

Using Information Theory to Observe Natural Intelligence and Artificial Intelligence

Lipeng Zhang, Limei Li, Yanming Pearl Zhang

Abstract—This paper takes a philosophical view as axiom, and reveals the relationship between information theory and Natural Intelligence and Artificial Intelligence under real world conditions. This paper also derives the relationship between natural intelligence and nature. According to communication principle of information theory, Natural Intelligence can be divided into real part and virtual part. Based on information theory principle that Information does not increase, the restriction mechanism of Natural Intelligence creativity is conducted. The restriction mechanism of creativity reveals the limit of natural intelligence and artificial intelligence. The paper provides a new angle to observe natural intelligence and artificial intelligence.

Keywords—Natural intelligence, artificial intelligence, creativity, information theory.

I. INTRODUCTION

THE research on Natural Intelligence and Artificial Intelligence has a long history. Generally, there are three main areas for the method of its research:

- 1) Computer simulation – AI: This method is to take both human and computer as black box and observe the results of input and output. [8]-[10] And to establish a computer mathematical model based on the results. Then modify their mathematical model further, until the results of input and output are closest to the results of human. The most famous example is the Turing test [7]
- 2) Study human intelligence through psychology: This method is to study human intelligent behavior by applying the measurement of IQ and EQ. [11] It is to take only human as black box and observe the results of input and output [12], [13].
- 3) Biological neural network research: This method is to study the response of neural network and observe the reaction of the biological neural networks (the results of output) when there is input [14], [15].

All these just belong to experimental theories. If we want to explain what are the limitations of natural intelligence and what is the relationship between natural and artificial intelligence, it is necessary to go further to build a solid base of theory. In order to do so, it needs an axiom to support this theory by using the imagination given by nature to reach a higher level of study.

If we say that natural intelligent is the procedure and results

Lipeng Zhang is with Hong Kong Applied Science and Technology Research Institute, Science Park, Shatin, Hong Kong (corresponding author to provide phone: 852-68064216; e-mail: l_p_zhang@yahoo.com).

Limei Li is with MMI The Corporate Headquarters of MMI, MMI, Kaki Bukit Industrial Estate, Kaki Bukit, 417942, Singapore (e-mail: lm_li@yahoo.com).

Yanming Pearl Zhang is with Faculty of Medicine, Hong Kong University, Hong Kong (e-mail: zhangympearl@gmail.com).

of information processing, no one would say no. Therefore, information theory should be considered as an optional tool to research Natural Intelligence. A number of researchers have tried to do related research. Some of them tried to take care of measurement for both Natural Intelligence and Artificial Intelligence to reveal the intelligent behavior. [16] However, the calculation is too complicated. Most researchers just pursue to implement Artificial Intelligence more close to Natural Intelligence.

There are many different views on the Natural Intelligence. We try to put it aside and make basic concepts more clear in order to push the research of Natural Intelligence and Artificial Intelligence to a higher level.

The purpose of this paper is to explore learning mechanism of natural intelligent system and the limit of natural intelligent system. The basic features of the natural intelligent system and origin of Natural Intelligence are studied. This paper is to observe Natural Intelligence and Artificial Intelligence from another angle.

II. QUESTIONS

There are some questions as follows: Origin of Natural Intelligence; The relationship between NI and nature; The restriction mechanism for NI and creativity of NI; The relationship among those; The suitable tools for the related researches. In order to answer these questions well, some philosophical points have to be considered. Because NI system can be described as an information processing system, information theory should be a suitable tool to explain the questions, including the relationships and restrictions.

III. PHILOSOPHY (AXIOM)

The existence of NI is resulted by the truth that universe and nature has its own rules

- 1) The causal relationship between NI and nature: If there is no rule in universe and nature, means the universe and nature is chaotic and random, there will be no NI. This is because the NI will be meaningless under this situation. The rules of nature are the cause and NI is the result.
- 2) NI cannot change the rules of nature: What NI can do is to observe, to learn and to apply the rules of nature.

NI cannot create or change rules of nature because it is not God.

IV. REVIEW OF INFORMATION THEORY

Information source is described by Information Theory as follows [1]-[3].

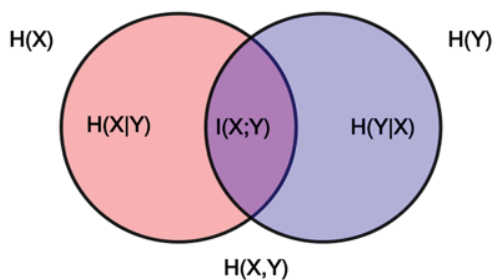


Fig. 1 Venn diagram

A. Entropy

For a discrete random variable X with possible values $\{x_1, \dots, x_n\}$ and probability mass function $P(X)$ as:

$$H(X) = E[I(X)] = E[-\ln(P(X))] \quad (1)$$

Here E is the expected value operator, and I is the information content of X . $I(X)$ is itself a random variable.

The entropy can explicitly be written as

$$H(X) = \sum_i P(x_i) I(x_i) = - \sum_i P(x_i) \log_b P(x_i) \quad (2)$$

where b is the base of the logarithm used.

B. Condition Entropy

One may also define the conditional entropy of two events X and Y taking values x_i and y_j respectively, as

$$H(X/Y) = \sum_{i,j} p(x_i y_j) \log \frac{p(y_j)}{p(x_i y_j)} \quad (3)$$

where $p(x_i, y_j)$ is the probability that $X = x_i$ and $Y = y_j$. The expectation value of condition self-information is conditional entropy.

C. Maximum

The measure should be maximal if all the outcomes are equally likely (uncertainty is highest when all possible events are equiprobable).

$$H_n(p_1, \dots, p_n) \leq H_n\left(\frac{1}{n}, \dots, \frac{1}{n}\right) = \log_b(n) \quad (4)$$

D Additive Property

For positive integers b_i where $b_1 + \dots + b_k = n$,

$$H_n\left(\frac{1}{n}, \dots, \frac{1}{n}\right) = H_n\left(\frac{b_1}{n}, \dots, \frac{b_k}{n}\right) + \sum_{i=1}^k \frac{b_i}{n} H_{b_i}\left(\frac{1}{b_i}, \dots, \frac{1}{b_i}\right) \quad (5)$$

If $k = n, b_1 = \dots = b_n = 1$ That means that the entropy of a certain outcome is zero: $H_1(1) = 0$.

E Joint Entropy

The joint entropy and condition entropy have relationship can be written as:

$$H(X,Y) = H(X/Y) + H(Y) = H(Y/X) + H(X) \quad (6)$$

When $Y=f(X)$ and f is deterministic, then $H(f(X)|X) = 0$. Using the previous formula to get $H(X, f(X))$

$$H(X) + H(f(X)/X) = H(f(X)) + H(X/f(X)) \quad (7)$$

When $H(f(X)) \leq H(X)$, the entropy of a variable can only decrease when the latter is passed through a deterministic function or system. That means information is not increased after information comes through any processing.

If X and Y are two independent events, then knowing the value of Y does not influence our knowledge of the value of X .

$$H(X/Y) = H(X) \quad (8)$$

The joint entropy should be not more than the sum of entropies of each individual events. Then

$$H(X,Y) \leq H(X) + H(Y) \quad (9)$$

It is easy to prove the equation from the previous two properties of entropy.

If X and Y are not independent, there will be:

$$H(X | Y) \leq H(X) \quad (10)$$

In the information processing, the more conditions there are, the smaller information entropy is [4]-[6].

V.NI SYSTEM AND INFORMATION PROCESSING SYSTEM

NI system is absolutely an information processing system. It has the following relationship:

NI systems belong to information processing system. This relationship should be no objection.

$$NI\text{-system} \in \text{Information-system}$$

Here, NI-system is natural intelligent systems; Information-system is information processing system.

According to information theory: for any information processing system, the output information is not more than the amount of input information.

$$\text{Information-system-output} \leq \text{Information-system-input} \quad (11)$$

Here, Information-system-output is output information of information processing system; Information-system-input is input information of information processing system.

$$\therefore NI\text{-system} \in \text{Information-system}$$

Here, NI-system is natural intelligent systems; Information-system is information processing system.

$$\therefore NI\text{-system-output} \leq NI\text{-system-input} \quad (12)$$

Here, NI-system-output is output information of natural intelligence system; NI-system-input is input information of natural intelligence system.

VI. THEOREMS

A. Theorem 1

For any natural intelligent system, output information is not more than the amount of input information when the natural intelligent system is in the certain natural environment.

Follow (12)

$$\text{NI-system-output} \leq \text{NI-system-input}$$

This theorem may bring controversy and doubt. Some people may think that the creativity of intelligent systems would break this rule. According to the principle of information communication, it is provable that output entropy is not more than input entropy, under a condition that the information processing system is in a closed environment.

Let's start with the basic concept of traditional information theory. The source and receiver of any information communication processing system are fixed. It means that both the receiver (output) and the source (input) have a unified protocol specification. Otherwise, communication is meaningless.

When a natural intelligence system tries to understand (communicates with) the nature, the protocol specification is just from one-side. It means that only nature gives the protocol specification to NI system unilaterally. Referring to the above philosophy axiom, NI system can just observe and learn the protocol specification given by nature. NI system is not the creator of the protocol specification. However, when the NI system change the protocol specification internally, which means the NI is source and receiver at the same time, wonderful things will happen, that is creativity. Therefore, NI system has duality. NI system can reflect not only the real outside world but also the internal virtual world.

- 1) Nature is source and NI system is receiver: When NI system is only a receiver to solve external problems, it has to keep consistence with the external natural environment by following the rules of nature. Otherwise, it is impossible to solve the external problems.
- 2) NI system is source and NI system is receiver too: When NI system is source and receiver, NI system can define the source by itself. When it increases its own entropy of the source, creativity will be appeared.

B. Theorem 2

It is possible using information entropy to measure the maximum entropy for creativity of NI system.

When information source elements and source entropy are confirmed, maximum source entropy minus source entropy is equal to maximum entropy of NI system creativity. Maximum entropy of NI system creativity is limited under the above conditions.

According to the features of information theory

$$H_n(p_1, \dots, p_n) \leq H_n\left(\frac{1}{n}, \dots, \frac{1}{n}\right) = H_n \text{ max} \quad (13)$$

Here, H_n is entropy of source; H_n max is the maximum

entropy.

$$H_n \text{ max} - H_n = H_n \text{ NI-creativity-max}$$

Here, H_n NI-creativity-max is the maximum entropy of NI system creativity. That is the maximum creative space.

Because nature has its own rules, the occurrence for the source elements will not be an equal probability. The source entropy should not be more than the maximum source entropy (equal probability source).

If NI systems do not comply with the theorem 1 that is mentioned in this paper: For any NI systems, output information is greater than the amount of input information. Some strange things will happen. In this way, if the NI system is just being input the initial information, the NI systems will be able to grasp all the information and knowledge without inputting any new information (learning new things). So far in our real world, we have not yet seen such NI system. This shows that NI system comes from the rules of nature. If there is no rule in nature, NI will not exist, because it is not necessary to predict anything.

VII. CONCLUSIONS

Based on the philosophy axiom that is proposed in this paper, and applying information theory and its basic concepts, we derive basic concepts and theorems for natural intelligence and creativity of natural intelligence.

The theorems in this paper answer some questions about natural intelligence and creativity of natural intelligence. A theory frame is set for nature, natural intelligence and creativity of natural intelligence. It will be very helpful for artificial intelligence research and could be used as a guide for the research direction. From this paper we already know there are some limit for natural intelligence and creativity. Following the restriction mechanism of natural intelligence creativity and the concept of virtual world of natural intelligence system, the creativity could be duplicated. We believe this paper will bring inspiration for natural intelligence and artificial intelligence research.

REFERENCES

- [1] Shannon, Claude E. (July–October 1948). "A Mathematical Theory of Communication". *Bell System Technical Journal* 27 (3): 379–423.
- [2] Borda, Monica (2011). *Fundamentals in Information Theory and Coding*. Springer. p. 11. ISBN 978-3-642-20346-6.
- [3] Han, Te Sun & Kobayashi, Kingo (2002). *Mathematics of Information and Coding*. American Mathematical Society. pp. 19–20. ISBN 978-0-8218-4256-0.
- [4] Schneider, T.D, *Information theory primer with an appendix on logarithms*, National Cancer Institute, 14 April 2007.
- [5] Carter, Tom (March 2014). *An introduction to information theory and entropy (PDF)*. Santa Fe. Retrieved Aug 2014. Young, "Synthetic structure of industrial plastics (Book style with paper title and editor)," in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.
- [6] Tian BaoYu edited Zhou JiongPan and Wu WeiLing review <Engineering Information Theory> Beijing University of Posts and Telecommunications Press
- [7] Turing, Alan (October 1950), "Computing Machinery and Intelligence", *Mind* LIX (236): 433–460, doi:10.1093/mind/LIX.236.433, ISSN 0026-4423, retrieved 2008-08-18.

- [8] Poole, David; Mackworth, Alan; Goebel, Randy (1998). *Computational Intelligence: A Logical Approach*. New York: Oxford University Press. ISBN 0-19-510270-3.
- [9] Russell, Stuart J.; Norvig, Peter (2003), *Artificial Intelligence Artificial Intelligence: A Modern Approach* (2nd ed.), Upper Saddle River, New Jersey: Prentice Hall, ISBN 0-13-790395-2.
- [10] McCarthy, John; Minsky, Marvin; Rochester, Nathan; Shannon, Claude (1955). "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence Artificial Intelligence". Archived from the original on 26 August 2007. Retrieved 30 August 2007\
- [11] Loehlin, J. C.; Perloff, R.; Sternberg, R. J.; Urbina, S. (1996). "Intelligence: Knowns and unknowns". *American Psychologist* 51 (2): 77. doi:10.1037/0003-
- [12] Maich, Aloysius (1995). "A Hobbes Dictionary". Blackwell. p. 305
- [13] Nidditch, Peter. "Foreword". *An Essay Concerning Human Understanding*. Oxford University Press. p. xxii
- [14] Gottfredson, Linda S. (1997). "Mainstream Science on Intelligence". *Intelligence* 24: 13–23. doi:10.1016/s0160-2896(97)90011-8. ISSN 0160-2896.
- [15] Neisser, Ulrich; Boodoo, Gwyneth; Bouchard, Thomas J.; Boykin, A. Wade; Brody, Nathan; Ceci, Stephen J.; Halpern, Diane F.; Loehlin, John C.; Perloff, Robert; Sternberg, Robert J.; Urbina, Susana (1996). "Intelligence: Knowns and unknowns". *American Psychologist* 51: 77–101. ISSN 0003-066X. Retrieved 9 October 2014.
- [16] Legg, Shane and Hutter, Marcus (2006) *A Formal Measure of Machine Intelligence*. Technical Report. UNSPECIFIED. Proceedings of the 16th Annual Machine Learning Conference of Belgium and The Netherlands, Belearn (2006)