $\bar{Q}(A)$. But from

(39) $(\neg\mathcal{F}', \bar{m}')$ is contained in $(\neg\mathcal{F}, \bar{m})$ (remember that Q is a decreasing set-function for set-inclusion), in the sense of a new kind of inclusion, which makes sense only for conjunctive evidence, namely $\overline{\subset}$ such that $(\mathcal{F}, m)\overline{\subset}(\mathcal{F}', m')$ if and only if

$$\forall\ A, Q\left(A\right) \leqq Q'\left(A\right). \tag{40}$$

$\overline{\subset}$ can be called ‘conjunctive inclusion’ while $\subseteq$ is called ‘disjunctive inclusion’, respectively abbreviated as c-inclusion and d-inclusion.

Note that c-inclusion is transitive, that mutual c-inclusion means equality (since Q determines m as well). Moreover

$$(\mathcal{F}, m) \cap (\mathcal{F}', m')\,\overline{\subset}\,(\mathcal{F}, m)\,\overline{\subset}\,(\mathcal{F}, m) \cup (\mathcal{F}', m') \tag{41}$$

which is simply (37) transformed by complementation. Similarly, c-inclusion applied to possibility measures is equivalent to Zadeh's[35] inclusion of fuzzy sets, i.e. a counterpart of (38) holds. A necessary condition to get (40) is given now:

Proposition 5. *If* $(\mathcal{F}, m)\overline{\subset}(\mathcal{F}', m')$ *then*

i) $S(\mathcal{F}) \subseteq S(\mathcal{F}'),\ C(\mathcal{F}) \subseteq C(\mathcal{F}')$,
ii) $\forall\ A \in \mathcal{F}, \exists\ A' \in \mathcal{F}',\ A \subseteq A'$.

Proof. (i) is easily seen due to $S(\neg\mathcal{F}) = \overline{C(\mathcal{F})},\ C(\neg\mathcal{F}) = \overline{S(\mathcal{F})}$ using complementation to turn $\overline{\subset}$ into $\overline{\supset}$. Now let $A \in \mathcal{F}$ be contained in no focal element in $\mathcal{F}'$ then

$$Q(A) \geqq m\left(A\right) > 0 = Q'\left(A\right)$$

which contradicts (40). Q.E.D.

At this point it is natural to define a third concept of inclusion which requires both (36) and (40) to hold:

Definition 1. $(\mathcal{F}, m)$ *is said to be included in* $(\mathcal{F}', m')$*, denoted* $(\mathcal{F}, m) \Subset (\mathcal{F}', m')$ *if and only if* $(\mathcal{F}, m)$ *is both c-included and d-included in* $(\mathcal{F}', m')$*.*

Inclusion is transitive, mutual inclusion is equality, (37) and (38) hold for $\Subset$. (Note that c-inclusion and d-inclusion are already equivalent for possibility measures). Moreover the complementation is order-reversing for $\Subset$.

Strong Inclusion

Yager[33] has introduced a fourth definition of inclusion in $\mathcal{B}(\Omega)$, which, for reasons to be clarified below, can be called *strong inclusion*, and will be denoted $\subset\!\!\!\subset\!\!\!\subset$. This concept can be presented as follows.

Definition 2. $(\mathcal{F}, m) \subset\!\!\!\subset\!\!\!\subset (\mathcal{F}', m')$ *if and only if the three following statements are valid:*

i) $\forall\ A_i \in \mathcal{F},\ \exists\ A'_j \in \mathcal{F}',\ A_i \subseteq A'_j$,
ii) $\forall\ A'_j \in \mathcal{F}',\ \exists\ A_i \in \mathcal{F},\ A_i \subseteq A'_j$,

iii) there exists a matrix W *with size* $m \times n$, $m = |\mathcal{F}|$, $n = |\mathcal{F}'|$, *whose entries are* $W_{ij} \in [0,1]$ *such that* $W_{ij} > 0 \Rightarrow A_i \subseteq A'_j$, $\sum_{ij} W_{ij} = 1$ *and the basic assignments* m *and* m' *can be expressed in terms of the* W_{ij}*'s as follows:*

$$\forall\ A_i \in \mathcal{F}, m(A_i) = \sum_{\substack{j \\ A_i \subseteq A'_j}} W_{ij}, \tag{42}$$

$$\forall\ A'_j \in \mathcal{F}', m'\left(A'_j\right) = \sum_{\substack{i \\ A_i \subseteq A'_j}} W_{ij}. \tag{43}$$

Note that (42) and (43) look like flow conservation equations in a flow network (Ford and Fulkerson[11]). This analogy is explained in the appendix and is useful to make Definition 2 work. The name 'strong inclusion' is justified by the following result:

Proposition 6. *Strong inclusion implies inclusion i.e.*

$$(\mathcal{F}, m) \Subset (\mathcal{F}', m') \Rightarrow \forall\ A, Q(A) \leqq Q'(A),\ \mathrm{Pl}(A) \leqq \mathrm{Pl}'(A).$$

The converse does not hold.

Proof. Assume $(\mathcal{F}, m) \Subset (\mathcal{F}', m')$.

$$\mathrm{Pl}'(B) = \sum_{A'_j \cap B \neq \varnothing} m'\left(A'_j\right) = \sum_{i,j} \left\{W_{ij} | A_i \subseteq A'_j; A'_j \cap B \neq \varnothing\right\}$$

but

$$\left\{(i,j)\,|A_i \subseteq A'_j; A'_j \cap B \neq \varnothing\right\} \supseteq \left\{(i,j)\,|A_i \subseteq A'_j; A_i \cap B \neq \varnothing\right\}$$

hence

$$\mathrm{Pl}'(B) \geqq \mathrm{Pl}(B) = \sum_{i,j} \left\{W_{ij} | A_i \subseteq A'_j; A_i \cap B \neq \varnothing\right\}.$$

A similar proof holds for the commonality function. Q.E.D.

That the converse does not hold is indicated by the following:

Counter-example 2 $\Omega = \{a, b, c, d, e\}$. Consider the two normal bodies of evidence:

$$(\mathcal{F}, m) = (\{a,b\}, 0.3); (\{a,c\}, 0.3); (\{c,d\}, 0.3); (\{e\}, 0.1)$$
$$(\mathcal{F}', m') = (\{a,b,c\}, 0.4); (\{a,b,d\}, 0.3); (\{a,c,d\}, 0.2); (\{c,d,e\}, 0.1).$$

To check that $(\mathcal{F}, m) \subseteq (\mathcal{F}', m')$ see on Table 1. To see that no matrix W satisfying (42)–(43) exists, it is enough to verify that the following system of equations has no solution in [0,1]:

$$\begin{cases} 0.4 = 0.3m_{11} + 0.3m_{21}\,(= m'(\{a,b,c\})) \\ 0.3 = 0.3\,(1 - m_{11})\,(= m'(\{a,b,d\}) \\ 0.2 = 0.3\,(1 - m_{21}) + 0.3m_{33}\,(= m'(\{a,c,d\})) \\ 0.1 = 0.1 + 0.3\,(1 - m_{33})\,(= m'(\{c,d,e\}))\,. \end{cases}$$

This system is equivalent to (42)–(43) where the W_{ij}'s have been changed into $W_{ij} = m(A_i)m_{ij}$, with $\sum_j m_{ij} = 1$, which eliminates (42). Deeper

Table 1. Counter-example 2

Events	$\mathrm{Cr}(A)$	$\mathrm{Cr}'(A)$	$Q(A)$	$Q'(A)$
$\{a\}$	0	0	0.6	0.9
$\{b\}$	0	0	0.3	0.7
$\{c\}$	0	0	0.6	0.7
$\{d\}$	0	0	0.3	0.6
$\{e\}$	0.1	0.	0.1	0.1
$\{a,b\}$	0.3	0	0.3	0.7
$\{a,c\}$	0.3	0	0.3	0.6
$\{a,d\}$	0	0	0	0.5
$\{a,e\}$	0.1	0	0	0
$\{b,c\}$	0	0	0	0.4
$\{b,d\}$	0.	0	0	0.3
$\{b,e\}$	0.1	0	0	0
$\{c,d\}$	0.3	0	0.3	0.3
$\{c,e\}$	0.1	0	0	0.1
$\{d,e\}$	0.1	0	0	0.1
$\{a,\ b,\ c\}$	0.6	0.4	0	0.4
$\{a,\ b,\ d\}$	0.3	0.3	0	0.3
$\{a,\ b,\ e\}$	0.4	0	0	0
$\{a,\ c,\ d\}$	0.6	0.2	0	0.2
$\{a,\ c,\ e\}$	0.4	0	0	0
$\{a,\ d,\ e\}$	0.1	0	0	0
$\{b,\ c,\ d\}$	0.3	0	0	0
$\{b,\ c,\ e\}$	0.1	0	0	0
$\{b,\ d,\ e\}$	0.1	0.	0	0
$\{c,\ d,\ e\}$	0.4	0.1	0	0.1
$\{a,\ b,\ c,\ d\}$	0.9	0.9	0	0
$\{a,\ b,\ c,\ e\}$	0.7	0.4	0	0
$\{a,\ b,\ d,\ e\}$	0.4	0.3	0	0
$\{a,\ c,\ d,\ e\}$	0.7	0.3	0	0
$\{b,\ c,\ d,\ e\}$	0.4	0.1	0	0

(The use of Cr or Pl to check the inclusion is indifferent because the bodies of evidence are normal).

understanding about the reasons why this system has no solution is gained in the appendix. Q.E.D.

The nice feature of Definition 2 is that it provides a construction method to build two bodies of evidence $(\mathcal{F}, m)$ and $(\mathcal{F}', m')$ such that one is strongly included in the other. It may act as a sufficient condition for having inclusion in the sense of Definition 1. Namely note that letting

$$\mathcal{F}'(A) = \{A' \in \mathcal{F}', A \subseteq A'\}\,; \qquad \mathcal{F}(A') = \{A \in \mathcal{F}, A \subseteq A'\}$$

then

$$\mathcal{F} = \bigcup_{A' \in \mathcal{F}'} \mathcal{F}(A')\,; \qquad \mathcal{F}' = \bigcup_{A \in \mathcal{F}} \mathcal{F}'(A)\,. \tag{44}$$

Given $(\mathcal{F}', m')$, *all* bodies of evidence $(\mathcal{F}, m) \subset\!\!\!\subset\!\!\!\subset (\mathcal{F}', m')$ can be obtained by the following procedure:

Procedure a $\forall\ A'_j \in \mathcal{F}'$ dispatch the weight $m'(A'_j)$ among any family $\mathcal{F}(A'_j)$ of subsets of A'_j, letting W_{ij} be the share of $m'(A'_j)$ allocated to $A_i \in \mathcal{F}(A'_j)$.

Define $\mathcal{F}$ and m by (44) and (42) respectively.

Similarly, given $(\mathcal{F}, m)$ all bodies of evidence $(\mathcal{F}, m) \subset\!\!\!\subset\!\!\!\subset (\mathcal{F}', m')$ can be obtained by the dual procedure.

Procedure b $\forall\ A_i \in \mathcal{F}$, dispatch the weight $m(A_i)$ among any family $\mathcal{F}'(A_i)$ of supersets of $A_i(\mathcal{F}'(A_i) \subseteq \{A | A_i \subseteq A\})$ letting W_{ij} be the share of $m(A_i)$ allocated to $A'_j \in \mathcal{F}'(A_i)$.

Define $\mathcal{F}'$ and m' by (44) and (43) respectively.

Note that a particular case of Procedure (a) is obtained by forcing $\mathcal{F}(A'_j)$ to contain only singletons (provided that $(\mathcal{F}', m')$ is normal). We then recover Dempster's[1] procedure to generate the set $\mathcal{C}(\mathcal{F}')$ of probability measures satisfying (11) as recalled in *1.2*. Procedure (a) thus generalizes Dempster's procedure, but cannot produce *all* bodies of evidence $(\mathcal{F}, m)$ satisfying (34), as indicated in Proposition 6. Procedure (b) was first suggested by Yager[33] who gives it as the very definition of inclusion in $\mathcal{B}(\Omega)$.

Inclusion $\subset\!\!\!\subset\!\!\!\subset$ is transitive. To see it, rewrite (42), (43) under the form

$$\forall\ A'_j \in \mathcal{F}', m'(A'_j) = \sum_{i=1}^{m} m(A_i) \cdot m_{ij}$$

as done in the proof of Proposition 6. Let M be the matrix with coefficient m_{ij}, $\mathbf{m}$ and $\mathbf{m}'$ be the column vectors expressing the basic assignments. Then using matrix notation:

$$\mathbf{m}' = M\mathbf{m}. \tag{45}$$

Now $(\mathcal{F}, m) \subset\!\!\!\subset\!\!\!\subset (\mathcal{F}', m')$ and $(\mathcal{F}', m') \subset\!\!\!\subset\!\!\!\subset (\mathcal{F}'', m'')$ translate into $\mathbf{m}' = M\mathbf{m}, \mathbf{m}'' = M'\mathbf{m}'$, where M, M' belong to the class $\mathcal{M}$ of Markovian matrices, i.e. with positive entries summing to 1 on each row. Hence $\mathbf{m}'' = M'M\mathbf{m}$, and thus $(\mathcal{F}, m) \subset\!\!\!\subset\!\!\!\subset (\mathcal{F}'', m'')$ since $\mathcal{M}$ is closed under matrix product, and (44) holds between $\mathcal{F}$ and $\mathcal{F}''$ as is straightforwardly checked.

Of course, mutual strong inclusion of two bodies of evidence means their equality. Strong inclusion applied to possibility measures is consistent with Zadeh's inclusion of fuzzy set:

Proposition 7. *If* $(\mathcal{F}, m)$ *and* $(\mathcal{F}', m')$ *are consonant then*

$$(\mathcal{F}, m) \subset\!\!\subset\!\!\subset (\mathcal{F}', m') \text{ if and only if } \mu_F \leqq \mu_{F'}$$

where μ_F *and* $\mu_{F'}$ *are the contour functions of* $(\mathcal{F}, m)$ *and* $(\mathcal{F}', m')$.

Proof. The difficult part is to prove that Zadeh's inclusion of fuzzy sets implies the existence of a matrix W satisfying (42) and (43). The proof is given through network flow theory arguments in the appendix, which gives a constructive procedure to build W. Q.E.D.

Lastly $\subset\!\!\subset\!\!\subset$ is order-reversing in $\mathcal{B}(\Omega)$ since Procedures (a) and (b) exchange via complementation. Inequalities (37) hold for the strong inclusion. Note that (i) and (ii) of Definition 2 hold between $\mathcal{F}$ and $\mathcal{F}'' = \{A \cup B' | A \in \mathcal{F}, B' \in \mathcal{F}'\}$. Moreover define $W_{ij} = m(A_i) \cdot m'(B'_j)$ as the share of $m(A_i)$ allocated to the focal element $A_i \cup B'_j$.

Properties of $\mathcal{B}(\Omega)$ under Inclusions

Any of the introduced inclusions equips $\mathcal{B}(\Omega)$ with a partial ordering structure (reflexive, transitive and weakly antisymmetric, that is xRy and yRx implies $x = y$). $\subset\!\!\subset\!\!\subset$ is able to compare less elements in $\mathcal{B}(\Omega)$ than $\subset\!\!\subset$, which in turn is able to compare less elements in $\mathcal{B}(\Omega)$ than any of $\subseteq$ and $\overline{\supset}$. However on $\pi(\Omega) = [0,1]^\Omega$, the set of possibility measures (or fuzzy sets), all four inclusions collapse into Zadeh's fuzzy set inclusion.

The greatest element in $\mathcal{B}(\Omega)$ in the sense of any inclusion is the total ignorance function ($m(\Omega) = 1$) and the least element is the empty body of evidence ($m(\varnothing) = 1$). The least elements in $\mathcal{B}^+(\Omega)$, i.e. normal bodies of evidence, are the probability measures. This is in the sense of disjunctive inclusion $\subseteq$. Indeed, because $\sum_\omega P(\{\omega\}) = 1$, probability measures are not comparable using (34) or (35). This is consistent with the idea that probability measures are sort of 'fuzzy points' (Höhle[16]) for which inclusion is meaningless (there is equality or disjointness!). Moreover given a normal body of evidence $(\mathcal{F}, m)$ any probability measure in $\mathcal{C}(\mathcal{F})$ is contained in $(\mathcal{F}, m)$ in the sense of disjunctive inclusion.

Probability measures are always interpreted in the disjunctive information framework (an event A occurs if and only if $\exists\, \omega \in A$ which is observed, and not only if *all* $\omega \in A$ are observed at the same time). Hence the commonality function Q is not interesting for probabilistic bodies of evidence ($Q(A) = 0$ as soon as $|A| > 1$). Hence probability measures have no interesting role in $(\mathcal{B}^+(\Omega), \overline{\subset})$. However they are still the least elements in $(\mathcal{B}^+(\Omega), \subset\!\!\subset\!\!\subset)$, because any probability measure in $\mathcal{C}(\mathcal{F})$ is strongly included in $(\mathcal{F}, m)$, from Dempster's[1] construction.

Lastly there is an interesting convexity property related to the inclusions:

Proposition 8. *The following subsets of* $\mathcal{B}(\Omega)$ *are convex:*

$$\{(\mathcal{F}, m) | (\mathcal{F}, m) R (\mathcal{F}', m')\}$$
$$\{(\mathcal{F}', m') | (\mathcal{F}, m) R (\mathcal{F}', m')\}$$

with $R = \subseteq,\ \overline{\subset},\ \Subset,\ \subset\!\!\!\subset\!\!\!\subset$.

Proof. Using the definition of the convex combination of two bodies of evidence $(\mathcal{F}, m)$ and $(\mathcal{G}, n)$ i.e. $\alpha(\mathcal{F}, m) + (1-\alpha)\ (\mathcal{G}, n) = (\mathcal{F} \cup \mathcal{G}, \alpha m + (1-\alpha)n)$ with credibility measure $\mathrm{Cr} = \alpha \mathrm{Cr}_m + (1-\alpha)\mathrm{Cr}_n$, it is obvious that Proposition 8 holds for $R = \subseteq$. Now $Q = \alpha Q_m + (1-\alpha)Q_n$ as well, so that Proposition 8 holds for $R = \overline{\subset}$ and $\Subset$. Lastly if $(\mathcal{F}, m)$ and $(\mathcal{G}, n)$ are strongly included in $(\mathcal{F}', m')$ then conditions (i) and (ii) in Definition 2 hold for $\mathcal{F} \cup \mathcal{G}$ with respect to $\mathcal{F}'$. Moreover $\mathbf{m} = M\mathbf{m}'$ and $\mathbf{n} = N\mathbf{m}'$ implies $\alpha\mathbf{m} + (1-\alpha)\mathbf{n} = (\alpha M + (1-\alpha)N)\mathbf{m}'$ where $\alpha M + (1-\alpha)N$ is still a Markovian matrix consistent with the conditions (i) and (ii) in Definition 2. Hence Proposition 8 holds for $\subset\!\!\!\subset\!\!\!\subset$. Q.E.D.

2.4 Projections and Cartesian Product

In this section only normal bodies of evidence are considered.

Let $(\mathcal{F}, m)$ be a body of evidence on a Cartesian product $\Omega = U \times V$. If S is a subset of Ω its projection on U (resp.: V) is denoted $U(S)$ (resp.: $V(S)$) and defined by

$$U(S) = \{u \in U | \exists\ v \in V, (u, v) \in S\}.$$

More generally the projection of $(\mathcal{F}, m)$ on U is $(\mathcal{F}_U, m_U)$ such that (Shafer[25])

$$\forall\ A \subseteq U, m_U(A) = \sum_{S: A = U(S)} m(S). \tag{46}$$

It is easy to check that $(\mathcal{F}_U, m_U)$ induces a plausibility measure Pl_U on U such that $\mathrm{Pl}_U(A) = \mathrm{Pl}(A \times V)$, and a credibility measure Cr_U such that $\mathrm{Cr}_U(A) = \mathrm{Cr}(A \times V)$, which sounds consistent.

Proof.

$$\mathrm{Pl}(A \times V) = \sum_{(A \times V) \cap S \neq \varnothing} m(S) = \sum_{A \cap U(S) \neq \varnothing} m(S) \triangleq \mathrm{Pl}_U(A).$$

Now $\mathrm{Cr}(A \times V) = 1 - \mathrm{Pl}(\bar{A} \times V)$. Q.E.D.

As a consequence, if Pl is a possibility measure Π i.e. its contour function π is a fuzzy relation on $U \times V$, Pl_U is the possibility measure based on the projection of the fuzzy relation (in the sense of Zadeh[36]), since

$$\pi_U(u) = \mathrm{Pl}_U(\{u\}) = \Pi(\{u\} \times V\} = \sup_{v \in V} \pi(u, v).$$

Conversely, given two bodies of evidence $(\mathcal{F}_U, m_U)$ and $(\mathcal{F}_V, m_V)$ on U and V respectively, we can define their cylindrical extensions and define the product of these extensions via Dempster rule (Shafer[25]). Namely, the *cylindrical extension* of $(\mathcal{F}_U,\ m_U)$ is $(c\mathcal{F}_U,\ cm_U)$ such that

$$\forall\ B \subseteq, U, cm_U\,(B \times V) = m_U\,(B)$$

and

$$cm_U\,(A) = 0 \text{ for other } A \subseteq \Omega = U \times V.$$

From $(\mathcal{F}_U, m_U)$, $(\mathcal{F}_V, m_V)$ on U and V respectively, $(\hat{\mathcal{F}},\ \hat{m}) \triangleq (\mathcal{F}_U,\ m_U) \times (\mathcal{F}_V,\ m_V)$, denotes a Cartesian product of bodies of evidence. $\hat{m}$ is calculated by:

$$\begin{aligned} \forall\ A \subseteq \Omega, \hat{m}\,(A) &= m_U\,(B) \cdot m_V\,(C) \quad \text{if} \quad A = B \times C \\ &= 0 \text{ otherwise.} \end{aligned} \tag{47}$$

Note that $\{(B, C) | A = B \times C\} = \{(U(A),\ V(A))\}$ and $B \times C = \varnothing$ only if B or $C = \varnothing$ so that Dempster rule really boils down to (47), and $\hat{\mathcal{F}} = \{B \times C | B \in \mathcal{F}_U,\ C \in \mathcal{F}_V\}$. Note that $(\hat{\mathcal{F}},\ \hat{m})$ is always normal since $(\mathcal{F}_U,\ m_U)$ and $(\mathcal{F}_V,\ m_V)$ are supposed to be so.

If $(\mathcal{F}_U,\ m_U)$ and $(\mathcal{F}_V,\ m_V)$ reduce to sets B and C, then their products in the sense of (47) is their Cartesian products. (47) is however not in accordance with Zadeh's[36] definition of the Cartesian product of fuzzy sets since if $(\mathcal{F}_U,\ m_U)$ and $(\mathcal{F}_V,\ m_V)$ are possibility measures, with contour functions μ_F and μ_G respectively then the fuzzy Cartesian product is the possibility measure with contour function $\min(\mu_F,\ \mu_G)$. Rather, (47) implies that $(\hat{\mathcal{F}},\ \hat{m})$ is generally not a fuzzy Cartesian product since it is consistent with $\mu_F \cdot \mu_G$, an operation previously introduced by the authors[3]; moreover $(\hat{\mathcal{F}},\ \hat{m})$ defines no possibility measure, generally.

The natural thing to do is now to project $(\mathcal{F},\ m)$ on U and V and recombine their projections. One may expect some relationship between $(\mathcal{F},\ m)$ and $(\hat{\mathcal{F}},\ \hat{m})$ in terms of specificity, namely that $(\mathcal{F}, m)$ is included in $(\hat{\mathcal{F}},\ \hat{m})$; unfortunately this property does not hold as shown below.

Counter example 3 $\mathcal{F} = \{S_1,\ S_2\}$ with $S_1 \cap S_2 = \varnothing$, $U(S_1) \cap U(S_2) = \varnothing$, $V(S_1) \cap V(S_2) = \varnothing$. $\hat{\mathcal{F}} = \{U(S_i) \times V(S_j) | i = 1, 2;\ j = 1, 2\}$. $\hat{\mathcal{F}}$ is made of four disjoint focal elements.

Now since $S = U(S_1) \times V(S_2) \notin \mathcal{F}$, $\mathrm{Cr}(S) = 0$ while $\hat{\mathrm{C}}\mathrm{r}(S) = m(S_1) \cdot m(S_2) > 0$. Moreover $\mathrm{Cr}(S_1 \cup S_2) = 1$ while $\hat{\mathrm{C}}\mathrm{r}(S_1 \cup S_2) \leqq m(S_1)^2 + m(S_2)^2 < 1$ (the equality holds if S_1 and S_2 are Cartesian products).

More particularly, if $(\mathcal{F},\ m)$ is a probabilistic body of evidence then $(\hat{\mathcal{F}},\ \hat{m})$ also generates a probability measure, and no inclusion must be expected, relating these two bodies of evidence.

However, if $(\mathcal{F},\ m)$ is consonant, this relationship might be expected to hold. The following result leaves no hope about it for the d-inclusion.

Proposition 9. *Even if* $(\mathcal{F},\ m)$ *is consonant, the property* $(\mathcal{F},\ m) \subseteq (\hat{\mathcal{F}},\ \hat{m})$ *does not hold.*

Counter example 4 $\mathcal{F} = \{S_1, S_2\},\ S_1 \subset S_2$, with $U(S_1) \neq U(S_2),\ V(S_1) \neq V(S_2)$. Let $\alpha = m(S_1)$. Hence, $\hat{m}(U(S_1) \times V(S_1)) = \alpha^2,\quad \hat{m}(U(S_1) \times V(S_2)) = \hat{m}(U(S_2) \times V(S_1)) = \alpha(1-\alpha),\ \hat{m}(U(S_2) \times V(S_2)) = (1-\alpha)^2$.

Now assume S is such that:

$$(U(S_1) \times V(S_2)) \cup (U(S_2) \times V(S_1)) \subset S \subset S_2$$

where the inclusions are strict. It is easy to figure out that such a set S may exist. Then we have:

$$\mathrm{Cr}(S) = \alpha < \hat{\mathrm{Cr}}(S) = \alpha(2-\alpha),\ \forall \alpha < 1. \quad \text{Q.E.D.}$$

Note that the c-inclusion does not hold either. Indeed assume that $S_i \neq U(S_i) \times V(S_i)$ for $i = 1, 2$ in the above counter example. Clearly,

$$Q(U(S_1) \times V(S_1)) = 1 - \alpha < \hat{Q}(U(S_1) \times V(S_1)) = 1$$
$$Q(S_2) = 1 - \alpha > \hat{Q}(S_2) = (1-\alpha)^2$$

since $S_2 \not\subset U(S_i) \times V(S_j),\ i \neq j$.

Proposition 9 contrasts with a well-known result in fuzzy set theory, due to Zadeh[36]. Namely, a fuzzy relation R on $U \times V$ is included in the Cartesian product of its projections. The inclusion turns into an equality if and only if $\mu_R(u,\ v)$ is of the form $\min(\mu_A(u)\mu_B(v))$ where A and B are fuzzy sets on U and V respectively. In the *possibilitistic case* it is interesting to specify conditions under which $(\mathcal{F},\ m) = (\hat{\mathcal{F}},\ \hat{m})$.

First any $S \in \mathcal{F}$ must be of the form $A \times B$. Then $\mathcal{F} \subseteq \hat{\mathcal{F}}$ is ensured. Now $\mathcal{F} = \{A_i \times B_i | i = 1, p\}$ and $\hat{\mathcal{F}} = \{A_i \times B_j | i = 1, p, j = 1, p\}$. Let $A_i \neq A_j$ and $B_i \neq B_j$, then one of $A_i \times B_j,\ A_j \times B_i$ should not be in $\hat{\mathcal{F}}$ since there is no inclusion relationship between them. So, to preserve a nested structure in $\hat{\mathcal{F}}$ we must have $\forall\ i,\ j,\ A_i = A_j$ or $\forall\ i,\ j,\ B_i = B_j$. Hence the following result, stated in the case when $\exists\ B_i \neq B_j$:

Proposition 10. *If* $(\mathcal{F},\ m)$ *is consonant,* $(\hat{\mathcal{F}},\ \hat{m}) = (\mathcal{F},\ m)$ *if and only if*

$$\exists\ A \subseteq U,\ B_1 \subset B_2 \subset \cdots \subset B_p \subset V$$

such that

$$\mathcal{F} = \{A \times B_i | i = 1, p\}.$$

This is equivalent to state that $m_U(A) = 1$, i.e. the projection of $(\mathcal{F},\ m)$ on U is a set. So that $m_V(B_i) = m(B_i)\ \forall\ i$.

In the general case, the concept of inclusion, even the weaker one proves too strong to be able to compare $(\mathcal{F}, m)$ and the product of its projections. Such a comparison can be however carried out using the measures of uncertainty and specificity respectively introduced by Higashi and Klir[15] and Yager[30]. Then some interesting inequalities can be obtained expressing that $(\hat{\mathcal{F}}, \hat{m})$ is not more specific than $(\mathcal{F}, m)$ (see Dubois and Prade[9]). Note that we may have $(\mathcal{F}, m) = (\hat{\mathcal{F}}, \hat{m})$ in the general case, since when $\mathcal{F}$ is not consonant the requirement $\mathcal{F} = \hat{\mathcal{F}}$ does not induce the same constraints on $\mathcal{F}$, as in the consonant case.

3 Consonant Approximation of a Body of Evidence

It is easier to deal with a possibility measure or a probability measure rather than with a general plausibility measure. The main reason is that in both cases, the body of evidence is completely characterized by its contour function, i.e. a probability allocation or a fuzzy set. The question of approximation of a body of evidence by either a probability measure or a possibility measure is thus worth considering.

3.1 The Approximation Problem

A body of evidence $(\mathcal{F}', m')$ can be viewed as a valid substitute of $(\mathcal{F}, m)$ as soon as $(\mathcal{F}, m) \subseteq (\mathcal{F}', m')$ (here we assume bodies of evidence are disjunctive). This is a generalized version of Zadeh's entailment principle[39], and it encompasses Yager's[33] proposal based on the strong inclusion. Moreover the knowledge of another body of evidence $(\mathcal{F}'', m'') \subseteq (\mathcal{F}, m)$ enables the plausibility measure associated with $(\mathcal{F}, m)$ to be located in an interval, i.e.

$$\forall\ A, \mathrm{Pl}''(A) \leqq \mathrm{Pl}(A) \leqq \mathrm{Pl}'(A)\,. \tag{48}$$

A related inequality holds for the credibility function, of course. Whenever (48) holds the pair $(\mathrm{Pl}'',\ \mathrm{Pl}')$ is said to be an approximation of Pl. Pl'' is the lower approximation, Pl' the upper approximation.

The approximation problem[2] for bodies of evidence can then be stated as follows: Let $\mathcal{A}$ be a suitable subset of $\mathcal{B}(\Omega)$ containing 'simple' bodies of evidence, in the sense that it is easy to deal with them for some reason. Given any body of evidence $(\mathcal{F},\ m) \notin \mathcal{A}$, find two bodies of evidence $(\mathcal{F}^*,\ m^*)$ and $(\mathcal{F}_*,\ m_*)$ in $\mathcal{A}$, upper and lower approximations of $(\mathcal{F},\ m)$ i.e.

[2] An example of this approximation methodology can be found in the recent paper by J. Gordon and E. H. Shortliffe: "A method for managing evidential reasoning in a hierarchical hypothesis space," *Artificial Intelligence*, **26**, 1985, pp. 323–357. In this paper the authors are looking for an approximation of the result of the combination of several bodies of evidence by means of Dempster rule because the exact result would be too difficult to compute.

$$(\mathcal{F}_*, m_*) \subseteq (\mathcal{F}, m) \subseteq (\mathcal{F}^*, m^*) \,. \tag{49}$$

Moreover $(\mathcal{F}_*, m_*)$ and $(\mathcal{F}^*, m^*)$ should be best approximations in the following sense: denote $\mathcal{A}^+(\mathcal{F}, m)$ and $\mathcal{A}_-(\mathcal{F}, m)$ the sets

$$\mathcal{A}^+ (\mathcal{F}, m) = \{(\mathcal{F}', m') \mid (\mathcal{F}, m) \subseteq (\mathcal{F}', m')\} \cap \mathcal{A}$$
$$\mathcal{A}^- (\mathcal{F}, m) = \{(\mathcal{F}'', m'') \mid (\mathcal{F}'', m'') \subseteq (\mathcal{F}, m)\} \cap \mathcal{A}$$

and let $\mathcal{A}^+_*(\mathcal{F}, m)$ (resp.: $\mathcal{A}^*_-(\mathcal{F}, m)$) be the set of minimal (resp.: maximal) elements in $\mathcal{A}^+(\mathcal{F}, m)$ (resp.: $\mathcal{A}_-(\mathcal{F}, m)$). Then we should require $(\mathcal{F}_*, m_*) \in \mathcal{A}^*_-(\mathcal{F}, m)$ and $(\mathcal{F}^*, m^*) \in \mathcal{A}^+_*(\mathcal{F}, m)$.

Clearly it is meaningless to choose $\mathcal{A}$ as being the set $\mathcal{P}(\Omega)$ of probability measures because an upper approximation will never exist when $(\mathcal{F}, m)$ is normal (except if $(\mathcal{F}, m)$ generate a probability measure) and a lower approximation only exists if $(\mathcal{F}, m)$ is normal. In such a case $\mathcal{A}^*_-(\mathcal{F}, m) = \mathcal{A} - (\mathcal{F}, m) = \mathcal{C}(\mathcal{F})$ since probability measures do not compare with one another via $\subseteq$. So all probability measures are equally candidate as lower approximations.

A member of $\mathcal{C}(\mathcal{F})$ is especially interesting and has been suggested by the authors[4, 5] previously. It is obtained by equally sharing the weights $m(A)$ among elements of A; we then have

$$\forall\, \omega \in \Omega, P\left(\{\omega\}\right) = \sum_{\omega \in A} \frac{m\left(A\right)}{|A|}. \tag{50}$$

(50) is in accordance with Laplace's principle of modeling a lack of information by uniformly distributed probability allocations. When $(\mathcal{F}, m)$ is consonant (50) defines a bijection between probability measures and possibility measures on a finite set, and the converse mapping can be useful to derive a possibilistic interpretation of histograms as explained in Dubois and Prade[4, 5].

3.2 Possibilistic Approximations of Normal Bodies of Evidence

A more satisfactory approach is to consider the set $[0,1]^\Omega$ of consonant bodies of evidence as the approximation set $\mathcal{A}$. In this section we derive best upper and lower approximations of $(\mathcal{F}, m)$ when $\mathcal{A} = [0,1]^\Omega$. The best lower approximation Π_* is first derived. The following result was already obtained in Dubois and Prade[4].

Proposition 11. *The best lower approximation in* $[0,1]^\Omega$ *of a body of evidence* $(\mathcal{F}, m)$ *is unique and is the possibility measure* Π_**, whose possibility distribution* π_* *is the contour function of* $(\mathcal{F}, m)$.

Proof. [4]

$$\forall\, A, \mathrm{Pl}\left(A\right) = \sum_{B \subseteq \Omega} m\left(B\right) \cdot \sup_{\omega \in A} \mu_B\left(\omega\right),$$

where μ_B is the characteristic function of B. Hence

$$\mathrm{Pl}\,(A) \geqq \sup_{\omega \in A} \sum_{B \in \Omega} \mu_B\,(\omega)\, m\,(B) \triangleq \sup_{\omega \in A} \mathrm{Pl}\,(\{\omega\})\,.$$

Let Π_* be the possibility measure such that $\Pi_*(\{\omega\}) = \mathrm{Pl}(\{\omega\})$, clearly $\Pi_* \in \mathcal{A}_-(\mathcal{F}, m)$. Let Π be a possibility measure such that $\Pi \leqq \mathrm{Pl}$. Then $\forall\, \omega \in \Omega,\ \pi(\omega) \leqq \mathrm{Pl}(\{\omega\}) = \pi_*(\omega)$. Hence $\Pi \leqq \Pi_*$. Q.E.D.

Π_* is defined for any $(\mathcal{F}, m) \in \mathcal{B}(\Omega)$. However if $(\mathcal{F}, m)$ is not consistent (i.e. the core $C(\mathcal{F})$ is empty), Π_* is not normal, while it is always normal otherwise, since $\mathrm{Pl}(\{\omega\}) = 1, \forall\, \omega \in C(\mathcal{F})$. As a consequence the lower approximation is completely meaningful for consistent bodies of evidence. Obviously, if $(\mathcal{F}, m)$ is consonant, then $\mathrm{Pl} = \Pi_*$. At the opposite if $(\mathcal{F}, m)$ defines a probability measure then $\pi_*(\omega) = P(\{\omega\})$, which is not very interesting.

The use of the contour function of $(\mathcal{F}, m)$ has been suggested by Zhang[43, 44] and Wang[27] to derive the membership of a fuzzy set from statistical data made of error intervals.

The set of focal elements of the lower approximation is $\mathcal{F}_*$ defined by

$$\mathcal{F}_* = \{\{\omega | \mathrm{Pl}\,(\{\omega\}) \geqq \alpha\}\, | \alpha \in]0, 1]\}$$

and letting $\alpha_1 = 1 > \alpha_2 \cdots > \alpha_p > 0$ be the elements of the set $\{\mathrm{Pl}(\{\omega\}) | \omega \in \Omega\} \cup \{1\}$. $\mathcal{F}_*$ contains p focal elements $A_1 \subset A_2 \subset \cdots \subset A_p$ with $A_i = \{\omega | \mathrm{Pl}(\{\omega\}) \geqq \alpha_i\}$. $A_1 \neq \varnothing$ if and only if $(\mathcal{F}, m)$ is consistent. Indeed it is easy to see that A_1 is the core of $(\mathcal{F}, m)$ i.e.

$$C\,(\mathcal{F}) = \{\omega, \forall\ A \in \mathcal{F}, \omega \in A\}$$

and A_p is the support of $(\mathcal{F}, m)$, i.e.

$$S\,(\mathcal{F}) = \{\omega, \exists\ A \in \mathcal{F}, \omega \in A\}\,.$$

Hence $(\mathcal{F}, m)$ and $(\mathcal{F}_*, m_*)$ have the same core and support. This remark enables a member of $\mathcal{A}^+(\mathcal{F}, m)$ to be constructed from the knowledge of $\mathcal{F}_*$; to do it we use a technique described in Dubois and Prade[6], which is an alternative way of deriving a membership function from a set of statistical data consisting of error intervals. This technique, which contrasts with Zhang and Wang's approach goes as follows.

i) Define a mapping $f : \mathcal{F} \to \mathcal{F}_*$ where $f(A)$ is the smallest A_i containing A, i.e.

$$f(A) = A_{\mathrm{i}} \quad \text{such that } A \subseteq A_{\mathrm{i}}, A \not\subset A_{\mathrm{i}-1}.$$

ii) Let $(\mathcal{F}^*, m^*)$ be such that $\mathcal{F}^* = f(\mathcal{F}) \subseteq \mathcal{F}_*$

$$\forall\ A_i, m^*\,(A_{\mathrm{i}}) = \sum_{A_{\mathrm{i}} = f(A)} m\,(A)\,.$$

Note that $f(A)$ is never empty since $\forall\, A,\ A \subseteq A_p$. Moreover f defines a partition of $\mathcal{F}$ through the equivalence relation $\sim$: $A \sim B \Leftrightarrow f(B) = f(A)$. Hence

$$\sum_{A_i \in \mathcal{F}_*} m^*\left(A_{\mathrm{i}}\right) = \sum_{A \in \mathcal{F}} m\left(A\right).$$

It is easy to check that $(\mathcal{F}, m)$ is strongly included in $(\mathcal{F}^*, m^*)$ since the above technique is a particular case of Procedure (b) of 2.3.2. where the whole mass $m(A)$ is allocated to $f(A)$. In Dubois and Prade,[6] however, the sets $A_1 \ldots A_p$ are given independently of $(\mathcal{F}, m)$ except that $A_1 = C(\mathcal{F}),\ A_p = S(\mathcal{F})$, and $A_1 \neq \varnothing$ i.e. the procedure was defined only for consistent bodies of evidence. Here we improve it by prescribing what are the focal elements A_i for $1 < i < p$. We now prove that $(\mathcal{F}^*, m^*)$ is a best approximation in some sense. Let π and π' be two possibility distributions on Ω, π and π' are said to be *order-equivalent* if and only if

$$\forall\ \omega, \omega',\ \pi(\omega) > \pi\left(\omega'\right) \Leftrightarrow \pi'\left(\omega\right) > \pi'\left(\omega'\right). \tag{51}$$

Order-equivalence can be nicely characterized in terms of focal elements:

Lemma 1. *π and π' are order-equivalent, if and only if their associated sets of focal elements are equal.*

Proof. Let $\{\alpha_1, \ldots, \alpha_p\} = \{\pi(\omega) > 0 | \omega \in \Omega\}$. It is well-known that $\mathcal{F} = \{A_1, \ldots, A_p\}$ where $i = 1, p,\ A_i = \{\omega | \pi(\omega) \geqq \alpha_i\}$.[4, 7] Let $\omega_i \in A_i$ such that $\pi(\omega_i) = \alpha_i$. Now from order-equivalence $A_i = \{\omega | \pi(\omega) \geqq \pi(\omega_i)\} = \{\omega | \pi'(\omega) \geqq \pi'(\omega_i)\}$. Hence $A_i \in \mathcal{F}'$, the set of focal elements of π'. Hence $\mathcal{F} \subseteq \mathcal{F}'$ and $\mathcal{F}' \subseteq \mathcal{F}$ since π and π' play the same role. The converse proposition is obvious. Q.E.D.

Note that π_* and π^* are generally not order-equivalent but satisfy the weaker statement

$$\forall\ \omega, \omega', \pi_*\left(\omega\right) > \pi_*\left(\omega'\right) \Rightarrow \pi^*\left(\omega\right) \geqq \pi^*\left(\omega'\right). \tag{52}$$

This is because generally $\mathcal{F}^* \subset \mathcal{F}_*$.

Proposition 12. *$(\mathcal{F}^*, m^*)$ is the best upper approximation of $(\mathcal{F}, m)$ among its order-equivalent consonant bodies of evidence.*

Proof. Let N^* and Π^* be the necessity and possibility measures induced by $(\mathcal{F}^*, m^*)$. First note that

$$\forall\ A_i^* \in \mathcal{F}^*, \mathrm{Cr}\left(A_i^*\right) = \sum_{A \subseteq A_i^*} m\left(A\right) = \sum_{j=1}^{i} m^*\left(A_{\mathrm{j}}^*\right) = N^*\left(A_i^*\right).$$

Let $(\mathcal{F}', m')$ be an upper approximation of $(\mathcal{F}, m)$. Because π' and π^* are order-equivalent, the Lemma yields $\mathcal{F}' = \mathcal{F}^*$. Now the inequality $N'(A) \leqq \mathrm{Cr}(A),\ \forall\ A$ implies $\forall\ A_i^* \in \mathcal{F}^*,\ N'(A_i^*) \leqq N^*(A_i^*)$ which also reads

$$\forall\, i \neq p, \max\left\{\pi'\left(\omega\right)|\omega \notin A_i^*\right\} \geqq \max\left\{\pi^*\left(\omega\right)|\omega \notin A_i^*\right\}. \qquad (53)$$

Now the maximum in both sides of (53) is reached by any element ω in $A_{i+1}^* \cap \bar{A}_i^*$ since A_i^* is a focal element in both $\mathcal{F}'$ and $\mathcal{F}^*$. Hence (53) translates into:

$$\forall\, \omega \notin A_1^*, \pi'\left(\omega\right) \geqq \pi^*\left(\omega\right).$$

Moreover $\forall\, \omega \in A_1^*,\ \pi'(\omega) = \pi^*(\omega) = 1$, since $C(\mathcal{F}') = C(\mathcal{F}^*)$. Q.E.D.

The condition of order-equivalence is a necessary one to get optimality. Indeed (53) implies only the existence of some ω' in $\bar{A}_i^*$ such that $\forall\, \omega \in A_{i+1}^* \cap \bar{A}_i^*,\ \pi'(\omega') \geqq \pi^*(\omega)$.

Counter example 5 $\Omega = \{a, b, c, d, e\}$

$$\mathcal{F} = \left\{\{c\}, \{c, d\}, \{b, c\}, \{c, d, e\}, \{a, b, c\}\right\}$$

with a uniformly distributed basic assignment $m(A) = \frac{1}{5}, \forall\, A \in \mathcal{F}$.

We have the following results, where π_* and π^* are calculated from m and π' is given:

	a	b	c	d	e
π_*	0.2	0.4	1	0.4	0.2
π_*	0.4	0.8	1	0.8	0.4
π'	0.4	0.6	1	0.8	0.4

Note that $\Pi^*(A) = \Pi'(A)$ except for $A = \{b\},\ \{a, b\},\ \{b, e\}$, for which $\Pi'(A) = 0.6 < \Pi^*(A) = 0.8$. But $\Pi' \geqq \mathrm{Pl}$ since $\mathrm{Pl}(\{a\}) = 0.2,\ \mathrm{Pl}(\{a, b\}) = 0.4$ and $\mathrm{Pl}(\{b, e\}) = \mathrm{Pl}(\{a, b, e\}) = 0.6$. But the distribution π' possesses a dissymmetry which does not look natural since neither Pl, nor π^* have such a dissymmetry. Assuming $\pi'(d) = \pi'(b) = 0.6$ is not possible since then (53) is violated. Q.E.D.

It is clear that if $(\mathcal{F}, m)$ is consonant then $(\mathcal{F}^*, m^*) = (\mathcal{F}, m)$ which shows a good behavior of $(\mathcal{F}^*, m^*)$.

The possibility distribution associated with $(\mathcal{F}^*, m^*)$ is π^* defined by

$$\forall\, \omega \notin A_p^*, \pi^*\left(\omega\right) = 0,$$

$$\forall\, \omega \in A_1^*, \pi^*\left(\omega\right) = 1,$$

$$\forall\, \omega \in A_i^* - A_{\mathrm{i}-1}^*, \pi^*\left(\omega\right) = 1 - N^*\left(A_{i-1}^*\right) = 1 - \sum_{A \subseteq A_{i-1}^*} m\left(A\right)$$

$$= \sum_{j=i}^{p} m^*\left(A_j^*\right).$$

When $(\mathcal{F}, m)$ generates a probability measure, the formula becomes

$$\forall\, \omega, \pi^*\left(\omega\right) = \sum \left\{P\left(\{\omega'\}\right)) | P\left(\{\omega'\}\right) \leqq P\left(\{\omega\}\right)\right\}$$

and if $p_1 \geqq p_2 \ldots \geqq p_n$ are the probability weights on $\omega_1, \ldots, \omega_n$ we get values $\pi_1^* \geqq \cdots \geqq \pi_n^*$ such that

$$\pi_i^* = \sum_{j=i}^{p} p_j, \forall\ i. \tag{54}$$

(54) defines a bijection between probability and possibility measures on Ω since it is equivalent to

$$p_i = \pi_i^* - \pi_{i+1}^*, \quad \forall\ i \tag{55}$$

with $\pi_{n+1}^* = 0$. Equation (54) provides the best possibilistic approximation of a probability measure in the sense of the consistency condition:

$$\forall\ A, N^*(A) \leqq P(A) \leqq \Pi^*(A)$$

and under order-equivalence assumption.

Particularly this transformation provides a more specific result than the converse of (50), proposed in a previous paper[5]. Note that in the case of $(\mathcal{F}, m)$ being a probability allocation, the result was already given in Dubois and Prade[4].

Conclusion

Shafer's theory of evidence seems to make measure theory and logic interfere with each other. Mathematical beings, living in $\mathcal{B}(\Omega)$ have a dual nature: they are kinds of sets (more precisely convex combinations of sets) and as such can combine via logical connectives such as union, intersection and complementation, and consequently *any* connective of classical logic can have an extension in $\mathcal{B}(\Omega)$. But they are also kinds of measures, and concepts of expectations can be defined from them as Dempster did[1]. However, because bodies of evidence first emerged as upper and lower probabilities, the possibility of constructing a logic calculus on them was not really pointed out by Dempster or Shafer, but by the people working in random set theory[12, 14].

Logical operations cannot be introduced in the setting of probability measures because they are generalized *points*, not sets. This is why, may be, logic and probability theory seem to ignore each other. Contrastedly, possibility theory, first discussed in terms of fuzzy sets, was naturally equipped with a logic calculus. The measure-theoretic point of view came afterwards when possibility measures were also viewed as upper probabilities[4].

Similarly, Shafer's book[24], due to a probabilistic background, always assume information is disjunctive, as it must be in probability theory. On the contrary a set is classically viewed as a conjunctive of values as often as a restriction on the value of a variable. The framework of credibility and plausibility measures enables the conjunctive point of view to enter the probabilistic arena, and this is very important for knowledge representation issues. The existence of logical connectives in $\mathcal{B}(\Omega)$ has been exploited by Yager[34] to define new patterns of reasoning which generalize the modus ponens. However, contrary to approximate reasoning, based on fuzzy sets, the choice of the implication connective is very much restricted by the unicity of basic operations such as the union and complementation. This unicity stems from the unicity of Dempster rule under decomposability conditions[10].

As noted by Zadeh, fuzzy set theory is not a particular case of Shafer's theory, although a possibility measure (i.e. a fuzzy set) is a special kind of a body of evidence, where focal element are consonant. The reason is that Shafer's theory needs Dempster rule of combination to perform intersection in $\mathcal{B}(\Omega)$ while fuzzy sets are conjunctively combined by means of triangular norms[8, 23]. Shafer's rationale for Dempster rule stems from probabilistic independence between two basic assignments m and m', viewed as probability allocations on 2^Ω. As a consequence the intersection of two consonant bodies of evidence is generally no longer consonant. On the contrary Zadeh's approach starts from the requirements that any logical combination of fuzzy sets should be a fuzzy set again. This requirement is linked to the fact that possibility distributions model the meaning of imprecise statements, and that the meaning of complex statements should be expressed as some combination of simpler statements that they involve.

Shafer theory is based on the re-interpretation of results by Dempster, results which were cast in a frequentist framework. And indeed Dempster rule has a frequentist flavor, and the development of a frequentist theory of upper and lower probabilities receives attention in the literature[26]. Such attempts combined with results of Sect. 3, can provide grounds for statistical estimation of membership functions[6, 27]. However the possibility of a frequentist interpretation of fuzzy set-theoretic operations seems to be very unlikely, while the connections between these operations and the theory of conjoint measurement[8, 18] are more promising. In other words fuzzy set theory seems to be closer to research in psychological measurement than to statistics, although possibility measures may have frequentist interpretation. The rules of combination of frequentist possibility measures will be dictated by independence-like arguments deriving from the study of statistical experiments, while the rules of combination of subjectivist possibility measures may turn out to be those of fuzzy set theory. The core of the debate is the relevance of subjective probability theory. If subjective probability theory is acknowledged as being too restrictive to model uncertainty judgments, then Shafer's subjectivist interpretation of upper and lower probabilities can be questioned on the same grounds. From a mathematical point of view, the theory of evidence is nothing but the rules of probability theory applied to imprecise statements, while classical probability theory leaves no room to imprecision. As a consequence the rules of combination of bodies of evidence are given by the rules of probability theory, and what is behind the problem of validating Shafer's theory as a theory of measurement of subjective uncertainty is the validity of the rules of (subjective) probability theory (and especially the rule of additivity). From this point of view fuzzy set theory seems to be far less normative than the theory of evidence, although both provide tools for modeling imprecision and uncertainty in a unique setting.

Acknowledgments

This paper has benefited from the help of several people. We wish to thank Prof. Zadeh for his remarks about conjunctive information; Ron Yager for discussions about inclusion of bodies of evidence; and also J. B. Cavaillé for information about network flows, which was useful getting some of the results.

References

1. A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping." *Annals of Mathematical Statistics*, **38**, 1967, pp. 325–339.
2. D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications.* Academic Press, New York, 1980.
3. D. Dubois and H. Prade, "Additions of interactive fuzzy numbers," *IEEE Trans. on Automatic Control*, **26**, 1981, 926–936.
4. D. Dubois and H. Prade, "On several representations of an uncertain body of evidence," *Fuzzy Information and Decision Processes*, edited by M. M. Gupta, E. Sanchez, North-Holland, Amsterdam, 1982, pp. 167–181.
5. D. Dubois and H. Prade, "Unfair coins and necessity measures. Towards a possibilistic interpretation of histograms." *Fuzzy Sets and Systems*, **10**, 1983, pp. 15–20.
6. D. Dubois and H. Prade, "Fuzzy sets and statistical data," *European Journal of Operational Research*, **25**(3), 1986, 345–356.
7. D. Dubois and H. Prade, *Théorie des Possibilités. Applications à la Représentation des Connaissances en Informatique*, Masson, Paris, 1985.
8. D. Dubois and H. Prade, "A review of fuzzy set aggregation connectives," *Information Sciences*, **36**, 1985, pp. 85–121.
9. D. Dubois and H. Prade, "Additivity and monotonicity of measures of information defined in the setting of Shafer's evidence theory." *BUSEFAL* (Universitė P. Sabatier, Toulouse), No. 24, 1985, pp. 64–76.
10. D. Dubois, R. Giles and H. Prade, "On the unicity of Dempster's rule of combination." *International Journal of Intelligent Systems*, **1**(2), 1986, 133–142.
11. L. R. Ford, Jr., D. R. Fulkerson, *Flows in Networks.* Princeton University Press, Princeton, N.J., 1962.
12. R. Fortet and M. Kambouzia, "Ensembles aléatoires et ensembles flous," *Publications Econométriques*, **IX**, fascicule 1, 1976, pp. 1–23.
13. I. R. Goodman, "Fuzzy sets as equivalence classes of random sets." In: *Fuzzy Set and Possibility Theory: Recent Developments*, edited by R. R. Yager, Pergamon Press, Oxford, 1982, pp. 327–342.
14. I. R. Goodman, "Characterization of n-ary fuzzy set operations which induce homomorphic random set operations," In: *Fuzzy Information and Decision Processes*, edited by M. M. Gupta and E. Sanchez, North-Holland, Amsterdam, 1982, pp. 203–212.
15. M. Higashi and G. Klir, "Measures of uncertainty and information based on possibility distributions." *International Journal of General Systems*, **9**, 1983, pp. 43–58.
16. U. Höhle, "Fuzzy filters: a generalization of credibility measures." *Proc. IF AC Symposium on Fuzzy Information, Knowledge Representation, and Decision Analysis*, Marseille, June 1983, pp. 111–114.

17. J. Kampé de Feriet, "Interpretation of membership functions of fuzzy sets in terms of plausibility and belief." In: *Fuzzy Information and Decision Processes*, edited by M. M. Gupta and E. Sanchez, North-Holland, Amsterdam, 1982, pp. 93–98.
18. D. H. Kranz, R. D. Luce, R. Suppes and A. Tversky, *Foundations of Measurement*. Vol. I, Academic Press, New York, 1971.
19. H. T. Nguyen, "On random sets and belief functions." *Journal of Mathematical Analysis and Applications*, **65**, 1978, pp. 531–542.
20. E. M. Oblow, "A hybrid uncertainty theory," *Proc. 5th International Workshop: "Expert Systems and their Applications*," Avignon, May 13–15, 1985, pp. 1193–1201, published by ADI Paris.
21. H. Prade and C. Testemale, "Representation of soft constraints and fuzzy attribute values by means of possibility distributions in databases." In: *The Analysis of Fuzzy Information*, edited by J. C. Bezdek, CRC Press, Boca Raton, Fl., 1986, to appear, Vol. II.
22. T. Sales, "Fuzzy sets as set classes," *Stochastica* (Barcelona, Spain), **VI**, 1982, pp. 249–264.
23. B. Schweizer and A. Sklar, "Associative functions and abstract semi-groups." *Publicationes Mathematicae Debrecen*, **10**, 1963, pp. 69–81.
24. G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, N.J., 1976.
25. G. Shafer, *Belief Functions and Possibility Measures*, Working Paper No. 163, School of Business, The University of Kansas, Lawrence, 1984.
26. P. Walley and T. Fine, "Towards a frequentist theory of upper and lower probability." *The Annals of Statistics*, **10**, 1982, pp. 741–761.
27. Wang Pei-Zhuang, "From the fuzzy statistics to the falling random subsets." In: *Advances in Fuzzy Sets, Possibility Theory and Applications*, edited by P. P. Wang, Plenum Press, New York, 1983, pp. 81–96.
28. Wang Pei-Zhuang and E. Sanchez, "Treating a fuzzy subset as a projectable random subset." In: *Fuzzy Information and Decision Processes*, edited by M. M. Gupta and E. Sanchez, North-Holland, Amsterdam, 1982, pp. 213–220.
29. R. R. Yager, "Hedging in the combination of evidence." *Journal of Information and Optimization Science*, **4**, No. 1, pp. 73–81, 1983.
30. R. R. Yager, "Entropy and specificity in a mathematical theory of evidence." *Int. Journal of General Systems*, **9**, 1983, pp. 249–260.
31. R. R. Yager, "On different classes of linguistic variables defined via fuzzy subsets," *Kybernetes*, **13**, 1984, pp. 103–110.
32. R. R. Yager, "Arithmetic and other operations on Dempster–Shafer structures," *Tech. Report* MII-508, Iona College, New Rochelle, N.Y., 1985.
33. R. R. Yager, The entailment principle for Dempster–Shafer granule," *Tech. Report* MII-512, Iona College, New Rochelle, N.Y., 1985.
34. R. R. Yager, Reasoning with uncertainty in expert systems." *Proc. 9th Int. Joint Conf. on Artificial Intelligence*, Los Angeles, 1985, pp. 1295–1297.
35. L. A. Zadeh, "Fuzzy sets," *Information and Control*, **8**, 1965, pp. 338–353.
36. L. A. Zadeh, "Calculus of fuzzy restrictions." In: *Fuzzy Sets and their Applications to Cognitive and Decision Processes*, edited by L. A. Zadeh, K. S. Fu, K. Tanaka, M. Shimura, Academic Press, New York, 1975, pp. 1–39.
37. L. A. Zadeh, "PRUF: A meaning representation language for natural languages." *Int. J. Man–Machine Studies*, **10**, 1978, pp. 395–460.

38. L. A. Zadeh, “Fuzzy sets as a basis for a theory of possibility.” *Fuzzy Sets and Systems*, **1**, 1978, pp. 3–28.
39. L.A. Zadeh, “A theory of approximate reasoning.” In: *Machine Intelligence*, Vol. 9, edited by J. Hayes, D. Michie and L. I. Mikulich, John Wiley, New York, 1979, pp. 149–194.
40. L. A. Zadeh, “Fuzzy sets and information granularity.” In: *Advances in Fuzzy Set Theory and Applications*, edited by M. M. Gupta, R. K. Ragade and R. R. Yager, North-Holland, Amsterdam, 1979, pp. 3–18.
41. L. A. Zadeh, “On the validity of Dempster’s rule of combination of evidence.” *Memo UCB/ERL* No. 79/24, University of California, Berkeley, 1979.
42. L. A. Zadeh, “A simple view of the Dempster–Shafer theory of evidence,” *Berkeley Cognitive Science Report* No. 27, University of California, Berkeley, 1984.
43. Zhang Nan-Lun, “A preliminary study of the theoretical basis of the fuzzy set.” *BUSEFAL* (Univ. P. Sabatier, Toulouse), No. 19, 1984, pp. 58–67.
44. Zhang Nan-Lun, “The membership and probability characteristics of random appearance,” *Journal of Wuhan Institute of Building Materials*, No. 1, 1981.

Appendix

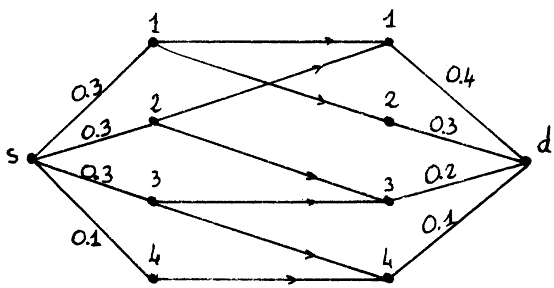
Flows in Networks and Inclusion

Let $(\mathcal{F}, m)$ and $(\mathcal{F}', m')$ be two bodies of evidence such that $(\mathcal{F}, m)$ is strongly included in $(\mathcal{F}', m')$ (cf. Definition 2). Let $A_1, \ldots, A_p$ (resp.: $A'_1, \ldots, A'_q$) be the elements of $\mathcal{F}$ (resp.: $\mathcal{F}'$). It is then possible to build a bipartite graph $(V, V', \mathcal{E})$ where V and V' are disjoint sets of nodes, and $\mathcal{E}$ is a set of arcs (v, v') where $(v \in V,\ v' \in V'$, defined as follows:

$$\text{each element of } V \text{ (resp.: } V'\text{) represents a focal element in } \mathcal{F} \text{ (resp.: } \mathcal{F}'\text{)},$$
$$\text{arc } (v_i, v'_j) \text{ exists if and only if } A_i \subseteq A'_j.$$

Note that $\forall\ v_i,\ \exists(v_i, v'_j) \in \mathcal{E}; \forall\ v'_j,\ \exists(v_i, v'_j) \in \mathcal{E}$ from strong inclusion.

Let s be a source node and d a sink node, which do not belong to $V \cup V'$. Build the arcs $(s, v_i) \forall\ v_i \in V$, with an associated capacity $a_i = m(A_i)$, and the arcs (v'_i, d) with an associated capacity $a'_i = m'(A_i)$. The graph corresponding to counter-example 2 is as follows:



Arcs in $\mathcal{E}$ are supposed to have infinite capacity. It is clear that the strong inclusion of $(\mathcal{F}, m)$ in $(\mathcal{F}', m')$ is equivalent to the existence of a flow of value 1 in the graph whose set of nodes is $N = V \cup V' \cup \{s, d\}$ and arcs, $E = \mathcal{E} \cup \{(s, v_i) | v_i \in V\} \cup \{(v'_j, d) | v'_j \in V'\}$. This fact is expressed by (42)–(43).

Now a *cut* in the graph is a partition $(X, \bar{X})$ of the nodes such that $s \in X,\ d \in \bar{X}$, and its capacity is the sum of the capacities of the arcs (i, j) such that $i \in X,\ j \in \bar{X}$. The max flow min-cut theorem[11] states that the maximal flow value from s to d is equal to the minimal cut capacity of all cuts separating s and d.

Obvious finite capacity cuts in (N, E) are obtained by stating $X = \{s\}$ or $\bar{X} = \{d\}$, and their capacity is 1. Hence the flow value through the graph is at most 1. Moreover if a cut involves an arc in $\mathcal{E}$, it has infinite capacity and is useless in the computation of the maximal flow. Hence interesting cuts are such that

if $S \subseteq V$ is a part of X then the set $\Gamma(S) \subseteq V'$ of successors of nodes in S is also in X,

if $T \subseteq V'$ is a part of $\bar{X}$ then the set $\Gamma^{-1}(T) \subseteq V$ of predecessors of nodes in T is also in $\bar{X}$.

Hence the set of cuts can be described as the set $\{(S, T) | S \subseteq V, T \subseteq V', \Gamma(S) \cap T = \varnothing, \Gamma^{-1}(T) \cap S = \varnothing\}$. The capacity of cut (S, T) is easily found as

$$C\,(S,T) = \sum_{v_i \in \bar{S}} a_i + \sum_{v_i \in \bar{T}} a'_i$$

since $X = \{s\} \cup S \cup \bar{T},\ \bar{X} = \{t\} \cup T \cup \bar{S}$.

Now consider the cut $(\Gamma^{-1}(T), \Gamma(S))$. It is clear that

$$C\,(S,T) \geqq C\left(\Gamma^{-1}\,(T)\,, \Gamma\,(S)\right).$$

So that the set of interesting cuts for the computation of the maximal flow is $\{(S, T) | S \subseteq V, T \subseteq V' \Gamma(S) = \bar{T}, \Gamma^{-1}(T) = \bar{S}\}$. A necessary and sufficient condition for the existence of a flow of value 1 through the network is thus

$$\forall\ S \subseteq V, \sum_{v_i \in \bar{S}} a_i + \sum_{v'_i \in \Gamma(S)} a'_i \geqq 1$$

$$\forall\ T \in V', \sum_{v_i \in \Gamma^{-1}(T)} a_i + \sum_{v'_i \in \bar{T}} a'_i \geqq 1$$

which also reads

$$\forall\ S \in V, \sum_{v_i \in S} a_i \leqq \sum_{v'_1 \in \Gamma(S)} a'_i \qquad \text{(I)}$$

$$\forall\ T \in V', \sum_{v'_i \in T} a'_i \leqq \sum_{v_i \in \Gamma^{-1}(T)} a_i. \qquad \text{(II)}$$

It is easy to check that the condition $Q(A) \leqq Q'(A)$ applied with $A = A_i$ gives (I) with $S = \{v_i\}$, and the condition $\mathrm{Cr}(A) \geqq \mathrm{Cr}'(A)$ applied with $A = A'_j$ gives (II) with $T = \{v'_j\}$. But generally it is possible to find $S \subseteq V$ such that

$$\nexists\ A, Q(A) = \sum_{u_i \in S} a_i$$

and T such that

$$\nexists\ A, \mathrm{Cr}'(A) = \sum_{u'_i \in T} a'_i.$$

This is why inclusion does not imply strong inclusion. In the above example if $S = \{v_3, v_4\}$ then $\Gamma(S) = \{v'_3, v'_4\}$ and (I) is violated i.e.

$$a_3 + a_4 = 0.4 > a'_3 + a'_4 = 0.3.$$

However, as the Table 1 shows, $Q(A) \leqq Q'(A),\ \mathrm{Cr}(A) \geqq \mathrm{Cr}'(A) \forall\ A$.

Now assume $(\mathcal{F}, m)$ and $(\mathcal{F}', m')$ are consonant. $\mathcal{F}$ and $\mathcal{F}'$ are ordered such that $A_1 \subset A_2 \cdots \subset A_p,\ A'_1 \subset A'_2 \cdots \subset A'_q$. The bipartite graph $(V, V', \mathcal{E})$ has a special structure since if $A_i \subseteq A'_j$ then $A_i \subseteq A'_k,\ \forall\ k \geqq j$.

We now prove that the flow equations always have a solution if the fuzzy set F associated to $(\mathcal{F}, m)$ is included in F' associated to $(\mathcal{F}', m')$.

$\forall\ v_i \in V$, let $\sigma(i)$ be the index such that

$$\sigma(i) = \min\left\{j | A_i \subseteq A'_j\right\}.$$

Similarly $\forall v'_j \in V'$, let $\tau(j)$ be the index such that

$$\tau(j) = \max\left\{i | A_i \subseteq A'_j\right\}.$$

Then the flow (42) and (43) reads

$$a_i = \sum_{j \geqq \sigma(i)} w_{ij} \qquad \forall\ i = 1, p \tag{III}$$

$$a'_j = \sum_{i \leqq \tau(j)} w_{ij} \qquad \forall\ j = 1, p. \tag{IV}$$

Let $n = |\sigma(V)|,\ k \in \sigma(V),\ \sigma^{-1}(k) = \{A_i | \sigma(i) = k\}$ and $i_k = \max\{i, A_i \in \sigma^{-1}(k)\}$. Because $(\mathcal{F}, m) \subseteq (\mathcal{F}',\ m'),\ A_1 \subseteq A'_1$ and $A_p \subseteq A'_q$. Moreover $\forall\ i \in \sigma^{-1}(k),\ A_i \subseteq A'_k$ but $A_i \not\subset A'_{k-1}$. Hence $\mathcal{F}$ can be partitioned into n groups of consecutive focal elements, and this partition creates a partition of $\mathcal{F}'$, also in n groups $\tau^{-1}(i_k),\ k \in \sigma(V)$ with

$$\tau^{-1}(i_k) = \{B_j | k \leqq j < \sigma(i_k + 1)\} \qquad \text{(see Fig. 1)}$$

and $\forall\ B_j \in \tau^{-1}(i_k),\ B_j \supseteq A_i$ for all $i \in \sigma^{-1}(\sigma(i_k))$, but $A_i \not\subset B_j,\ j < k$. Note that $\max_{k \in \sigma(V)} i_k = p$ so that $\sigma(p+1) = q+1$ by convention. It is clear that $i_k = \tau(k+1) - 1$ (see Fig. 1). Pairs $(\sigma^{-1}(k), \tau^{-1}(i_k))$ are ranked along increasing k's and can be renumbered as $\{(G_i,\ G'_i)\ i = 1,\ n\}$ as in Fig. 1. For all $(A_i,\ A'_j) \in G_k \times G'_l$ we define $x_{kl} = \Sigma\{w_{ij} | (A_i,\ A'_j) \in G_k \times G'_l\}$.

$$\bar{a}_k = \sum_{A_i \in G_k} a_i, \qquad \bar{a}'_l = \sum_{A'_j \in G'_l} a'_j.$$

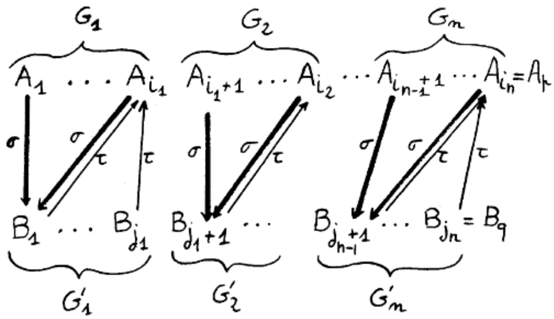


FIGURE 1

Now (III) and (IV) imply

$$\bar{a}_k = \sum_{\substack{l \\ l \geqq k}} x_{kl} \qquad k = 1, n, \tag{V}$$

$$\bar{a}'_l = \sum_{\substack{k \\ l \geqq k}} x_{kl} \qquad l = 1, n. \tag{VI}$$

Similarly, let Φ and Φ' be the fuzzy sets derived from F and F' as follows: Φ and Φ' have n α-cuts which are respectively the smallest set in each G_k and the greatest set in each G'_k, the mass allocated to the set from G_k (resp.: G'_k) being $\bar{a}_k$ (resp.: $\bar{a}'_k$). It is easy to check that $\Phi \subseteq F \subseteq F' \subseteq \Phi'$.

System (V) and (VI) always have solutions. Let

$$\mu_i = \sum_{k=i}^{n} \bar{a}_k, \qquad \mu'_i = \sum_{k=i}^{n} \bar{a}'_k.$$

$\Phi \subseteq \Phi'$ implies $\mu_i \leqq \mu'_i\ \forall\ i = 1,\ n$. Then let

$$\begin{aligned} x_{11} &= 1 - \mu'_2 \\ x_{ii} &= \mu_i - \mu'_{i+1} \qquad 1 < i < n \\ x_{i,i+1} &= \mu'_{i+1} - \mu_{i+1} \qquad 1 \leqq i < n \\ x_{nn} &= \mu_n \\ x_{ij} &= 0 \text{ otherwise.} \end{aligned}$$

This is a solution of (V–VI). Indeed, it is demonstrated as follows:

$$\bar{a}_k = x_{kk} + x_{kk+1} = \mu_k - \mu_{k+1}$$
$$\bar{a}'_k = x_{kk} + x_{k-1k} = \mu'_k - \mu'_{k+1}.$$

From this solution, a solution to (III) and (IV) is easily deduced letting

$$w_{ij} = \frac{a_i \cdot a'_j}{\bar{a}_k \cdot \bar{a}'_l} x_{kl} \quad \text{whenever} \quad (A_i, A_j) \in G_{k.l}.$$

Hence if $F \subseteq F'$ then $(\mathcal{F}, m)$ is strongly included in $(\mathcal{F}', m')$.

Example $\Omega = \{a, b, c, d, e\}$
$F = \{1/a, 0.5/b, 0.4/c, 0.2/d\}$
$F' = \{1/a, 1/b, 0.5/c, 0.3/d, 0.3/e\}$.

Then

$A_1 = \{a\}$	$a_1 = 0.5$	$A'_1 = \{a, b\}$	$a'_1 = 0.5$
$A_2 = \{a, b\}$	$a_2 = 0.1$	$A'_2 = \{a, b, c\}$	$a'_2 = 0.2$
$A_3 = \{a, b, c\}$	$a_3 = 0.2$	$A'_3 = \Omega$	$a'_3 = 0.3$
$A_4 = \{a, b, c, d\}$	$a_4 = 0.2$		
$G_1 = \{A_1, A_2\}$	$\bar{a}_1 = 0.6$	$G'_1 = \{A'_1\}$	$\bar{a}'_1 = 0.5$
$G_2 = \{A_3\}$	$\bar{a}_2 = 0.2$	$G'_2 = \{A'_2\}$	$\bar{a}'_2 = 0.2$
$G_3 = \{A_4\}$	$\bar{a}_3 = 0.2$	$G'_3 = \{A'_3\}$	$\bar{a}'_3 = 0.3$

$x_{11} = 0.5$

$$x_{12} = 0.5 - 0.4 = 0.1 \Rightarrow \begin{cases} w_{11} = 2.5/6 \\ w_{21} = 0.5/6 \end{cases}$$
$x_{22} = 0.1$ $\quad w_{12} = 0.5/6$
$x_{23} = 0.1$ $\quad w_{22} = 0.1/6$ etc....
$x_{33} = 0.2$

Hence the flow