Article

Unity-Based Diversity: System Approach to Defining Information

Yixin Zhong

University of Posts and Telecommunications Beijing, Beijing 100876, China;
E-Mail: yxzhong@ieee.org

Abstract: What is information? This is the first question that information science should answer clearly. However, the definitions of information have been so diversified that people are questioning if there is any unity among the diversity, leading to a suspicion on whether it is possible to establish a unified theory of information or not. To answer this question, a system approach to defining information is introduced in this paper. It is proved that the unity of information definitions can be maintained with this approach. As a by-product, an important concept, the information eco-system, was also achieved.

1. Introduction

As the most fundamental basis of information science study, the issue of precisely understanding and properly defining the concept of information has received a huge amount of attention from academic circles in information-related fields. Due to the extremely popular existence and wide applicability of information, hundreds of definitions of information, different from each other in one way or another, have been successively proposed in various fields of research, among which Shannon [1], Wiener [2], Brillouin [3], Ashby [4], Longo [5], and others, are the most influential contributors.

The diversity of definitions of information shows the importance of information studies on one hand but on the other hand, a question has been naturally raised as to whether the definitions which people talked about in different backgrounds have anything in common? In other words, people are concerned as to whether a unified theory of information is feasible or not [6,7].

Many scholars have proposed various ways to answer the question, either positive or negative. Mark Burgin recently gave an approach, the parametric approach [8], to the unification of information theory. The parametric definition of information utilizes a system parameter. It is said that an infological system plays the role of a parameter that discerns different kinds of information, e.g., social, personal, chemical, biological, genetic, or cognitive, and combines all existing kinds and types of information in one general concept information. This seems an interesting approach.

The author of the paper presented a system approach in his monographic book titled "Principles of Information Science" (published in Chinese in 1988) [9], for the purpose of organizing the various definitions of information in good order and establishing a unified theory of information. The system approach can not only organize the different, yet not erroneous, definitions of information into an orderly spectrum, but can also integrate all information processes into a harmonious eco-system: from information acquisition (sensing), to information transmission (communication), information processing (computation), information understanding (knowledge producing), information regeneration (strategy formation), and strategy execution (control). Among them, the process of information understanding and regeneration is at the core of the eco-system, and is termed as Artificial Intelligence.

The basic concept of the system approach to the definition of information had been very briefly reported to the International Conference on Foundation of Information Science, FIS2010, in Beijing. It was requested to present the approach and the major results in somewhat more detail here in this paper for having further exchange and discussions.

The rest of the paper is organized as follows. The system approach to define complicated phenomena and the concept of spectrum of definitions are explained in general terms in Section 2. The most fundamental, and also most significant, definitions of information derived from the system approach are elaborated in Section 3. An interesting concept of information eco-system based on the information spectrum is introduced in Section 4. Brief conclusions and remarks are made in Section 5.

2. System Approach to Defining Information and the Spectrum of Definition

It may be good to start with a well-known fable, Blind Men and the Elephant. The fable says that, having touched different parts of the same elephant body, different blind men got a different impression about it. Some said the elephant felt like a wall, other said it felt like the trunk of a tree, the third said it felt like a snake, and the fourth said it felt like a spear, and the fifth said it felt like a big fan, and so on.

This perhaps is a typical example showing that people from different angles of observation (locations of touching) may obtain different impressions of the same object. It is common knowledge that the elephant image (as a system) may well be formed from the different descriptions (as elements) given by all the blind men, although their descriptions are so diversified. The only thing that really matters for the restoration of the system from the elements (or for finding the unity from the diversity) is "relation among the conditions" (or the angles of viewing) under which the elements were derived. Consequently, people realized that the wall felt by the blind man is the image of elephant's belly, the tree trunk is the image of elephant's leg, the snake is the image of elephant's trunk, the spear is the image of the elephant's tooth, the big fan is the elephant's ear, *etc.*, and all of them form an entire image of the elephant.

Hence, it is reasonable to set up a spectrum of the formal descriptions about the elephant by properly listing the conditions under which different descriptions were derived. Specifically, putting the conditions and the corresponding descriptions (definitions) in a systematic order, the spectrum in question can then be formed as shown in Table 1, below. This is the system approach to definition.

Table 1. System Approach to Defining: An Example.

Conditions for Defining	Levels of Definitions	Names of Definitions
None	Highest	Elephant as a whole (Original System)
Ignoring Legs	Second Highest	Main Body of Elephant (Big Element)
Ignoring Legs, Head and Tail	Third Highest.	Belly of Elephant (Small Element)
More conditions Constrained	Lower	Smaller Element
….	….	….
The strictest Condition	Lowest	Smallest Element

What we would like to emphasize here is that as long as the descriptions of the parts of elephant are correct, the unity of the elephant image can be restored no matter how many kinds of different descriptions about the partial image of the elephant existed, and regardless of how different the descriptions of the different parts of the elephant may be. Above all, each of the different descriptions has resulted from the real elephant, and the relation among each of the angles is clearly described.

Similarly, information may be understood in different ways by different observers who may have their own knowledge backgrounds or angles of view. They could make different descriptions on—and thus give different definitions to—information and thus may lead to the diversity. The only thing that does matter for keeping the unity among various information definitions is the correctness of the definition for each kind of the information and the clearness of the relation among the conditions under which the different definitions of information are derived, as was seen in the case of Blind Men and the Elephant. Hence, the system approach for defining and the concept of spectrum of definitions, explained above, can be employed to deal with the definition of information as shown in Table 2, below.

Table 2. System Approach to Information Defining and the Related Spectrum.

Conditions for Defining	Levels of Definitions	Names of Definitions	Area of Applicability
None	Highest	Ontological Information	Widest
Human subject as observer	Second Highest	Epistemological Information	Second Widest
….	….	….	….
More conditions added	Lower	Definition k	Narrower
….	….	….	….
All possible conditions added	Lowest	Definition K	Narrowest

As can be seen in Table 2 above, the definition of information with no constraining conditions is named ontological information, while the definition of information with one single (the smallest in number, yet the greatest in significance) condition of "with human subject as observer" is named epistemological information. The former can be regarded as the unique root of definitions of information

from which all other definitions of information (including epistemological one) can result, whereas the latter can be regarded as the only main trunk of definitions of information from which all other definitions of information with various conditions added (except the ontological one) can be derived.

Because of no restriction from any condition, the definition of ontological information is really the origin and must be placed at the top of the table, and must have the widest area for its applicability too. In contrast, if a definition is derived under all the possible constraining conditions, the definition must be placed at the bottom of the table and must have the narrowest area of applicability. Other cases will be placed between the two extremes. As a consequence, if the conditions for defining can be arranged in good order according to certain criterion, then the various definitions thus derived can also be arranged in good order under the same criterion, forming a spectrum of information definitions.

It needs to be pointed out that the word "order" in the arrangement of conditions and that of definitions can either be quantitative order (such as longer or shorter, and so on), or qualitative order (like better or worse, and so forth), or order according to any possible criteria.

This methodology is named system approach because it establishes the definitions by systematically considering all possible synergetic conditions. The system approach makes all definitions, excluding the mistaken ones, which might seem chaotic in order and arbitrary in content, become systematically arranged in good order and with harmonious interrelationship.

Apparently, it is neither possible nor necessary to list all the possible conditions and definitions of each component of the spectrum in the table, due to the limited space. However, whenever a specific condition for defining was given, it is possible to find a proper position for it within the table with the same criterion for arrangement. On the other hand, although some definitions may have not yet appeared so far, the proper positions within the table can be reserved for them respectively.

It is interesting to note that when the condition of "with human subject as observer" is added to the definition of ontological information, this definition will be down to the definition of epistemological information. On the other hand, when the condition of "with human subject as observer" is deleted from the definition of epistemological information, this definition will be upgraded to the one of ontological information. So, the two definitions can go up or down when the condition is deleted or added.

The rule, adding or deleting a condition will move a definition up or down, is also valid for any other definitions in the spectrum of Table 2. For example, if adding the condition of "the human subject is dealing with natural science study" to the definition of epistemological information, this definition will be down to the definition of information in natural science. If adding the condition of "the human subject is dealing with physics" to the definition of information in natural science, the latter will be down to the definition of information in physics. Similarly, if adding the condition of "the human subject is dealing with biology" to the definition of natural science, the latter will be down to the definition of information in biology. Furthermore, if adding a condition of "the human subject is dealing with the research in the field of communication theory" to the definition of information in physics, the latter will be down to the definition of Shannon information, and so on and so forth.

Therefore, in order to keep the unity from the diversified definitions of information via system approach, the most important issues, among many others, consist of two folds. The first issue is to have a correct description (definition) for the ontological information and epistemological information, the top two in the spectrum. The second issue is to make clear interrelationships among the various

definition conditions, including the background of knowledge and the angle of observations under which the various kinds of definitions of information are established. Due to the limited space of the paper, only the first issue, the most important one, will be discussed below.

3. Fundamental Definitions of Information

As can be seen in Table 2, the definitions of ontological information and epistemological information are the most fundamental ones. This is because of the basic fact that all other definitions originate from these two. In other words, these two definitions are respectively the roots and trunk of the big tree of information definitions.

There is one basic point that must be clarified before being able to talk about how we could properly define the concept of information. The point is about the question "What should people mean by saying that one obtained information?" Without common understanding over this point, the unity of definitions will really be impossible—people all talk about "information" but they mean different things in their minds.

To the author's understanding, having obtained information about a thing means having observed its appearance including the state at which it is at the time and the way by which its state varies from others. In other words, information corresponds to the description of a phenomenon which is able to answer the question of "What the thing is".

It is "the state in which the things is and the ways by which it varies to others" that can be used to remove the uncertainty concerning the things. Also, "the state in which the thing is and the ways by which it varies from others" is information, but not the things themselves (the matter), and not the energy either.

Note that behind, and/or beyond, the description of phenomenon which can be sensed by humans or machines, there should have been something else which would explain why the phenomenon behaves in such a way and this, in fact, has been the concept of knowledge.

Thus, another issue that we need to make clear before talking about the definition of information is about the interrelationship among the concepts of information, knowledge and intelligence. People will all agree that the three concepts have essential linkages on one hand, but they are all different on the other hand. We cannot consider information as the same thing as knowledge and intelligence, and *vs*. To clarify the similarities and differences among the concepts of information, knowledge, and intelligence, one can say that <u>information is a description of a phenomenon and able to answer the question of</u> "**what**"; <u>knowledge is a description of the essence of the phenomenon and able to answer the question of</u> "**why**"; <u>and intelligence is a description of the problem-solving strategy and able to answer the question of</u> "**how**". Thus, information is the basic resource, knowledge can be produced from information mainly through induction-like operations, and intelligence can be produced from knowledge and information mainly through deduction-like operations.

Having said the above, the definition of ontological information can be stated as follows.

Ontological Information of a thing is its self-presentation about "the state in which the thing is and the way by which the state varies from others".

Here, the word "thing" can be any phenomena, including human beings in human society, living beings and non-organic things in real world, and also spiritual events in human mind or animal mind, if any. Thus, the definition of ontological information covers everything and can be applied in natural sciences and social sciences to describe the thing's "state in which the thing is and the way by which the state varies from others".

Obviously, this definition of ontological information is highly coincident with our intuition in daily life as well as in scientific research, and therefore can be acceptable. Ontological information thus defined is really a resource provided by nature and society.

Note that the definition of ontological information is in good agreement with the one given by Wiener because "the state and the way of state varying" *per se* is neither matter, nor energy. Also, it has no contradiction with the one given by Shannon because it is just "the state and the way of state varying" that can be used for removal of uncertainty in communication.

Noticing again that the ontological information of a thing is defined without any conditions needed to obey and therefore the definition of ontological information can be unconditionally applied to anything and to whatever fields, like natural science, social science, cognitive science, *etc.*

According to the principle indicated in Table 2 (or Table 1 equivalently), definitions of other kinds of information can effectively be derived from the definition of ontological information by adding respective conditions to it. There is one condition, among others, that possesses the most significance, which says "there must be human subject who is concerned with the information". If this condition is added to the ontological information, it will become the "Epistemological Information" which will thus have a level lower than the one of ontological information in Table 2.

Epistemological information is of high significance; this is because of the vital fact that information itself is a kind of resource and, therefore, making good use of information for humans is just the major, and even the unique, purpose of the study of information science. Hence, the study in depth of the interrelationship between humans and information can never be overemphasized.

Any human subject normally possesses three categories of abilities in dealing with information, namely, the ability to observe the form, the ability to understand the content, and the ability to evaluate the utility of the information he/she faces. The three categories of ability are indispensable. Therefore the definition of epistemological information should become more complicated than that of ontological information, and can be expressed as follows.

The epistemological information for a human subject about a thing is the subject's description on the form, the content, and the utility of "the state in which the thing is and the way by which the state varies from others", and is respectively named form, content, and utility information.

Note that the trinity of the three components is named Comprehensive Information, which is another name for epistemological information, and will be used preferably in what follows.

It is easy to see that the form component of comprehensive information can be regarded as the carrier of information content and utility, while the content and utility components can be regarded as what are carried by the form component and as what the subject really wants to know. As a trinity, the carrier (the form), the content, and the utility cannot be isolated from each other. As a matter of fact, there is no such content and utility that exists without form. There is also no such a kind of form

which has no content and utility. In other words, the three components, the form information, the content information, and the utility information, are also indispensable.

Practically, if someone has obtained all three components of comprehensive information concerning a thing, this means that he/she knew the form, understood the content, and was aware of the utility of the thing. As a consequence, he/she should be able to make decisions about the thing. In contrast, if there is any component(s) of the comprehensive information unavailable, then the decision-making would be difficult, if not impossible. In view of the extremely high importance of decision-making in human activities, the greatest value should be attached to the study of comprehensive information.

Imagine that what we received from the environment is just "the states in which things are and the ways by which the states vary from others"—ontological information and nothing else. On the other hand, what we gave to the environment (things, people, or machines) is just the description about "the states in which things are, and the ways by which the states vary from others" with their form, content, and utility in our mind—comprehensive/epistemological information.

Comparing the two definitions above, it is clear that the descriptions of ontological and comprehensive information are different. The description of the former is the thing itself whereas for the latter it is the human subject. Nonetheless, both definitions are concerned with "the state in which the thing is and the way by which the state varies from others". Consequently, adding the condition of "with human subject as observer" will make the definition of ontological information become the one of comprehensive information, while deleting the same condition will make the definition of comprehensive information become the one of ontological information. This is the harmonious relationship between the two fundamental definitions.

Note that the form, content, and utility information can respectively be called the syntactic, semantic, and pragmatic information, which are terms bought from semiotics. Due to controversies still associated with the semiotics, however, great care should be taken in the use of such terms.

As for the relationship between comprehensive information and Shannon information, it is apparent that the latter is a special case of the former. More specifically, adding the condition of "the subject is dealing with statistical communication and, therefore, the content and utility components are ignored" will make the definition of comprehensive information being the one of Shannon information. Hence, Shannon theory does deal only with a kind of carrier (statistical one) and not with content and utility. As mentioned above, Shannon theory is an indispensable branch of comprehensive information theory.

It should also be emphasized here that the definition of Shannon information as a special case of comprehensive information by ignoring content and utility components cannot properly be applied to such fields as economic, social, and psychological sciences. Yet there is no such obstacle in principle for applying the definition of comprehensive information to these fields.

The difference between Shannon theory of information and comprehensive information theory in dealing with economic, social, and psychological sciences can well be explained by mentioning the following fact.

It is well known that the most concerned, and most crucial, parameters in such studies of economic, social, psychological sciences as well as arts, poetry, and the like, are the benefits and harmfulness, gains and losses, appreciation and hatred, to the decision-makers. Due to ignoring content and utility components, Shannon theory of information has only the means of a formal description and therefore

cannot describe such kinds of parameters. On the other hand, however, the definition of comprehensive information contains the form, content and utility components so that both the beneficial and harmful parameters in economic, social, as well as psychological science can properly be described by the content and utility components.

It is believed that there should be many more points, in addition to those above, related to the definition of information which needs to be discussed since this is really a complicated issue. However, it may be wise to stop here for the time being as the space of the paper is limited and it is impossible to deal with all issues within one single contribution. Anyway, it has been proved throughout the discussions we have had in the paper so far, that the unity of information definitions has been evidently maintained from its diversity, provided the system approach, the spectrum of definitions, and the definitions of ontological and comprehensive information are observed and accepted.

It is undeniable, however, that due to the high complexity of the information problem, not all definitions of information attempted are correct. Instead, there have been quite a number of incorrect, or imprecise, "definitions" of information put forward in history, and even today. These erroneous definitions cannot find proper positions in the spectrum of information definition, of course.

One of the most frequently seen items in such erroneous "definitions" is that "uncertainty" has been regarded as the same thing as information. This is clearly a misunderstanding. As a matter of fact, information is something that can be used to remove uncertainty. If entropy, to a certain extent, is employed to measure the amount of the uncertainty, then the amount of information is somewhat negative entropy. So, information and uncertainty are opposite concepts. Another example of imprecise "definition" of information is the claim that information should be defined as "pure abstraction without form/carrier". This conception is incorrect because such a kind of information cannot exist in reality.

It is, of course, not possible to list all examples of such incorrect and imprecise definitions here in the paper. However, it can be pointed out again that all the incorrect and imprecise definitions of information will be unable to find their positions in the spectrum of information definition shown in Table 2.

4. Information Eco-System
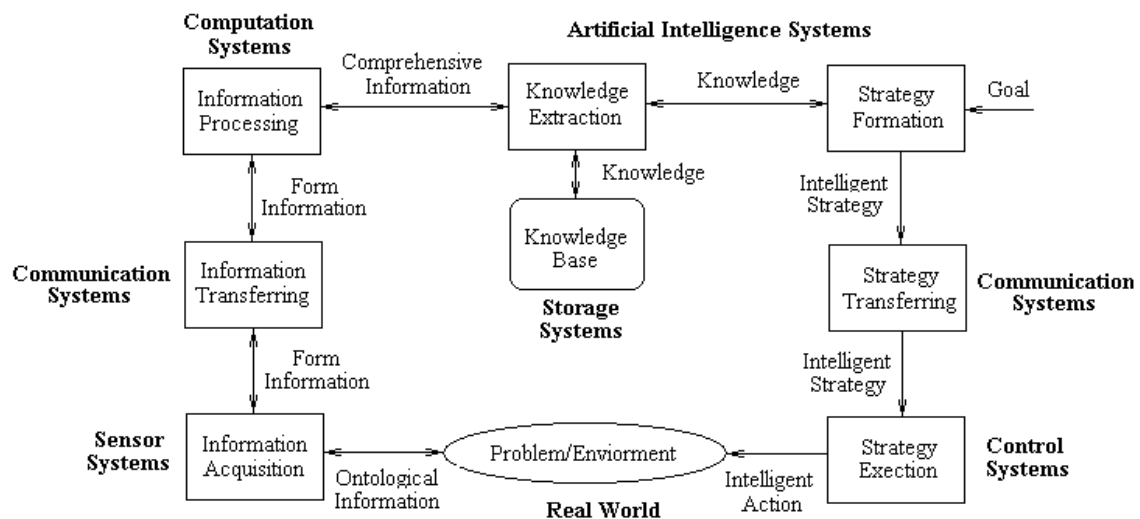
Based on the definitions of ontological and epistemological/comprehensive information discussed above, an interesting concept, the concept of information eco-system, can be established. In view of its importance, a brief introduction will necessarily be provided in this section of the paper.

Everyone knows that there have been various kinds of information systems working in the field of information technology, such as sensing systems, communication systems, computation systems, artificial intelligence systems, control systems, and so on. Each of them has been growing into different disciplines in history although they are all named as some kind of "information" system. For example, Sensor is named as the system for information acquiring, Communication is named as the system for information transferring, Computation as the system for information processing, Artificial Intelligence as the system for information refining and re-generation and Control as the system for strategic information execution. They can work together via protocols in practice; but there has been a

lack of an information theory that can effectively unify all theories of these kinds of information systems into a harmonious whole, until the present time.

After in depth investigations on the essence of information processes and recalling the definitions of ontological and comprehensive information, it is discovered that these kinds of information systems constitute a well-organized family in terms of both theory and technology, as is shown in Figure 1, below.

Figure 1. Conceptualized Model of Information Eco-system.

This is an information eco-system which clearly indicates how information is growing from its very origin (ontological information) to its final destination (intelligent strategy and intelligent action).

Referring to the definitions of ontological and comprehensive information and the model in Figure 1, it can be seen that ontological information is growing into comprehensive information (with form information as intermediate product) through information acquisition and processing systems; then comprehensive information is growing into knowledge through knowledge extraction system (which can also be named as information understanding system); and further, knowledge is growing into intelligence (intelligent strategy and intelligent action) through strategy formation and execution systems (which can also be named strategic information regeneration and execution systems).

More briefly, the information ecological system described above can be characterized as the growing process from information to knowledge and intelligence, or as the information-knowledge-intelligence chain. The original resource of the eco-system is information, the intermediate product is knowledge, and the final product is intelligence (intelligent strategy and intelligent action). For more detail, see reference [10].

We must say that it is the system approach and the spectrum of information definitions that not only make the diversified definitions of information unified but also integrates the already separated theories of information systems. This will make great sense both in theory and practice.

5. Conclusions and Remarks

A system approach, the methodology for information defining, the spectrum of information definitions, and the definitions of ontological and epistemological/comprehensive information in

particular, were proposed in the present paper for discussion. A new concept of high importance called information eco-system was also briefly presented.

It has been noticed that among the great number of useful definitions of information that have appeared earlier or later in various cases, there have been some definitions that do not at all make sense in scientific standards. With the exception of these incorrect ones, all valid definitions can find their positions in the spectrum of information definition introduced in the paper according to the conditions under which they have been derived. This demonstrates the unification between unity and diversity, and thus provides the necessary foundation to the establishment of information science.

However, it was also noticed that there are still some issues that people have frequently been addressing.

One such issue is that people would very much like to have "only one single definition of information" than to have the concept of spectrum. It is a quite understandable and acceptable demand. In this case, the definition of ontological information can be the best candidate because of the fact that this definition has no restriction in any sense and can generally be applied wherever. However, as usual, there is a sensitive relationship between "general" and "specific". If one is concerned only with the general study, like the study of philosophy of information, the definition of ontological information would be sufficient; if one is concerned with some more specific studies, like the study in biology, intelligence, sociology, *etc.*, the definition of ontological information alone will be insufficient. Therefore, the spectrum of information definition will be helpful.

Another such issue is the argument over Shannon theory of information. Some people insist on the viewpoint that information is concerned only with meaning, or content, and has nothing to do with the carrier, or the form. Therefore, they do not accept Shannon theory of information as a branch of information theory. Some of them even strongly denounced Shannon theory as a mistaken theory of information. As was mentioned in a previous section, the precise concept of information should be the comprehensive information which consists of three components, *i.e.*, the form, the content, and the utility. There is no such information in reality that is pure content without form and utility. Therefore, as the statistical form component of comprehensive information, Shannon theory of information is a necessary, but not sufficient, branch of the comprehensive information theory.

The third issue is the argument as to whether information is measurable or not. Some people, who insist the view of "information is concerned only with the pure content", consider that information cannot be measured quantitatively. Generally speaking, whenever people properly set up parameters for the description of information, it should be, in principle, a measurable entity. Even if for the time being there is a lack of suitable mathematical tools or instruments for measuring it, one can reasonably expect to have such mathematics and instruments in the near future, as long as the mathematical and technological progress continues.

It is realized, of course, that the open issues related to information science must be much more than those herein clarified. Due to the high complexity and the great importance these issues possess, further exchange and discussion among researchers are absolutely essential and very valuable for future achievements. It is believed that the collective efforts will be much more effective than those achieved individually and separate from each other.

Lastly, but not necessarily the least, it may also be worth mentioning that the concept of information ecological system can make great sense in the study of information science and

technology. Only when all the information systems (sensing, communication, computation, artificial intelligence, and control system) work as a harmonious entity, could information systems as a whole become real intelligent systems, providing human society with more and more intelligent services.

Acknowledgement

References

1. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *47*, 379–632.
2. Wiener, N. *Cybernetics, Second Edition: or the Control and Communication in the Animal and the Machine*, 2nd ed.; MIT Press: Cambridge, MA, USA, 1961.
3. Brillouin, L. *Science and Information Theory*; Academic Press Inc.: New York, NY, USA, 1956.
4. Ashby, W.R. *Introduction to Cybernetics*; Wiley, New York, NY, USA, 1956.
5. Longo, G. *Information Theory: New Trends and Open Problems*; Springer-Verlag: Berlin, Germany, 1975.
6. Marijuan, P.C. The advancement of information science. *TripleC* **2009**, *7*, 369–375.
7. Capurro, R. Is Unified Theory of Information Feasible? In *The Quest for a Unified Theory of Information*, Proceedings of the 2nd International Conference on Foundation of Information Science; Hofkirchner, W., Ed.; Overseas Publishing Association: Amsterdam, The Netherlands, 1999; pp. 9–30.
8. Burgin, M. *Theory of Information: Fundamentality, Diversity and Unification*; World Scientific Publishing Company: Singapore, 2010.
9. Zhong, Y.X. *Principles of Information Science (in Chinese)*; BUPT Press: Beijing, China, 1988.
10. Zhong, Y.X. *Principles of Cognetics in Machine (in Chinese);* Science Press: Beijing, China, 2007.