Uncertainty and Information Measures for Imprecise Probabilities: An Overview

George J. Klir
Center for Intelligent Systems
and
Department of Systems Science & Industrial Engineering
Thomas J. Watson School of Engineering and Applied Science
Binghamton University - SUNY
Binghamton, New York 13902-6000, USA

Abstract

The paper deals with basic issues regarding the measurement of relevant types of uncertainty and uncertainty-based information in theories that represent imprecise probabilities of various types. Existing results and encountered difficulties regarding these issues, primarily in evidence theory and possibility theory, are presented. Some important open questions and unexplored areas of research in this domain are also discussed.

Keywords. Imprecise probabilities, uncertainty measures, uncertainty-based information, evidence theory, possibility theory, Hartley-like measure, Shannon-like measure

1 Uncertainty and Uncertainty-Based Information

The concepts of *uncertainty* and *information*, as understood in this paper, are intimately connected. Uncertainty involved in any problem-solving situation is a result of some information deficiency. Information pertaining to the situation can be obtained via any action by which the uncertainty is reduced. Viewing the concept of uncertainty as primary, the concept of information may then be defined in terms of uncertainty reduction. This particular conception of information is usually referred to as *uncertainty-based information* [8].

To develop a fully operational theory for dealing with uncertainty of some conceived type, we must address a host of issues at four distinct levels. At the *first level*, we need to find an appropriate mathematical representation of the conceived type of uncertainty. At the *second level*, we need to develop a calculus by which this type of uncertainty can be properly manipulated. At the *third level*, we need to find a meaningful way of measuring relevant uncertainty in any situation formalizable within the theory. At the *fourth level*, we need to develop methodological aspects of the theory, including procedures for making the various uncertainty principles operational within the theory. The scope of applicability of any particular uncertainty theory, assumed to be sufficiently developed at each of the four levels, is determined by its generality.

Traditionally, it was taken for granted that uncertainty is adequately captured by *probability theory*, as axiomatically developed by Kolmogorov [12]. However, it is more and more recognized that the concept of uncertainty is too broad to be captured by probability theory alone. As is well known, a number of alternative uncertainty theories have been proposed since at least the 1970s. A common difference of these theories from probability theory is that they do not abide by the *additivity axiom* of probability theory. This means, in turn, that they cannot be based on the classical measure theory, which deals with additive measures. They require a more general measure theory, which can deal with *nonadditive measures*. It is usually required that the measures be monotonic with respect to set inclusion and continuous or semicontinuous; a theory dealing with measures of this kind is usually called a *fuzzy measure theory* [22].

2 Precise and Imprecise Probabilities

In classical probability theory, elementary events are required to be pairwise disjoint and the probability of each is required to be expressed precisely by a real number in the unit interval [0, 1]. This precision requirement of classical probability theory, which is a consequence of the additivity axiom, is often difficult to satisfy. This difficulty may be a result of unavoidable measurement errors, insufficient statistical information, missing data, conflicting evidence, etc., and it is particularly severe when instead of measurements we are dependent on assessments based on subjective human judgements. To alleviate this difficulty of precise probabilities, we need to relax on the precision requirements and allow us to express the individual probabilities imprecisely.

The imprecision in expressing probabilities introduces a new dimension into the formalization of uncertainty and uncertainty-based information. The first throughout study of imprecise probabilities was taken by Walley [21]. His principal result is a demonstration that reasoning and decision making based on imprecise probabilities satisfy the principles of coherence and avoidance of sure loss, which are generally viewed as principles of rationality. Hence, the requirement of precision (or, equivalently, the additivity axiom) can not be justified as inevitable for rationality, as previously believed. The soundness of using imprecise probabilities is based on this important result.

It is recognized that imprecise probabilities of different types exist and they require different methodological treatments. The imprecision of probabilities can be expressed, for example, by closed intervals of real numbers or by fuzzy intervals with various special properties. Some particular theories of imprecise probabilities are already well developed at the first two levels. They include the evidence theory developed by Shafer [17], possibility theory [1], interval-valued probability distributions [13], fuzzy probability distributions [14], and a theory based on λ-measures introduced by Sugeno [20]. However, none of these theories is sufficiently developed at the third level -- the level of measuring uncertainty and the associated uncertainty-based information. The purpose of this paper is to present an overview of obtained results, encountered difficulties, and open problems in this area.

3 Classical Measures of Uncertainty

Two of the well established measures of uncertainty are now considered as classical. One of them, called a *Hartley measure*, applies to uncertainty formalized in terms of classical (crisp) possibility theory. The other one, called a *Shannon measure* (or Shannon entropy), applies to uncertainty formalized in terms of classical (precise) probability theory.

The classical Hartley measure quantifies the most fundamental type of uncertainty, one expressed in terms of a finite set of possible alternatives. To describe this type of uncertainty, let X denote the set of all alternatives under consideration (predictions, retrodictions, diagnoses, etc.). In each situation, only one of the alternatives is true, but we do not necessarily know which one. However, we may know, on the basis of all available evidence, that the only possible candidates for the true alternative are in a subset A of X. It was shown by Hartley [2] that the only sensible way to measure the amount of uncertainty in this situation is to use function H defined by the simple formula

$$H(A) = \log_2 |A|, \tag{1}$$

provided that the measurement unit is a *bit*; $|A|$ denotes the cardinality of set A. The uniqueness of H(A) is also well established on axiomatic grounds [8].

The type of uncertainty quantified by the Hartley measure is well captured by the term *nonspecificity*. Uncertainty in the described situation results from the lack of specificity in characterizing the true alternative. The larger the set of *possible alternatives*, the less specific is the characterization. Full specificity is obtained when only one alternative is possible.

The classical Shannon measure applies to uncertainty formalized in terms of a probability distribution $\mathbf{p} = \langle p(x) | x \in X \rangle$ defined on a finite set X of alternatives (elementary events) under consideration. Is was shown by Shannon [18] that the only meaningful way to measure the amount of uncertainty in any situation characterized by the probability distribution **p** is to use the functional S defined by the formula

$$S(\mathbf{p}) = -\sum_{x \in X} p(x) \log_2 p(x), \tag{2}$$

provided that the measurement unit is the bit. The uniqueness of S(**p**) has been well established in numerous ways on axiomatic grounds [8].

The Shannon measure may also be expressed in the form

$$S(\mathbf{p}) = -\sum_{x \in X} p(x) \log_2 [1 - \sum_{y \neq x} p(y)]. \tag{3}$$

The term

$$c(x) = \sum_{y \neq x} p(y)$$

in Eq. (3) represents the total evidential claim that fully conflicts with the one focusing on alternative x. The function

$$-\log_2 [1 - c(x)],$$

which is employed in Eq. (3) is monotonic increasing with c(x) and extends its range from [0, 1] to [0, ∞]; hence, it represents the same conflict on a different scale. The use of the logarithmic function is a consequence of the axiomatic requirements that function S must satisfy.

It follows directly from (3) that S(**p**) may be viewed as the weighted average of the conflict among evidential claims expressed by **p**. As discussed later in this paper, this view of the Shannon measure was accepted as a base for deriving Shannon-like measure of uncertainty in evidence theory and possibility theory.

4 Other Established Measures of Uncertainty

Using the classical measures of uncertainty introduced in Sec.3 as a basis, a search for their counterparts in evidence theory and possibility theory has been pursued by numerous researchers. Thus far, only the counterparts (or generalizations) of the Hartley measure in these theories have been discovered and fully justified on intuitive grounds as well as by rigorous axiomatic treatments. These counterparts of the Hartley measure are conveniently called *Hartley-like measures*. To discuss them, the following notation and terminology is used; basic knowledge of evidence theory, at least as covered by Klir and Wierman in [8], is assumed.

The universal set under consideration (often called a frame of discernment in evidence theory) is denoted in this paper by X. The three basic functions of evidence theory -- *belief functions*, *plausibility functions*, and *basic probability assignment functions* -- are denoted by Bel, Pl, and m respectively. These functions map subsets of X to the unit interval [0, 1]. Each subset A of X for which $m(A) > 0$ is called a *focal element*, and the set of all focal elements induced by m together with the associated values of function m is called a *body of evidence*. The set of all focal elements induced by m is denoted by F(m).

Each situation for which the Hartley measure is applicable may be viewed from the standpoint of evidence theory as a simple body of evidence that consists of a single focal element, the set A. To measure nonspecificity of arbitrary bodies of evidence, we need a function by which values of the Hartley measure for all focal elements are properly aggregated. The most natural way of aggregation in this case is to take the average of these values, weighted by the associated values of the basic probability assignment function m. This is perfectly meaningful since the values of function m are required to add to 1.The *Hartley-like measure* in evidence theory, HL, is then for each m defined by the formula

$$\mathrm{HL}(m) = \sum_{A \in F(m)} m(A) \log_2 |A| . \tag{4}$$

This measure is justified not only on intuitive grounds, but also on mathematical grounds. It satisfies all the following essential mathematical properties: (i) *subadditivity* — the value of HL for any joint body of evidence is always smaller than or equal to the sum of the values of HL for the associated marginal bodies of evidence; (ii) *additivity* — the equality in (i) is obtained iff the joint body of evidence is based on noninteractive marginal bodies of evidence; (iii) *monotonicity* — if $\langle m_1, F_1 \rangle \subseteq \langle m_2, F_2 \rangle$ (as defined in evidence theory), then $\mathrm{HL}(m_1) \leq \mathrm{HL}(m_2)$; (iv) *minimum* — $\mathrm{HL}(m) = 0$ iff $m(\{x\}) = 1$ for some $x \in X$; (v) *maximum* — $\mathrm{HL}(m) = \log_2 |X|$ iff $m(X) = 1$. Moreover, HL was proven unique on axiomatic grounds by Ramer [16].

The *Hartley-like measure in possibility* theory may also be expressed by Eq.(4). However, since possibilistic bodies of evidence are based on nested sets of focal elements, various special and computationally more efficient forms of possibilistic Hartley-like measure can easily be derived [8]. An important property of the Hartley-like measure given by Eq.(4) is that it is additive and subadditive under the calculi of both evidence theory and possibility theory. The uniqueness of the possibilistic Hartley-like measure was proven on axiomatic grounds by Klir and Mariano in [6].

The Hartley measure H is applicable only to finite sets. Its counterpart, H_n, for convex subsets of the n-dimensional Euclidean space $\mathbb{R}^n$ ($n \geq 1$) was proposed by Klir and Yuan in [11]. For any convex subset A of $\mathbb{R}^n$ and some $n \geq 1$, H_n is defined by the formula

$$H_n(A) = \min_{t \in T} \ln \left[\prod_{i=1}^{n} [1 + \mu(A_{i_t})] + \mu(A) - \prod_{i=1}^{n} \mu(A_{i_t}) \right], \tag{5}$$

where μ denotes the Lebesgue measure, T denotes the set of all transformations from one orthogonal coordinate system to another, and A_{i_t} denotes the i-th projection of A in coordinate system t. It has been proven that function H_n possesses all properties that a measure of uncertainty is expected to possess: *monotonicity* with respect to set inclusions, *subadditivity*, *additivity* (proven only for $n \leq 3$ by Klir and Yuan [11], but proven later for any $n \geq 1$ by Ramer [15]), *continuity*, and *coordinate invariance*.

In practical applications, as is argued in [23], the universal set involved is usually not the full space $\mathbb{R}_n$, but a finite n-dimensional hypercube X whose sides represent specified ranges $[\underline{v}_i, \overline{v}_i]$ of the numerical variables v_i ($i = 1, 2, \ldots, n$) employed in each application. In such cases, it is useful to compute normalized versions, $\hat{v}_i$, of variables v_i by the formula

$$\hat{v}_i = \frac{v_i - \underline{v}_i}{\overline{v}_i - \underline{v}_i} \quad (i = 1, 2, \ldots, n),$$

and rescale thus X to the unit hypercube $[0,1]^n$. Then, we have $H_n(A) \in [0, 1]$ for any $A \subset [0,1]^n$.

A counterpart, HL_n, of the Hartley-like measure HL for the n-dimensional Euclidean space is obviously expressed by the formula

$$\mathrm{HL}_n(m) = \sum_{A \in F(m)} m(A) H_n(A) , \tag{6}$$

provided that F(m) is finite and all sets in F(m) are convex sets.

The application of uncertainty measure H_n can be extended to convex fuzzy sets via the coherent fuzzy-set interpretation of possibility theory introduced by Klir [5].

For each convex fuzzy set F defined on a convex universal set $X \subset \mathbb{R}^n$, H_n is expressed by the formula

$$H_n(F) = \int_0^{h_F} H_n({}^{\alpha}F)d\alpha + (1 - h_F)H_n(X), \qquad (7)$$

where ${}^{\alpha}F$ denotes the α-cut of F ($\alpha \in [0, 1]$) and h_F is the height of F [9].

5 Encountered Difficulties

Considering probabilistic bodies of evidence, whose focal elements are singletons, it is obvious that HL(m) = 0 for any probability measure. This clearly demonstrates that probability theory is not capable of representing nonspecificity. It represents a different type of uncertainty, one associated with the conflict among evidential claims within a given probability distribution, for which the Shannon measure is well established.

So far, the search for a Shannon-like measure in evidence theory and possibility theory has not been successful, even though this issue is extensively addressed in the literature. An historical overview of this unsuccessful search is given in the book by Klir and Wierman [8].

From several functions proposed in the literature as candidates for the Shannon-like measure in evidence theory, two eventually emerged as the best justified on intuitive grounds. These measures, referred to as *strife* and *discord* and denoted by ST and D, respectively, are defined by the formulas

$$ST(m) = -\sum_{A \in F(m)} m(A) \log_2 \sum_{B \in F(m)} m(B) \frac{|A \cap B|}{|A|}, \qquad (8)$$

$$D(m) = -\sum_{A \in F(m)} m(A) \log_2 \sum_{B \in F(m)} m(B) \frac{|A \cap B|}{|B|}, \qquad (9)$$

As shown by Klir and Yuan [10], the distinction between strife and discord reflects the distinction between disjunctive and conjunctive set-valued propositions, respectively.

Strife and discord seem to be perfectly justified on intuitive grounds as measures of conflict among evidential claims within each given body of evidence when dealing with either disjunctive or conjunctive set-valued propositions, respectively. They also satisfy all the required mathematical properties except one -- *subadditivity* -- one of the essential properties of uncertainty measures. In fact, none of the other proposed candidates for the Shannon-like measure satisfy subadditivity either.

The long, unsuccessful, and often frustrating search for the Shannon-like measure of uncertainty in evidence theory was in the early 1990s replaced with the search for a justifiable measure of total uncertainty. The first attempt was to add the well-justified Hartley-like measure with one of the proposed candidates for the Shannon-like measure. Again, the sums HL + ST and HL + D emerged as the best candidates on intuitive grounds, one for disjunctive set-valued propositions and one for conjunctive set-valued proposition. And, again, these functions were found to possess all mathematical properties required for uncertainty measures except subadditivity.

A measure of total uncertainty in evidence theory that possesses all the required mathematical properties was eventually found (independently by several authors at about the same time), but not as a composite of measures of uncertainty of the two types. This aggregate uncertainty measure, AU, is defined for each belief function Bel on the power P(X) set of X by the formula

$$AU(Bel) = \max_{P_{Bel}} \left[-\sum_{x \in X} p_x \log_2 p_x \right], \qquad (10)$$

where the maximum is taken over the set P_{Bel} of all probability distributions $\langle p_x \mid x \in X \rangle$ that are consistent with the given belief measure Bel, which means that they satisfy the constraint

$$Bel(A) \le \sum_{x \in A} p_x \text{ for all } A \in P(X), \qquad (11)$$

in addition to the usual axiomatic constraints of probability distributions. An efficient algorithm for computing this aggregate uncertainty measure is available [8].

Although the function AU is acceptable on mathematical grounds as an aggregate measure of uncertainty in evidence theory and possibility theory, some fundamental questions regarding the measurement of uncertainty in these theories still remain to be answered.

It is conceptually well understood that situations described in terms of evidence theory involve, in general, two types of uncertainty. However, only one of them, which is appropriately called nonspecificity, is well understood on both intuitive and mathematical grounds, and it is quantified by the well-justified Hartley-like measure. The second type of uncertainty in evidence theory, the one obtained by generalizing the uncertainty quantified in probability theory by the Shannon measure, is still ill-understood.

All attempts to find a Shannon-like measure in evidence theory were based on the assumption that this measure should quantify the conflict among evidential claims in each given body of evidence. This is a reasonable assumption suggested by the Shannon measure itself. However, none of the proposed measures of evidential conflict in evidence theory, some of which are conceptually well-justified, is acceptable on mathematical grounds. All candidates for the Shannon-like measure that were proposed in the literature violate some mathematical properties essential for any measure of uncertainty.

Hence, no Shannon-like measures in evidence theory and possibility theory have been found so far. While function AU is acceptable as an aggregate measure of uncertainty in these theories, we do not know how to decompose it into measures of the two types of uncertainty that in these theories manifestly coexist. This remains an open question.

Although function AU is acceptable on mathematical grounds as an aggregate measure of uncertainty in evidence theory and possibility theory, it is rather insensitive to changes in evidence that seem significant on intuitive grounds. To illustrate this undesirable feature of the function, let us examine a very simple example. Let $X = \{x_1, x_2\}$ and $m(\{x_1\}) = a$, $m(X) = 1-a$. Then, $Bel(\{x_1\}) = a$, $Bel(\{x_2\}) = 0$, $Bel(X) = 1$, and $AU(Bel) = 1$ for all $a \in [0, 0.5]$. Increasing evidence focusing on the alternative x_1 from 0 to 0.5 is thus not captured by the value of AU.

One additional undesirable feature of the aggregate measure AU should be mentioned. The measure does not take into account differences in convex sets of probability distributions that are consistent with the various bodies of evidence. Thus, for example, the situation of total ignorance, when $m(X) = 1$, has the same value of AU as the situation characterized by the uniform probability distribution $m(\{x\}) = 1/|X|$ for all $x \in X$. However, these two situations are associated with very different sets of probability distributions. In the first situation, the set consists of all probability distributions that can be defined on X; in the second situation, the set consists of a single probability distribution, the uniform one. This is an important difference, at least from the behavioral point of view. While the second situation contains information for rational betting, no such information is available in the first situation. This difference was recently examined (independently of evidence theory) by Kapur et al.[3], who suggested to express it in terms of the difference between the maximal and minimal values of the Shannon measure under given constraints. It seems reasonable that the minimal values of the Shannon measure for given bodies of evidence could be utilized in some way to formulate a more sensitive aggregate measure of uncertainty in evidence theory. However, whether minimizing the Shannon measure can actually be utilized for this purpose is another open question.

6 Open Questions in Previous Research

In spite of the rather long quest for the Shannon-like measures in evidence theory and possibility theory, we have not been successful in finding these measures. Each of the many proposed candidates has some defect, be it conceptual or mathematical. Hence, the nature of the Shannon-like measures in these two theories remains an open question.

It is quite surprising that none of the proposed candidates for the Shannon-like measures satisfies the essential requirement of subadditivity. This consistent violation of subadditivity raises a fundamental question: It is really necessary to require that the Shannon-like measure alone be subadditive? Perhaps we need to require that only the total uncertainty be subadditive, as suggested by the following example.

Let the frame of discernment be the Cartesian product $X \times Y$, where $X = \{1,2\}$ and $Y = \{a, b\}$, and let us consider a joint body of evidence that consists of two disjoint focal elements with

$$m(\{1a, 2b\}) = m(\{2a, 1b\}) = 0.5.$$

Then, clearly,

$$m_X(\{0, 1\}) = 1 \text{ and } m_Y(\{a, b\}) = 1$$

are basic probability assignments of the associated marginal bodies of evidence. In this case, it is obvious that the joint body (two disjoint elements with equal probabilities) contains 1 bit of conflict in evidence, while the marginal bodies do not contain any conflict. Every Shannon-like measure should give these results (in fact, all the proposed candidates give them). Hence, no Shannon-like measure can be subadditive in this case. However, we have

$$HL(m) = 1 \text{ and } HL(m_X) = HL(m_Y) = 1$$

in this example and, consequently, the subadditivity requirement is satisfied if we consider both types of uncertainty that coexist in evidence theory. Unfortunately, none of the proposed candidates satisfies this more reasonable formulation of the subadditivity requirement.

A new candidate for the Shannon-like measure, SL, is defined by the formula

$$SL(Bel) = -\frac{1}{c}\sum_{x\in X}[Bel(\{x\})\log_2 Bel(\{x\}) + Pl(\{x\})\log_2 Pl(\{x\})], \tag{12}$$

where

$$c = \sum_{x\in X}[Bel(\{x\}) + Pl(\{x\})],$$

and

$$Pl(\{x\}) = 1 - Bel(X - \{x\})$$

for all $x \in X$. This candidate seems to behave exactly as it should, and it has so far withstood extensive testing of the subadditivity requirement involving the total uncertainty SL + HL. However, the subadditivity of this total uncertainty has not been proven as yet. Moreover, bodies of evidence theory are not fully represented by beliefs and plausibilities on singletons.

One possible direction, worth of pursuing, is to explore the linear combination, C_β, of the aggregate measure AU and the Hartley-like measure of nonspecificity HL,

$$C_\beta(Bel_m) = \beta AU(Bel_m) + (1-\beta)\, HL(m), \qquad (13)$$

for various values of the parameter $\beta \in [0, 1]$; Bel_m in (13) denotes the belief measure associated with the basic probability assignment m. Clearly, C_β collapses to HL or AU when $\beta = 0$ or 1, respectively. Since both HL and AU are additive and subadditive functions, the additivity and subadditivity of C_β is guaranteed. Since the range of both AU and HL is $\left[0, \log_2|X|\right]$ for any m defined on P(X), the range of their linear combinations (for any β) is the same.

It is obvious that function C_β defined by (13) is an aggregate measure of uncertainty for any $\beta \in (0, 1)$. This measure is more sensitive to changes in evidence than the current measure AU. This can be illustrated by using the same body of evidence by which the insensitivity of measure AU was exemplified: $X = \{x_1, x_2\}$, $m(\{x_1\}) = a$, $m(X) = 1-a$. In this case, clearly, $C_\beta(Bel_m) = 1-a(1-\beta)$ while $AU(Bel_m) = C_1(Bel_m) = 1$ for all $a \in [0, 0.5]$.The difference is even more pronounced when $m(\{x_1\}) = a$, $m(\{x_2\}) = b$, and $m(\{x_1, x_2\}) = 1-a-b$. Then, $C_\beta(Bel_m) = 1-a(1-\beta) - b(1-\beta)$, while $AU(Bel_m) = C_1(Bel_m) = 1$ for all $a \in [0, 0.5]$ and $b \in [0, 0.5]$.

Measure C_β is obviously preferable to AU from the standpoint of its sensitivity to changes in evidence. However, its meaning is not transparent. It seems that different values of β will be needed for different interpretations of DST. How to determine the right values of β for the various interpretations is an issue to be investigated.

When all focal elements are singletons and, hence, m is a probability distribution function, C_β assumes the form

$$C_\beta(Bel_m) = -\beta \sum_{x \in X} m(\{x\}) \log_2 m(\{x\}).$$

Clearly, this is the Shannon measure, as in the case of AU, but scaled down by the factor of β. The value of β, when properly determined in the context of each application, may conveniently capture the difference between the total ignorance and the evidence expressed in terms of the uniform probability distribution.

7 Unexplored Areas

One concern about the proposed candidate (12) for the Shannon-like measure in evidence theory is that it is fully formulated in terms of singletons or their complements, but a belief function of evidence theory is not determined, in general, by its values on singletons and their complements. However, as is well known [22], a significant subset of belief functions and plausibility functions can be represented by the λ-measures proposed by Sugeno [20]. For this subset of belief functions, any belief function is uniquely determined by its values on singletons.

Uncertainty represented in terms of λ-measures is one area in which the measurement of uncertainty has not been investigated as yet. Another area that have not been explored in this sense is the area of imprecise probabilities in which imprecise probabilities are represented by feasible interval-valued probability distributions [13] or fuzzy probability distributions [14]. It seems that requirements for the measure of uncertainty in this area may be formulated in the same way as in classical probability theory, but applying constrained interval or fuzzy arithmetic [4,7]. This will inevitably lead to challenging mathematical and computational issues.

There are of course other areas of imprecise probabilities. The issue of how to measure uncertainty and uncertainty-based information has not been even raised in these areas. This is thus a fertile field for future research.

A measure of uncertainty-based information for evidence theory, which is fundamentally different from the described measures, was proposed by Smets [19]. He formulated requirements for such a measure without considering bodies of evidence on Cartesian products. Hence, his formulation does not contain the usual requirements of additivity and subadditivity. Instead, he requires that the information content of two non-contradictory bodies of evidence be the same as the information content of the combined body of evidence obtained by the Dempster rule of combination [17], in addition to monotonicity and appropriate boundary conditions. Under these requirements, Smets shows that the only meaningful measure of information in bits is a function denoted by I and defined by the formula

$$I(Bel) = -\sum_{A \in P(X)} \log_2 Q(A) \qquad (14)$$

for each given non-dogmatic belief function Bel on P(X), (i.e., $m(X) > 0$ and, hence, $Q(X) > 0$), where Q denotes the commonality function [17]. Since the kind of normalization employed in the Dempster rule of combination is rather controversial [24], the justification of the measure of information given by (14) is controversial as well. Nevertheless, the issue of connecting a measure of information in evidence theory with a rule of combination (not necessarily the Dempster rule) is worth investigating.

REFERENCES

[1] De Cooman, G., "Possibility theory." *Intern. J. of General Systems*, **25**(4): 291-371, 1997.

[2] Hartley, R. V. L., "Transmission of information." *The Bell System Technical J.*, **7**(3):535-563, 1928.

[3] Kapur, J. N., G.Baciu and H.K.Kesavan, "The MinMax information measure." *Intern. J. of Systems Science*, **26**(1):1-12, 1995.

[4] Klir, G.J., "Fuzzy arithmetic with requisite constraints." *Fuzzy Sets and Systems*, **91**(2):165-175, 1997.

[5] Klir, G. J., "On fuzzy-set interpretation of possibility theory." *Fuzzy Sets and Systems*, **104**, 1998.

[6] Klir, G. J. and M. Mariano, "On the uniqueness of possibilistic measure of uncertainty and information." *Fuzzy Sets and Systems*, **24**(2):197-219,1987.

[7] Klir, G. J. and Y. Pan, "Constrained fuzzy arithmetic: Basic questions and some answers." *Soft Computing* , **2**(2):100-108, 1998.

[8] Klir, G. J. and M.J. Wierman, *Uncertainty-Based Information: Elements of Generalized Information Theory.* Physica-Verlag/ Springer-Verlag, Heidelberg and New York, 1998.

[9] Klir, G. J. and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications.* Prentice Hall, PTR, Upper Saddle River, NJ, 1995.

[10] Klir, G. J. and B. Yuan, "On measures of conflict among set-valued statements." *Proc. 1993 World Congress on Neural Networks*, Portland, Oregon: 627-630, 1993.

[11] Klir, G.J. and B.Yuan, "On non-specificity of fuzzy sets with continuous membership functions." *Proc. 1995 Intern. Conf. on Systems, Man, and Cybernetics*, Vancouver: 25-29, 1995.

[12] Kolmogorov, A. N., *Foundations of the Theory of Probability*. Chelsea, New York, 1950. (First published in German in 1933.)

[13] Pan, Y. and G.J.Klir , "Bayesian inference based on interval probabilities." *J. of Intelligent and Fuzzy Systems*, **5**(3):193-203, 1997.

[14] Pan, Y. and B.Yuan, "Bayesian inference of fuzzy probabilities." *Intern. J. of General Systems*, **26**(1-2):73-90, 1997.

[15] Ramer, A., "Euclidean specificity: two solutions and few problems." *Proc. World Congress of the Intern. Fuzzy Systems Assoc., Prague*, (Vol.IV): 268-271, 1997.

[16] Ramer, A., "Uniqueness of information measure in the theory of evidence." *Fuzzy Sets and Systems*, **24**(2): 183-196, 1987.

[17] Shafer G., *A Mathematical Theory of Evidence*. Princeton Univ. Press, Princeton, NJ, 1976.

[18] Shannon, C. E., "The mathematical theory of communication." *The Bell System Technical J.*, **27**(3&4): 379-423, 623-656, 1948.

[19] Smets, P., "Information content of an evidence." *Intern. J. of Man-Machine Studies*, **19** (1):33-43, 1983.

[20] Sugeno, M., "Fuzzy measures and fuzzy integrals : A survey." In: Gupta, M. M., Saridis, G. N. and Gaines, B. R., eds., *Fuzzy Automata and Decision Processes*. North-Holland, Amsterdam and New York, pp. 89-102, 1977.

[21] Walley, P., *Statistical Reasoning With Imprecise Probabilities*. Chapman and Hall, London, 1991.

[22] Wang, Z. and G. J. Klir, *Fuzzy Measure Theory*. Plenum Press, New York, 1992.

[23] Wu, E.N. and G.J. Klir, "Optimal redundancy management in reconfigurable control systems based on normalized nonspecificity." *Intern. J. of Systems Science* (submitted).

[24] Zadeh, L.A., "A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination." *AI Magazine*, **7**(2):85-90, 1986.