

## Entropic Explanation of Power Set

Yutong Song, Yong Deng

### Yutong Song

Institute of Fundamental and Frontier Science  
University of Electronic Science and Technology of China, Chengdu, 610054, China  
songyutong@std.uestc.edu

### Yong Deng\*

1. Institute of Fundamental and Frontier Science  
University of Electronic Science and Technology of China, Chengdu, 610054, China
2. School of Education  
Shannxi Normal University, Xi'an, 710062, China
3. School of Knowledge Science  
Japan Advanced Institute of Science and Technology  
Nomi, Ishikawa 923-1211, Japan Department of Management Technology and Economics
4. ETH Zurich, Zurich, Switzerland

\*Corresponding author: dengentropy@uestc.edu.cn, prof.deng@hotmail.com

### Abstract

A power set of a set  $S$  is defined as the set of all subsets of  $S$ , including set  $S$  itself and empty set, denoted as  $P(S)$  or  $2^S$ . Given a finite set  $S$  with  $|S| = n$  hypothesis, one property of power set is that the amount of subsets of  $S$  is  $|P(S)| = 2^n$ . However, the physical meaning of power set needs exploration. To address this issue, a possible explanation of power set is proposed in this paper. A power set of  $n$  events can be seen as all possible  $k$ -combination, where  $k$  ranges from 0 to  $n$ . It means the power set extends the event space in probability theory into all possible combination of the single basic event. From the view of power set, all subsets or all combination of basic events, are created equal. These subsets are assigned with the mass function, whose uncertainty can be measured by Deng entropy. The relationship between combinatorial number, Pascal's triangle and power set is revealed by Deng entropy quantitatively from the view of information measure.

**Keywords:** Power set, Combinatorial number, Pascal's triangle, Dempster-Shafer evidence theory, mass function, Shannon entropy, Deng entropy

## 1 Introduction

Set theory is the basis of mathematics [7]. Power sets play an important role in many applications. Power set is a family of sets consisting of all subsets from a set. For a set  $A = \{a, b, c\}$ , the corresponding power set is  $\{\phi\}, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}$  [9]. Probability can be viewed as the likelihood of a possible outcome occurring in an experiment. The undefined concepts of probability theory are experiment, event and probability mass function. In probability theory, probabilities are assigned to single subsets with an event [15]. While in Dempster-Shafer evidence theory [1, 17], the framework of discernment is extended to power

set and mass functions are assigned not only on single subset but also on multi-subsets [8, 20, 21]. Though evidence theory is widely used, the meaning of power set is neglected to some degree. To address this issue, an explanation of power set from the entropy view is represented in this paper. In this explanation, a power set can be seen as the set with all possible event combinations, each combination is represents as a subsets. For a quantitative description of the question and methods, different entropy measures [23] is introduced in this paper. The proposed explanation is supported by Deng entropy [3, 11, 13] measure quantitatively.

## 2 Preliminaries

Some preliminaries on power set [7, 10], combinatorial number [19], Pascal’s triangle [16] and basic probability assignment [12, 17] are briefly introduced, including entropy measures such as Shannon entropy [12, 18] in probability theory and Deng entropy [3] in Dempster-Shafer evidence theory.

- Combinatorial number

Taking  $m$  ( $m \leq n$ ) elements non-repeating from  $n$  different elements and forms a group is called a combination of  $m$  elements from  $n$  different elements, noted as  $C(n, m)$  [19]. The combinatorial number can describe the amount of all possible combinations of  $m$  ( $m \leq n$ ) elements from  $n$  different elements.

- Pascal’s triangle

In mathematics, Pascal’s triangle is a triangular array of the binomial coefficients that arises in probability theory [7]. It is a geometric arrangement of binomial coefficients in a triangle.

For a Pascal’s triangle with  $n$  rows and  $x$  of maximum column number, some of its properties are listed:

1. In this triangle, the amount of numbers in the  $n - th$  row is  $n$  and numbers are displayed symmetrically.
2. The  $x - th$  number in the  $n - th$  row can be expressed as  $C(n - 1, x - 1)$ , meaning the combinatorial number of drawing  $x - 1$  elements from  $n - 1$  elements.
3. The sum of all numbers in the  $n - th$  row is  $2^n$ .

The first five rows of Pascal’s triangle is show in Figure 1.

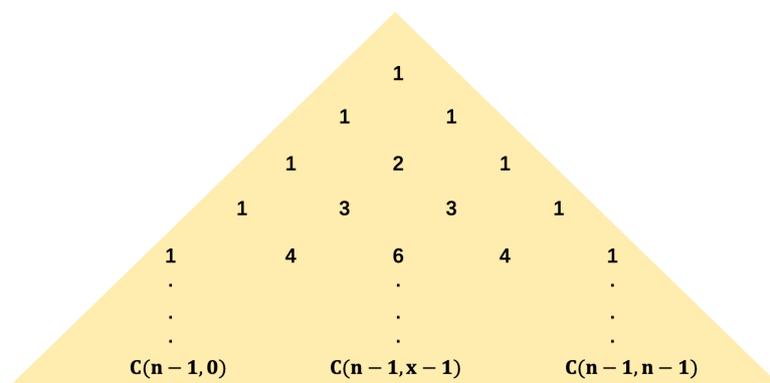


Figure 1: The first five rows of Pascal’s traingle

- Shannon entropy

In information theory, entropy of a random variable is the average level of information, inherent in the variable’s possible outcomes. Shannon entropy is a widely used information entropy [12, 14, 18].

**Definition 1.** Given a discrete random variable  $X$ , with  $x_1, \dots, x_n$ , which occur with probability  $p(x_1), p(x_2), \dots, p(x_n)$ , the entropy of  $X$  is formally defined as [18]:

$$E = - \sum_{i=1}^n p_{x_i} \log p_{x_i} \tag{1}$$

- Power set

Power set is a set family composed of all subsets (including complete set and empty set) in the original set [10]. For any as set  $a$ , there exists the set whose members are just all subsets of  $a$ . The set is called power set of  $a$ .

- Basic probability assignment

In this probabilistic view evidence is seen as more or less probable arguments for certain hypotheses and they can be used to support those hypotheses to certain degrees. Dempster-Shafer evidence theory is considered as the extension of probability theory, when the frame of discernment is extended to power sets. In Dempster-Shafer evidence theory, belief is assigned in subsets as basic probability assignments.

**Definition 2.** For a frame of discernment  $\Omega$ , a basic probability assignment, or called as mass function, is a mapping  $m$  from  $2^\Omega$  to  $[0, 1]$ , formally defined as [12, 17, 22]

$$m : 2^\Omega \rightarrow [0, 1] \tag{2}$$

which satisfies the following condition:

$$\sum_{A \in 2^\Omega} m(A) = 1 \tag{3}$$

If  $m(A) > 0$ ,  $A$  is called a focal element. The union of all focal elements is called the core of basic probability assignment.

- Deng entropy

Similar to Shannon entropy playing an important role for uncertainty measure, Deng entropy [3] is proposed for measuring the uncertainty of basic probability assignment [5] and is further developed to deal with the information volume with respect to different techniques [2, 4].

**Definition 3.** Assume in the frame of discernment  $X$ ,  $x_i$  is the focal element of basic probability assignment  $m(x_i)$ . The corresponding Deng entropy is defined as follows [3]:

$$E_d = - \sum_{x_i \subseteq X} m(A) \log_2 \frac{m(x_i)}{2^{|x_i|} - 1} \tag{4}$$

### 3 A possible explanation of power set in evidence theory

In probability theory, events that represented by discrete random variables with probability are exclusive with each other. The set of all possible outcomes of an experiment is called the sample space and any an event is a subset of sample space. The function that gives probabilities associated with all possible values of a discrete random variable is called probability mass function [6]. Since all events happen exclusively to each other, all subsets in sample space are single subsets containing one element. For example, when you roll a die once, there are six outcomes (sample space is  $\{1, 2, 3, 4, 5, 6\}$ ), yet you can get only one possible outcome from the sample space in one time. In this experiment, the event "rolling once but two outcomes" is impossible to happen so no description in sample space. From this respective, the set that represents a sample space in probability theory can only describe all exclusive events, not non-exclusive events combinations.

However, in the real world, a probability mass function can describe an event such as  $P(A)$ ,  $P(B)$ , and  $P(C)$ , the mass function of exclusive event combination can be describe as  $P(A \cap B)$ . But the exclusive event combinations can not describe two events happen in the same time.

Compared with probability distribution in probability theory, basic probability assignment in evidence theory can handle more uncertain information. The identification framework of basic probability assignment is power set, which means that basic probability assignment can describe not only the belief allocation of a separated event and the combination of non-exclusive events. In the identification framework of evidence theory, subsets describe all possible outcomes as well. The outcome power set is formed by three kind of subsets: empty subset represents "none possible independent events happen", single-element-subset with one element represents "an independent event happen", multi-subset represents "more than one independent event happen". A subset of may be an event or an events-combination with non-exclusive events like  $\{A, B\}$ .

In this view, a power set is explained as a family of all exclusive and non-exclusive events combinations. The single subsets in a power set family represent as exclusive events, while the multiple subsets represent as non-exclusive events happening along the time. The amount of each kind of event combination can be counted and represented as a combinatorial number. Size of power set is equal to the amount of all possible events combinations.

#### 3.1 An example to explain power set from the view of entropy

Below is an example explaining the difference between a set family consisting of single subsets only and a power set from the perspective of entropy. The example can explain the physical meaning of power set quantitatively.

- Question A:

If there are 32 numbered football teams (from team 1 to team 32) playing with each other and only one team get the champion. If we can only ask yes / no questions, how many times at least we ask, can we certainly know the champion?

Answer: 5 times.

Based on the information we get, the probability to get champion for every team is the same. If we want to certainly know the champion team even, bisection method is recommended.

Question	Is champion team	Answer	Information
1	from 1 to 16?	Yes	Champion team is in 1 to 16.
2	from 1 to 8?	No	Champion team is in 9 to 16.
3	from 9 to 12?	No	Champion team is in 13 to 16.
4	from 13 to 14?	Yes	Champion team is 13 or 14.
5	13?	No	Champion team is 14.

Figure 2: Possible asking options for knowing the champion team

In Figure 2, there is a possible asking to knowing the champion team, showing how bisection method works.

- Question B:

If there are 32 numbered students (from student 1 to student 32) taking math exam. If we can only ask yes / no question, how many times at least we ask, we can certainly know the student who got the highest score in the exam?

Answer: 32 times.

For question A, we can give probability as  $p(i) = 1/32$ , for  $i = 1, 2, \dots, 32$ .

$$\text{Shannon entropy} = -1 \times \log_2 \frac{1}{32} = 5$$

From the coding point of view, this problem can be interpreted as giving different 32 codes to 32 teams with 0 / 1, which, means 5 bits are required.

For question B, we can give basic probability assignment as  $m(x_1, \dots, x_{32}) = 1$

$$\text{Deng entropy} = -1 \times \log_2 \frac{1}{2^{32} - 1} = 32$$

Different from the football game, there may be several students getting the highest score, the most special case is that all students got the same score as well as the highest score. For example, all students got the full marks. The event "Student 1 gets the highest score" is independent to event "Student 2 gets full marks" and there can be tied highest scores more so several non-exclusive events may happen in this problem. In another word, the exclusive assumption is not correct in this situation. Both Student 1 and Student 2 can get the highest goal. The solution is to asking from student 1 to student 32 without order non-repeatedly, amounting to 32 times.

The explanation of this example in the Pascal's triangle is illustrated in the Figure 4. In the 32 row, all the composition of power set is given, and the amount of "getting the highest score" event-combinations is given as the numbers in this row. For example, the first number "1" tells the amount of events that "none of students getting the highest score" and the second number "32" tells that "one of students getting the highest score".

In this example, we can find that, for mutually exclusive events, we can use Shannon entropy to measure the information volume in the system. While For systems whose events are not mutually exclusive, Deng entropy can explain the information volume and uncertainty better. Because basic probability assignment could describe the belief in the framework of events including mutually exclusive and non-exclusive events better. In this perspective, the physical meaning of the power set could be better explained as collection of all alternative events by Deng entropy. Compared with Question B, framework of events for Question A is consisted of single subsets so Shannon entropy is useful for information measure, while the framework of events for Question B is extended to a power set so Deng entropy is better for information measure.

### 3.2 The connection with power set to combinatorial number

Comparing the two questions from the perspective of power set, in the first example, all possible champion is a power set  $A$  as  $\{\{\phi\}, \{1\}, \{2\}, \dots, \{32\}\}$ , while the second example, all possible students with full marks is

also a power set  $B$  as  $\{\{\phi\}, \{1\}, \{2\}, \dots, \{32\}, \{1, 2\}, \{1, 3\}, \dots, \{31, 32\}, \dots, \{1, 2, \dots, 32\}\}$ . The size of  $A$  is 32, which can also be represents as  $2^5$ , while the size of  $B$  is the sum of  $C(32, 0), C(32, 1), \dots, C(32, 32)$ , that is  $2^{32}$ .

For a combinatorial number like  $C(n, m)$ , represents the combinations amount of picking  $m$  ( $m \leq n$ ) elements from  $n$  different elements. It can also considered as the event "there are  $m$  independent events non-exclusively happening".  $C(n, 0)$  is 1, which can be understand as only one combination for the event "none events are picked from the possible event set". This is the connection with power set and combinatorial number.

Here is an example to the explanation of power set based on a combinatorial number.

A box contains three balls, one in red, one in blue and one in green. If you have one chance to drawn ball(s) from the box. How many different outcomes would you get?

Selections	Solution 1	Solution 2	Solution 3	Combinational number	Sum
$\{B\}$ $\{R\}$ $\{G\}$				$C(3, 1)$	3
$\{B, R\}$ $\{B, G\}$ $\{R, R\}$	 	 	 	$C(3, 2)$	3
$\{B, R, G\}$	  			$C(3, 3)$	1
$\{\emptyset\}$				$C(3, 0)$	1

Figure 3: An example of selecting balls

Figure 3 gives all possible outcomes to this ball game. You can get one ball, two balls, three balls or none of them, since every ball has two possibilities (drawn and not drawn) and the status of each ball is independent. All possible solutions consists of a power set. From this example, we can explain power set as a set with all possible "drawing ball(s)" combinations. In this situation, there are four kinds of event combinations. The first one is "none of balls is selected", that can be represented as a set  $\{S1 : \{\phi\}\}$ , the size of  $S1$  is 1. The second event combination is "one of balls is selected", that can be represented as a set  $\{S2 : \{R\}, \{G\}, \{B\}\}$ , the size of  $S2$  is 3. The third event combination is "two of balls are selected", that can be represented as a set  $\{S3 : \{R, B\}, \{R, G\}, \{B, G\}\}$ , the size of  $S3$  is 3. The last event combination is "three balls are selected", that can be represented as the set  $\{S4 : \{R, G, B\}\}$ .

The meaning of the empty subset of the power sets in this example is drawing no balls from the box, the outcomes amount of it is  $C(3, 0)$ . The meaning of single subsets is drawing one ball from the box, this outcomes amount is  $C(3, 1)$ . The meaning of multiple subsets in the power sets of this example is drawing more than one ball, whose amount is the sum of  $(C(3, x), x$  is bigger than 1 and not bigger than 3  $(C(3, 2) + C(3, 3) = 4)$ .

The combinatorial number describe the amount of each event combination. A single subset with only one element is considered as an independent event happening alone, whose amount is the combinatorial number  $C(N, 1)$ . Multi-subsets with several elements are seen as possible non-exclusive events happening, the amount of each event combination is  $C(N, X)$ , in which  $1 < X \leq N$ . All the subsets make up a power set, representing a whole event family set, the amount of all subsets is the same of all combinatorial number  $(2^N)$ .

### 3.3 Relationship between power set and Pascal's triangle

From the previous knowledge, the properties of it is known. The connection between Pascal's triangle and combinatorial number is also explored. In the previous part, the connection with power set and combinatorial number is built. So, in this part, power set is used to explain the logic and meaning of Pascal's triangle.

From Figure 1, the row sum is 1, 2, 4, 8 and 16, which can be represented as  $2^0, 2^1, 2^2, 2^3$  and  $2^4$ . The  $x - th$  number in the  $n - th$  row can be expressed as  $C(n - 1, x - 1)$ , the same as the combinatorial number of  $x - 1$  elements from  $n - 1$  different elements. The numbers in the  $n - th$  row represent the amount of power  $x$ -size subset in a set with  $n - 1$  elements respectively. This is the relationship between combinatorial number and Pascal's triangle. From the perspective of power set, Pascal's triangle shows amount of different size of subsets in a power set family.

The relationship between power set and Pascal's triangle is illustrated in Figure 4. We can find that, for any single row in the Pascal's triangle, each number represents the amount of an event-combination. For example, there is four numbers in the forth row as 1, 3, 3, 1, which can be explained as the amount of event-combination "none event happening", "one event happening", "two events happening" and "three events happening". In this explanation, a power set gives all possible events combinations of exclusive and non-exclusive events, concretely, single subsets means an event happening along and multi-subsets means events happening simultaneously.

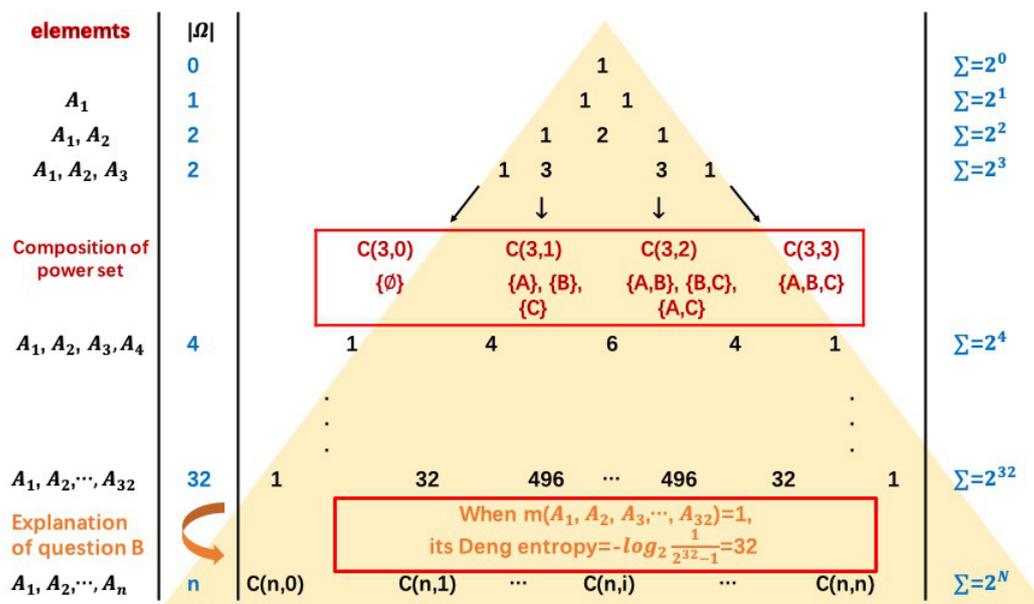


Figure 4: Explanation of power set in the Pascal's triangle

It is remarkable that the event "none of students getting the highest score" is impossible to happen in Question B, which is reflected by Deng entropy formula in Figure 4. Each number in Pascal' triangle represents the amount of a kind of event combination but not all event combinations are possible to happen for different situations.

## 4 Conclusion

In this paper, an entropic explanation of power set is illustrated in the paper. In Dempster-Shafer evidence theory, the event space in probability theory is extended into all possible combinations of the single basic event by power set. A power set is explained as an events collection set of all exclusive events and non-exclusive events combinations, the size of power set is the amount of all possible events. Single subsets of the power set family represent exclusive events, while multiple subsets represent non-exclusive events combinations possibly happening in the same time, such as two balls (green one and red one) drawn simultaneously, denoted as  $\{G, R\}$ . Combinatorial number, Pascal's triangle and information measures are illustrated to explain the meaning of power set. In addition, the relationship between Pascal's triangle and power set is given by Deng entropy quantitatively from the view of information measure. The proposed explanation may have the potential applications in evidence theory.

## Acknowledgment

The work is partially supported by National Natural Science Foundation of China (Grant No. 61973332), JSPS Invitational Fellowships for Research in Japan (Short-term).

## References

- [1] Dempster, A.P., 1967. Upper and lower probabilities induced by a multivalued mapping, 57–72.
- [2] Deng, J., Deng, Y., 2021. Information volume of fuzzy membership function. International Journal of Computers Communications & Control 16, 4106.
- [3] Deng, Y., 2016. Deng entropy. Chaos, Solitons & Fractals 91, 549–553.
- [4] Deng, Y., 2020a. Information volume of mass function. International Journal of Computers Communications & Control 15, 3983.
- [5] Deng, Y., 2020b. Uncertainty measure in evidence theory. SCIENCE CHINA Information Sciences 63, 210201.

- [6] Fenton, G.A., Griffiths, D.V., 2008. Review of Probability Theory. Risk Assessment in Geotechnical Engineering.
- [7] Fraenkel, A.A., Bar-Hillel, Y., Levy, A., 1973. Foundations of set theory. Elsevier.
- [8] Gao, X., Pan, L., Deng, Y., 2021. Quantum pythagorean fuzzy evidence theory (qpfet): A negation of quantum mass function view. IEEE Transactions on Fuzzy Systems .
- [9] Jaynes, E.T., 2004. Probability theory: The logic of science. Mathematical Intelligencer 57, 76–77.
- [10] Jech, T., 2013. Set theory. Springer Science & Business Media.
- [11] Kazemi, M.R., Tahmasebi, S., Buono, F., Longobardi, M., 2021. Fractional deng entropy and extropy and some applications. Entropy 23, 623.
- [12] Kullback, S., 1997. Information theory and statistics. Courier Corporation.
- [13] Liao, H., Ren, Z., Fang, R., 2020. A deng-entropy-based evidential reasoning approach for multi-expert multi-criterion decision-making with uncertainty. International Journal of Computational Intelligence Systems 13, 1281–1294.
- [14] Lin, J., 1991. Divergence measures based on the shannon entropy. IEEE Transactions on Information theory 37, 145–151.
- [15] Ross, S., 2009. A first course in probability. Upper Saddle River 6.
- [16] Ryser, H.J., 1963. Combinatorial mathematics. volume 14. American Mathematical Soc.
- [17] Shafer, G., 1976. A mathematical theory of evidence. volume 42. Princeton university press.
- [18] Shannon, C.E., 1948. A mathematical theory of communication. The Bell system technical journal 27, 379–423.
- [19] Van Lint, J.H., Wilson, R.M., Wilson, R.M., 2001. A course in combinatorics. Cambridge university press.
- [20] Xiao, F., 2021a. CaFtR: A fuzzy complex event processing method. International Journal of Fuzzy Systems, 1–14.
- [21] Xiao, F., 2021b. CEQD: A complex mass function to predict interference effects. IEEE Transactions on Cybernetics .
- [22] Xinyang, D., Yebi, C., Jiang, W., 2021. An ecr-pcr rule for fusion of evidences defined on a non-exclusive framework of discernment. Chinese Journal of Aeronautics .
- [23] Xue, Y., Deng, Y., 2021. Interval-valued belief entropies for dempster–shafer structures. Soft Computing, 1–9.



Copyright ©2021 by the authors. Licensee Agora University, Oradea, Romania.

This is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International License.

Journal's webpage: <http://univagora.ro/jour/index.php/ijccc/>



This journal is a member of, and subscribes to the principles of,  
the Committee on Publication Ethics (COPE).

<https://publicationethics.org/members/international-journal-computers-communications-and-control>

*Cite this paper as:*

Song Y.; Deng Y. (2021). Entropic Explanation of Power Set, *International Journal of Computers Communications & Control*, 16(4), 4413, 2021.

<https://doi.org/10.15837/ijccc.2021.4.4413>