Which is the best belief entropy?

Liguo Fei, Yong Deng, Sankaran Mahadevan

Abstract—**In this paper, many numerical examples are designed to compare the existing different belief functions with the new entropy, named Deng entropy. The results illustrate that, among the existing belief entropy functions, Deng entropy is the best alternative due to its reasonable properties.**

Index Terms—**Dempster-Shafer evidence theory, Uncertainty measure, belief entropy, Deng entropy.**

I. Introduction

EVIDENCE theory is becoming more and more important in the field of information fusion. However, it is an open issue about how to measure the uncertainty degree in evidence theory. This is a very important work to measure the uncertainty for estimating the conflict and combining information among different evidence.

To measure the uncertainty degree in evidence theory, a number of measurement methods have been developed by researchers. In [1], Dubois and Prade analyzed about Yager's specificity index and the "possibilistic entropy" introduced by Higashi and Klir. A probabilistic interpretation of Yager's index is provided in that, so, there is a conclusion that both indices can be applied to evaluate the amount of imprecision of Shafer's belief functions. Höhle [2] proposed the concept of Confusion to measure the uncertainty degree. Yager [3] introduced the concepts of entropy and specificity in the framework of Dempster-Shafer evidence theory. Then to indicate the quality of evidence. In [4], Klir and Ramer proposed a new method called a measure of discord, then they also discussed a measure of total uncertainty, which is defined as the sum of nonspecificity and discord. Klir and Parviz [5] introduced a new entropy-like measure as well as a new measure of uncertainty in Dempster-Shafer evidence theory. And they argued that this method is the best one comparing with all the existing measure. In [6], George and Pal established the need for a new measure of conflict and followed a fresh approach to achieve the same. In their work, the average of conflict between propositions provoids a measure of total conflict in an evidence. Maluf [7] examined one form of current development regarding the entropy measure induced from the measure of dissonance. Entropy measure as a monotonically decreasing function is proposed in his paper, symmetrical to the measure of dissonance. In [8], existing methods and encountered difficulties regarding these issues are proposed in evidence theory and possibility theory. In the recent times, Deng [9] proposed a new belief entropy named Deng entropy to measure the uncertainty of the mass functions. In this paper, all the above methods are applied to several numerical examples to compare the effectiveness among them.

L. Fei and Y. Deng are with the School of Computer and Information Science, Southwest University, Chongqing, 400715, China e-mail: (ydeng@swu.edu.cn,prof.deng@hotmail.com). S. Mahadevan is with the School of Engineering, Vanderbilt University, 37212, USA

The paper is organized as follows. The preliminaries Dempster-Shafer evidence theory and uncertainty measurement methods are briefly introduced in Section 2. Section 3 presents some numerical examples and compares the different measurement methods. Finally, this paper is concluded and discussed in Section 4.

II. Preliminaries

A. Dempster-Shafer evidence theory

Dempster-Shafer evidence theory (D-S theory) is proposed by Dempster and developed later by Shafer [10], [11]. This theory extends the elementary event space in probability theory to its power set named as frame of discernment and constructs the basic probability assignment(BPA) on it. In addition, there is a combination rule presented by Dempster to fuse different BPAs. In particular, D-S theory can definitely degenerate to the probability theory if the belief is only assigned to single elements. Therefore, the D-S theory is the generalization of probability theory with the purpose of handling uncertainty and is widely used to uncertainty modeling [12], decision making [13], [14], information fusion and uncertain information processing [15]. The basic definitions about D-S theory is shown as follows:

1) Frame of discernment: D-S theory supposes the definition of a set of elementary hypotheses called the frame of discernment, defined as:

$$\theta = \{H_1, H_2, ..., H_N\} \tag{1}$$

That is, θ is a set of mutually exclusive and collectively exhaustive events. Let us denote 2^θ the power set of θ.

2) Mass functions: When the frame of discernment is determined, a mass function m is defined as follows.

$$m : 2^\theta \to [0, 1] \tag{2}$$

which satisfies the following conditions:

$$m(\phi) = 0 \tag{3}$$

$$\sum_{A \in 2^\theta} m(A) = 1 \tag{4}$$

In D-S theory, a mass function is also called a basic probability assignment (BPA).

3) Belief and plausibility functions : The belief function Bel is defined as:

$$Bel(A) = \sum_{B \in A} m(B) \tag{5}$$

The plausibility function Pl is defined as:

$$Pl(A) = \sum_{B \bigcap A \neq \phi} m(B) \tag{6}$$

and

$$Pl(A) = 1 - Bel(\bar{A}), Pl(\phi) = 0 \tag{7}$$

where $\bar{A}$ = Ω - A. Obviously, $Pl(A) \geq Bel(A)$, these functions Bel and Pl are the lower limit function and upper limit function of the probability to which proposition A is supported, respectively.

4) Dempster's rule of combination: In a real system, there may be many evidence originating from different sensors, so we can get different BPAs. Dempster [11] proposed orthogonal sum to combine these BPAs. Suppose m_1 and m_2 are two mass functions. The Dempster's rule of combination denoted by $m = m_1 \bigoplus m_2$ is defined as follows:

$$m(A) = \frac{\sum_{B \bigcap C = A} m_1(B) m_2(C)}{1 - K} \tag{8}$$

with

$$K = \sum_{B \bigcap C = \phi} m_1(B) m_2(C) \tag{9}$$

Note that the Dempster's rule of combination is only applicable to such two BPAs which satisfy the condition $K < 1$.

B. Weighted Hartley entropy

The non-specifity in D-S theory can be described by Cardinality-Based based on Hartley entropy. Dubois *et al.* defined a kind of consonant belief structure, and they proofed the plausibility of this structure is a possibility measure. The consonant means the focal elements in D-S theory is a group nested Sets, such as $B_1 \supset B_2 \supset ... \supset B_q$. Let's suppose that μ is possibility measure, then $\mu(\{x_i\}) = Poss(\{x_i\}) = \alpha_i$, if x_i has the order $\alpha_i \geq \alpha_j, (i > j)$. So, $\alpha_n = 1$. Let's suppose $B_j = \{x_i | i = j, j = 1, ..., n\}$, then it's the consonant structure, if the BPA is: $m(B_i) = \alpha_j - \alpha_{j-1}, j = 1, 2, ..., n$, supposing that $\alpha_0 = 0$, we have that:

$$Pl(\{x_i\}) = \sum_{i=1}^{j} m(B_i) = \sum_{i=1}^{j} (\alpha_j - \alpha_{j-1}) = \alpha_j \tag{10}$$

Based on this work, Dubois and Prade [1] generalized the uncertainty into D-S theory:

$$N_{DP}(m) = \sum_{A \subseteq X} m(A) log(|A|) \tag{11}$$

It is a kind of weighted Hartley entropy of total focal elements. m will degenerate to a probability distribution when the BPA is only assigned to single elements.

C. Conflicting model

It becomes more and more important to measure the conflict in D-S theory between different mass functions. Some uncertainty models are proposed based on the conflicting model, including the confusion proposed by Höhle [2], Dissonance introduced by Yager [3], Discord presented by Klir and Ramer [4], and Strife defined by Klir and Pariz [5]. Their definitions are shown as follows.

$$Hohle: [2] \quad C_H(m) = -\sum_{A \in F} m(A) \log Bel(A) \tag{12}$$

$$Yager: [3] \quad C_Y(m) = -\sum_{A \in F} m(A) \log Pl(A) \tag{13}$$

$$Klir\&Ramer: [4] \quad C_{KR}(m) = -\sum_{A \in F} m(A) \log \sum_{B \in F} m(B) \frac{|A \bigcap B|}{|B|} \tag{14}$$

$$Klir\&Parviz: [5] \quad C_{KP}(m) = -\sum_{A \in F} m(A) \log \sum_{B \in F} m(B) \frac{|A \bigcap B|}{|A|} \tag{15}$$

George and Pal [6] proposed a similar model:

$$C_{GP}(m) = \sum_{A \in F} m(A) \sum_{B \in F} m(B) [1 - \frac{|A \bigcap B|}{|A \bigcup B|}] \tag{16}$$

D. Shannon entropy-like model

There exist some researchers constructed uncertainty model based on Shannon entropy-like model, for example, Maluf [7] determined the evidence entropy model as follows.

$$C_M = -\sum_{A \in F} Pl(A) \log Bel(A) \tag{17}$$

In addition, Klir [8] presented a new model as follows.

$$C_K(Bel) = -\frac{1}{c} \sum_{A \in F} Bel(A) \log Bel(A) + Pl(A) \log Pl(A) \tag{18}$$

where $\forall A \in F$, $c = \sum_{A \in F} [Bel(\{A\}) + Pl(\{A\})]$.

E. Deng entropy

Deng entropy [9] is presented to measure the uncertainty degree of basic probability assignment as a generalized Shannon entropy in D-S evidence theory. Deng entropy can be described as follows

$$E_d = -\sum_i m(F_i) \log \frac{m(F_i)}{2^{|F_i|} - 1} \tag{19}$$

where F_i is a proposition in mass function m, and $|F_i|$ is the cardinality of F_i. Deng entropy is similar with Shannon entropy in form. The difference is that the belief for each proposition F_i is divided by a term $(2^{F_i} - 1)$ which represents the potential number of states in F_i (The empty set is not included). So Deng entropy is the generalization of Shannon entropy, which is used to measure the uncertainty degree of BPA [9].

Specially, Deng entropy can definitely degenerate to the Shannon entropy if the belief is only assigned to single elements. The process is shown as follows.

$$E_d = -\sum_i m(\theta_i) \log \frac{m(\theta_i)}{2^{|\theta_i|} - 1} = -\sum_i m(\theta_i) \log m(\theta_i) \tag{20}$$

III. Examples and Comparison

Example III.1. *Given a frame of discernment X with 15 elements which are denoted as element 1, element 2, etc. A mass function is shown as follows.*

$$m(\{3,4,5\}) = 0.05, m(\{6\}) = 0.05, m(A) = 0.8, m(X) = 0.1$$

Figure 1 lists various uncertainty when A changes, which is graphically shown in Figure 2. The results shows that the Deng entropy and the Weighted Hartley entropy of m increases monotonously with the rise of the size of subset A. However, others uncertainty measurement methods are declining or changing irregularly. The dash area of Figure 1 represents the irrationality of uncertainty measurement methods. It is rational that the entropy increases when the uncertainty involving a mass function increases. So, the conclusion can be come to from Example III.1 that Deng entropy and Weighted Hartley entropy have a good performance in measuring the uncertainty.

Uncertainty \ Case	Case1 A={1}	Case2 A={1,2}	Case3 A={1,2,3}	Case4 A={1,...,4}	Case5 A={1,...,5}	....	Case11 A={1,...,11}	Case12 A={1,...,12}	Case13 A={1,...,13}	Case14 A={1,...14}
Weighted Hartley entropy[1]	0.4699	1.2699	1.7379	2.0699	2.3275		3.2375	3.3379	3.4303	3.5158
Confusion[2]	0.6897	0.6897	0.6897	0.6897	0.6198		0.5538	0.5538	0.5538	0.5538
Dissonance[3]	0.3953	0.3953	0.1997	0.1997	0.1997		0.0074	0.0074	0.0074	0.00074
Discord[4]	0.6469	0.6374	0.4914	0.4356	0.3934		0.2685	0.2687	0.2682	0.2670
Strife[5]	0.6538	0.6137	0.4876	0.4157	0.3639		0.1574	0.1487	0.1404	0.1326
Conflict between propositions[6]	0.3317	0.3210	0.2943	0.2677	0.2410		0.1959	0.1877	0.1791	0.1701
Monotonically decreasing function[7]	1.5863	1.5863	5.0600	5.0600	4.9769		8.3637	8.3637	8.3637	8.3637
Imprecise probabilities[8]	0.4019	0.4019	0.2507	0.2507	0.2365		0.1203	0.1203	0.1203	0.1203
Deng entropy[9]	2.6623	3.9303	4.9082	5.7878	6.6256		11.4617	12.2620	13.0622	13.8622

Fig. 1. Uncertainty of different methods by changing the Size of A, and the dash area denotes the irrationality of corresponding uncertainty measurement methods

Example III.2. *Given a frame of discernment $X = \{a_1, a_2, \cdots, a_N\}$, let us consider three special cases of mass functions as follows.*

$$m(X) = 1.$$

Figure 3 lists the uncertainty change with N, as shown in Figure 4.

It can be seen that all the listed method except Deng entropy and Weighted Hartley entropy obtain the uncertainty is 0 from Figure 3. The irrationality also can be found in Figure 4. So, the same conclusion is that Deng entropy and Weighted Hartley entropy do a good job in measuring the uncertainty in Example III.2.

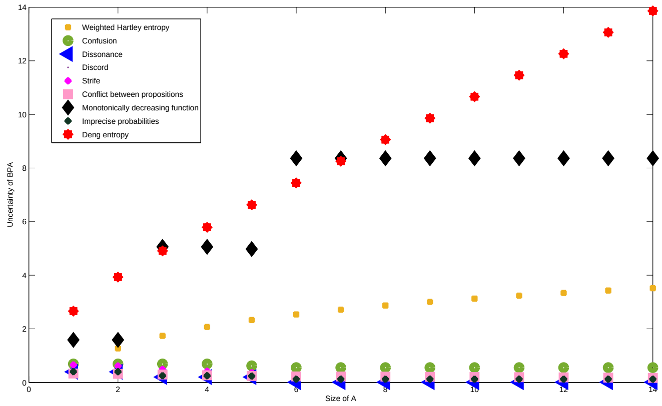

Fig. 2. Comparing uncertainty measurement methods by changing the Size of A

Uncertainty \ N	N = 1	N = 2	N = 3	N = 4	N = 5	...	N = 16	N = 17	N = 18	N= 19	N = 20
Weighted Hartley entropy[1]	0	1	1.5850	2	2.3219		4	4.0875	4.1699	4.2479	4.3219
Confusion[2]	0	0	0	0	0		0	0	0	0	0
Dissonance[3]	0	0	0	0	0		0	0	0	0	0
Discord[4]	0	0	0	0	0		0	0	0	0	0
Strife[5]	0	0	0	0	0		0	0	0	0	0
Conflict between propositions[6]	0	0	0	0	0		0	0	0	0	0
Monotonically decreasing function[7]	0	0	0	0	0		0	0	0	0	0
Imprecise probabilities[8]	0	0	0	0	0		0	0	0	0	0
Deng entropy[9]	0	1.5850	2.8074	3.9069	4.9542		16.00	17.00	18.00	19.00	20.00

Fig. 3. Uncertainty of different methods by changing N, and the dash area denotes the irrationality of corresponding uncertainty measurement methods

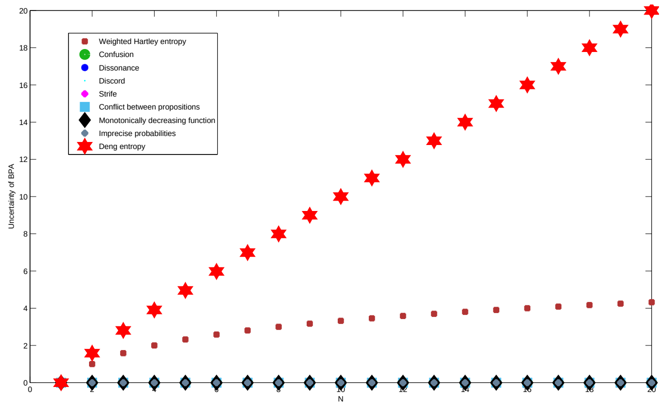

Fig. 4. Comparing uncertainty measurement methods by changing N

Example III.3. *Given a frame of discernment $X = \{a_1, a_2, \cdots, a_N\}$, let us consider three special cases of mass functions as follows.*

$$m(a_1) = m(a_2) = \cdots = m(a_N) = 1/N.$$

Their associated Deng entropies change with N, as shown in Figure 6.

It is obvious that the Weighted Hartley entropy method obtains the uncertainty keeps 0 with the change of N from Figure 5. But others method obtain the reasonable results from Figure 5-6. From this example, the conclusion can be reached that Weighted Hartley entropy method losed efficacy in some *cases. However, only the Deng entropy keeps good performance without exception.*

N / Uncertainty	N = 1	N = 2	N = 3	N = 4	N = 5	...	N = 16	N = 17	N = 18	N= 19	N = 20
Weighted Hartley entropy[1]	0	0	0	0	0		0	0	0	0	0
Confusion[2]	0	1	1.5850	2	2.3219		4	4.0875	4.1699	4.2479	4.3219
Dissonance[3]	0	1	1.5850	2	2.3219		4	4.0875	4.1699	4.2479	4.3219
Discord[4]	0	1	1.5850	2	2.3219		4	4.0875	4.1699	4.2479	4.3219
Strife[5]	0	1	1.5850	2	2.3219		4	4.0875	4.1699	4.2479	4.3219
Conflict between propositions[6]	0	0.5	0.6667	0.7500	0.8000		0.9375	0.9412	0.9444	0.9474	0.9500
Monotonically decreasing function[7]	0	1	1.5850	2	2.3219		4	4.0875	4.1699	4.2479	4.3219
Imprecise probabilities[8]	0	1	1.5850	2	2.3219		4	4.0875	4.1699	4.2479	4.3219
Deng entropy[9]	0	1	1.5850	2	2.3219		4	4.0875	4.1699	4.2479	4.3219

Fig. 5. Uncertainty of different methods by changing N, and the dash area denotes the irrationality of corresponding uncertainty measurement methods

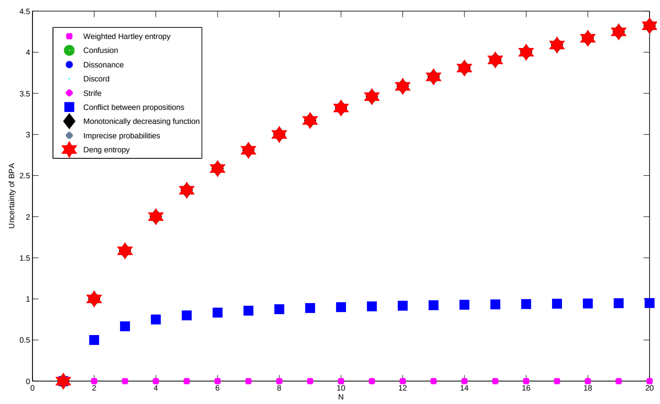

Fig. 6. Comparing uncertainty measurement methods by changing N

IV. Conclusion and Discussion

According to the problem how to measure uncertainty in evidence theory, in this paper, the main existing methods are introduced and analyzed. Some numerical examples are defined to comparing the effectiveness of different approaches. The conclusion can be obtained that the new entropy, named Deng entropy is a effective means for measuring the uncertainty in evidence theory from the experimental results.

Acknowledgment

The work is partially supported by National High Technology Research and Development Program of China (863 Program) (Grant No. 2013AA013801), National Natural Science Foundation of China (Grant Nos. 61174022,61573290), China State Key Laboratory of Virtual Reality Technology and Systems, Beihang University (Grant No.BUAA-VR-14KF-02).

References

[1] D. Dubois and H. Prade, “A note on measures of specificity for fuzzy sets,” *International Journal of General System*, vol. 10, no. 4, pp. 279–283, 1985.
[2] U. Höhle, “Entropy with respect to plausibility measures,” in *Proceedings of the 12th IEEE International Symposium on Multiple-Valued Logic*, 1982, pp. 167–169.
[3] R. R. Yager, “Entropy and specificity in a mathematical theory of evidence,” *International Journal of General System*, vol. 9, no. 4, pp. 249–260, 1983.
[4] G. J. Klir and A. Ramer, “Uncertainty in the dempster-shafer theory: a critical re-examination,” *International Journal of General System*, vol. 18, no. 2, pp. 155–166, 1990.
[5] G. J. Klir and B. Parviz, “A note on measure of discord,” *San Mateo, CA.:Morgan Kaufmann*, pp. 138–141, 1992.
[6] T. George and N. R. Pal, “Quantification of conflict in dempster-shafer framework: a new approach,” *International Journal Of General System*, vol. 24, no. 4, pp. 407–423, 1996.
[7] D. A. Maluf, “Monotonicity of entropy computations in belief functions,” *Intelligent Data Analysis*, vol. 1, no. 1, pp. 207–213, 1997.
[8] G. J. Klir, “Uncertainty and information measures for imprecise probabilities: An overview.” in *ISIPTA*, 1999, pp. 234–240.
[9] Y. Deng, “Deng entropy: a generalized shannon entropy to measure uncertainty,” *viXra*, p. 1502.0222, 2015. [Online]. Available: http://vixra.org/abs/1502.0222
[10] G. Shafer *et al.*, *A mathematical theory of evidence*. Princeton university press Princeton, 1976, vol. 1.
[11] A. P. Dempster, “Upper and lower probabilities induced by a multivalued mapping,” *The annals of mathematical statistics*, pp. 325–339, 1967.
[12] P. Smets and R. Kennes, “The transferable belief model,” *Artificial intelligence*, vol. 66, no. 2, pp. 191–234, 1994.
[13] R. R. Yager and N. Alajlan, “Decision making with ordinal payoffs under dempster–shafer type uncertainty,” *International Journal of Intelligent Systems*, vol. 28, no. 11, pp. 1039–1053, 2013.
[14] Y. Deng, “Generalized evidence theory,” *Applied Intelligence*, vol. 43, no. 3, pp. 530–543, 2015.
[15] ——, “D numbers: Theory and applications,” *International Journal of Computer & Information Sciences*, vol. 9, no. 9, pp. 2421–2428, 2012.