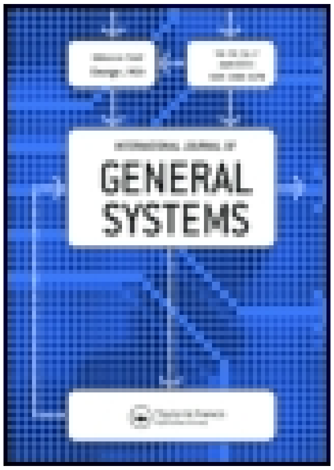

International Journal of General Systems
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ggen20

UNCERTAINTY IN THE DEMPSTER-SHAFER THEORY: A CRITICAL RE-EXAMINATION

GEORGE J. KLIR [a] & ARTHUR RAMER [b]

[a] Department of Systems Science , Thomas J. Watson School, State University of New York , Binghamton, New York, 13901, USA

[b] School of Electrical Engineering and Computer Science, University of Oklahoma , Norman, Oklahoma, 73019, USA

Published online: 06 Apr 2007.

Int. J. General Systems, Vol. 18, pp. 155–166
Reprints available directly from the publisher
Photocopying permitted by license only

UNCERTAINTY IN THE DEMPSTER–SHAFER THEORY: A CRITICAL RE-EXAMINATION

GEORGE J. KLIR

Department of Systems Science, Thomas J. Watson School, State University of New York, Binghamton, New York 13901, USA

ARTHUR RAMER

School of Electrical Engineering and Computer Science, University of Oklahoma, Norman, Oklahoma 73019, USA

(*Received 24 August 1990*)

Measures of two types of uncertainty that coexist in the Dempster–Shafer theory are overviewed. A measure of one type of uncertainty, which expresses nonspecificity of evidential claims, is well justified on both intuitive and mathematical grounds. Proposed measures of the other types of uncertainty, which attempt to capture conflicts among evidential claims, are shown to have some deficiencies. To alleviate these deficiencies, a new measure is proposed. This measure, which is called a measure of discord, is not only satisfactory on intuitive grounds, but has also desirable mathematical properties. A measure of total uncertainty, which is defined as the sum of nonspecificity and discord, is also discussed. The paper focuses on conceptual issues. Mathematical properties of the measure of discord are only stated; their proofs are given in a companion paper.[15]

INDEX TERMS: Dempster–Shafer theory, uncertainty, nonspecificity, dissonance, confusion, discord, conflict, belief measure, plausibility measure, possibility theory.

1. INTRODUCTION

From the mid-seventeenth century (when the formal concept of numerical probability emerged[3]) until the mid-twentieth century, uncertainty was conceived solely in terms of probability theory. This seemingly unique connection between uncertainty and probability theory, which was taken for granted for three centuries, has only recently been challenged. The challenge came from several mathematical theories, distinct from probability theory, which are demonstrably capable of characterizing situations under uncertainty. They include: Choquet's theory of capacities,[1] theory of fuzzy sets,[22] theory of fuzzy measures,[20] theory of random sets,[11] possibility theory,[23] theory of rough sets,[13] *O*-theory,[12] and the Dempster–Shafer theory.[16]

An important aspect of every mathematical theory by which we conceptualize uncertainty is the capability to quantify the uncertainty involved. This requires that we can measure, in a unique and adequately justified way, the amount of uncertainty involved in each possible characterization of uncertainty within the theory.

Assume that we can measure the amount of uncertainty involved in a problem-solving situation conceptualized in a particular mathematical theory. Assume further that the amount of uncertainty can be reduced by obtaining relevant information as a result of some action (finding a relevant new fact, designing a relevant experiment and observing the experimental outcome, receiving a requested message, discovering a relevant historical record). Then, the amount of information obtained by the action may be measured by the reduction of uncertainty that results from the action. In this sense, the amount of uncertainty (pertaining to a problem-solving situation) and the amount of information (obtained by a relevant action) are intimately connected. Furthermore, the amount of information contained in a mathematical description of a problem solving situation may be measured by the difference between the maximum and actual amounts of uncertainty pertaining to the situation.

The nature of uncertainty (and the associated information) depends on the mathematical theory within which problem-solving situations are formalized. Each formalization of uncertainty in a problem-solving situation is a mathematical model of the situation. When we commit ourselves to a particular mathematical theory, our modelling becomes necessarily limited by the constraints of the theory. Clearly, a more general theory is capable of capturing uncertainties of some problem situations more faithfully than its less general competitors.

A measure of probabilistic uncertainty was established by Shannon in 1948.[17] The issue of how to measure uncertainty in the various alternative theories was investigated mostly in the 1980s.[8,18] It became clear by these investigations that the Dempster–Shafer theory is capable of formalizing simultaneously two distinct types of uncertainty. This contrasts with probability theory, within which only one of these two types of uncertainties can be captured.

The purpose of this paper is to critically re-examine previous results regarding the measurement of uncertainty in the Dempster–Shafer theory and propose a solution to some unresolved questions. This re-examination applies also to possibility theory, which is viewed here as a special case of the Dempster–Shafer theory.

2. TERMINOLOGY AND NOTATION

In order to introduce relevant terminology and notation, we briefly overview in this section basic properties of the Dempster–Shafer theory, possibility theory, and probability theory. The overview is rather concise; we assume that the reader is familiar with fundamentals of these theories at least to the extent at which they are covered in the text by Klir and Folger.[8]

Let X denote a universal set under consideration, assumed here to be finite for the sake of simplicity, and let $P(X)$ denote the power of set of X. Then, the *Dempster–Shafer theory* is based upon a function

$$m: P(X) \rightarrow [0, 1]$$

such that

$$m(\varnothing) = 0 \quad \text{and} \quad \sum_{A \subseteq X} m(A) = 1.$$

This function is called a *basic assignment*; the value $m(A)$ represents the degree of belief (based on relevant evidence) that a specific element of X belongs to set A, but not to any particular subset of A. Every set $A \in P(X)$ for which $m(A) \neq 0$ is called a *focal element*. The pair (F, m), where F denotes the set of all focal elements of m, is called a *body of evidence*.

Associated with each basic assignment m is a pair of measures, a *belief measure*, Bel, and a *plausibility measure*, Pl, which are determined for all sets $A \in P(X)$ by the equations

$$\mathrm{Bel}(A) = \sum_{B \subseteq A} m(B), \tag{1}$$

$$\mathrm{Pl}(A) = \sum_{B \cap A \neq \varnothing} m(B). \tag{2}$$

These equations and the definition of the basic assignment form the core of the Dempster–Shafer theory. This theory is most completely described by Shafer.[16]

Belief and plausibility measures are connected by the equation

$$\mathrm{Pl}(A) = 1 - \mathrm{Bel}(\bar{A}) \tag{3}$$

for all $A \in P(X)$, where $\bar{A}$ denotes the complement of A. Furthermore,

$$\mathrm{Bel}(A) \leqq \mathrm{Pl}(A) \tag{4}$$

for all $A \in P(X)$.

A belief measure (or a plausibility measure) becomes a *probability measure*, Pr, when all focal elements are singletons. In this case, $\mathrm{Pr}(A) = \mathrm{Bel}(A) = \mathrm{Pl}(A)$ for all $A \in P(X)$, which follows immediately from Eqs. (1) and (2), and we obtain the *additivity property*

$$\mathrm{Pr}(A \cup B) = \mathrm{Pr}(A) + \mathrm{Pr}(B) - \mathrm{Pr}(A \cap B) \tag{5}$$

of probability measures. Any probability measure, Pr, on a finite set X can be uniquely determined by a *probability distribution function*

$$p: X \to [0, 1]$$

via the formula

$$\mathrm{Pr}(A) = \sum_{x \in A} p(x). \tag{6}$$

From the standpoint of the Dempster–Shafer theory, clearly

$$p(x) = m(\{x\}).$$

When some focal elements are not singletons, Eq. (5) bifurcates into the inequalities

$$\mathrm{Bel}(A \cup B) \geqq \mathrm{Bel}(A) + \mathrm{Bel}(B) - \mathrm{Bel}(A \cap B), \tag{7}$$

$$\mathrm{Pl}(A \cup B) \leqq \mathrm{Pl}(A) + \mathrm{Pl}(B) - \mathrm{Pl}(A \cap B). \tag{8}$$

When all focal elements are nested (ordered by set inclusion), we obtain special plausibility measures, which are called *possibility measures* (or *consonant plausibility measures*), and the corresponding special belief measures, which are called *necessity* measures. A possibility measure, Pos, is conveniently (and uniquely) determined by a *possibility distribution function*

$$r: X \to [0, 1]$$

via the formula

$$\mathrm{Pos}(A) = \max_{x \in A} r(x) \tag{9}$$

for all $A \in P(X)$. The corresponding necessity measure, Nec, is then determined for all $A \in P(X)$ by a formula equivalent to Eq. (3),

$$\mathrm{Nec}(A) = 1 - \mathrm{Pos}(\bar{A}). \tag{10}$$

As shown later, possibility distributions and basic assignments of nested bodies of evidence are uniquely connected via Eqs. (13) and (14).

A theory that deals with nested bodies of evidence in terms of possibility and necessity measures is usually called a *possibility theory*. Possibility and necessity measures satisfy the equations

$$\mathrm{Pos}(A \cup B) = \max\,[\mathrm{Pos}(A), \mathrm{Pos}(B)], \tag{11}$$

$$\mathrm{Nec}(A \cap B) = \min\,[\mathrm{Nec}(A), \mathrm{Nec}(B)]. \tag{12}$$

Assume that $X = \{x_1, x_2, \ldots, x_n\}$ and let $A_1 \subset A_2 \subset \cdots \subset A_n$, where $A_i = \{x_1, x_2 \ldots, x_i\}$, $i = 1, 2, \ldots, n$, be a complete sequence of nested subsets that contains all focal elements of a possibility measure Pos. That is, if $m(A) \neq 0$ then $A \in \{A_1, A_2, \ldots, A_n\}$. Let $m_i = m(A_i)$ and $r_i = r(x_i)$ for all $i = 1, 2, \ldots, n$. Then, the n-tuples

$$\mathbf{m} = (m_1, m_2, \ldots, m_n),$$

$$\mathbf{r} = (r_1, r_2, \ldots, r_n)$$

fully characterize the basic assignment and the possibility distribution, respectively, by which the possibility measure Pos is defined. The nested structure implies that $r_i \geqq r_{i+1}$ for all $i = 1, 2, \ldots, n-1$. Furthermore,

$$r_i = \sum_{k=i}^{n} m_k, \tag{13}$$

$$m_i = r_i - r_{i+1} \tag{14}$$

for all $i = 1, 2, \ldots, n$, where $r_{n+1} = 0$ by convention.[8]

Possibility theory can be formulated not only in terms of nested bodies of evidence within the Dempster–Shafer theory, but also in terms of fuzzy sets. It was introduced in this latter manner by Zadeh.[23] A *fuzzy set* is a set whose boundary is not sharp. That is, the change from nonmembership to membership in a fuzzy set is gradual rather than abrupt. This gradual change is expressed by a *membership grade function*, μ_A, of the form

$$\mu_A: X \to [0, 1],$$

where A is a label of the fuzzy set defined by this function within the universal set X. The value $\mu_A(x)$ expresses the grade of membership of element x of X in the fuzzy set A or, in other words, the degree of compatibility of x with the concept represented by the fuzzy set. A fuzzy set A is called *normalized* when $\mu_A(x) = 1$ for at least one $x \in X$. If $\mu_A(x) \leqq \mu_B(x)$ for all $x \in X$, then A is called a *fuzzy subset* of B.

An important concept associated with fuzzy sets is an α-cut. Given a fuzzy set A and a specific number $\alpha \in [0, 1]$, the α-*cut*, A_α, is a crisp (non-fuzzy) set

$$A_\alpha = \{x \in X \mid \mu_A(x) \geqq \alpha\}.$$

The set of all elements of X for which $\mu_A(x) > 0$ is called a *support* of the fuzzy set A; it is usually denoted by $\text{supp}(A)$.

Given a regular fuzzy set A with membership grade function μ_A (the range of μ_A is $[0, 1]$), Zadeh[23] defines a possibility distribution function, r_A, associated with A as numerically equal to μ_A, i.e.,

$$r_A(x) = \mu_A(x)$$

for all $x \in X$; then, he defines the corresponding possibility measure, Pos_A, by the equation

$$\text{Pos}_A(B) = \max_{x \in B} r_A(x)$$

for all $B \in P(X)$. In this interpretation of possibility theory, focal elements correspond to distinct α-cuts A of the fuzzy set A. This follows from the property that $A_\alpha \subset A_\beta$ when $\alpha > \beta$.

The Dempster–Shafer theory can be fuzzified. In its fuzzified form, the basic assignment is a function

$$\tilde{m}: \tilde{P}(X) \to [0, 1],$$

where $\tilde{P}(X)$ denotes the set of all fuzzy subsets of X. This function must satisfy the same requirements for the extended domain $\tilde{P}(X)$ as function m does for the domain $P(X)$. Plausibility and belief based upon $\tilde{m}$ are expressed by the following generalized counterparts of Eqs. (1) and (2),

$$\widetilde{\mathrm{Bel}}(A)=\sum_{B\in F}\tilde{m}(B)\left[1-\max_{x\in X}\min(1-\mu_A(x),\mu_B(x))\right], \tag{1'}$$

$$\widetilde{\mathrm{Pl}}(A)=\sum_{B\in F}\tilde{m}(B)\left[\max_{x\in X}\min(\mu_A(x),\mu_B(x))\right], \tag{2'}$$

where $\mu_A(x)$ and $\mu_B(x)$ are degrees of membership of element x in fuzzy sets A and B, respectively, and F is the set of all focal elements (fuzzy sets) associated with $\tilde{m}$.

3. MEASURES OF UNCERTAINTY IN THE DEMPSTER–SHAFER THEORY

It follows from the nature of the Dempster–Shafer theory that it subsumes two distinct types of uncertainty. One of them is well characterized by the name *nonspecificity*. It is now well established that this type of uncertainty is properly measured by a function N defined by the formula

$$N(m)=\sum_{A\in F}m(A)\log_2|A|, \tag{15}$$

where $|A|$ denotes the cardinality of the focal element A. This function, which was proven unique under appropriate requirements,[8] measures nonspecificity of a body of evidence in units that are called *bits*: one bit of uncertainty expresses the total ignorance regarding the truth or falsity of one proposition. The range of the function is

$$0\leqq N(m)\leqq\log_2|X|. \tag{16}$$

Function N is connected with a simple measure of information (and uncertainty) that was proposed within the classical set theory by Hartley in 1928.[4] He showed that, given a finite set of *possible* alternatives, A, the amount of information (in bits), $I(A)$, needed to characterize one of the alternatives is given by the simple formula

$$I(A)=\log_2|A|. \tag{17}$$

Function N can clearly be viewed as a weighted average of the Hartley information for all focal elements.

We can easily see that function N has no connection with the probabilistic measure of uncertainty, the Shannon entropy H, which assumes in the Dempster–Shafer theory the form

$$H(m)=-\sum_{x\in X}m(\{x\})\log_2 m(\{x\}). \tag{18}$$

Indeed, all focal elements of any probability measure are singletons, which means that $|A|=1$ for all $A \in F$ in Eq. (15). Consequently, if m defines a probability measure, then $N(m)=0$. That is, probability theory is devoid of nonspecificity and, hence, the Shannon entropy measures uncertainty that is of a different type.

It is well understood now that the probabilistic uncertainty results from evidential claims focusing on disjoint subsets (singletons) and thus conflicting with one another. The greater the lack of discrimination among the evidential claims expressed by a probability distribution, the greater the conflict and, as a measure of the conflict, the greater is the value of the Shannon entropy.

What is the generalized counterpart of the Shannon entropy in the Dempster–Shafer theory? Thus far, two candidates were proposed:

$$E(m) = -\sum_{A \in F} m(A) \log_2 \mathrm{Pl}(A), \tag{19}$$

$$C(m) = -\sum_{A \in F} m(A) \log_2 \mathrm{Bel}(A). \tag{20}$$

Function E defined by Eq. (19), which is usually called a *measure of dissonance* (or *entropy-like measure*) was proposed by Yager.[21] Function C given by Eq. (20), which is usually called a *measure of confusion*, was proposed by Höhle.[6] It is obvious that either of the functions collapses into the Shannon entropy when m defines a probability measure.

What do functions E and C actually measure? From Eq. (2) and the general property of basic assignments (satisfied for every $A \in P(X)$),

$$\sum_{A \cap B = \varnothing} m(B) + \sum_{A \cap B \neq \varnothing} m(B) = 1,$$

we obtain

$$E(m) = -\sum_{A \in F} m(A) \log_2 \left[1 - \sum_{A \cap B = \varnothing} m(B) \right]. \tag{21}$$

The term

$$K(A) = \sum_{A \cap B = \varnothing} m(B)$$

in Eq. (21) represents the total evidential claim pertaining to focal elements that are disjoint with the set A. That is, $K(A)$ expresses the sum of all evidential claims that fully conflict with the one focusing on the set A. Clearly, $K(A) \in [0, 1]$. The function

$$-\log_2[1 - K(A)],$$

which is employed in Eq. (21), is monotonic increasing with $K(A)$ and extends its range from $[0, 1]$ to $[0, \infty)$. The choice of the logarithmic function is motivated in the same way as in the classical case of the Shannon entropy.

It follows from these facts and the form of Eq. (21) that $E(m)$ is the mean (expected) value of the conflict among evidential claims within a given body of evidence (F, m); it measures the conflict in bits and its range is $[0, \log_2 |X|]$.

Functional E is not fully satisfactory since we feel intuitively that $m(B)$ conflicts with $m(A)$ whenever $B \nsubseteq A$, not only when $B \cap A = \varnothing$. This broader view of conflict is expressed by the measure of confusion C given by Eq. (20). Let us demonstrate this fact.

From Eq. (1) and the general property of basic assignments (satisfied for every $A \in P(X)$),

$$\sum_{B \subseteq A} m(B) + \sum_{B \nsubseteq A} m(B) = 1,$$

we get

$$C(m) = -\sum_{A \in F} m(A) \log_2 \left[1 - \sum_{B \nsubseteq A} m(B) \right]. \tag{22}$$

The term

$$L(A) = \sum_{B \nsubseteq A} m(B)$$

in Eq. (22) expresses the sum of all evidential claims that conflict with the one focusing on the set A according to the broader view of conflict: $m(B)$ conflicts with $m(A)$ whenever $B \nsubseteq A$. The reasons for using the function

$$-\log_2 [1 - L(A)]$$

instead of $L(A)$ in Eq. (22) are the same as already explained in the context of function E. The conclusion is that $C(m)$ is the mean (expected) value of the conflict, viewed in the broader sense, among evidential claims within a given body of evidence (F, m).

Function C is also not fully satisfactory as a measure of conflicting evidential claims within a body of evidence, but for a different reason than function E. Although it employs the broader, and more satisfactory, view of conflict, it does not properly scale each particular conflict of $m(B)$ with respect to $m(A)$ according to the degree of violation of the subsethood relation $B \subseteq A$. We feel intuitively that the more this subsethood relation is violated the greater the conflict.

Hence, neither of the two functions, E and C, hitherto proposed as candidates for measuring the mean value of conflict among evidential claims within each given body of evidence is really acceptable on intuitive grounds. To alleviate this frustrating situation, we propose a new measure of conflict that is both intuitively and mathematically sound.

4. NEW MEASURE OF CONFLICT

On the basis of our conclusions regarding the deficiencies of functions E and C as adequate measures of conflict in the Dempster–Shafer theory, we propose to replace them with the following function:

$$D(m) = -\sum_{A \in F} m(A) \log_2 \left[1 - \sum_{B \in F} m(B) \frac{|B - A|}{|B|} \right]. \tag{23}$$

Observe that the term

$$\mathrm{Con}\,(A) = \sum_{B \in F} m(B) \frac{|B - A|}{|B|}$$

in Eq. (23) expresses the sum of individual conflicts of evidential claims with respect to a particular set A, each of which is properly scaled by the degree to which the subsethood $B \subseteq A$ is violated. This conforms exactly to the intuitive idea of conflict that emerged from the critical re-examination of functions E and C in the previous section. Let function Con be called a *conflict*. Clearly, $\mathrm{Con}(A) \in [0, 1]$ and, furthermore,

$$K(A) \leqq \mathrm{Con}\,(A) \leqq L(A). \tag{24}$$

The reason for using the function

$$-\log_2 [1 - \mathrm{Con}\,(A)]$$

instead of Con in Eq. (23) is exactly the same as previously explained in the context of function E. This monotonic transformation extends the range of $\mathrm{Con}\,(A)$ from $[0, 1]$ to $[0, \infty)$.

Function D, which we propose to call a *measure of discord*, is clearly a measure of the mean conflict (expressed by the logarithmic transformation of function Con) among evidential claims within each given body of evidence. It follows immediately from (24) that

$$E(m) \leqq D(m) \leqq C(m). \tag{25}$$

Observe that $|B - A| = |B| - |A \cap B|$ and, consequently, Eq. (23) can be rewritten as

$$D(m) = -\sum_{A \in F} m(A) \log_2 \sum_{B \in F} m(B) \frac{|A \cap B|}{|B|}. \tag{26}$$

From (25), it is obvious that

$$\mathrm{Bel}(A) \leqq \sum_{B \in F} m(B) \frac{|A \cap B|}{|B|} \leqq \mathrm{Pl}(A). \tag{27}$$

Function D is applicable equally well to the fuzzified Dempster–Shafer theory provided that the cardinality $|A|$ of a fuzzy set A is defined by the formula

$$|A| = \sum_{x \in X} \mu_A(x)$$

and the set intersection is determined by the minimum operator. In this sense,

$$S(B, A) = \frac{|A \cap B|}{|B|}$$

in Eq. (26) expresses the degree of subsethood, $S(B, A)$, of a fuzzy set B in a fuzzy set A, as shown by Kosko.[9]

It is easy to see that function D measures the conflict of evidential claims within each body of evidence in bits: $D(m) = 1$ is equivalent to a full conflict between the evidential claims regarding the truth or falsity of a single proposition. We have also managed to prove that the function is additive (in the same sense as functions E and C)[2] and that its range is $[0, \log_2 |X|]$. The minimum is obtained for all bodies of evidence with a single focal element; its maximum is obtained only when m defines the uniform probability distribution on X. Proofs of these properties, which are rather technical, are a subject of another paper.[15]

When we specialize to possibility theory and, hence, deal only with nested bodies of evidence, the measure of discord, D, is still applicable, while the measure of dissonance, E, is not. That is, nested bodies are consonant (in the sense of function E), but they are not, in general, conflict-free (in the sense of the more general function D). They are conflict-free only if they contain one focal element. Since such bodies of evidence correspond to crisp sets, it seems that the measure of discord might also play a role as a measure of fuzziness of normalized fuzzy sets. This issue requires further investigation.

Using the notation introduced in Section 2, which is based upon the notion of a complete sequence of nested subsets that contains all focal elements, it is easy to derive the following possibilistic form of the measure of discord:

$$D(m) = -\sum_{i=1}^{n-1} m_i \log_2\left(\sum_{j=1}^{i} m_j + \sum_{j=i+1}^{n} m_j \frac{i}{j}\right). \tag{28}$$

Furthermore, using Eq. (14), we obtain

$$D(m) = -\sum_{i=1}^{n-1} (r_i - r_{i+1}) \log_2\left[1 - i \sum_{j=i+1}^{n} \frac{r_j}{(j-1)j}\right]. \tag{29}$$

5. TOTAL UNCERTAINTY

As previously suggested by Lamata and Moral,[10] we may define the amount of total uncertainty in the Dempster–Shafer theory as the sum of amounts of the two types of uncertainty that coexist in the theory. While Lamata and Moral proposed to add the values of nonspecificity and dissonance, we propose, as a natural consequence of our discussion in Section 4, to add the values of nonspecificity and discord. Hence, we define the total uncertainty, $T(m)$, for any given body of evidence (F, m) as the sum

$$T(m) = N(m) + D(m) = \sum_{A \in F} m(A) \log_2 \frac{|A|}{\sum\limits_{B \in F} m(B) \frac{|A \cap B|}{|B|}}. \tag{30}$$

It is our conjecture that the range of this function is $[0, \log_2 |X|]$ and that it possesses desirable properties such as symmetry, expansibility, additivity, and subadditivity. Although there is ample evidence to support this conjecture, relevant proofs necessary to convert it to a theorem are yet to be done. These proofs, if accomplished, will make function T as well justified as a measure of uncertainty in the Dempster–Shafer theory as the Shannon entropy is in probability theory.

Observe that $T(m) = 0$ whenever $m(\{x\}) = 1$ for some $x \in X$. The presumed maximum, $T(m) = \log_2 |X|$, is not unique; it is obtained not only for $m(X) = 1$ and for the uniform probability distribution on X, but also for other bodies of evidence that seem to possess certain symmetries. For example, the maximum is obtained for the body of evidence defined on $X = \{1, 2, 3, 4\}$ whose focal elements are $\{1, 2\}$, $\{2, 3\}$, $\{3, 4\}$, $\{1, 4\}$ and the values of the basic assignment are uniform (1/4 for each focal element).

6. CONCLUSIONS

The Dempster–Shafer theory and one of its special branches, possibility theory, have become viable alternatives to probability theory, which appear to be more appropriate for dealing with some problem-solving situations involving uncertainty.[5,19] We have argued elsewhere that the proper way of managing uncertainty in problem-solving situations is to employ two complementary principles, the principle of minimum uncertainty and the principle of maximum uncertainty.[7,8] However, to make these principles operational, relevant uncertainty must be adequately quantified.

We are sufficiently confident that the measure of discord, which is proposed in this paper, together with the well established measure of nonspecificity, quantify adequately the two types of uncertainty that are captured by the Dempster–Shafer theory. These measures are now justified, both on intuitive and mathematical grounds, equally well as the Hartley measure in classical set theory and the Shannon entropy in probability theory.

The measure of total uncertainty, given by the sum of discord and nonspecificity, requires further mathematical investigation. If our conjecture regarding mathematical properties of this function is proven correct, which we consider highly plausible, the function will undoubtedly play a particularly important role in the principles of minimum and maximum uncertainty.

REFERENCES

1. G. Choquet, "Theory of capacities." *Annales de L'Institut Fourier*, **5**, 1953/54, pp. 131–295.

2. D. Dubois and H. Prade, "Properties of measures of information in evidence and possibility theories." *Fuzzy Sets and Systems*, **24**, No. 2, 1987, pp. 161–182.
3. I. Hacking, *The Emergency of Probability*. Cambridge University Press, Cambridge, 1975.
4. R. V. L. Hartley, "Transmission of information." *The Bell System Technical J.*, **7**, 1928, pp. 535–563.
5. S. J. Henkind and M. C. Harrison, "An analysis of four uncertainty calculi." *IEEE Trans. on Systems, Man, and Cybernetics*, **18**, No. 5, 1988, pp. 700–714.
6. U. Höhle, "Entropy with respect to plausibility measures." *Proc. 12th IEEE Symp. on Multiple-Valued Logic*, Paris, 1982, pp. 167–169.
7. G. J. Klir, "The role of uncertainty principles in inductive systems modelling." *Kybernetes*, **17**, No. 2, 1988, pp. 24–34.
8. G. J. Klir and T. A. Folger, *Fuzzy Sets, Uncertainty, and Information*. Prentice Hall, Englewood Cliffs (N.J.), 1988.
9. B. Kosko, "Fuzzy entropy and conditioning." *Information Sciences*, **40**, No. 1, 1987, pp. 1–10.
10. M. T. Lamata and S. Moral, "Measures of entropy in the theory of evidence." *Intern. J. of General Systems*, **14**, No. 4, 1988, pp. 297–305.
11. G. Matheron, *Random Sets and Integral Geometry*. John Wiley, New York, 1975.
12. E. M. Oblow, "O-theory: a hybrid uncertainty theory." *Intern. J. of General Systems*, **13**, No. 2, 1987, pp. 95–106.
13. Z. Pawlak, "Rough sets." *Intern. J. of Computer and Information Sciences*, **11**, 1982, pp. 341–356.
14. A. Ramer, "Uniqueness of information measure in the theory of evidence." *Fuzzy Sets and Systems*, **24**, No. 2, 1987, pp. 183–196.
15. A. Ramer and G. J. Klir, "Measures of conflict and discord." *Information Sciences* (to appear).
16. G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, 1976.
17. C. E. Shannon, "The mathematical theory of communication." *The Bell System Technical Journal*, **27**, 1948, pp. 379–423, 623–656.
18. *Special Issue of Fuzzy Sets and Systems on Measures of Uncertainty*, **24**, No. 2, November 1987, pp. 139–254.
19. H. E. Stephanou and A. P. Sage, "Perspectives on imperfect information." *IEEE Trans. on Systems, Man, and Cybernetics*, **17**, No. 5, 1987, pp. 780–798.
20. M. Sugeno, "Fuzzy measures and fuzzy integrals: a survey." In: *Fuzzy Automata and Decision Processes*, edited by M. M. Gupta, G. N. Saridis, and B. R. Gaines, North-Holland, Amsterdam and New York, 1977, pp. 89–102.
21. R. Yager, "Entropy and specificity in a mathematical theory of evidence." *Intern. J. of General Systems*, **9**, No. 4, 1983, pp. 249–260.
22. L. A. Zadeh, "Fuzzy sets." *Information and Control*, **8**, No. 3, 1965, pp. 338–353.
23. L. A. Zadeh, "Fuzzy sets as a basis for a theory of possibility." *Fuzzy Sets and Systems*, **1**, No. 1, 1978, pp. 3–28.

For biographies and photographs of the authors, please see Vol. 17 (1990), Nos. 2–3, pp. 248 and 275 of this journal.