An Intelligent Semantic Agent for e-Learning Message Communication

Ying-Hong Wang Wen-Nan Wang Yi-Hsiang Yen
Department of Computer Science and Information Engineering
Tamkang University, Tamsui, Taipei 251, Taiwan
inhon@mail.tku.edu.tw 892190033@s92.tku.edu.tw 692192239@s92.tku.edu.tw

Abstract

This paper presents an English chat room system in which students discuss course contents and ask questions to and receive from teachers and other students. A novel mechanism is also proposed that does not require that supervisors constantly remain online to handle queries. The mechanism contains an agent that detects syntax errors in sentences written by the online users and also checks the semantics of each sentence. Learners may make a semantic level mistake imply that they do not understanding the course topic or particular issue. The agent can thus offer recommendations to the user and analyze the data of the learner corpus. When users query the system or others, this system will attempt to find the answers from the knowledge ontology or learner corpus that is stored in the records of previous user comments. If sufficient numbers of QA pairs can be obtained, the FAQ can be provided for the learners.

Keywords: Semantic Agent, Link Grammar, and E-learning

1. Introduction

In these years, Distance learning is the hot issue in computer science. Online learning through the web has become popular in the decade [1]. By allowing the use of text, graphics, audio, animation and video, WWW has become an ideal media for distance learning [2]. With rapid, technological advance and the pervasiveness of the Internet, Distance Learning allows students to enroll courses on the Internet and acquire new knowledge. With the aid of voice, video and data on the Internet, people can learn at anytime and anywhere. Therefore, education on the net will become popular in the decade [3 - 4].

However, it is difficult that instructors need to keep watch the activities and behaviors of learners from Internet. For examples, instructors may want to know:

- Do learners understand the context of teaching?
- Do learners talk about the indicated issues that instructors set up?
- Do learners really understand the issues of the course?

Therefore, it is quite useful if there are some automatic supervisors of these systems. The supervisors will monitor the talking messages and correct the mistakes or problems of learners on grammar and understanding of courses. It helps people obtain education credits or complete training they might otherwise not have done without a central classroom. Thus, people can teach or learn at anywhere and any time.

However, some instructors cannot control the learners' activities in network, Instructors cannot on-line forever, Instructors don't keep the frequently answers and questions (FAQ) or problems of learners cannot create extra tutorial to assist learners, and Learners cannot learn from ancestral learners and other learners.

The goal of this paper is to build up a Semantics-based Intelligent Agents system that provides the supervision and learning-assistance over textual message communication systems. This system applies Learning_Angel to provide the monitoring and on-line error checking to solve the instructor-off problems and collect/analyze the statistic frequently mistakes and problems and build up the learning corpora.

2. Theoretical Backgrounds

2.1 Link grammar

Link grammar is an English grammar parser system that was proposed by The School of Computer Science of Carnegie Mellon University (CMU). Link grammar is a context-free formula to

Proceedings of the 19th International Conference on Advanced Information Networking and Applications (AINA'05)
1550-445X/05 $20.00 © 2005 IEEE

describe natural language [5]. Link grammar consists of a set of words, which are the terminal symbols of the grammar, and each has a linking requirement. The linking requirements of each word are gathered in a dictionary. The linkage can be perceived as a graph and the words can be treated as vertices, which are connected by labeled arcs. The labeled arcs that connect word to others on either their left or right are links. A valid parse is called a linkage. A sequence of words is a sentence defined by grammar, if links can be established among the words so as to satisfy the formula of each word. It includes the following meta-rules: **1.Planarity:** Links are drawn above the sentence and do not cross. **2.Connectivity:** Links connect all the words in the sequence. **3.Ordering:** When the connectors of a formula are traversed from left to right, the words to which they connect proceed from near to far. Namely, consider a word, and consider two links connecting that word to the word on its left. The link connecting the closest word must satisfy a connector that appears to the left of that connector in the other word. **4.Exclusion:** No two links may connect the same pair of words.

The use of a formula to specify a link grammar dictionary is convenient for creating natural language grammars, however it is cumbersome for mathematical analysis there of, and as well as in describing algorithms to parse link grammar. An alternate method of expressing link grammar is known as disjunctive form, in which each word has an associated set of disjuncts.

2.2 Ontology

Ontology is a popular research issues in various communities such as knowledge engineering, natural language processing, intelligent information integration and knowledge management. It provides a shared and common understanding of a domain that can be communicated between people and heterogeneous and widely spread application systems. Ontology has been developed in AI to facilitate knowledge sharing and reuse. Ontology provides an explicit conceptualization that describes the semantics of the data. [7]

Depending on their generality level, different types of ontology may be identified that fulfill different roles such as:

- Domain ontology: capture the knowledge valid for a particular type of domain.
- Metadata ontology: like Dublin Core [6] provides a vocabulary for describing the content of on-line information sources.
- Generic or common sense ontology: aim at capturing general knowledge about the world.
- Representational ontology: do not commit themselves to any particular domain.

3. System Architecture

The follows figure 1 shows the system workflow and functions. The left side diagram shows the Augmentative Chat Room components and the flow of Chat Room Supervision and the flow of Chat Room Grading, and the Ontology Definition process. The right side is the database that includes Distance Learning Ontology, Learner Corpus Database, and User Profile Database.

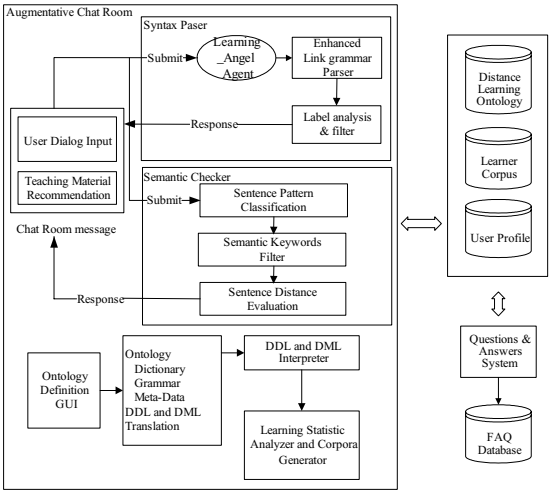

Fig. 1 The system architecture and operation flow

When the augmentative chat room learner submits the sentence to the Learning_Angel Agent, Learning_Angel will forward the message to Link Grammar Parser to query ontology to get the tag about the input sentence. Link Grammar Parser will send the tags and the sentence to Label analysis & filter. Label analysis & filter will find if there are wrong linkages occurred. If the input message has grammar error, the Label analysis & filter will also find the correctly sentence from Learner Corpus database and send to the learner user interface to correct the grammar error which learner made.

Link Grammar Parser will send the tags and the sentence to Label analysis & filter. Label analysis & filter analyze the links of input words sequence to see if the links of words satisfy the meta-rules such like planarity, connectivity, ordering, and exclusion. If the

input words sequences don't satisfy the meat-rules, the analysis & filter will records to Learning Corpus, record the score, and send the correction message to the learner interface.

When the learner submits the sentence to the Learner_Angel Agent, Learner_Angel Agent will forward the message to Link Grammar Parser to query distance-learning ontology to get the tags about the input sentence. Link Grammar Parser will send the tags and the sentence to Label analysis & filter. Label analysis & filter analyze the links of input words sequence to see if the links of words satisfy the meta-rules such like planarity, connectivity, ordering, and exclusion. If the input words sequences have special tag from Learninn_Angel, the analysis & filter will records to Learning Corpus, record the score, and send the suggestion message to the learner interface.

In the chat room or other discussion system, the learner often asks questions to each other or asks questions to the system. In this system, the domain knowledge that is in the Distance learning ontology and learning corpus can provide answers for the users. When collecting enough Questions and Answers pairs, the frequency will be statistics.

4. Semantic Checker and FAQ

4.1 Semantic Checker

The syntax of this sentence: "The data is pushed in this heap." is correct. But there is semantic wrong with this sentence because heap doesn't have push method. So the meaning of the sentence could be wrong or correct in different situation. For this reason, the domain must be restricted. So the proposed semantic checker is deal with only "Data Structure" domain.

We propose the Semantic Relation of Knowledge Ontology. This methodology is to evaluate the distance of these keywords such as: 1.Sentence Pattern Classification. 2.Semantic Keywords Filter. 3.Sentence Distance Evaluation.

In the first step, the sentence pattern classification will check the sentence pattern. If there is a question sentence, the system will ignore this one. In our opinion, maybe the user doesn't know the semantic mean of this keyword. And he needs to ask question to others. For this reason this system doesn't care the semantic of the question. If there is not a question sentence, this sentence will be send to next step. In this step, the system can detect five kinds of sentence patterns. There are Simple sentence pattern, Negative sentence pattern, Interrogative sentence pattern, WH question sentence pattern and Imperative sentence pattern.

In the next step, semantic keyword filter extracts the keywords of Data Structure course Knowledge Ontology from the sentence. For instance, "the tree doesn't have pop method," these two keywords "tree" and "pop" will be detected that is the keywords of knowledge ontology. And the system can search the id of these words.

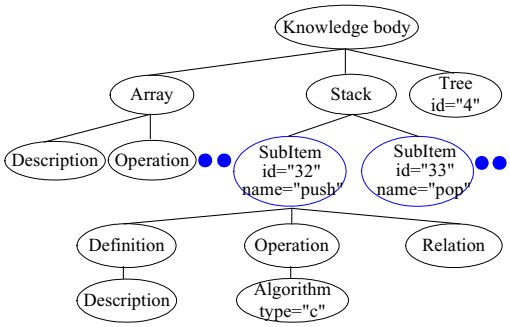

Fig. 2: The knowledge ontology of Data Structure

As figure 2 shows, the id of the keywords "tree" and "pop" is 4 and 33. The system can evaluate the distance of these two words. The semantic checker will think this sentence is correct. In this simple sentence pattern, the ids of the keywords stack and pop are 4 and 34, these two words are not in the same branch, so there are some wrong with this sentence.

4.2 Questions and Answers System

In this section, we use the knowledge ontology based approach- Semantic Relation of Knowledge Ontology construct semantic checker. This methodology can detect the sentence pattern and find the position of the keywords in the knowledge ontology. According to these characteristics, that is a new ideal to design a Questions and Answers system (QA system). The architecture QA system is shown in Figure 3.

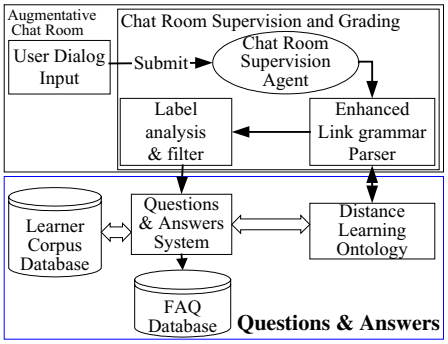

Fig. 3: The Architecture of questions and answers system

When QA system gets a Question pattern sentence, the system will check semantic meaning of the question by the semantic keywords. And then find the location of keywords in the ontology. Eventually, try to answer the question to the learner. If the QA system gets the question sentence "What is Stack" from the learners, it will extract the keyword "stack" and find the positions of the knowledge ontology after that to match the question temple "What is". The system will understand the meaning of this question wants to ask the definition of stack and try to find the definition or description of stack to the user. The follows shows the example of Data Structure ontology with "Stack":

```
- <KeyItem id="3" name="stack">
  - <Definition>
      <Description>A stack is a Last In, First Out
        (LIFO) data structure in which all insertions
        and deletions are restricted to one end called
        a top. There are three basic stack operations:
        push, pop, and stack top.</Description>
      <Symbol name="underflow">When the last
        item in the stack is deleted, it must be set to
        its empty state. If pop is called when the
        stack is empty, then it is in an underflow
        state.</Symbol>
    </Definition>
```

If there are enough questions and answer pairs in the FAQ database the FAQ system can statistic the questions and answers and find the most frequency Question and Answer pairs and provide QA service to learners.

5. Conclusion

In the proposed chat room system, students can talk and send messages to each other in this English environment and they can ask questions and discuss to teacher or the others. In this paper, we propose the Learning_Angel, which can help the learner automatically. The learner can practice English conversation and discuss about the studying course that takes Data Structure as example course. The teacher and supervisor won't be always waiting for the question from the students. Learning_Angel can automatic either detect syntax error when the student make some mistakes or check semantic of the sentence when the learner misunderstanding the course or the topic of the chat room. The proposed framework can parse sentences that the original Link Grammar cannot, and the modified dictionary has progressed steadily. It appears that the idea proposed herein can be applied to other natural language processing and its corresponding applications. The system parses input sentences on-line. It is important to philologists to analysis sentences generated by students, so that they can easily locate common or special mistakes. This system provides a better and more interactive environment for both teachers and students. Furthermore, it combines the humanistic education with technology and advances these two sciences for superior achievement.

In the future, we will focus on evaluating the accuracy of the two proposed methodologies of semantic checker. Finally, try to follow famous distance-learning standards.

Reference:

[1] Tsang, H.W.; Hung, L.M.; Ng, S.C. "A multimedia distance learning system on the Internet", proceedings of IEEE International Conference on Systems, Man, and Cybernetics (SMC), Vol. 2, Page(s): 243 –246

[2] Adhvaryu, S., and Balbin, I., "How Useful is Multimedia on the WWW for Enhancing Teaching and Learning?", Proceedings of International Conference on Multimedia Computing and Systems (ICMCS'98), Austin, USA, June 28-July 31, 1998

[3] Harris, D., Cordero, C., and Hsieh, J., "High-Speed Network for Delivery of Education-on-Demand", Proceeding of Multimedia Computing and Networking Conference (MCN'96), San Jose, California, January 29-31, 1996

[4] Willis, Barry, "Distance Education at a Glance", Engineering Outreach at the University of Idaho, http://www.uidaho.edu/evo/distglan.html

[5] D. K. Sleator, Davy Temperley "Parsing English with a Link Grammar" October 1991, CMU-CS-91-196

[6] S. Weibel, J. Gridby, and E. Miler: OCLC/NCSA Metadata Workshop Report, Dublin, EUA, 1995. http://www.oclc.org:5046/oclc/research/conferences /metadata/dublin_core_report.html

[7] Deborah L. McGuinness and Frank van Harmelen, "OWL Web Ontology Language Overview" W3C Recommendation 10 February 2004