# An information volume measure

Yong Deng[a,*]

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

**Abstract**

How to measure the volume of uncertainty information is an open issue. Shannon entropy is used to represent the uncertainty degree of a probability distribution. Given a generalized probability distribution which means that the probability is not only assigned to the basis event space but also the power set of event space. At this time, a so called meta probability space is constructed. A new measure, named as Deng entropy, is presented. The results show that, compared with existing method, Deng entropy is not only better from the aspect of mathematic form, but also has the significant physical meaning.

*Keywords:* Information volume, Uncertainty measure, Entropy, Deng entropy, Shannon entropy, Meta probability

## 1. Question

There are many entropy functions to measure uncertainty [1, 2]. One of the most famous entropy functions is the Shannon entropy [3], which is the

*Corresponding author: Institute of Integrated Automation, School of Electronic and Information Engineering, Xi'an Jiaotong University, Xian, Shaanxi, 710049, China.

*Email address:* ydeng@xjtu.edu.cn; prof.deng@hotmail.com (Yong Deng)

base of information theory. To some degree, Shannon entropy can be used as the information volume. Here is the example to illustrate. Suppose there are 32 football teams with a straight knockout. We want to know the winner of them. We can ask one person who knows the result. This person can only answers “Yes” or “No” to our questions. The question is how many times can we know the winner.

Assume the times is $t$, it is easy to answer the problem through calculating the information volume by using information entropy

$$t = \log_2 32 = 5 \tag{1}$$

For example, the winner, denoted as $x$, is assumed to be No.2 among the 32 teams ordered as No.1, No.2,...,No.32. According to the process shown in Fig.1, the winner No.2 can be found through five times asking at most.

That means, if we use Shannon entropy as follows [3],

In this case, obviously

$$t \geq \log_2 32 \tag{2}$$

the information volume is 5 BIT as follows.

$$H = -\sum_{i=1}^{N} p_i \log_b p_i \tag{3}$$

The further question is that, if we replace the football team as the student. There are 32 students to take an exam. How many times can we know the student with the highest score. Is it 5 times in this situation? The

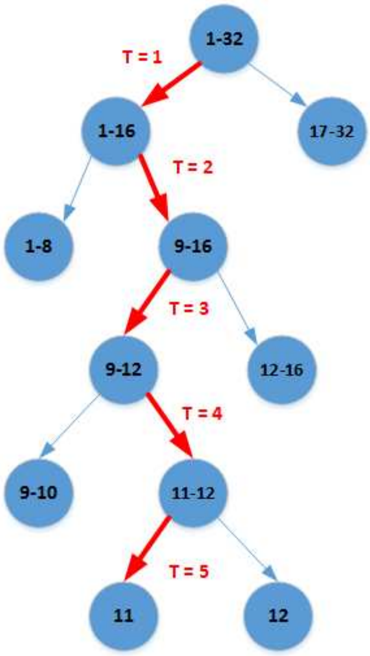


Figure 1: Guess which is the first one in 32 football teams

answer is obvious "NO" since some students may be tied with the highest score. It seems that we needs more than 5 times to know the results. How many times to determine the highest score students is an open issue, which is the motivation of this paper.

## 2. Solution

In this section, let's assume a generalized probability with the so called power set space. Let $X$ be a set of mutually exclusive and collectively exhaustive events, indicated by

$$X = \{\theta_1, \theta_2, \cdots, \theta_i, \cdots, \theta_{|X|}\} \tag{4}$$

where set $X$ is the basic event space. The power set of $X$ is indicated by $2^X$, namely [4, 5]

$$2^X = \{\emptyset, \{\theta_1\}, \cdots, \{\theta_{|X|}\}, \{\theta_1, \theta_2\}, \cdots, \{\theta_1, \theta_2, \cdots, \theta_i\}, \cdots, X\} \quad (5)$$

A generalized probability can be assigned as follows

$$P_G : \quad 2^X \to [0, 1] \quad (6)$$

which satisfies the following condition:

$$P_G(\emptyset) = 0 \quad and \quad \sum_{A \in 2^X} P_G(A) = 1 \quad (7)$$

It should be noted that, unlike the so called basis probability assignment in evidence theory [4, 5], the generalized probability means that the exclusive events can be happened AT THE SAME TIME. In this situation, it can be seen as a meta probability. Compared with many existing explanations and argumentations about BPA, this paper regards it as a probability, which is the possibility degree of some exclusive events (defined in classical probability theory) can be happened SIMULTANEOUSLY.

For example, the generalized probability

$$P_G(1, 2, 3) = 0.8; P_G(1, 2, 3, ..., 32) = 0.2$$

means that No.1, No.2 and No.3 students can be the highest students with the probability 0.8 and all the students have the highest score with the probability 0.2.

The following entropy, named as Deng entropy, is give as follows,

$$E_d(p_G) = -\sum_{A \subseteq X} \left(\frac{p_G(A)}{2^{|A|} - 1}\right) \log_2 \frac{p_G(A)}{2^{|A|} - 1} \quad (8)$$

where $p_G$ is a generalized probability defined in the power space of $X$, and $|A|$ is the cardinality of $A$. As shown in the above definition, Deng entropy, formally, is similar with the classical Shannon entropy, but the belief for each set $A$ is divided by a term $(2^{|A|} - 1)$ which represents the potential number of states in $A$ (of course, the empty set is not included).

Given the above Deng entropy, the maximum Deng entropy from 2 to 32 can be illustrated in Figure (2).

As a result, if 32 students have an exam, we need 32 times to determine the highest students. The way is to ask the teacher: is No.1 student has the highest score? Is No.2 student has the highest score and so on.

It should be noted that a former Deng entropy is presented in [6], which is not satisfied the conjectures in [7], where we think that the time to determine the highest score students is 32. In this paper, the maximum Deng entropy not only well supports the conjectures in [7], but also has the significant physical meanings to our intuition.

## 3. Conclusion

In classical probability theory, when the probability measure is only assigned to single event, the Shannon entropy is efficient to measure the uncertain information volume. However, when the exclusivity can not be satisfied, the so called Dent entropy is presented to determine the volume of the generalized probability information. In short, given a set with $N$ events, the information volume is $\log_2 N$, determined by Shannon entropy, if the event is exclusive with each other . However, if the event is

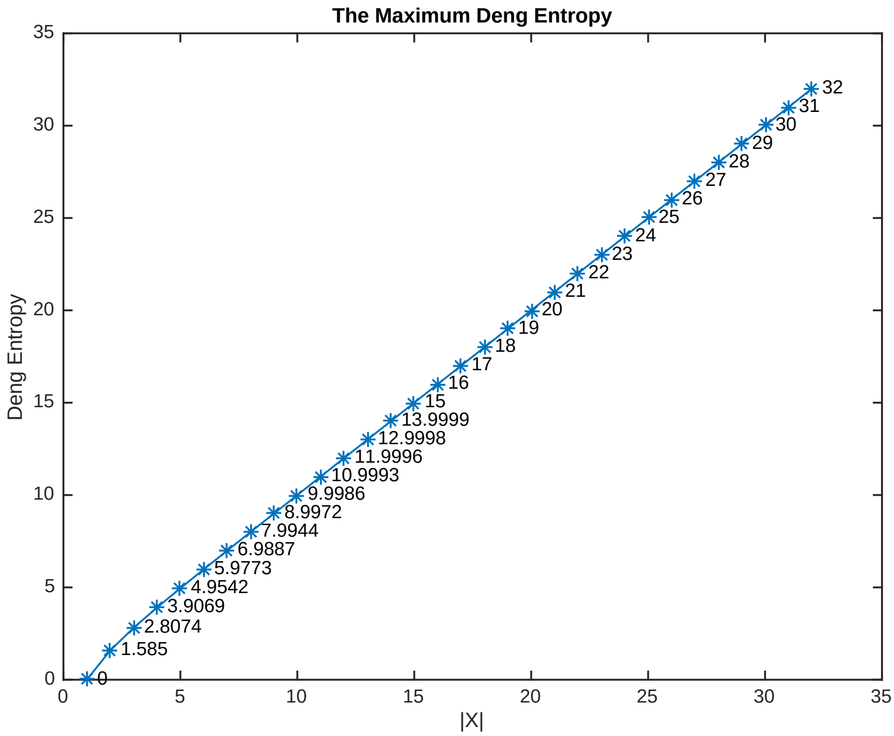


Figure 2: Deng entropy changes with scale of frame of discernment $|X|$

not exclusive with each other, means they can happen at the same time, the information volume is $N$, determined by Deng entropy. Undoubtedly, Deng entropy can also be the measure of information uncertainty. One of the ongoing works is to use Deng entropy to quantum computation.

## 4. Acknowledgement


The author greatly appreciates Professor Shan Zhong, the China academician of the Academy of Engineering, for his encouragement to do this research. The author greatly appreciates Professor Yugeng Xi in Shanghai

Jiao Tong University, Professor Rehan Sadiq in The University of British Columbia for their support to this work. The author's Ph.D students in Shanghai Jiao Tong University, Peida Xu and Xiaoyan Su, discussed the uncertain measure. The author's Ph.D students in Southwest University, Xinyang Deng and Bingyi Kang, graduate student in Southwest University, Liguo Fei discussed some properties of the entropy. The author greatly appreciates the continuous funding for the last ten years. This work is partially supported by National Natural Science Foundation of China, Grant Nos. 30400067, 60874105, 61174022 and 61573290, the Chenxing Scholarship Youth Found of Shanghai Jiao Tong University, Grant No. T241460612, Program for New Century Excellent Talents in University, Grant No. NCET-08-0345, Shanghai Rising-Star Program Grant No.09QA1402900, Chongqing Natural Science Foundation for distinguished scientist, Grant No. CSCT, 2010BA2003.

**References**

[1] 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.

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

[3] C. E. Shannon, A mathematical theory of communication, ACM SIG-

MOBILE Mobile Computing and Communications Review 5 (1) (2001) 3–55.

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

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

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

[7] X. Deng, Y. Deng, On the axiomatic requirement of range to measure uncertainty, Physica A Statistical Mechanics & Its Applications 406 (10) (2014) 163–168.