A new belief entropy: possible generalization of Deng entropy, Tsallis entropy and Shannon entropy

Bingyi Kang[a], Yong Deng[a,b,*]

[a]*School of Computer and Information Science, Southwest University, Chongqing, 400715, China*

[b]*Institute of Integrated Automation, School of Electronic and Information Engineering, Xian Jiaotong University, Xi'an, Shaanxi, 710049, China*

Abstract

Shannon entropy is the mathematical foundation of information theory, Tsallis entropy is the roots of nonextensive statistical mechanics, Deng entropy was proposed to measure the uncertainty degree of belief function very recently. In this paper, A new entropy H was proposed to generalize Deng entropy, Tsallis entropy and Shannon entropy. The new entropy H can be degenerated to Deng entropy, Tsallis entropy, and Shannon entropy under different conditions, and also can maintains the mathematical properity of Deng entropy, Tsallis entropy and Shannon entropy.

Keywords: Uncertainty measure, Entropy, Belief entropy, Deng entropy, Tsallis entropy, Shannnon entropy, Dempster-Shafer evidence theory

---

*Corresponding author: School of Computer and Information Science, Southwest University, Chongqing, 400715, China.

Email address: `ydeng@swu.edu.cn,prof.deng@hotmail.com` (Yong Deng)

1. Introduction

How to measure the uncertainty has attracted much attention [1, 2]. A lot of theories has been developed, such as probability theory [3], fuzzy set theory [4], possibility theory [5], Dempster-Shafer evidence theory [6, 7], rough sets[8], DSmT[9, 10], generalized evidence theory [11] and D numbers[12, 13].

Since firstly proposed by Clausius in 1865 for thermodynamics [14], various types of entropies are presented, such as information entropy [15], Tsallis entropy [16], nonadditive entropy [17, 18, 19]. Information entropy [15], derived from the Boltzmann-Gibbs (BG) entropy [20] in thermodynamics and statistical mechanics, has been an indicator to measures uncertainty which is associated with the probability density function (PDF).

Dempster-Shafer theory evidence theory[6, 7] is mainly proposed to handle such uncertainty. In Dempster-Shafer evidence theory, the epistemic uncertainty simultaneously contains nonspecificity and discord [21] which are coexisting in a basic probability assignment function (BPA). Several uncertainty measures, such as AU [22, 23], AM [21], have been proposed to quantify such uncertainty in Dempster-Shafer theory. What's more, five axiomatic requirements have been further built in order to develop a justifiable measure. These five axiomatic requirements are range, probabilistic consistency, set consistency, additivity, subadditivity, respectively [24]. Existing methods are not efficient to measure uncertain degree of BPA. To address this issue, a new entropy, named as Deng entropy [25], is proposed to measure the uncertainty of basic probability assignment for the evidence theory. In this paper, a discussin of the maximal value of Deng entropy has been discussed and proofed, which is useful for the real application of Deng entropy.

The paper is organized as follows. The preliminaries Dempster-Shafer evidence theory and Deng entropy and Tsallis entropy are briefly introduced in Section 2. Section 3 makes some discussions about the new entropy. Finally, this paper is concluded in Section 4.

2. Preliminaries

In this section, some preliminaries are briefly introduced.

2.1. Dempster-Shafer evidence theory

Dempster-Shafer theory (short for D-S theory) is presented by Dempster and Shafer [6, 7]. This theory is widely applied to uncertainty modeling [26, 27, 28, 29, 30], decision making [31, 32, 33, 34, 35, 36, 37, 38], information fusion [39, 40] and uncertain information processing [41, 42]. D-S theory has many advantages to handle uncertain information. First, D-S theory can handle more uncertainty in real world. In contrast to the probability theory in which probability masses can be only assigned to singleton subsets, in D-S theory the belief can be assigned to both singletons and compound sets. Second, in D-S theory, prior distribution is not needed before the combination of information from individual information sources. Third, D-S theory allows one to specify a degree of ignorance in some situations instead of being forced to be assigned for probabilities. Some basic concepts in D-S theory are introduced.

Let X be a set of mutually exclusive and collectively exhaustive events, indicated by

$$X = \{\theta_1, \theta_2, \cdots, \theta_i, \cdots, \theta_{|X|}\} \tag{1}$$

where set X is called a frame of discernment. The power set of X is indicated by 2^X, namely

$$2^X = \{\emptyset, \{\theta_1\}, \cdots, \{\theta_{|X|}\}, \{\theta_1, \theta_2\}, \cdots, \{\theta_1, \theta_2, \cdots, \theta_i\}, \cdots, X\} \tag{2}$$

For a frame of discernment $X = \{\theta_1, \theta_2, \cdots, \theta_{|X|}\}$, a mass function is a mapping m from 2^X to $[0, 1]$, formally defined by:

$$m: \quad 2^X \to [0, 1] \tag{3}$$

which satisfies the following condition:

$$m(\emptyset) = 0 \quad and \quad \sum_{A \in 2^X} m(A) = 1 \tag{4}$$

In D-S theory, a mass function is also called a basic probability assignment (BPA). Assume there are two BPAs indicated by m_1 and m_2, the Dempster's rule of combination is used to combine them as follows:

$$m(A) = \begin{cases} \frac{1}{1-K} \sum\limits_{B \cap C = A} m_1(B) m_2(C) \ , & A \neq \emptyset; \\ 0 \ , & A = \emptyset. \end{cases} \tag{5}$$

with

$$K = \sum_{B \cap C = \emptyset} m_1(B) m_2(C) \tag{6}$$

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

2.2. Deng entropy

With the range of uncertainty mentioned above, Deng entropy [25] can be presented as follows

$$E_d = -\sum_i m(F_i) \log \frac{m(F_i)}{2^{|F_i|} - 1} \tag{7}$$

where, F_i is a proposition in mass function m, and $|F_i|$ is the cardinality of F_i. As shown in the above definition, Deng entropy, formally, is similar with the classical Shannon entropy, but 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 (of course, the empty set is not included).

Specially, Deng entropy can definitely degenerate to the Shannon entropy if the belief is only assigned to single elements. Namely,

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

2.3. Tsallis entropy

For a discrete random variable $X = \{X_i, i = 1, 2, ..., N\}$ that has a probability distribution $P = \{p_i, i = 1, 2, ..., N\}[p_i$ is the probability of $X = x_i]$. Scaling p_i to ${p_i}^m$, where m is any real number, Tsallis Entropy [16] H_m can be denoted as

$$H_m = k\frac{1 - \sum_{i=1}^N p_i^m}{m - 1} = \frac{k}{m - 1} \sum_{i=1}^N [p_i - p_i^m] \tag{8}$$

where k is often taken as unity. For $m \to 1$, the Tsallis entropy reduces to Shannon entropy

3. Proposed new Entropy: possible generalization of Deng entropy, Tsallis Entropy, Shannon entropy

Assume F_i is the focal element and $m(F_i)$ is the basic probability assignment for F_i, then the possible generalization of Deng entropy and Tsallis entropy can be defined as H

$$H = \frac{1 - \sum\limits_i \left[m\left(F_i\right)\right]^m}{m-1} + \sum_i m\left(F_i\right) \log\left(2^{|F_i|} - 1\right) \tag{9}$$

where $i = 1, 2, ..., 2^X - 1$, and X is the scale of the frame of discernment.

(1) For $m \to 1$, the new entropy H reduces to Deng entropy, namely

$$E_d = -\sum_i m(F_i) \log \frac{m(F_i)}{2^{|F_i|} - 1} \tag{10}$$

Proof. For $m \to 1$, the new entropy can be shown as

$$H_{m\to1} = \lim_{m\to1}\left\{\frac{1 - \sum\limits_i \left[m\left(F_i\right)\right]^m}{m-1} + \sum_i m\left(F_i\right) \log\left(2^{|F_i|} - 1\right)\right\} \tag{11}$$

Then

$$H_{m\to1} = \lim_{m\to1}\left\{\frac{1 - \sum\limits_i \left[m\left(F_i\right)\right]^m}{m-1}\right\} + \sum_i m\left(F_i\right) \log\left(2^{|F_i|} - 1\right) \tag{12}$$

Because

$$\begin{aligned}
&\lim_{m\to1}\left\{\frac{1-\sum\limits_i [m(F_i)]^m}{m-1}\right\} \\
&= \lim_{m\to1} \frac{\frac{\partial}{\partial m}\left\{1-\sum\limits_i [m(F_i)]^m\right\}}{\frac{\partial}{\partial m}(m-1)} \\
&= -\lim_{m\to1} \sum_i \left[m\left(F_i\right)\right]^m \log\left[m\left(F_i\right)\right] \\
&= -\sum_i m\left(F_i\right) \log\left[m\left(F_i\right)\right]
\end{aligned} \tag{13}$$

Then

$$
\begin{aligned}
&H_{m\to 1} = -\sum_i m\left(F_i\right)\log\left[m\left(F_i\right)\right] + \sum_i m\left(F_i\right)\log\left(2^{|F_i|}-1\right) \\
&= -\sum_i \left\{m\left(F_i\right)\log\left[m\left(F_i\right)\right] - m\left(F_i\right)\log\left(2^{|F_i|}-1\right)\right\} \\
&= -\sum_i m\left(F_i\right)\log\frac{m(F_i)}{2^{|F_i|}-1}
\end{aligned}
\tag{14}
$$

Hence

$$
H_{m\to 1} = E_d = -\sum_i m\left(F_i\right)\log\frac{m\left(F_i\right)}{2^{|F_i|}-1} \tag{15}
$$

End proof $\square$

(2) For the belief is only assigned to single elements, the new entropy H reduces to Tsallis entropy, namely

$$
H_m = k\frac{1-\sum_{i=1}^{N} p_i^m}{m-1} = \frac{k}{m-1}\sum_{i=1}^{N}\left[p_i - p_i^m\right] \tag{16}
$$

Proof. For the belief is only assigned to single elements, the $|F_i| = 1$, we can easily get that

$$
\sum_i m\left(F_i\right)\log\left(2^{|F_i|}-1\right) = 0 \tag{17}
$$

Hence

$$
\begin{aligned}
&H = \frac{1-\sum\limits_i \left[m(F_i)\right]^m}{m-1} + \sum_i m\left(F_i\right)\log\left(2^{|F_i|}-1\right) \\
&= \frac{1-\sum\limits_i \left[m(F_i)\right]^m}{m-1} = \frac{1-\sum\limits_i \left[m(\theta_i)\right]^m}{m-1} = H_s
\end{aligned}
\tag{18}
$$

$\square$

End proof

(3) For $m \to 1$ and the belief is only assigned to single elements, the new entropy H reduces to Shannon entropy, namely

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

Proof. When $m \to 1$, from Eq. (15), we can get

$$H_{m\to 1} = E_d = -\sum_i m\left(F_i\right)\log\frac{m\left(F_i\right)}{2^{|F_i|}-1} \tag{20}$$

When the belief is only assigned to single elements, the $|F_i| = 1$, we can easily get that

$$H = -\sum_i m(F_i)\log\frac{m(F_i)}{2^{|F_i|}-1} = -\sum_i m(\theta_i)\log m(\theta_i) \tag{21}$$

End proof $\square$

4. Conclusion

Shannon entropy is the mathematical foundation of information theory, Tsallis entropy is the roots of nonextensive statistical mechanics, Deng entropy was proposed to measure the uncertainty degree of belief function very recently. In this paper, A new entropy H was proposed to generalize Deng entropy, Tsallis entropy and Shannon entropy. The new entropy H can be degenerated to Deng entropy, Tsallis entropy, and Shannon entropy under different conditions, and also can maintains the mathematical properity of Deng entropy, Tsallis entropy and Shannon entropy.

Acknowledgments

This work is partially supported by National Natural Science Foundation of China, Grant Nos. 30400067, 60874105 and 61174022, Chongqing Natural Science Foundation for distinguished scientist, Grant No. CSCT, 2010BA2003, Program for New Century Excellent Talents in University, Grant No. NCET-08-0345, Shanghai Rising-Star Program Grant No.09QA1402900, the Chenxing Scholarship Youth Found of Shanghai Jiao Tong University, Grant No. T241460612, the open funding project of State Key Laboratory of Virtual Reality Technology and Systems, Beihang University (Grant No.BUAA-VR-14KF-02).

References

[1] G. J. Klir, Uncertainty and information: foundations of generalized information theory, John Wiley & Sons, 2005.

[2] S. Sankararaman, S. Mahadevan, Model validation under epistemic uncertainty, Reliability Engineering & System Safety 96 (9) (2011) 1232–1241.

[3] W. Feller, An introduction to probability theory and its applications, Vol. 2, John Wiley & Sons, 2008.

[4] L. A. Zadeh, Fuzzy sets, Information and control 8 (3) (1965) 338–353.

[5] D. Dubois, H. M. Prade, H. Farreny, R. Martin-Clouaire, C. Testemale, Possibility theory: an approach to computerized processing of uncertainty, Vol. 2, Plenum press New York, 1988.

[6] A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Annals of Mathematics and Statistics 38 (2) (1967) 325–339.

[7] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, 1976.

[8] Z. Pawlak, Rough sets, International Journal of Computer & Information Sciences 11 (5) (1982) 341–356.

[9] J. Dezert, Foundations for a new theory of plausible and paradoxical reasoning, Information and Security 9 (2002) 13–57.

[10] F. Smarandache, J. Dezert, Advances and Applications of DSmT for Information Fusion (Collected works), second volume: Collected Works, Vol. 2, Infinite Study, 2006.

[11] Y. Deng, Generalized evidence theory, Applied Intelligence (2015) doi 10.1007/s10489–015–0661–2.

[12] Y. Deng, D numbers: Theory and applications, International Journal of Computer & Information Sciences 9 (9) (2012) 2421–2428.

[13] X. Deng, Y. Hu, Y. Deng, S. Mahadevan, Supplier selection using AHP methodology extended by D numbers, Expert Systems with Applications 41 (1) (2014) 156–167.

[14] R. Clausius, The mechanical theory of heat: with its applications to the steam-engine and to the physical properties of bodies, J. van Voorst, 1867.

[15] C. E. Shannon, A mathematical theory of communication, ACM SIGMOBILE Mobile Computing and Communications Review 5 (1) (2001) 3–55.

[16] C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, Journal of Statistical Physics 52 (1-2) (1988) 479–487.

[17] C. Tsallis, Nonadditive entropy: the concept and its use, The European Physical Journal A 40 (3) (2009) 257–266.

[18] C. Tsallis, Nonadditive entropy and nonextensive statistical mechanics-an overview after 20 years, Brazilian Journal of Physics 39 (2A) (2009) 337–356.

[19] C. Tsallis, U. Tirnakli, Nonadditive entropy and nonextensive statistical mechanics–some central concepts and recent applications, in: Journal of Physics: Conference Series, Vol. 201, 2010, p. 012001.

[20] J. L. Lebowitz, Boltzmann's entropy and time's arrow, Physics today 46 (1993) 32–32.

[21] A.-L. Jousselme, C. Liu, D. Grenier, E. Bosse, Measuring ambiguity in the evidence theory, IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans 36 (5) (2006) 890–903.

[22] Y. Maeda, H. T. Nguyen, H. Ichihashi, Maximum entropy algorithms for uncertainty measures, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 1 (01) (1993) 69–93.

[23] D. Harmanec, G. J. Klir, Measuring total uncertainty in dempster-shafer theory: A novel approach, International journal of general system 22 (4) (1994) 405–419.

[24] G. J. Klir, M. J. Wierman, Uncertainty-based information: elements of generalized information theory, Springer, 1999.

[25] Y. Deng, Deng entropy, Chaos, Solitons & Fractals 91 (2016) 549–553.

[26] P. Smets, R. Kennes, The transferable belief model, Artificial Intelligence 66 (2) (1994) 191–234.

[27] T. L. Wickramarathne, K. Premaratne, M. N. Murthi, Toward efficient

computation of the dempster–shafer belief theoretic conditionals, Cybernetics, IEEE Transactions on 43 (2) (2013) 712–724.

[28] J. Abellán, Éloi Bossé, Drawbacks of uncertainty measures based on the pignistic transformation, IEEE Transactions on Systems, Man, and Cybernetics: Systems (2016) 1–7doi:10.1109/TSMC.2016.2597267.

[29] J. Abellán, S. Moral, Upper entropy of credal sets. applications to credal classification, International Journal of Approximate Reasoning 39 (2) (2005) 235–255.

[30] J. Abellán, S. Moral, Difference of entropies as a non-specificity function on credal sets, International journal of general systems 34 (3) (2005) 201–214.

[31] L. V. Utkin, A new ranking procedure by incomplete pairwise comparisons using preference subsets, Intelligent Data Analysis 13 (2) (2009) 229–241.

[32] V.-N. Huynh, T. T. Nguyen, C. A. Le, Adaptively entropy-based weighting classifiers in combination using dempster–shafer theory for word sense disambiguation, Computer Speech & Language 24 (3) (2010) 461–473.

[33] Y. Deng, F. Chan, A new fuzzy dempster MCDM method and its application in supplier selection, Expert Systems with Applications 38 (8) (2011) 9854–9861.

[34] W. Jiang, B. Wei, C. Xie, D. Zhou, An evidential sensor fusion method in fault diagnosis, Advances in Mechanical Engineering 8 (3) (2016) 1–7.

[35] W. Jiang, C. Xie, B. Wei, D. Zhou, A modified method for risk evaluation in failure modes and effects analysis of aircraft turbine rotor blades, Advances in Mechanical Engineering 8 (4) (2016) 1–16. doi:10.1177/1687814016644579.

[36] R. R. Yager, N. Alajlan, Decision making with ordinal payoffs under dempster–shafer type uncertainty, International Journal of Intelligent Systems 28 (11) (2013) 1039–1053.

[37] J.-B. Yang, D.-L. Xu, Evidential reasoning rule for evidence combination, Artificial Intelligence 205 (2013) 1–29.

[38] J. Abellán, S. Moral, Maximum of entropy for credal sets, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 11 (05) (2003) 587–597.

[39] F. Cuzzolin, A geometric approach to the theory of evidence, Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 38 (4) (2008) 522–534.

[40] J. Schubert, Conflict management in Dempster-Shafer theory using the degree of falsity, International Journal of Approximate Reasoning 52 (3) (2011) 449–460.

[41] S. Rao, K. Annamdas, A comparative study of evidence theories in the modeling, analysis, and design of engineering systems, Journal of Mechanical Design 135 (6) (2013) 061006.

[42] B. Kang, Y. Deng, R. Sadiq, S. Mahadevan, Evidential cognitive maps, Knowledge-Based Systems 35 (2012) 77–86.