Knowledge Network Based on the Standard Model of Knowledge Representation¹

Yin Wensheng*

School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China wsyin@mail.hust.edu.cn

Keywords: knowledge representation; standard model; knowledge network; semantic network; parataxis network

Abstract. The basic unit of knowledge representation is the knowledge component. We can represent complex semantic through the construction of complex knowledge component or the combination of some knowledge components, where the latter is the common method. For the representation of complex semantic, the abilities and defects using the representation method of tradition semantic network and parataxis network are analyzed, and a knowledge network representation method based on the standard model of knowledge representation is proposed. This knowledge network has power ability to represent the complex semantic by the series-parallel connection of knowledge components. The running of this knowledge network is determinate, so it can express the complex semantic accurately without ambiguity. From the analysis to the knowledge network, it is pointed that knowledge network is not the copy of physiological structure of a human brain, but the logical description of knowledge. The knowledge network has relationship with the solving of knowledge, many parts of it are produced dynamically in the process of inference in agent, and using static knowledge network to construct the knowledge system is unable to avoid the explosion of knowledge network.

Introduction

Bacon said: "knowledge is power". How knowledge to generate power requires two conditions, one is the knowledge must be linked to form a system, that is, a large number of knowledge is connected to form a network; the second is the knowledge must be used, that is, transformed into force through human intelligence activity. A single piece of knowledge is difficult to use, while the knowledge not be used is meaningless.

Related Work

The precondition of applying knowledge is to express the knowledge. Knowledge representation method is often divided into two major categories: symbolic method and other methods, where symbolic method includes natural language, formal language, graphics language, etc.

In the field of artificial intelligence, the symbolism is the mainstream of artificial intelligence research. Symbolism considers that human cognitive primitive is symbol, and the cognitive process is a procedure of symbol operation. The core problem of artificial intelligence is knowledge representation, knowledge inference and knowledge application. Symbolism has done a lot of works on the operation method, such as the establishment, modification, copying and deletion of the physical symbol system and put forward a large number of knowledge representation and solving methods [1].

Natural Language. All kinds of language used by people to communicate in daily life are called natural language, for example, English, Chinese, etc. Natural language is the most basic, the most common, and the longest history knowledge representation method. It is a kind of knowledge representation method, which uses speech, writing, and special marks to carry out communication

¹ National Natural Science Foundation of China: Assembly sequence planning based on the knowledge of engineering connection (51175200)

and exchange. Natural language has conventions, ambiguity, thumbnail, implication, metaphors and other features, and can be used to represent very complex and obscure knowledge, and play a very important role in the knowledge representation.

In form, natural language is a one-dimensional language, although it can add some special markers, such as in the spoken language to add some notes, weight change, in the text to add some symbols and so on to express more complex semantics. In general, the understanding of natural language requires certain intelligence.

Formal Language. To make knowledge representation more convenient and more strictly, the formal language is often used. Formal language has a strict definition and description, which must be strictly observed in the process of using. Formal language includes mathematical language, computer language, professional language and graphic language and so on.

Common forms of knowledge representation languages include: State space representation, Predicate logic representation, Ontology, Semantic representation, Frame representation, Production representation, Petri net representation, Database, Mathematical model, Computer program, Object oriented representation and so on[2-11].

Graphics Language. According to ontology, everything in the world can be expressed by concepts and their relationships, and you can use symbols that represent these concepts and their relationships to represent the world. Symbols can be divided into two categories, one is ordinary symbols, such as letter, Chinese characters stroke, mathematic symbol, and so on, the other is graphic symbol, such as points/nodes, lines, etc. The former emphasizes the combination of the sequence of symbols, such as the natural language, in the sequence of one dimension, and the mathematical language according to the combination of mathematical formulas, while the latter emphasizes the combination of the basic graphics, to study the relationship between things through the two-dimensional graphics. The method of representing knowledge by means of a two-dimensional graph is called a graphics language.

Using graphics and graphic language for knowledge representation is very favourable, and it can be very clear to express the inherent relationship between the components of knowledge. In fact, in the using of other languages to express the knowledge, we often use both graphics and graphics language. For example, the various explanatory illustrations in books, the State space method of knowledge for solving the state diagram, Assembly sequence And-Or graph, Semantic networks, and so on.

Things are interrelated, in most cases, using a single piece of knowledge to express the relevant knowledge is not enough, so you have to use a lot of knowledge. For example, in natural language, more than one statement is needed to express the problem clearly. Similarly, a mathematical model may contain a large number of mathematical definitions and formulas. If the knowledge is represented by a graph language, a graph is formed. Because this kind of graph is a kind of network structure, which is composed of the basic graphic elements that represent the concepts and their relations, and form a kind of network structure, so it is often called the knowledge network.

Other Methods. In addition to represent by the symbols, knowledge can also be expressed and applied in other ways, such as a kind of physical structure. In the field of artificial intelligence, the other two kinds of famous knowledge representation methods are connectionism and behaviorism.

Connectionism holds that human intelligence comes from the structure of the human brain, and artificial intelligence can be achieved through the building of artificial brain model by studying the human brain structure. In 1943, psychologist McCulloch and mathematical logician Pitts proposed MP brain model, pioneered the study of connectionism artificial intelligence. At the beginning of 1980s, Hopfield proposed the method of using hardware to simulate the neural network [12]. After that, Rumelhart et al put forward a multi-layer network back-propagation (BP) algorithm.

Behaviorism derived from Cybernetics and perception—motion control system, it believed that AI depends on perception and action and put forward intelligent behaviour "perception—action" mode [13]. The main idea is: intelligence does not require knowledge representation and reasoning; human intelligence is gradually evolved, so artificial intelligence can also be gradually evolved as human beings; the evolution of intelligence can not be separated from the real world surrounding environment interaction.

Knowledge Components

The knowledge is in the form of semantic, semantic and knowledge has some synonymy, so you can study the expression of semantic, namely through the study of semantic expression language to study the knowledge representation. Then, when it is not specified, semantic is as same as knowledge. **Knowledge Representation Standard Model.** If k is any knowledge to describe the relationship among some things and their states, then, according to ontology, reductionism, system theory and control theory, a system can be used to encircle all the things and their status having relations with the knowledge, that is, the things and their status in this system and their relationship are combined to form the content of the knowledge, and the connection between interior and exterior of the system becomes the boundary, so that the knowledge representation with system view can be obtained. The knowledge representation can be show by Fig. 1, known as the standard model of knowledge representation.

Knowledge representation standard model can also be described with the following mathematical expression:



$$k = f(x_1, x_2, \mathbf{L}, x_n, y_1, y_2, \mathbf{L}, y_m)$$

Fig. 1 Standard model of knowledge representation

Knowledge representation standard model can be divided into three parts: the input part, the processing part and the output part. The meaning of each part is as follows:

Input part : $x_1, x_2, \mathbf{L}, x_n$

It is used to describe the starting state of the system, also known as the input section of the knowledge. Input includes input type and value.

Processing part : f

It is the current input data processing method to describe how the things changes in the system. It is the core part of the knowledge. The whole knowledge representation of the standard model is called as knowledge, but because the processing part is the core of knowledge, sometimes also call the processing part as knowledge of the whole, also called method in some systems.

Output part :

It is used to describe the end state system changed, that is, the final state of things in the system, also known as the output of current knowledge. Output is the result of the current input data processed. The output also includes output type and value.

Input and output parts are description of the concept or state of things. They can be determined, or it can be random. They can represent time changes or space changes, and can be a constant or a variable.

Semantic Topology Representation. From the viewpoint of ontology, any sentence is a representation of the concept or the relationship between them, so that can be represented by the relationship between the nodes and the nodes, namely, can be represented by a topological graph.

In natural language, sentence structure is the most basic SPO (Subject, Predicate, Object) structure, called a simple sentence. Topology of each simple sentence can be expressed in two ways, one is to express the concept or the thing with the node, using marked lines represent relationships between concepts, as shown in Fig. 2 (a); The other is represented by a bipartite graph, that is, concepts and conceptual relations are represented by nodes, while the directed line segment means there is a relationship between concepts, as shown in Fig. 2 (b); The former method is simple and clear, and the latter is more comprehensive and easy to show the relationship between single item and multiple item. As a matter of fact, the first method can also express the relationship among the single item and multi item, so the two expressions can express the topological relation of the sentence.



Fig. 2 Topological graph of sentence

Knowledge Components. Definition 1: a full expression of a piece of knowledge is called knowledge component.

It is clear that the knowledge represented in Fig. 1 is the knowledge component, because it can completely express a piece of knowledge. In semantic network representation, a semantic relation is also a piece of knowledge unit. In the language system, the unit of the complete expression of the meaning is sentence, which is the mapping or description of the knowledge component in language. The corresponding relationship between the knowledge element and the sentence is discussed as follows.

In language, the basic structure is SPO simple sentences and can be express by the standard model of knowledge representation as shown in Figure 3. Among them, s_1 and o_1 are input, s_2 and o_2 are

output, p is the predicate, s_1 and s_2 are corresponding, represent the before and after states of subject respectively, o_1 and o_2 are corresponding, represent the before and after states of object respectively. The model can be used to express that s_1 do action of p on o_1 , and finally s_1 and o_1 become the s_2 and o_2 .



Fig. 3 Model of simple sentence

The change of things is diverse, and the model showed in Fig. 3 is needed to modify. You can have only one output, for example, "he has a cat", the results of output was "one of his cats"; you can also have multiple outputs, such as "he broke the vase", the result of output is "the vase was broken into a lot of debris". Similarly, you can have only one input, for example, "he laughed"; you can also have more than one input, for example, "Lingling, Doudou, and Miaomiao go to kindergarten". The more complex things change, the more modification needed. In order to facilitate the discussion later, we call the standard model for each piece of knowledge as knowledge component. If the input and output of knowledge component is relatively simple, the sentence corresponding to the knowledge component is simple sentence, otherwise it is complex sentence.

Knowledge Network

The complex knowledge relations can be described in two ways: one is to construct complex knowledge component, that is, using more professional even the mathematical model to describe the input, output and processing of knowledge standard model. Another is to combine some simple knowledge components to form a complex network structure, that is, knowledge networks, which is the key point of our discussion. Obviously, a knowledge component corresponds to a sentence in a natural language that contains only one predicate. If a sentence contains more than one predicate, it needs to be represented by a number of knowledge components.

Expression of Complex Semantic. Complex semantics must be expressed in complex sentences. The complex sentence can be divided into two categories, one is the complex sentence adding modified ingredients and the other is the complex sentence composed with some multiple statements. The former is called complex modification sentence, while the latter is called compound sentences. Both of them expressing multidimensional relationships with one dimension statement and the compound sentence also has the typical hierarchical structure. Generally speaking, knowledge component is relatively simple, in order to express complex semantics, we must use a number of knowledge components to make the knowledge network.

For complex modification sentence, it can be divided into two categories: multivariate modification and composition modification. The multivariate modification means that some objects are added to the subject or object part. For example, "John and Marie play games together".

Multivariate modification does not add the predicate. Composition modification is that add modification to each composition of the SPO. For example, in Chinese, there is attribute, adverbial and complement to modify the sentence. The modifier of subject and object is called attribute, the modifier of predicate (including verb and adjective) is called adverbial, and the supplementary description of predicate is called complement. In Chinese, the words and phrases that can be acted as attribute, adverbial and complement are very much and the rules are very complex, here we do not describe in detail. Here we assumed that the attribute is acted by adjective or phrase, adverbial is acted by adverbs, complement is acted by a predicate or predicate phrases, so we think that there are more predicates in the complex modification sentence, so it can not be expressed by a single knowledge component when we use standard model of knowledge representation.

Traditional Semantic Network Graph. The research of semantic network method is generally carried out around the natural language. So it is not only easy to understand, but also easy to reveal the intrinsic nature of the knowledge system. The most typical one is Simmons's semantic network. In semantic network knowledge representation, knowledge units are generally defined as very simple semantic units and express complex semantics through their combination. The following is the

analysis of the statement: "John sent Marie 999 red roses".

Although the words of statement being analyzed are not much, we can not use a knowledge unit to express it completely. First, we must define a number of knowledge components, and then connect the knowledge components, finally also add some additional relevant knowledge components. For example, Fig. 4 is a

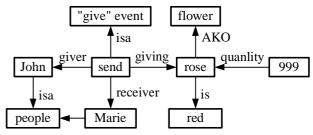


Fig. 4 Simmons's semantic network

result of the using of the semantic network representation method of Simmons to represent the statement. It not only expresses all literal content of the statement but also added some relevant significance, such as "John and Mary are people", "rose is a flower", so that the inference can naturally around the direction of the search and improving solving efficiency.

Adding nodes is very beneficial to reasoning, but it may bring some problems. Because everything is a part of a number of other things, that is, each thing is inextricably linked with others. If all these things exist in the semantic network as node, the number is quite amazing. For example, "Lulu" is a puppy, as a puppy, it has all attributes of dogs, such as has hair, likes meat, listens to the master and so on. As a specific pet, it has a name, a registration certificate, a rabies vaccination certificate and so on. Obviously, to solve complex problems, we must build up all the relationships. But the problem is that people do not know which nodes should be added, because in many cases only when you do inference you can know what nodes are you need. At the same time, the construction of the semantic network becomes very complex, and the number of nodes and connections increases very quickly, which will seriously affect the inference efficiency. Even more serious is that, even if increasing very many semantic relations, it can only reflect the part of the semantic relations, and it still can not meet the requirements when solving the problem. In short, increasing the nodes of the semantic network can increase the chances of success, but when the nodes are too many, it will affect the reasoning, and even lead to the failure of the reasoning and the system.

Parataxis Network Graph. To address the shortcomings of traditional semantic network, Lu Chuan proposed Parataxis network representation [14]. Fig. 5 is the description using parataxis network representation proposed by Lu Chuan of the statement. There are two kinds of networks in the parataxis network: the main parataxis network and the association network. Its characteristic is that the semantics is directly described in the main parataxis network,

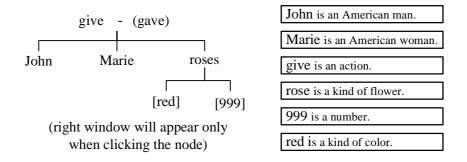


Fig. 5 Lu Chuan's parataxis network graph

and it is not particularly to supplement nodes and connections. We can click the node in the main network to activate and call the association network for the inference. The main parataxis network is dynamic, while the associative network is static, and dynamic network is generated by the application of knowledge, and the static sub network is the knowledge system itself. Expressing complex semantics by parataxis network statement is relatively simple and will not affect the reasoning ability.

To compound sentence, the semantic relationship is more complex, because it must integrate the semantic of multiple sentences. We extend the statement "John sent Marie 999 red roses", to a simple compound sentence: "John sent Marie 999 red roses, and Marie agreed to John's proposal", and then analyze the semantic representation of compound sentence.

If the semantic network of Simmons is used, the expression of the semantic can be accomplished by increasing nodes. As shown in Fig. 6, an increasing of two nodes, one is "agree", and the other is "proposal". Here it has been simplified to "proposal" as a concept, and not be disassembled Semantic by the network representation.

According to Fig. 6, it can be found that in the semantic network of Simmons, the compound sentence can be realized by adding

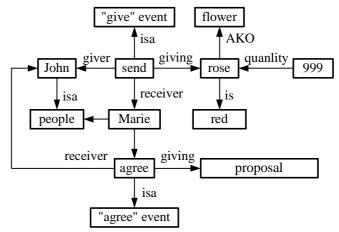


Fig. 6 Traditional semantic network of compound sentence

some nodes to realize the semantic representation, which means that the node of the predicate is indispensable. Obviously, the more complex sentence is, the more complex the semantic network is. Compared the compound sentence being analyzed, it will be found that the semantic network showed in Fig. 6 also has some defects and the most important is that it has no sequence of relations which contained in compound sentence, that is, the first "John send Marie rose", then the "Marie promised John".

If the parataxis network is used, we can get the graph of the compound sentence, as showed in Fig. 7. In Fig. 7, each sentence will generate a sub-parataxis network, and then according to the relationship between the sentences, combine this sub-parataxis network to form a complete Parataxis network.

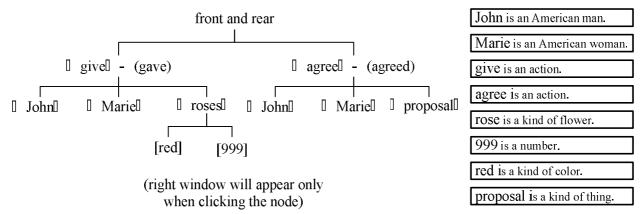


Fig. 7 Parataxis semantic network of compound sentence

Parataxis network method is a kind of hierarchical method, which is used to obtain a hierarchical network connection through semantic analysis. This hierarchical semantic network can reflect the surface semantics of the sentence completely, and it is simple and easy to handle. Compared to Simmons's semantic network representation, there are many reduplicative points, for example, one object appears once in one statement, it will also appear once in the network, and you must ensure the consistency of the object.

Standard Model Semantic Network Graph. For the standard model, all kinds of complex sentences are expressed as the combination of many elements, and the complexity of the complex sentence is related to the number and the application of the predicate. For example, "John sent Marie 999 red roses" can be shown in Fig. 8, while "John sent Marie 999 red roses, Marie promised John proposal" can be shown in Fig. 9.

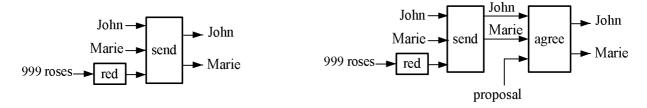


Fig. 8 Complex modification sentence represented with standard model

Fig. 9 Complex compound sentence represented with standard model

From Fig. 8 and Fig. 9, it can be seen that the semantics of the sentence is a combination of a number of knowledge components when the sentence is represented by the standard model. The more complex the sentence is, the more complex the combination form is. But it is worth noted that the using of the standard model of knowledge can also omit some of the content. For example, for predicate "send", the number of input and output of the knowledge component can be large, corresponding to the subject, object, attribute, adverbial, complement and so on in natural language, but only 3 inputs and 2 outputs are listed in Fig. 9 and Fig. 8, because other inputs and outputs are not changed. At the same time, there appeared more than one same content in Fig. 9 and Fig. 8, for example, John appeared 3 times, but the 3 "John" is actually not the same, the left is the input, the middle is "John who sent flowers", and the right is "John who propose marriage to Marie successfully", here in order to simplify we omit it. In fact, in the following discussion of the implementation of the system of knowledge, if the "John" is seen as an object, and the state change of the system is labelled with attribute value, it is beneficial to the reasoning of knowledge system.

Knowledge Network. Through the above analysis, it can be known that whether it is a traditional semantic network, or a standard model, it is bound to form a complex network, which requires the use of semantic networks or knowledge networks to express complex semantics. In turn, a complex knowledge network is mapped to a complex sentence or several sentences.

Many problems will inevitably arise when using multiple statements to express the meaning of a number of issues. It mainly includes two problems, one is the consistency problem, and the other is the complexity problem.

The so-called consistency problem is how to ensure the same meaning expressed by the same words or phrases in concept. A number of the same concepts may appear in a paragraph of text, which is a description of the static and dynamic state of the objective things. Because the objective things are changing, so that the concept of these objective things in different time, different locations are not the same, strictly speaking, it is not accurate to express these changes with the same concept. But if things are to be expressed by different concepts, the concept is not only a sharp expansion, but also very easy to lose the link between the various concepts. The traditional semantic network did not take the changes into consideration. It mainly expresses the relationship of the static, which is easy to build more complex relationship, but in the semantic understanding prone to error, needing to use reasoning to determine the specific semantics. For example, in Fig. 6 it is not clear which is first "Mary promised John's proposal", or "Mary took John's 999 roses". Parataxis network is to distinguish the semantic change of things, but not to consider whether or not the things themselves are changing. For example, Fig. 7 uses structure to express "Mary took John's 999 roses" before "Mary promised John's proposal", but it identifies that "Mary" in two statements is the same person. Standard model can not only accurately record the change sequence of things, but also the changes in the state of things, and it not implies more other semantics like the traditional semantic network. For example, we can infer that "Mary" is a person from Fig. 6, but not from Fig. 9.

The so-called complexity problem is the degree of semantic network to be understood by the man. There are many facts to increase the complexity of semantic network, and some of them are discussed as follows.

- 1) increase of the nodes and lines
- 2) multivariate relationship

Multivariate relationship exists in many places, because the situation that multiple things work together is very common, which must be described by using the multivariate relationship.

3) complex semantic relations

The traditional semantic network is essentially a hypergraph which different from the ordinary graph. It has self loop phenomenon, that is, the arc can be connected end-to-end; there are many arcs between two concepts; or there are many ends in one arc indicating the relationship among multiple concepts, and so on. Parataxis network and standard model are different from the traditional semantic network and they all reflect the change in the state.

4) process relationship

The traditional semantic network is difficult to describe the process relationship, and the sequence relationship can only be determined in inference. In parataxis network, the semantic network is divided into some sub-semantic networks by the segmentation and hierarchicalization at first, and then they are combined into a complete semantic network by the process relationship of time and location. While the standard model itself is a process model, that is, the sequence relationships are contained in the connections of the knowledge components.

From the above analysis, we can draw the following conclusions.

1) knowledge network is logical

Semantic network, such as the traditional semantic network, the parataxis network, the standard model and so on, are the logical description of the semantic or knowledge. In other words, the semantic network is not similar to the physiological structure of the human brain, and it is only a description of the knowledge. Because in essence, nature language is a one-dimensional structure, semantic network is a two-dimensional structure, the human brain is a three-dimensional structure, while the objective things are multi-dimensional. Knowledge network will not be certainly similar in structure for the low dimensional structure is used to describe multi-dimensional objective things.

- 2) knowledge is contained in the knowledge network
- 3) knowledge communication and spread through knowledge network

Knowledge is contained in the knowledge network, which is the manifestation that the agent can contact. Various kinds of knowledge are represented through the knowledge network, which is constructed with all kinds of representation methods and recorded in a variety of media, including text, sound, pictures, video, as well as other artificial or natural objects. Among them, the most typical

representation of the knowledge network is the natural language, for human knowledge and civilization is mainly inherited through natural language.

- 4) knowledge networks are generated dynamically
- 5) knowledge network is the application of knowledge

According to the standard model, it can be seen that the knowledge network is the combination of knowledge components, which aims to solve some problems.

6) knowledge network is partitioned

Knowledge network is the application of knowledge, so it is related to the field of application closely. From the philosophical level, all things in world are connected with each other, but the degree of contacting is different. We can divide the knowledge network into different blocks according to the different contacting degree of things, so that we can effectively improve the efficiency of knowledge inference.

7) the knowledge network needs to be systematic

Because a single knowledge component represented with the standard model records the process of objective things change, so the knowledge networks, which is the combination of knowledge components, will record more knowledge of the objective things change. In addition, the decision making knowledge is only different in using purposes, so it can also be recorded by the knowledge network. In the word, the standard model of knowledge representation and knowledge network constructed by this model can not only be used to represent the narrative knowledge, but also the process knowledge and decision knowledge.

Conclusions

The most basic unit of knowledge representation is the knowledge component, which corresponds to a sentence in a natural language. The knowledge is not isolated, so the performance to express knowledge with a single knowledge component is very limited. In order to express complex knowledge, it is necessary to construct a complex knowledge component, or combine several knowledge components to form a knowledge network. Knowledge network is the basis for the exchange of the agents, that is to say, the knowledge networks of the two agents must be consistent with each other when communicating. The knowledge network is not the representation of the physical structure of the human brain, but the logical description of the knowledge. In many cases, the knowledge network is related to the knowledge solving and generated dynamically. Using static knowledge network to construct the knowledge system is unable to avoid the explosion of knowledge network.

References

- [1] Shi Zhongzhi. Advanced Artificial Intelligence [M]. Beijing: Science Press, 2011.7
- [2] Newell A, Simon H. A. A Program That Simulates Human Thought [M]. New York, USA: McCraw-Hill, 1963
- [3] Feigenbaum E. A., Feldman J. Computers and Thought [M]. New York, USA: McGraw-Hill, 1963
- [4] Simon H. A. The Sciences of Artificial [M]. Cambridge, USA: MIT Press, 1969
- [5] Newell A., Simon H. A. Human Problem Solving [M]. Englewood Cliffs, USA: Prentice-Hall, 1972
- [6] Minsky M. L. The Society of Mind [M]. New York, USA: Simon and Schuster, 1986
- [7] Zhang Pan, Wang Bo, Qing Xiaoxia. Integrated application of multiple knowledge representation in expert system [J]. microcomputer application, , 2004,20(6):4-6

- [8] Wu Chaohui, Chen Huajun. The Semantic Grid: Fundamental Issues, Methodologies and Applications [M], Hangzhou: Zhejiang University Press, 2008.2
- [9] Lin Chuang. Application of Petri nets to logical inference of horn clauses [J]. Journal of software, 1993, 4(4): 32 37
- [10] Steven Walczak, Knowledge acquisition and Knowledge representation with class: the object-oriented paradigm [J]. Expert Systems with Applications, 1998,15:235-244
- [11] W. L. Xu, L. Kuhnert, K. Foster, et al, Object-oriented knowledge representation and discovery of human chewing behaviours [J]. Engineering Applications of Artificial Intelligence, 2007,20 :1000-1012
- [12] Rukmi Sari Hartati, M. E.El-Hawary. A summary of application of Hopfield neural network to economic load dispatch [J]. 2000 Canadian Conference on Electrical and Computer Engineering, March 7-10, 2000
- [13] Books R A. Intelligence without Representation [J]. Artificial Intelligence, 1991, 47(1/2/3): 139-160
- [14] Lu Chuan. Linguistics for Knowledge Engineering [M]. Beijing: Tsinghua Press, 2010.6