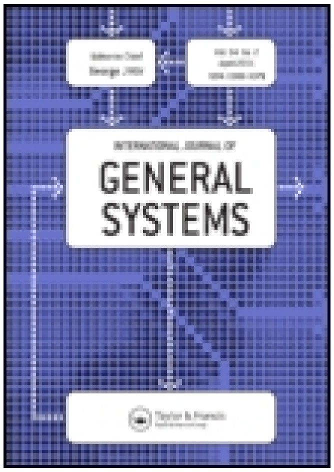

International Journal of General Systems
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ggen20

MEASURING TOTAL UNCERTAINTY IN DEMPSTER-SHAFER THEORY: A NOVEL APPROACH

DAVID HARMANEC [a] & GEORGE J. KLIR [a]

[a] Department of Systems Science , Thomas J. Watson School of Engineering and Applied Science, State University of New York at Binghamton , Binghamton, NewYork, 13902, U.S.A.
Published online: 06 Apr 2007.

Int. J. General Systems, Vol. 22, pp. 405–419
Reprints available directly from the publisher
Photocopying permitted by license only

MEASURING TOTAL UNCERTAINTY IN DEMPSTER-SHAFER THEORY: A NOVEL APPROACH†

DAVID HARMANEC AND GEORGE J. KLIR

Department of Systems Science, Thomas J. Watson School of Engineering and Applied Science, State University of New York at Binghamton, Binghamton, New York 13902, U.S.A.

(Received 6 December 1993)

A novel approach to measuring uncertainty and uncertainty-based information in Dempster-Shafer theory is proposed (independently also proposed by Maeda et al. [1993]). It is shown that the proposed measure of total uncertainty in Dempster-Shafer theory is both additive and subadditive, has a desired range, and collapses correctly to either the Shannon entropy or the Hartley measure of uncertainty for special probability assignment functions. The paper is restricted, for the sake of simplicity, to finite sets.

INDEX TERMS: Uncertainty, Dempster-Shafer theory, maximum entropy principle.

1. INTRODUCTION

Dempster-Shafer theory (DST), which is also referred to as *evidence theory*, is one of the many theories of uncertainty and uncertainty-based information that emerged during the last three decades or so [Klir, 1993]. The theory, which is best explicated in the books by Shafer [1976] and Guan and Bell [1991], has quickly become one of the leading theories of uncertainty, due primarily to its broad applicability.

In this paper, we deal only with one question pertaining to DST: how to measure uncertainty and uncertainty-based information in DST? This question, which is formulated more precisely in Sec. 3, has been investigated since the early 1980's. Although a considerable progress has been made by these investigations, as outlined in Sec. 4, no fully satisfactory answer was obtained. Proposing a novel approach to measuring uncertainty in DST in Sec. 5 (independently also proposed by Maeda et al. [1993]), we show in Sec. 6 that the question can now be answered fully satisfactorily. For the sake of simplicity, the paper is restricted to finite sets.

2. DST: RELEVANT CONCEPTS

Let X denote a nonempty *universal set* under consideration (often referred to in the literature on DST as a *frame of discernment*), which is assumed here to be finite for

†This research work was supported in part by the NSF Grant No. IST-90 15675.

the sake of simplicity, and let $\mathcal{P}(X)$ denote the power set of X. Then, DST may be formulated in terms of a function

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

such that $m(\varnothing) = 0$ and

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

This function is called a *basic probability assignment*. For each $A \in \mathcal{P}(X)$, the value $m(A)$ expresses the proportion to which all available evidence supports the claim that a particular element of X, whose characterization in terms of relevant attributes is deficient, belongs to the set A. This value, $m(A)$, pertains only to a single set, set A; it does not imply any additional claims regarding subsets of A. If there is some additional evidence supporting the claim that the element belongs to a subset of A, say $B \subset A$, it must be expressed by appropriate nonzero value $m(B)$.

Contrary to probability theory, which is based on a single additive measure, DST is based on two nonadditive measures, a *belief measure*, Bel, and a *plausibility measure*, Pl. Given a basic probability assignment m, the two measures are determined for all sets $A \in \mathcal{P}(X)$ by the formulas

$$\mathrm{Bel}(A) = \sum_{B \subseteq A} m(B) \tag{2}$$

$$\mathrm{Pl}(A) = \sum_{B \cap A \neq \varnothing} m(B). \tag{3}$$

Inverse procedure is also possible. Given, for example, a belief measure, Bel, the corresponding basic probability assignment m is determined for all $A \in \mathcal{P}(X)$ by the formula

$$m(A) = \sum_{B \subseteq A} (-1)^{|A-B|}\mathrm{Bel}(\mathrm{B}), \tag{4}$$

where $|A - B|$ denotes the cardinality of the set $A - B$ [Shafer, 1976].

It follows from Eqs. (2) and (3) that belief and plausibility measures are connected by the equation

$$\mathrm{Pl}(A) = 1 - \mathrm{Bel}(\bar{A}) \tag{5}$$

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

$$\mathrm{Bel}(A) \leq \mathrm{Pl}(A) \tag{6}$$

for all $A \in \mathcal{P}(X)$.

Given a basic probability assignment m, every set $A \in \mathcal{P}(X)$ for which $m(A) > 0$

is called a *focal element*. The pair $\langle \mathcal{F}, m \rangle$, where $\mathcal{F}$ denotes the set of all focal elements induced by m, is called a *body of evidence*.

Total ignorance is expressed in DST by $m(X) = 1$ and $m(A) = 0$ for all $A \neq X$. Full certainty is expressed by $m(\{x\}) = 1$ for one particular element of X and $m(A) = 0$ for all $A \neq \{x\}$.

DST is a special branch of *fuzzy measure theory* [Wang and Klir, 1992], a theory that deals with general measures that do not require additivity. Belief measures are *superadditive* in the sense of the inequality

$$\mathrm{Bel}(A_1 \cup A_2 \cup \cdots \cup A_N) \geq \sum_j \mathrm{Bel}(A_j) - \sum_{j<k} \mathrm{Bel}(A_j \cap A_k) + \cdots + (-1)^{N+1} \mathrm{Bel}(A_1 \cap A_2 \cap \cdots \cap A_N), \tag{7}$$

which holds for all possible families of subsets of X. Plausibility measures, on the other hand, are *subadditive* in the sense of the inequality

$$\mathrm{Pl}(A_1 \cap A_2 \cap \cdots \cap A_N) \leq \sum_j \mathrm{Pl}(A_j) - \sum_{j<k} \mathrm{Pl}(A_j \cup A_k) + \cdots + (-1)^{N+1} \mathrm{Pl}(A_1 \cup A_2 \cup \cdots \cup A_N), \tag{8}$$

which, again, holds for all possible families of subsets of X.

Two significant special types of bodies of evidence are recognized in DST. One type is obtained when all focal elements are *singletons* $\{x\}$ for some $x \in X$. In these cases, $\mathrm{Bel}(A) = \mathrm{Pl}(A)$ by Eqs. (2) and (3), superadditivity and subadditivity collapse into additivity, and we obtain the classical *probability measures*. The other type is obtained when focal elements are required to be *nested*. In these cases, belief measures and plausibility measures become, respectively, *necessity measures* and *possibility measures* of *possibility theory* [Dubois and Prade, 1988; Klir and Folger, 1988; Wang and Klir, 1992].

To investigate ways of measuring the amount of uncertainty represented by each body of evidence, it is essential to understand properties of bodies of evidence whose focal elements are subsets of the Cartesian product of two sets. That is, we need to examine basic probability assignments of the form

$$m: \mathcal{P}(X \times Y) \rightarrow [0, 1],$$

where X and Y denote frames of discernment pertaining to two distinct domains of inquiry (e.g., two investigated variables), which may be connected in some fashion. Let any basic probability assignment of this form be called a *joint basic probability assignment*.

Each focal element induced by a joint probability assignment m is a binary relation, R, on $X \times Y$. When R is projected on set X and on set Y, we obtain, respectively, the sets

$$R_X = \{x \in X \mid \langle x, y \rangle \in R \text{ for some } y \in Y\},$$

$$R_Y = \{y \in Y \mid \langle x, y \rangle \in R \text{ for some } x \in X\}.$$

These sets are instrumental in calculating *marginal basic probability assignments*, m_X and m_Y, from the given joint basic probability assignment m:

$$m_X(A) = \sum_{R|A=R_X} m(R) \text{ for all } A \in \mathcal{P}(X),$$

$$m_Y(B) = \sum_{R|B=R_Y} m(R) \text{ for all } B \in \mathcal{P}(Y).$$

Bodies of evidence $\langle \mathcal{F}_X, m_X \rangle$ and $\langle \mathcal{F}_Y, m_Y \rangle$ are, according to DST, *noninteractive* if and only if for all $A \in \mathcal{F}_X$ and all $B \in \mathcal{F}_Y$

$$m(A \times B) = m_X(A) \cdot m_Y(B) \tag{9}$$

and $m(R) = 0$ for all $R \neq A \times B$. That is, two marginal bodies of evidence are noninteractive if and only if the only focal elements of the joint body of evidence are Cartesian products of focal elements of the marginal bodies and m is determined from m_X and m_Y by Eq. (9).

3. MEASURING UNCERTAINTY IN DST

To measure the *amount of uncertainty* ingrained in any given belief measure, we need to find an appropriate function of the form

$$AU: \mathcal{B} \rightarrow [0, \infty),$$

where $\mathcal{B}$ denotes the set of all belief measures. Due to the one-to-one correspondences between belief measures, plausibility measures, and basic probability assignments, as expressed by Eqs. (2)–(4), the domain of this function may also be reinterpreted in terms of the corresponding plausibility measures or basic probability assignments.

To qualify as a meaningful measure of the total uncertainty ingrained in any given belief measure, function AU must satisfy certain requirements that are considered essential on intuitive grounds. It is generally agreed that at least none of the following requirements should be violated by the function AU:

Req. 1. When Bel is a probability measure defined on $\mathcal{P}(X)$, AU collapses to the *Shannon entropy* [Shannon, 1948; Klir and Folger, 1988] and has the form

$$AU(\text{Bel}) = -\sum_{x \in X} \text{Bel}(\{x\}) \log_2 \text{Bel}(\{x\}). \tag{10}$$

Req. 2. When Bel defined on $\mathcal{P}(X)$ focuses on a single set A(i.e., $\text{Bel}(A) = 1$ for one particular set A and $\text{Bel}(B) = 0$ for all $B \not\supseteq A$, where $A, B \subseteq X$), AU collapses to the *Hartley measure* of set-theoretic uncertainty [Hartley, 1928; Klir and Folger, 1988].

Req. 3. When Bel is defined on $\mathcal{P}(X)$ and the uncertainty is measured in bits [Klir and Folger, 1988], the *range* of AU is defined by the inequalities

$$0 \leq AU(\text{Bel}) \leq \log_2|X|. \tag{11}$$

The justification of this range is the same as for the classical measures of uncertainty, the Hartley measure and the Shannon entropy.

Req. 4. When Bel is an arbitrary joint belief measure defined on $\mathcal{P}(X \times Y)$ and Bel_X, Bel_Y are the associated marginal belief measures,

$$AU(\text{Bel}) \leq AU(\text{Bel}_X) + AU(\text{Bel}_Y). \tag{12}$$

This property is usually referred to as *subadditivity*.

Req. 5. When Bel, Bel_X, Bel_Y have the same meaning as in Req. 4 and Bel_X, Bel_Y are based on noninteractive bodies of evidence,

$$AU(\text{Bel}) = AU(\text{Bel}_X) + AU(\text{Bel}_Y) \tag{13}$$

This property is usually referred to as *additivity*.

The concept of uncertainty is intimately connected with the concept of information. The most fundamental aspect of this connection is that uncertainty involved in any situation is a result of information deficiency. Once we can measure uncertainty, we can measure the associated information as well. Assume, for example, that the amount of uncertainty can be reduced by obtaining relevant information as a result of some action (finding a relevant new fact, receiving a requested message, performing a relevant experiment and observing its outcome, discovering a relevant historical record, etc.). Then, the amount of information obtained by the action may be measured by the reduction of uncertainty that results from the action. When information is conceived and measured solely in terms of uncertainty reduction, it should be given a special name, such as *uncertainty-based information*, to distinguish it from other types of information.

4. PROPOSED UNCERTAINTY MEASURES: A CHRONOLOGY

In order to characterize the context in which our approach to measuring uncertainty in DST emerged, we outline in this section previous developments regarding this issue. The following is a concise chronology of these developments. We do not cover here the various proposed measures that do not attempt to satisfy the requirements stated in Sec. 3. Moreover, we do not attempt to list all relevant references. We list only references that, in our opinion, are essential for understanding the context in which the new measure of uncertainty in DST is introduced in this paper.

1. Measurement of uncertainty (and associated information) was first conceived in terms of classical set theory. It was shown by Hartley [1928] that using the function

$$I(A) = \log_2 |A| \tag{14}$$

is the only sensible way to measure the amount of uncertainty associated with a finite set of possible alternatives when we choose bits as measurement units.

The uniqueness of function I as a measure of uncertainty expressed in set-theoretic terms was also proven axiomatically by Rényi [1970].

2. A measure of uncertainty pertaining to probability theory was first conceived by Shannon [1948]. The measure, which is usually referred to as the *Shannon entropy*, assumes in DST the form given by Eq. (10). Its uniqueness has been proven axiomatically in numerous ways [Klir and Folger, 1988].

3. A natural generalization of the Hartley function from classical set theory to fuzzy set theory and possibility theory was proposed by Higashi and Klir [1983] under the name *U-uncertainty*. For any normal fuzzy set A defined on a finite universal set X, the U-uncertainty has the form

$$U(A) = \int_0^1 \log_2 |A_\alpha| d\alpha, \tag{15}$$

where $|A_\alpha|$ denotes the cardinality of the α-cut of A [Klir and Folger, 1988]. When the U-uncertainty is applied to possibility theory, α in Eq. (15) assumes the role of a possibility distribution and the α-cuts A_α become focal elements of a possibilistic body of evidence.

4. Attempts to generalize the Shannon entropy from probability theory to DST were made by Höhle [1982] and Yager [1983] resulting, respectively, in the *entropy-like measures*

$$C(m) = -\sum_{A \subseteq X} m(A) \log_2 \mathrm{Bel}(A), \tag{16}$$

$$E(m) = -\sum_{A \subseteq X} m(A) \log_2 \mathrm{Pl}(A). \tag{17}$$

These functions are usually referred to in the literature as a *measure of confusion* and a *measure of dissonance*, respectively. It was recognized that these functions, whose domain is the set of all basic probability assignments, measure certain aspects of conflicts among evidential claims within each body of evidence.

5. It was shown by Dubois and Prade [1985] that the U-uncertainty can be generalized to the form

$$N(m) = \sum_{A \subseteq X} m(A) \log_2 |A|, \tag{18}$$

which is applicable to any given basic probability assignment in DST. Function N is clearly a weighted average of the Hartley measure and, consequently, it is often referred to as a *Hartley-like measure* in DST. The type of uncertainty that is measured by function N is usually referred to as *nonspecificity*.

6. The uniqueness of the U-uncertainty as a measure of possibilistic nonspecificity was proven by Klir and Mariano [1987]. A formulation of a possibilistic counterpart of the well-known branching property of the Shannon entropy was crucial in this proof, as follows from the analysis made by Ramer and Lander [1987].

7. The uniqueness of the Hartley-like measure N as a measure of nonspecificity in DST was proven by Ramer [1987].
8. The three proposed measures of uncertainty in DST, functions N, C, E, were further analyzed by Dubois and Prade [1987] and Klir [1987]. The Hartley-like measure N emerged from these analyses as a fully justified measure of nonspecificity in DST. In addition to all essential properties, such as additivity, subadditivity, and proper range, function N was also shown to possess other desirable properties [Klir and Wierman, 1987]. The entropy-like measures, C and E, emerged as less justified. It was shown that neither of them is subadditive and, in addition, function C does not have the required range. Function E emerged thus as preferable to function C, but not fully justified.
9. The issue of measuring total uncertainty in DST was raised by Lamata and Moral [1988]. They proposed, as a measure of total uncertainty, NE, the sum

$$NE(m) = N(m) + E(m). \tag{19}$$

10. It was argued by Klir and Ramer [1990] that neither function C nor function E is fully satisfactory as a measure of conflicting evidential claims within a body of evidence. A new measure,

$$D(m) = -\sum_{A \subseteq X} m(A) \log_2 \sum_{B \subseteq X} m(B) \frac{|A \cap B|}{|B|}, \tag{20}$$

was introduced to overcome the identified conceptual deficiencies of functions C and E. Function D was given the name *discord* and was shown that

$$E(m) \leq D(m) \leq C(m). \tag{21}$$

A total uncertainty, ND, was then defined as the sum

$$ND(m) = N(m) + D(m), \tag{22}$$

and it was shown that ND (as well as D) is additive, has the required range, and collapses properly to the Shannon entropy and Hartley measure as required [Ramer and Klir, 1993]. Unfortunately, it was also shown that neither D nor ND is subadditive [Vejnarová, 1991].
11. Properties of function D in possibility theory were investigated by Geer and Klir [1991]. It was shown that $D(m)$ in possibility theory increases with $|X|$ and converges to a constant, estimated as 0.892 as $|X| \to \infty$. It was thus established that, contrary to previous views, possibilistic bodies of evidence also involve conflict (in spite of their nested structure), but its amount is bounded from above and negligible for large possibilistic bodies of evidence.
12. A new entropy-like measure,

$$S(m) = -\sum_{A \subseteq X} m(A) \log_2 \sum_{B \subseteq X} m(B) \frac{|A \cap B|}{|A|}, \tag{23}$$

was introduced by Klir and Parviz [1992] to overcome a perceived conceptual deficiency of function D involving focal elements that are proper subsets of other focal elements. Function S was given the name *strife*. The controversy regarding the differences between functions D and S were further analyzed by Klir and Yuan [1993]. It was concluded that S expresses the conflict among *disjunctive set-valued statements*, while D expresses the conflict among *conjunctive set-valued statements*. Both functions were shown to be bounded from above in exactly the same way in possibility theory. Mathematical properties of strife and the associated total measure of uncertainty,

$$NS(m) = N(m) + S(m), \tag{24}$$

were analyzed by Vejnarová and Klir [1993]. It was proven that functions S and NS have similar properties as the corresponding functions D and ND: they are additive, have the required range, and collapse in special cases to the Shannon entropy and Hartley measure as required. However, they are not subadditive.

5. A NOVEL APPROACH TO MEASURING TOTAL UNCERTAINTY IN DST

It follows from all previous work on the measurement of uncertainty in DST, as outlined in Sec. 4, that two types of uncertainty coexist in DST. These types of uncertainty are well captured by the names *nonspecificity* and *conflict*. Nonspecificity in DST is measured by function N given by Eq. (18). This function, which is a generalization of the Hartley measure, is now well justified on both intuitive and mathematical grounds.

A fully justified measure of conflict in DST, which should be based on an appropriate generalization of the Shannon entropy, has not emerged from the previous work. Although functions S and D are well justified on intuitive grounds as measures of conflict among, respectively, disjunctive set-values statements and conjunctive set-valued statements, they are deficient in the sense that neither of them is subadditive. Moreover, the associated measures of total uncertainty, NS and ND, are not subadditive either.

To break this impasse, we suggest a novel approach to measuring uncertainty in DST. In this approach, we do not distinguish the two components of uncertainty in DST, nonspecificity and conflict. Instead, we define a measure of total uncertainty in DST by using the principle of maximum entropy.

The crux of our approach lies in the recognition that any given body of evidence defined on $\mathcal{P}(X)$ may be viewed as a set of constraints that define which probability distributions on X are acceptable. Among the acceptable probability distributions, one has the largest value of the Shannon entropy. It is reasonable to take this value as a measure of total uncertainty associated with the body of evidence. It should be noted here, however, that this motivation is not intrinsic to the proposed measure of uncertainty. We may simply view this measure as the only one known thus far that satisfies all the requirements stated in Sec. 3.

Let the proposed measure of the total uncertainty be denoted by AU and, as suggested in Sec. 3, let its domain be the set of all belief measures, $\mathcal{B}$. The measure AU is then defined as follows.

DEFINITION 1 Let X, Bel denote a frame of discernment and a belief function defined on $\mathcal{P}(X)$, respectively, and let $\langle p_x \mid x \in X \rangle$ denote a probability distribution on X. Then, we define the amount of uncertainty contained in Bel, denoted as $AU(\text{Bel})$, by

$$AU(\text{Bel}) = \max\left\{-\sum_{x\in X} p_x \log_2 p_x\right\},$$

where the maximum is taken over all distributions $\langle p_x \mid x \in X\rangle$ that satisfy the constraints:

(a) $p_x \in [0, 1]$ for all $x \in X$ and $\Sigma_{x\in X}\, p_x = 1$;
(b) $\text{Bel}(A) \leq \Sigma_{x\in A}\, p_x \leq 1 - \text{Bel}(\bar{A})$ for all $A \subseteq X$.[††]

In the next section, we show that function AU introduced by this definition satisfies all the requirements stated in Sec. 3 and, consequently, qualifies fully as a measure of total uncertainty in DST.

It should be mentioned that, during our work on this paper, we found that the idea incorporated in Definition 1 has lately also been pursued by Maeda and Ichihashi [1993] and Maeda, Nguyen, and Ichihashi [1993]. However, their motivation is somewhat different. Moreover, they consider as total uncertainty in DST the sum of functions N and AU, which violates the required range and does not collapse to the Shannon entropy for probability measures (even when normalized).

6. PROPERTIES OF FUNCTION AU

In this section, we show that function AU given by Definition 1 satisfies the five requirements stated in Sec. 3 and, consequently, it qualifies as a meaningful measure of total uncertainty in DST. The following theorems are numbered in the same way as the requirements in Sec. 3. In Theorems 1–3, we assume that the frame of discernment is X.

THEOREM 1 $AU(\text{Bel})$ is equal to the Shannon entropy whenever Bel is a probability measure.

Proof: By Eq. (5) and Def. 1, we have

$$\text{Pl}(A) = 1 - \text{Bel}(\bar{A}) \geq 1 - \sum_{x\in\bar{A}} p_x = \sum_{x\in A} p_x$$

for all $A \subseteq X$. When Bel is a probability measure, $\text{Bel}(A) = \text{Pl}(A)$ for all $A \subseteq X$, and this implies that $\text{Bel}(\{x\}) = p_x = \text{Pl}(\{x\})$ for all $x \in X$. Hence, the maximum in Def. 1 is taken over one particular probability distribution $\langle p_x \mid x \in X\rangle$ and $AU(\text{Bel})$ is equal to the Shannon entropy of this distribution. ■

THEOREM 2 $AU(\text{Bel})$ is equal to the Hartley measure whenever Bel is based only upon one focal element.

[††]Observe that Constraint (b) in Definition 1 may be reduced to $\text{Bel}(A) \leq \Sigma_{x\in A} p_x$. Also notice that the nature of the constraints guarantees the existence of the maximum in Definition 1.

Proof: Let m denote the basic probability assignment corresponding to Bel by Eq. (4). Assume that $m(A) = 1$ for some nonempty set $A \subseteq X$. Then, by Eq. (2) and Def. 1, we have

$$1 = \mathrm{Bel}(A) \leq \sum_{x \in A} p_x \leq 1$$

and, consequently,

$$\sum_{x \in A} p_x = 1.$$

This implies that $p_x = 0$ for all $x \notin A$. Since $\mathrm{Bel}(B) = 0$ for all $B \subset A$, all probability distributions over A, each extended to X by zero probabilities on all elements $x \notin A$, satisfy the inequalities (b) in Def. 1, and these are the only distributions that satisfy the inequalities. This means that $AU(\mathrm{Bel}) = \log_2 |A|$ since the maximum of the Shannon entropy over all probability distributions on set A is $\log_2 |A|$. ■

THEOREM 3 $0 \leq AU(\mathrm{Bel}) \leq \log_2 |X|$.

Proof: By properties of the Shannon entropy,

$$0 \leq -\sum_{x \in X} p_x \log_2 p_x \leq \log_2 |X|.$$

Since $AU(\mathrm{Bel})$ is defined as the maximum of the Shannon entropy over sets of probability distributions defined on X by constraints (a) and (b) of Def. 1, it has the same range as the Shannon entropy. ■

THEOREM 4 Function AU is subadditive.

Proof: Let Bel be a belief function over $X \times Y$ and let $\langle \hat{p}_{xy} \mid \langle x, y \rangle \in X \times Y \rangle$ denote a probability distribution for which

$$AU(\mathrm{Bel}) = -\sum_{x \in X} \sum_{y \in Y} \hat{p}_{xy} \log_2 \hat{p}_{xy}$$

and

$$\mathrm{Bel}(A) \leq \sum_{\langle x, y \rangle \in A} \hat{p}_{xy}$$

for all $A \subseteq X \times Y$. Furthermore, let

$$\hat{p}_x = \sum_{y \in Y} \hat{p}_{xy} \text{ and } \hat{p}_y = \sum_{x \in X} \hat{p}_{xy}.$$

Using now the Gibbs inequality [Klir and Folger, 1988], we have

$$-\sum_{x\in X}\sum_{y\in Y}\hat{p}_{xy}\log_2\hat{p}_{xy} \le -\sum_{x\in X}\sum_{y\in Y}\hat{p}_{xy}\log_2(\hat{p}_x\cdot\hat{p}_y)$$
$$= -\sum_{x\in X}\hat{p}_x\log_2\hat{p}_x - \sum_{y\in Y}\hat{p}_y\log_2\hat{p}_y.$$

Observe that, for all $B \subseteq X$,

$$\mathrm{Bel}_X(B) = \mathrm{Bel}(B\times Y) \le \sum_{x\in B}\sum_{y\in Y}\hat{p}_{xy} = \sum_{x\in B}\hat{p}_x$$

and, analogically, for all $C \subseteq Y$,

$$\mathrm{Bel}_Y(C) \le \sum_{y\in C}\hat{p}_y.$$

Considering all these facts, we have

$$AU(\mathrm{Bel}) = -\sum_{x\in X}\sum_{y\in Y}\hat{p}_{xy}\log_2\hat{p}_{xy} \le -\sum_{x\in X}\hat{p}_x\log_2\hat{p}_x - \sum_{y\in Y}\hat{p}_y\log_2\hat{p}_y$$
$$\le AU(\mathrm{Bel}_X) + AU(\mathrm{Bel}_Y). \quad \blacksquare$$

THEOREM 5 Function AU is additive.

Proof: Let Bel be a belief function over $X \times Y$ that satisfies Eq. (9). Let $\langle\hat{p}_x \mid x \in X\rangle$ denote a probability distribution for which

$$AU(\mathrm{Bel}_X) = -\sum_{x\in X}\hat{p}_x\log_2\hat{p}_x$$

and

$$\mathrm{Bel}_X(\mathrm{B}) \le \sum_{x\in B}\hat{p}_x$$

for all $B \subseteq X$; similarly, let $\langle\hat{p}_y \mid y \in Y\rangle$ denote a probability distribution for which

$$AU(\mathrm{Bel}_Y) = -\sum_{y\in Y}\hat{p}_y\log_2\hat{p}_y$$

and

$$\mathrm{Bel}_Y(C) \le \sum_{y\in C}\hat{p}_y$$

for all $C \subseteq Y$. Furthermore, let $\dot{p}_{xy} = \hat{p}_x \cdot \hat{p}_y$ for all $\langle x, y \rangle \in X \times Y$. Clearly, $\dot{p}_{xy}$ is a probability distribution over $X \times Y$. Moreover, for all $A \subseteq X \times Y$,

$$\begin{aligned}\sum_{\langle x,y\rangle \in A} \dot{p}_{xy} &= \sum_{\langle x,y\rangle \in A} \hat{p}_x \cdot \hat{p}_y = \sum_{x \in A_X} \hat{p}_x \sum_{\langle x,y\rangle \in A} \hat{p}_y \geq \sum_{x \in A_X} \hat{p}_x \sum_{C \mid \{x\} \times C \subseteq A} m_Y(C) \\ &= \sum_{x \in A_X} \sum_{C \mid \{x\} \times C \subseteq A} \hat{p}_x \cdot m_Y(C) = \sum_{C \subseteq A_Y} \sum_{x \mid \{x\} \times C \subseteq A} m_Y(C) \cdot \hat{p}_x \\ &\geq \sum_{C \subseteq A_Y} m_Y(C) \sum_{B \times C \subseteq A} m_X(B) = \sum_{B \times C \subseteq A} m_X(B) \cdot m_Y(C) = \mathrm{Bel}(A).\end{aligned}$$

This implies that

$$\begin{aligned} AU(\mathrm{Bel}) &\geq -\sum_{x \in X} \sum_{y \in Y} \dot{p}_{xy} \log_2 \dot{p}_{xy} \\ &= -\sum_{x \in X} \hat{p}_x \log_2 \hat{p}_x - \sum_{y \in Y} \hat{p}_y \log_2 \hat{p}_y \\ &= AU(\mathrm{Bel}_X) + AU(\mathrm{Bel}_Y). \end{aligned}$$

By subadditivity of AU (Theorem 4), we have

$$AU(\mathrm{Bel}) \leq AU(\mathrm{Bel}_X) + AU(\mathrm{Bel}_Y).$$

Hence,

$$AU(\mathrm{Bel}) = AU(\mathrm{Bel}_X) + AU(\mathrm{Bel}_Y). \quad \blacksquare$$

7. CONCLUSIONS

We demonstrate in this paper that function AU, introduced by Def. 1, satisfies all requisite requirements (listed in Sec. 3) to qualify as a satisfactory measure of the amount of total uncertainty associated with any given body of evidence in DST. This is a significant result since none of the previously proposed measures satisfies one of the essential requirements, the requirement of subadditivity.

A principal difference between the previously proposed measures and function AU is that the former are conceptualized in terms of basic probability assignments, while the latter is formulated in terms of belief measures. This difference might be responsible for the failure of previous attempts to find a fully satisfactory measure of uncertainty in DST. While the basic probability assignment function plays an important role related to the quality or reliability of evidence, it is the belief function that actually carries the information content of given evidence. It seems, therefore, that uncertainty, which is an expression of information deficiency, should be connected with the belief function rather than with the basic probability assignment function. This conclusion is, of course, rather speculative and further research is needed to substantiate it.

Our results in this paper and the connected work by Maeda and Ichihashi [1993] and Maeda, Nguyen, and Ichihashi [1993] only expose an important new research area with many formidable open problems. Let us touch upon some of these open problems.

1. One of the most important questions is the question of *uniqueness* of function AU as a satisfactory measure of uncertainty in DST. While the requirements stated in Sec. 3 are generally considered as key for any acceptable measure of uncertainty in DST, other intuitively justifiable requirements may be added, if desirable, to facilitate a uniqueness proof. One such requirement is the monotonicity under random set inclusion, as argued by Maeda and Ichihashi [1993].
2. To be able to utilize function AU in practice will require to solve some very difficult *computational problems*. Some initial work in this direction, resulting in algorithms for computing function AU in five special classes of bodies of evidence, has already been done by Maeda, Nguyen, and Ichihashi [1993]. Due to the large computational complexity of the optimization problem involved in the definition of function AU, a study of suitable approximate methods will be of particular importance.
3. We need to clarify the relationship between function AU and function N, the well-justified Hartley-like measure of nonspecificity. Since function AU is viewed as a measure of total uncertainty in DST, we should be able to show that the value of N does not exceed the value of AU for any given belief measure. Then, AU could be interpreted as a composite of two types of uncertainty. One of them, measured by values of N, is nonspecificity. The other one, measured by the difference in values of AU and N, is another type of uncertainty whose exact meaning has yet to be determined.
4. The minimum of function AU, $AU(\text{Bel}) = 0$, is obtained if and only if Bel expresses the full certainty : $\text{Bel}(\{x\}) = 1$ for a particular element $x \in X$ and $\text{Bel}(A) = 0$ for each $A \not\supseteq \{x\}$. The maximum of AU, $AU(\text{Bel}) = \log_2 |X|$, is obtained for various belief measures, each of which expresses, in some particular way, the state of the highest information deficiency. One maximum is obtained when $\text{Bel}(X) = 1$ and $\text{Bel}(A) = 0$ for each $A \neq X$; in this case, the information deficiency is expressed purely in set-theoretic terms. Another maximum is obtained when $\text{Bel}(\{x\}) = 1/|X|$ for all $x \in X$; in this case, the information deficiency is expressed in purely probabilistic terms. In all remaining cases, the information deficiency is expressed by a particular combination of set-theoretic and probabilistic aspects. In order to develop a thorough understanding of the meaning of measurements obtained by function AU, we need to study all conditions under which the function reaches its maximum.
5. An important issue is to study the behavior of function AU within possibility theory. As learnt from a similar study performed for functions D and S [Geer and Klir, 1991], such a study may produce new insights into possibility theory itself.
6. Function AU should be utilized to analyze the Dempster rule and other prospective rules of combination in DST.
7. The meaning and properties of function AU should be investigated for fuzzified DST and DST not restricted to finite sets [Wang and Klir, 1992].

REFERENCES

1. Dubois, D. and H. Prade [1985], "A note on measures of specificity for fuzzy sets." *Intern. J. of General Systems*, **10**(4), pp. 279–283.

2. Dubois, D. and H. Prade [1988], *Possibility Theory*. Plenum Press, New York.
3. Dubois, D. and H. Prade [1987], "Properties of measures of information in evidence and possibility theories." *Fuzzy Sets and Systems*, **24**(2), pp. 161–182.
4. Geer, J. F. and G. J. Klir [1991], "Discord in possibility theory." *Intern. J. of General Systems*, **19**(2), pp. 119–132.
5. Guan, J. W. and D. A. Bell [1991], *Evidence Theory and Its Applications*, Vol. 1. North-Holland, New York.
6. Hartley, R. V. L. [1928], "Transmission of information." *The Bell Systems Technical Journal*, **7**, pp. 535–563.
7. Higashi, M. and G. J. Klir [1983], "Measures of uncertainty and information based on possibility distributions." *Intern. J. of General Systems*, **9**(1), pp. 43–58.
8. Höhle, U. [1982], "Entropy with respect to plausibility measures." In: *Proc. 12th IEEE Intern. Symp. on Multiple-Valued Logic*, Paris, pp. 167–169.
9. Klir, G. J. [1987], "Where do we stand on measures of uncertainty, ambiguity, fuzziness, and the like?" *Fuzzy Sets and Systems*, **24**(2), pp. 141–160.
10. Klir, G. J. [1993], "Developments in uncertainty-based information." In: Yovits, M. C., ed., *Advances in Computers*. Academic Press, San Diego, pp. 255–332.
11. Klir, G. and T. Folger [1988], *Fuzzy Sets, Uncertainty, and Information*. Prentice-Hall, Englewood Cliffs, NJ.
12. Klir, G. J. and M. Mariano [1987], "On the uniqueness of possibilistic measure of uncertainty and information." *Fuzzy Sets and Systems*, **24**(2), pp. 197–219.
13. Klir, G. J. and B. Parviz [1992], "A note on the measure of discord." In: Dubois, D. et al., eds., *Proceedings of the Eighth Conference on Artificial Intelligence*. Morgan Kaufmann, San Mateo, California, pp. 138–141.
14. Klir, G. J. and A. Ramer [1990], "Uncertainty in the Dempster-Shafer theory : A critical re-examination." *Intern. J. of General Systems*, **18**, pp. 155–166.
15. Klir, G. J. and M. Wierman [1987], "On properties of the V-uncertainty." In: *Proc. NAFIPS Meeting*, pp. 96–106.
16. Klir, G. J. and B. Yuan [1993], "On measures of conflict among set-valued statements." *Proc. 1993 World Congress on Neural Networks*; July 11–15, 1993, Portland, Oregon
17. Lamata, M. T. and S. Moral [1988], "Measures of entropy in the theory of evidence." *Intern. J. of General Systems*, **14**(4), pp. 297–305.
18. Maeda, Y. and H. Ichihashi [1993], "An uncertainty measure with monotonicity under the random set inclusion." *Intern. J. of General Systems*, **21**(4), pp. 379–392.
19. Maeda, Y., H. T. Nguyen and H. Ichihashi [1993], "Maximum entropy algorithms for uncertainty measures." *Intern. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, **1**(1), pp. 69–93.
20. Ramer, A. [1987], "Uniqueness of information measure in the theory of evidence." *Fuzzy Sets and Systems*, **24**(2), pp. 183–196.
21. Ramer, A. and G. J. Klir [1993], "Measures of discord in the Dempster-Shafer theory." *Information Sciences*, **67**(1–2), pp. 35–50.
22. Ramer, A. and L. Lander [1987], "Classification of possibilistic uncertainty and information functions." *Fuzzy Sets and Systems*, **24**(2), pp. 221-230.
23. Rényi, A. [1970], *Probability Theory*. North-Holland, Amsterdam (Chapter IX, *Introduction to Information Theory*, pp.540–616).
24. Shafer, G. [1976], *A Mathematical Theory of Evidence*. Princeton Univ. Press, Princeton.
25. Shannon, C. E. [1948], "The mathematical theory of communication." *The Bell System Technical Journal*, **27**, pp. 379–423, 623–656.
26. Vejnarová, J. [1991], "A few remarks on measures of uncertainty in Dempster-Shafer theory." In: *Proc. of Workshop on Uncertainty in Expert Systems*; Sept. 9–12, Alšovice, Czechoslovakia. Czechoslovak Academy of Sciences, Prague.
27. Vejnarová, J. and G. J. Klir [1993], "Measure of strife in Dempster-Shafer theory." *Intern. J. of General Systems*, **22**(1), pp. 25–42.
28. Wang, Z. and G. J. Klir [1992], *Fuzzy Measure Theory*. Plenum Press, New York.
29. Yager, R. R. [1983], "Entropy and specificity in a mathematical theory of evidence." *Intern. J. of General Systems*, **9**(4), pp. 249–260.

David Harmanec was born in Prague, Czechoslovakia, in 1967. He received his M.Sc. from the Charles University in Prague in 1990. Then he started his post-graduate study at the Institute of Computer and Information Sciences, Czechoslovak Academy of Sciences, with Dr. Petr Hájek. Now, he is in the PhD program at the Department of System Science, Thomas J. Watson School of Engineering and Applied Science, State University of New York at Binghamton. His research interests are in the area of Dempster-Shafer theory, where he is an author or coauthor of several articles.

For biography and photograph of George Klir, please see Vol. 21 (1992), No. 1, pp. 49–50 of this journal.