A Semantic-Aware Methodology Adapt to e-Learning Environment

Ying-Hong Wang[1], Chu-Chi Huang[2], and Wen-Nan Wang [3,*]

Department of Computer Science & Information Engineering

Tamkang University, Tamshui, Taipei County, Taiwan

[1]inhon@mail.tku.edu.tw, [2]895410131@s95.tku.edu.tw, and [3]892190033@s92.tku.edu.tw

Received 15 June 2006; Revised 1 August 2006 ; Accepted 15 September 2006

Abstract. Automatic assistant-learning and self-learning mechanisms are the primary objective in today's digital teaching environment. Previous research has focused on the questions-and-answers system of digital leaning field and addressed several aspects of how to possess the following characteristics: (1) Understand questioners' questions which are in nature language form, (2) Arise the accuracy of searching result, and (3) Establish automatic learning mechanism. In this study we develop the English questions-and-answers system in the distance learning platform and purpose a Semantic-Aware method to answer questions which are in nature language form effectively. This research use Link Grammar and combine Ontology and WordNet to set up the Semantic-Aware questions-and-answers system in distance learning environment. Furthermore, we choose Data Structure as the targeted course and design Ontology structure of Algorithm. This study also generalizes and analyzes the sentence structure of English Interrogative Sentence. Finally, and most important, this research sift the structure from Interrogative Sentence and address several corresponding Semantic-Aware Algorithm.

Keywords: Grammar-Based, Semantic Aware, Distance Learning, e-Learning

1 Introduction

Language is the tool for human to communicate and exchange signals. Human can record and spread knowledge through language and words.

Recently, Automatic assistant-learning and self-learning mechanisms are the primary objective in today's digital teaching environment.

Famous search engine such as Google and Yahoo use keyword search, statistics and ranking mechanisms. There are some drawbacks: (1) need to use keyword to tie in with special logic symbols, cannot use natural language to search. (2) Return too many search results. (3) There is much meaning for each keyword. We will get lots useless information from keyword search. (4) We may not get information which we really need because the different keywords represent the same meaning in the knowledge database.

In order to solve the above-mentioned problems, this paper use Link Grammar Parser, cooperate with Ontology and combine with WordNet to set up the Semantic-Aware questions-and-answers system.

2 Related Works

2.1 Link Grammar

Link grammar [1] which developed by Carnegie Mellon University is a graphical grammar analyzing tool. This system can produce all grammar linkage from English sentence which users input and determine the sentence correctness thought the linking result.

Figure 1 will show what Link Grammar is. First, input an English sentence into this system. For example, what is the algorithm of PUSH in the stack? We can get the system output (Linkages in Fig. 1). Each word has some curves and each curve has one label on it. The curve and label is a Link which to express a kind of linkage. After analyze and parse the sentence though Link Grammar, we will get a lots of Linkages.

* Correspondence Author

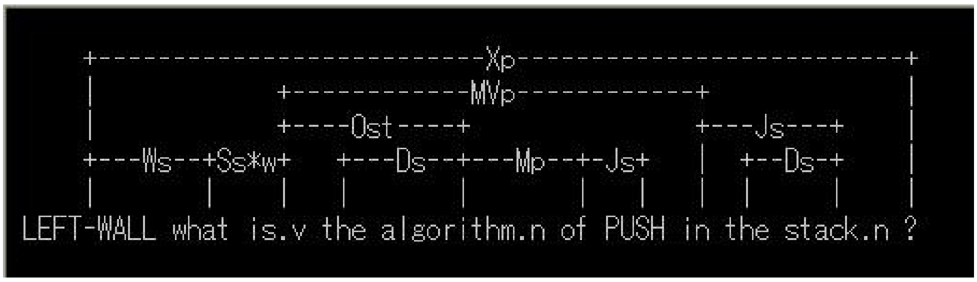

Fig. 1. Link Grammar output sample

It adds the Phrase Parser component after Link Grammar version 4. It recombines the sentence information analyzed by Link Grammar Parser to form phrase structure, such as noun phrase, verb phrase, and preposition phrase. For example, what is the algorithm of PUSH in the stack? We can get the output result showing as Fig. 2.

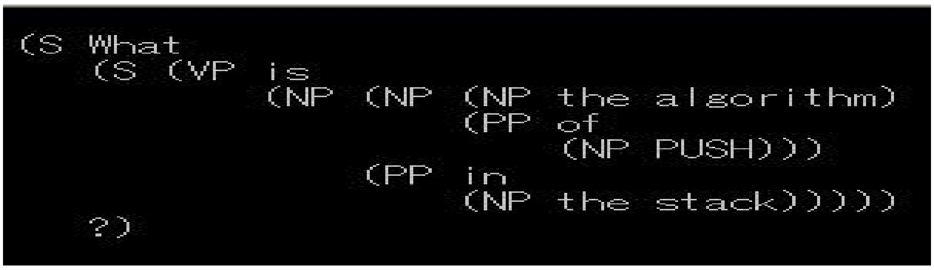

Fig. 2. Link Grammar output sample though Phrase Parser component

2.2 WordNet

WordNet [2] originated from Cognitive Science Laboratory in Princeton University. It is a vocabulary reference system designed by researchers who inspired by psychology theory. WordNet processed the first level classification according to part of speech (POS). Then base on different meaning of words and expressions, it forms several Synset. Each Synset symbolizes one vocabulary and takes down other words and expression with the same meaning.

Fig. 3 represents the Synset structure when we search the word, stack, in WordNet.

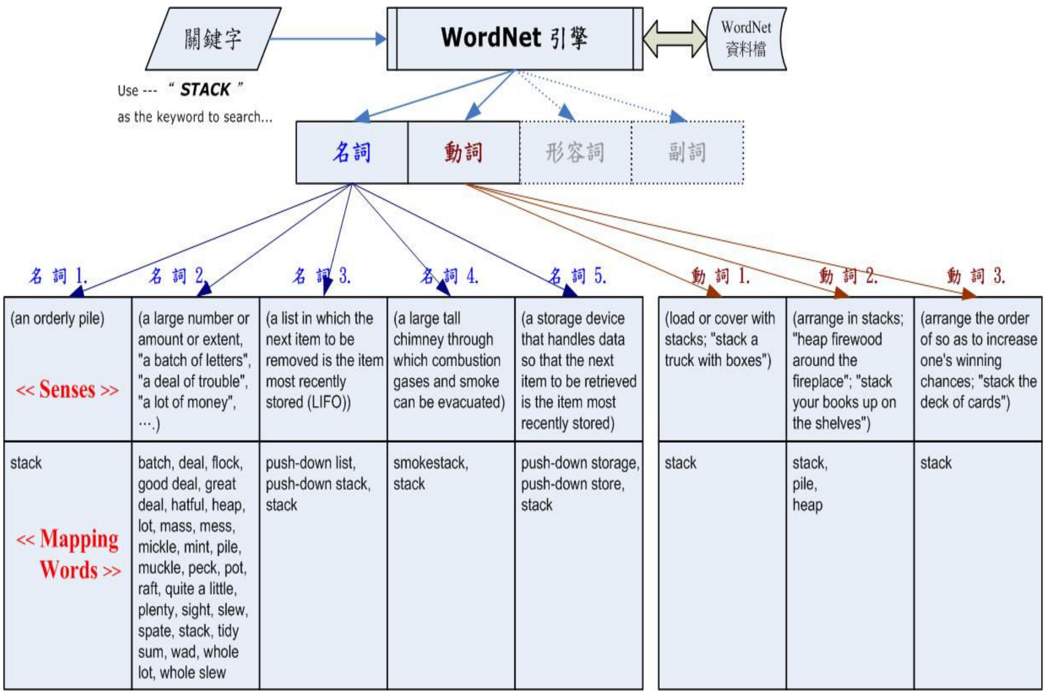

Fig. 3. A Synset structure sample of WordNet.

2.3 Ontology

People already explores some ways to show up the meaning of data thought the data processing in Information technology field although human explore the meaning of knowledge from Hellenistic Age to the present age.

There are many tools to support Ontology, such as DAML, KAON, etc. Fig. 4 is the example using KAON. It represents one Person component and attributes of Phone, title, email, name, and fax. Furthermore, we observe communicators. Person may have a communicator, and it can also be passive communicator. From here we can see the relationship between components is directional.

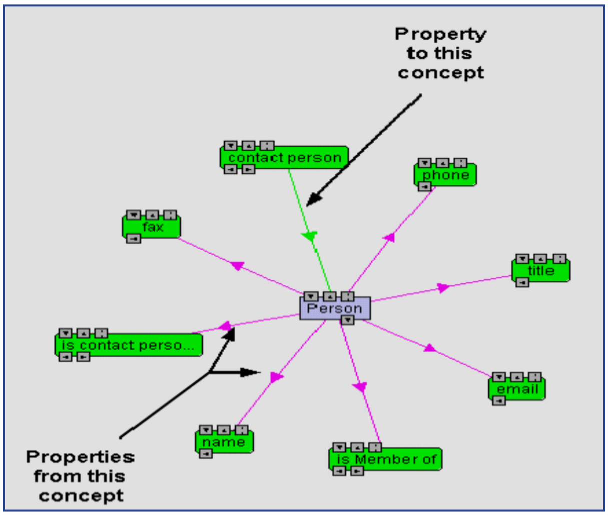

Fig. 4. An example of segment Ontology thought KAON tool.

2.4 Other Relate Works

Previous research [5] addressed the automatic QA exploring system that using Latent Semantic Analysis Latent Semantic Analysis (LSA) as the core technology. It employs the statistics method to calculate relation matrix and uses ranking to get the possible solution.

In [6], Daniel and James discussed the parallel processing method of SNAP (Semantic Network Array Processor). Using Concept Sequence Layer, Semantic Concept Hierarchy, Syntactic Layer, and Instance Layer to proceed the semantic knowledge.

Representing Meaning. Some article [7] pointed the importance of Representing Meaning Method. It is the communication standard which uses words, symbols and graphics to express the meaning of language. Using Representing Meaning Method can help automatic check, using variables and guide reasoning mechanism.

Fig. 5 lists four expression methods. (1) First Order Predicate uses words, operators and symbols, (2) Semantic Network uses words and graphics, (3) Conceptual Dependency and Frame-based Representation uses word structure expression method. These methods express the Semantic idea about lectors have cars.

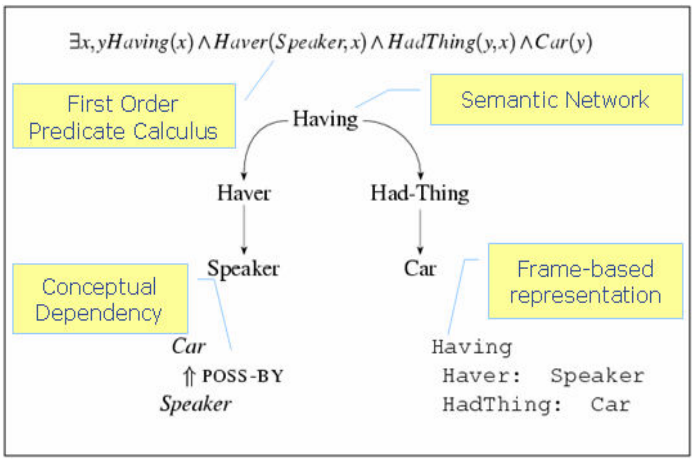

Fig. 5. Representing Meaning methods

3 System Architecture

3.1 System Functional Blocks Diagram

Fig. 6 is the System Functional blocks Diagram. The following illustrates each function of block.

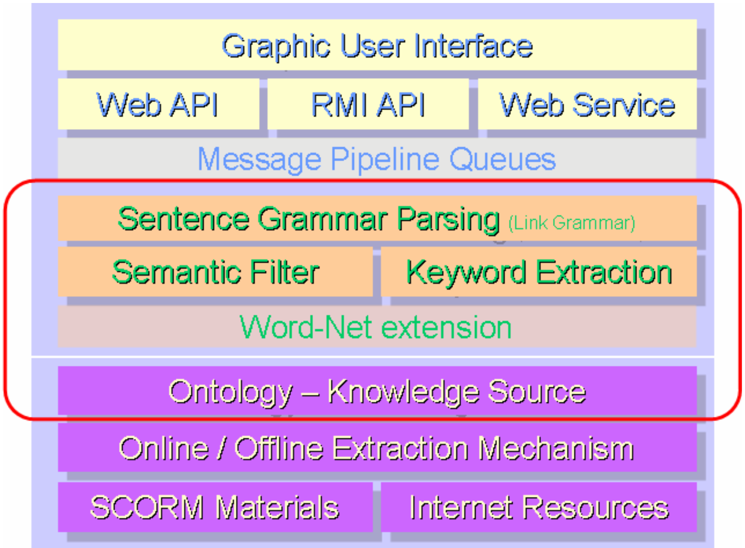

Fig. 6. System Functional Blocks

(a) User Interface:

Graphic User Interface: Provide Web base application an convenient operation user graphic interface.

Web API: Provide supporting other programs of distance learning system.

RMI API: Provide supporting other programs of distance learning system.

Web Service: Provide supporting famous across platform application service.

(b) Message Pipeline Queues: In order to satisfy a large number of requests though user input and every request can be processed effec-tively by system, we design the Message Pipeline Queues. The system stores each user's requests immediately and delivers the request to system to process sequentially.

(c) Intelligent Semantic Engine: This method focus on this part.

Sentence Grammar Parser: Using Sentence Grammar Parser of Link Grammar to process most sentence grammar check, distinguish the relationship between words and sentence grammar structure.

Semantic Filter: This engine cooperates with syntax analyzer to proceed keyword extraction. It applies to wrong grammar check.

Keyword Extraction: This method cooperates with syntax analyzer to proceed keyword extraction. It applies to wrong grammar check.

(d) Knowledge Extraction Engine:

Word-Net Extension: It carries out explanation and extension of the vocabulary to be aimed at the keywords of existing Ontology.

Ontology-Knowledge Source: Automatically established the specific knowledge structure fit in this research based on different courses.

Online / Offline Extraction Mechanism: When Ontology cannot provide enough information to users, system can provide information immediately though data Extraction mechanism. Moreover, it can proceed data collection job in away time.

SCORM Materials: he source of Knowledge Extraction includes SCORM (Sharable Course Object Reference Model). It is the e-learning standard material developed by XML.

Internet Resources: The source of Knowledge Extraction includes various internet data.

3.2 System Flow

After the user input Question Sentence in the Graphic User Interface (or API、Web Service), system will store this question into the Message Pipeline Queues (Please the Fig. 7).

Then system will get the Sentence from Message Pipeline Queues and deliver to Sentence Grammar Parser of Intelligent Semantic Engine to proceed to analyze the English Question Sentence.

If the grammar structure is correct after analyzing, system will pass the analyzed sentence structure information to Semantic Filter. Semantic Filter will extract useful Keyword Pair from sentences (according to the methods from Semantic Filter). If the grammar structure is wrong, system should use Keyword Extraction to extract useful keyword. We use the structure which Link Grammar can recognize the data as the base to deal with Keyword Pair.

Then we use the data-matching mechanism to match the extractive keyword and the Ontology which extended by WordNet to find out the answer of the question.

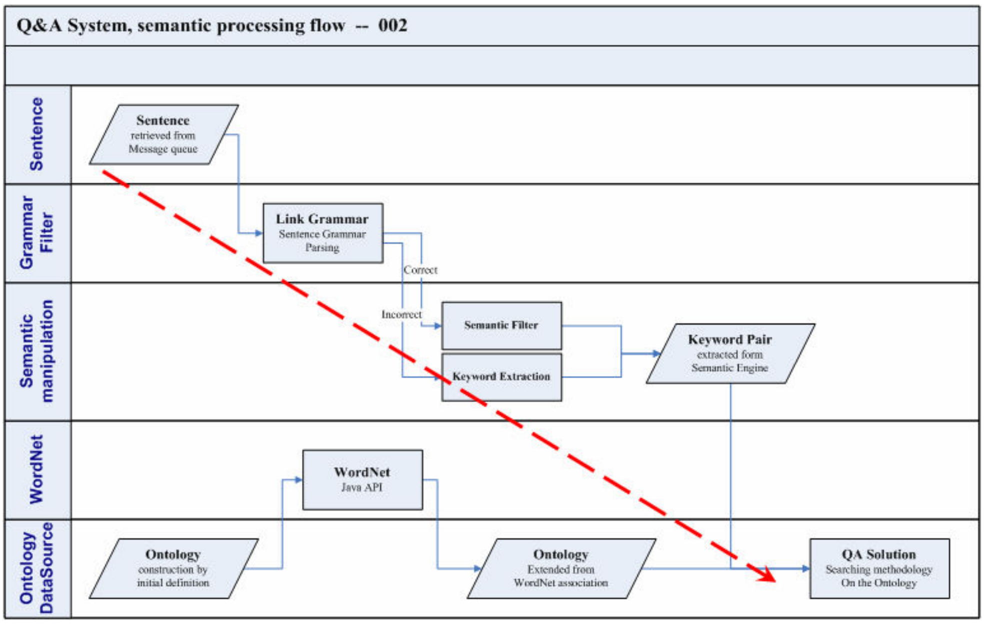

Fig. 7. System process flow

3.3 Definitions of Methodology Structure

This section we will introduce the definitions of Methodology Structure, such as related elements, linkage, and Semantic expression method.

Basic Elements:

Interrogative Sentence Pattern: After user input the interrogative sentence, this method will diagnose the basic sentence. We use the phrase string from Link Grammar and Pattern Matching Table to analyze the sentence pattern of the Interrogative Sentence and expect answer (Please see Table 1).

Table 1. Pattern Matching Table

QHeading	QN1	QN2	QN3	Tar-get	MType
What	is	{NP}		{NP}	1
What	are	{NP}		{NP}	1
How many {NP}	are	there	{PP}	{NP}	2
Which {NP}	is	there	{PP}	{NP}	4
Which {NP}	are	there	{PP}	{NP}	4
{VP}	{NP}	{PP}		{Y/N}	9

Word: Analyzing methods which deal with natural language mostly analyze the relationship between words. Link Grammar Parser use linkage to link two different words in the sentence, and then finish the analysis and linkage of the whole sentence.

Phrase: Besides the meaning of the word, we are interested in the phrase relationship structure from Phrase Parser module of Link Grammar Parser (please see Fig. 2). Take one sentence for example (A stack is a LIFO list). Through Phrase Parser, we will get the following components.

```
(S (NP A stack)
  (VP is
     (NP a LIFO list))
  .)
```

(S …) represents the clause; (NP A stack) is the noun phrase; (VP is ...) is the verb phrase; (NP a LIFO list) is the noun phrase.

Element: The above-mentioned word and Phrase.

Phrase Pair: The word set which are formed from two or more elements with particular relate linking. It also represents the meaning of some context.

Semantic Linking Unit:

We define two or more elements with specific relate linkage to be a Semantic Linking Unit, Please see the Fig. 8.

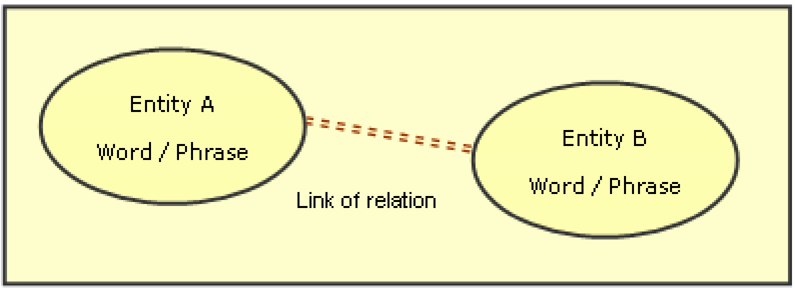

Fig. 8. Semantic Linking Unit chart

In Fig. 8, Link of relation, is the linkage between elements. We had defined an inclusive linkage which stands for modifiers (of, in or on). It represents as the straight line with arrowhead. Please see the Fig. 9.

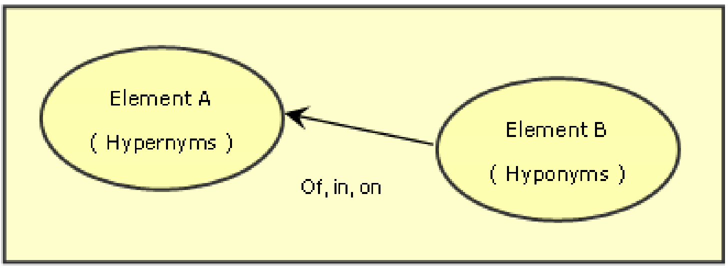

Fig. 9. Semantic Linking Unit with inclusive relationship

Element B includes element A. In other words, element B is the subset of element A. The possible preposition keyword is in, on, of, and so on. The form of the phrase is element B of element A, element B in element A or element B on element A.

Semantic Synonymous Unit:

As we mentioned before, one word may have different meaning. So we define Semantic Synonymous Unit which shows up in Fig. 10.

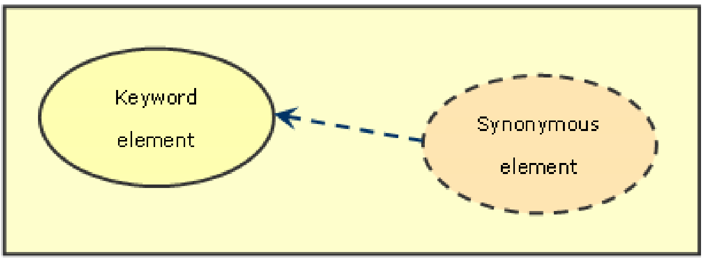

Fig. 10. Semantic Synonymous Unit chart

Using this synonymous words module, we expand the vocabulary from WordNet in Ontology and perform as Synonymous element. When Synonymous element is found in this system, this system will guide to Keyword element according to the dotted line in Figure 10. In other word, all existing Synonymous elements can find comparative Keyword elements.

Semantic Tree:

In order to get and express the meaning of English sentences, we define the Semantic Tree. We use the above-mentioned basic components to compose the Semantic Tree. We give an example, the algorithm of PUSH operation (writing) in C, represents in Fig. 11.

Ontology of Data Structure:

We choose data structure from many distance learning courses as the target course to set up Ontology. We use Concept Layer and Instance Layer of Ontology to express a whole Ontology structure.

Concept Layer of Ontology: As Fig. 12, The Concept Layer of Ontology is our system knowledge boundary for the specific domain knowledge, Data Structure. In other word, the system cannot understand the knowledge go beyond this Ontology.

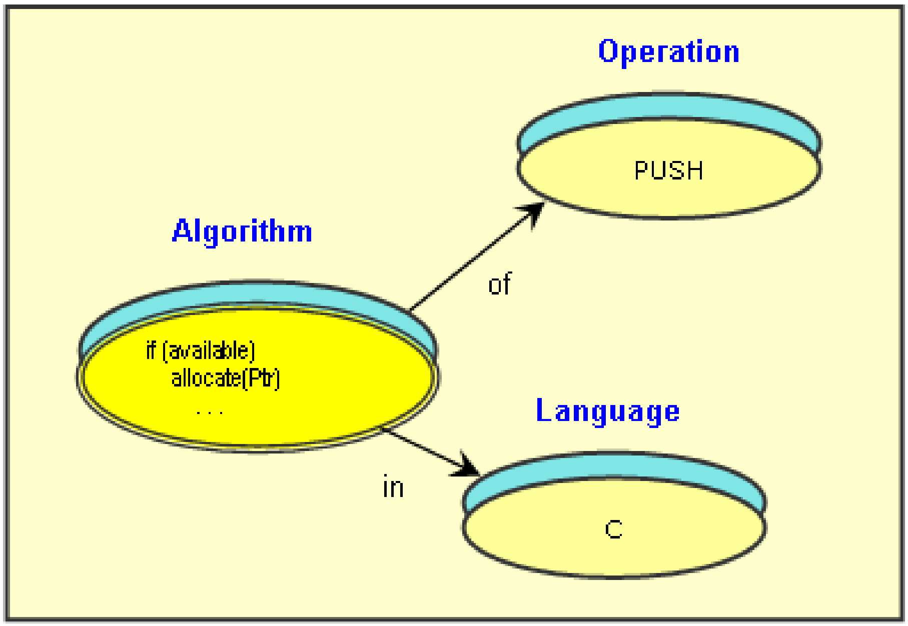

Fig. 11. Semantic Tree example

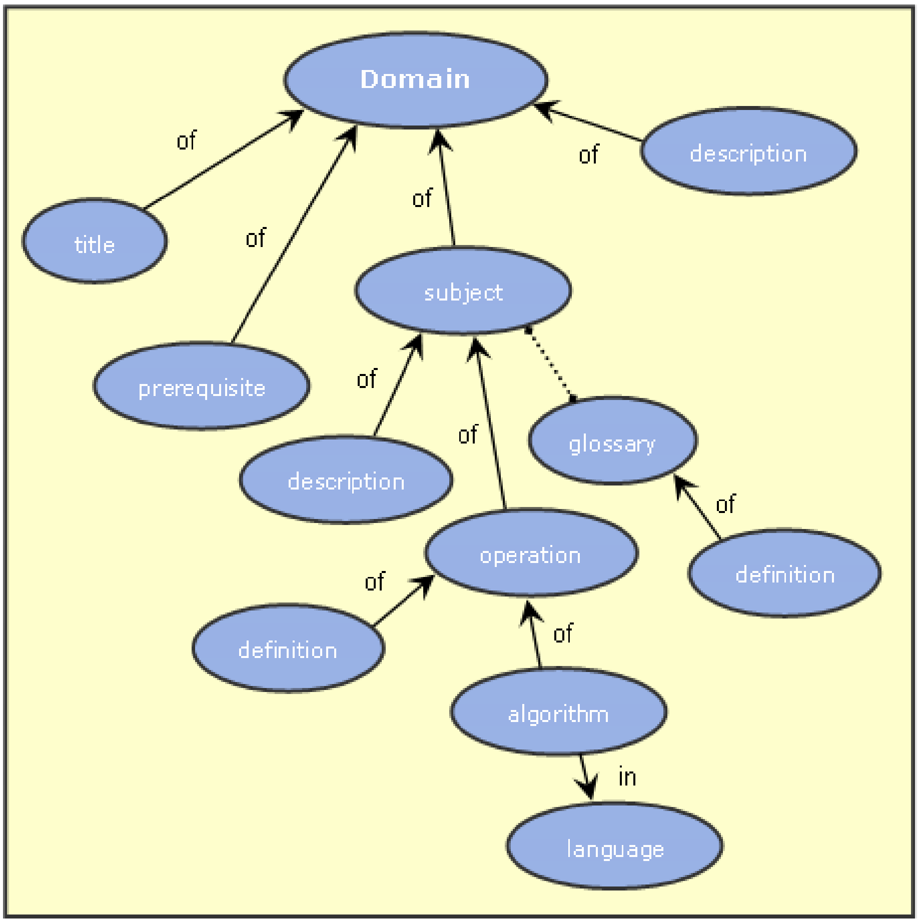

Fig. 12. Concept Layer of DS Ontology

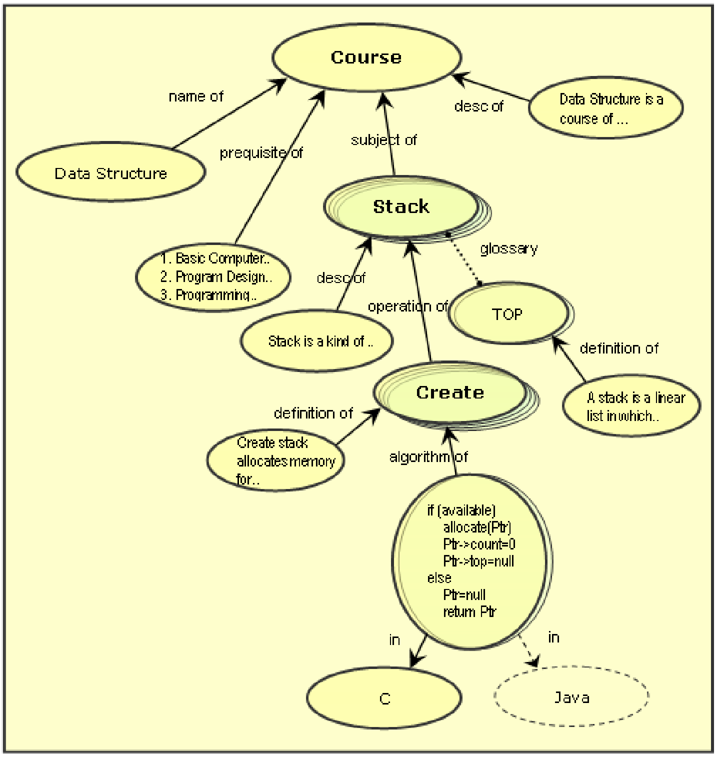

Fig. 13. Instance Layer of DS Ontology

Instance Layer of Ontology: Fig. 13 represents the corresponding instance data. There are two data type of Instances. (1) Searching keyword, such as Stack and Top;(2) No-searching data, such as "Stack is a kind of …".

DB Schema:

The DB Schema in Fig. 14 represents all kind of components, Semantic Linking Unit, Semantic Linking Unit, the Ontology framework of data structure, and the Ontology instances of data structure.

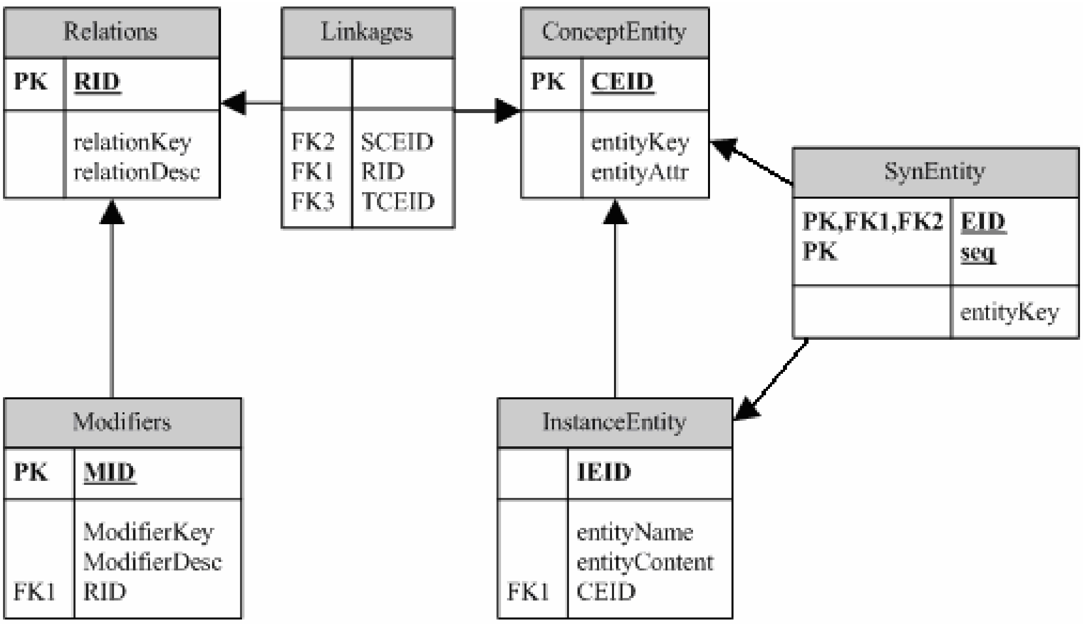

Fig. 14. The DB Schema for Semantic-Aware Ontology

4 Semantic Aware Algorithm

The following is the brief description of our method in this article.

Phrase Pair Acquirement: Use Phrase Parser of Link Grammar and set the parameter to be equal to 2 (It is suit for computer processing).

Interrogative Sentence Pattern Match: Analyzing the output string and contrast with Pattern Matching Table of Interrogative Sentence, we will find the sentence pair and target phrase set.

Target Element Extraction: Find the core none phrase from the phrase set and extract the target element.

Semantic Tree Construction: Beginning with the target element, set up the element and linkage of phrase set to form the Interrogative Sentence Semantic Tree.

Acquire and Answer with Ontology: The above-mentioned semantic tree needs to contrast with Ontology. Follow the existing linkage and element to rebuild the partial Sentence Semantic Tree to find the solution.

4.1 Phrase Pair Acquirement

First set up constituents and use Link Grammar to proceed the Interrogative Sentence. Take the following Interrogative Sentence for example. “What is the algorithm of PUSH in the stack?” After the processing from Link Grammar, we will get the output string (Please see Fig. 15).

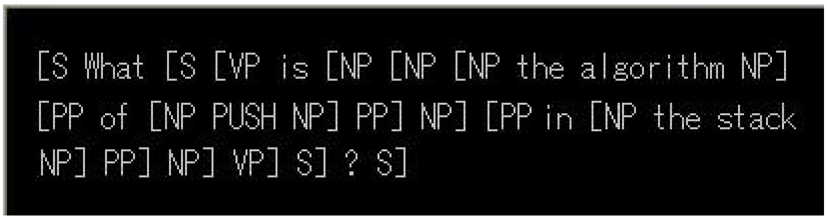

Fig. 15. The output string after parsing with Link Grammar

Then we use sting substitution. <S> replaces [S, <VP> replaces [VP, </S> replaces VP], and so on. At last, we will get an XML string and use DOM to analyze object. We can operate and read the content of each node of XML documents.

4.2 Interrogative Sentence Pattern Match

Analyze the XML object in previous step and contrast with Pattern Matching Table (Please see Table 1). After the analyzing and comparison, we can know the Interrogative Sentence Pattern is “What is” sentence. And the solution of the target is followed by the non phrase (Please see Fig. 16).

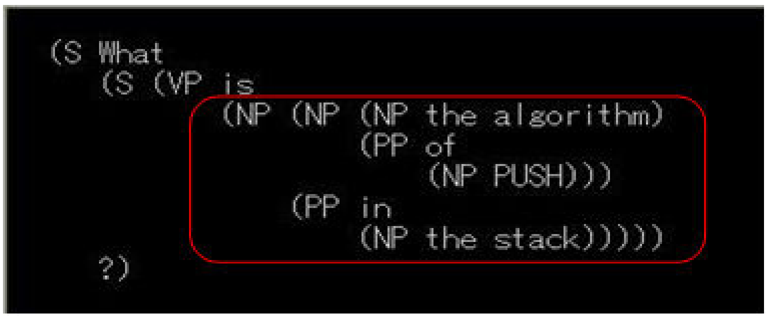

Fig. 16. The Target Element (NP) produced after Pattern Matching

In other words, the word set shuts in a red pen is the target phrase set from Interrogative Sentence Pattern Matching.

4.3 Target Element Extraction

In previous step, we know none phrase is the target phrase of the Interrogative Sentence..
We get the phrase set from the previous step (Please see figure 16). The target element is “(NP the algorithm)”. Next procedure, we will start to set up the Interrogative Sentence Target Semantic Tree upon this element.

4.4 Semantic Tree Construction

First, we use a starting cycle node to present the phrase, “(NP the algorithm)”, as Fig. 17.

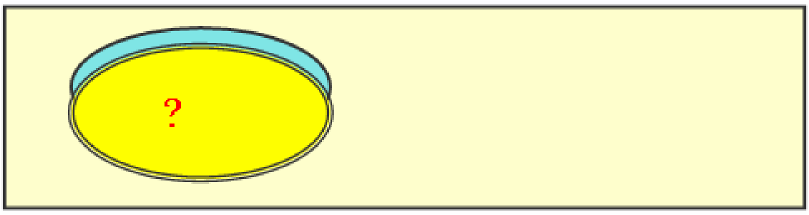

Fig. 17. Semantic Tree Construction - Beginning with “(NP the algorithm)”

Then, we use Fig. 18 to present the phrase pair, (NP (NP the algorithm) (PP of (NP Push))).

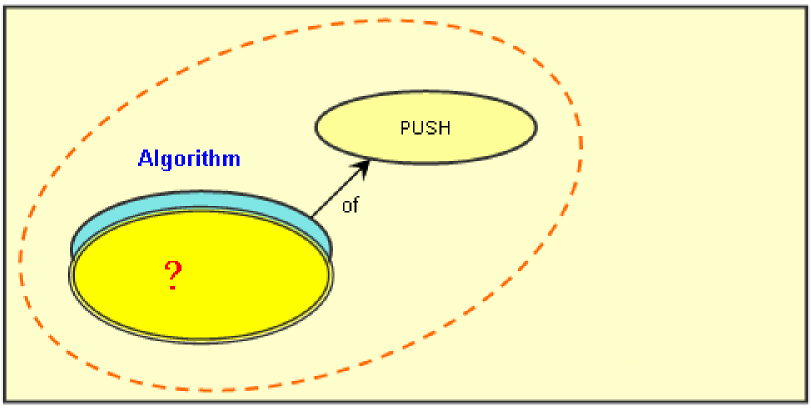

Fig. 18. Semantic Tree Construction - link to the next, second element

Continue to set up the Semantic Tree (NP (NP (NP the algorithm) (PP of (NP Push))) (PP in (NP the stack))), please see Fig. 19.

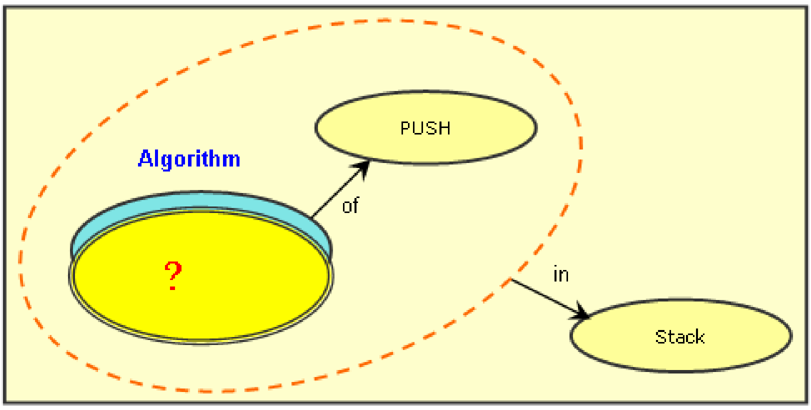

Fig. 19. Semantic Tree Construction - link to the next, third element

In Fig. 19, the element is formed from the algorithm of PUSH and is included in stack. The interrogative Sentence target Semantic Tree is show in Fig. 20.

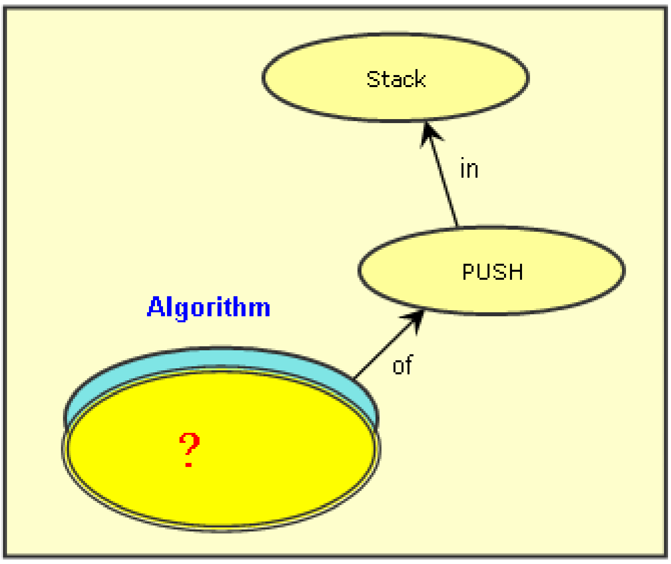

Fig. 20. Semantic Tree Construction – the target Semantic Tree

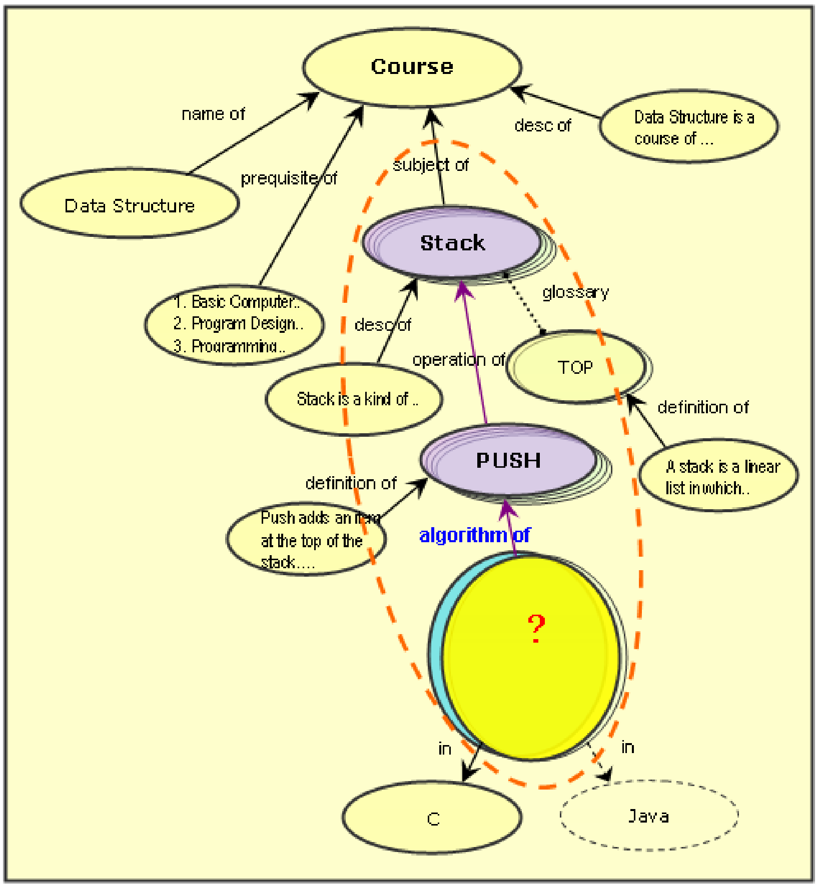

Fig. 21. Acquire answer with DS Ontology –- fulfillment of contrast

4.5 Acquire Answer with Ontology

The Semantic Tree we get in previous section contrasts with Data Structure Ontology. If we can find the same Semantic Tree from Ontology, The element data is the solution.

Different to previous section, we will find any Semantic Tree corresponds to the interrogative sentence Semantic Tree in the previous section in Data Structure Ontology.

First of all, we will find the beginning contrast element "the algorithm" and start to compare one by one which fellows the steps in the previous section. Finally, we find the same interrogative sentence Semantic Tree in Data Structure Ontology (Fig. 21).

So we can get the answer of this interrogative sentence (What is the algorithm of PUSH in the stack?). The answer is the data of this target element ((if (stack full) success = false …)).

5 Conclusion and future work

We will present the major contribution, constraint and future research in the following sections.

5.1 Conclusion

Our major contributions are described as follow:

- This paper presents the QA system architecture for the automatic learning platform and also uses Link Grammar as the basic method of Semantic cognition.
- We can input natural language (English) to the QA system. The accuracy of an answer is high. It is not alike searching engine to return a lot of useless reference resources.
- Ontology is expanded by related keyword. Even we do not input exactly the same keyword; we can search for what we want effectively.
- The Pattern Matching Table of interrogative sentence can expand more sentence pattern.

The system constraints are showing behind:

- The structure of Ontology and the design of relationship are partial to be a static state.
- The relationship between elements is containing relation. It will affect system dealing with the quantity of English interrogative sentence.
- This method need to satisfy all elements in the Semantic Tree and existing relationship. Different from keyword search of search engine, it may have the chance not to find the answer.
- The Semantic cognition method cannot produce ratiocination.

5.2 Future Work

Future research can plan and analyze the relation type between elements deeply. The Semantic cognition method can deal with more interrogative sentence type. For example, "How to get the data from the stack?". The target answer should be the data of algorithm. The meaning of "get the data" implies that the element corresponding to algorithm should be POP. Although we have the Pattern matching table to define the interrogative sentence type, it cannot put "get" (general verb) into the Pattern matching table. We need additional elements to modify or expand the Semantic cognition method in this paper.

Furthermore, the Pattern matching table needs to be analyzed and expanded in the future to adapt to deal with more interrogative sentence type. Finally, how to let Ontology learn automatically, combine and link the data correctly is worth for future research.

6 Acknowledgement

This research was funded by the National Science Council of the Republic of China under the Contracts No: NSC-NSC 94-2520-S-032-003 and NSC 95-2520-S-032-001. The authors also appreciate the instrument support of TamKang Software Engineering Group at TamKang University, Taiwan, ROC.

References

[1] D. Sleator, D. Temperley, and J. Lafferty, “Link Grammar,” *http://www.link.cs.cmu.edu/link/*

[2] G. Miller, “WordNet, a lexical database for the English language,” *http://www.cogsci.princeton.edu/~wn/*

[3] M. Denny, “Ontology Tools Survey, Revisited,” *http://www.xml.com/pub/a/2004/07/14/onto.html*

[4] FZI and AIFB, “KAON is an open-source ontology management infrastructure targeted for business applications,” *http://kaon.semanticweb.org*

[5] Y.-F. Chen, *Automatic Concept-Based Answer-Finding System*, Master's thesis in National Chiao Tung University, 2003

[6] M. Chung and D. Moldovan, “Parallel Memory-Based Parsing on SNAP,” *Proceedings of Seventh International on Parallel Processing Symposium*, 1993, pp. 680-684.

[7] D. Jurafsky and J. H. Martin. “Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition,” *Prentice Hall*, 2000