CSIEC: A Computer Assisted English Learning Chatbot Based on Textual Knowledge and Reasoning

Jiyou Jia
Department of Educational Technology, School of Education,
Peking University. Beijing, 100871, China.
jjy@pku.edu.cn

Abstract: CSIEC (Computer Simulation in Educational Communication) system with newly developed multiple functions for English instruction still focuses on supplying a virtual chatting partner (chatbot) which can chat in English with the English learners anytime anywhere. It generates communicative response according to the user input, the dialogue context, the user's and its own personality knowledge, common sense knowledge, and inference knowledge. All these kinds of knowledge are expressed in the form of NLML, an annotation language for natural language text. These NLMLs can either be automatically obtained through parsing the text, or easily authored with the help of GUI editors designed by us. So the CSIEC system suggests a naïve approach of logical reasoning and inference directly through syntactical and semantic analysis of textual knowledge. This approach has advantages over the old ELIZA-like keywords matching mechanism. The chatting log summarization of free Internet usage within six months demonstrates this advantage. In this paper we present the system architecture and underlying technologies, and the educational application results.

Keywords: Computer Assisted Language Learning; Chatbot; Natural Language Processing; Human-Computer Interaction; Reasoning; Inference

1. Introduction

English, as an international language, is treated as a key tool for the development and cultivation of the cross-cultural communication ability. In China, English language is now listed as one of three core courses in elementary and secondary education, and as a compulsory course in higher education[1].

One of the best ways to learn a foreign language is speaking with native speakers. But it is not practical in the classroom due to the one - to - one student/teacher ratio it implies, especially in China and other countries with English as a foreign language. A number of factors ranging from the lack of time to shyness or limited opportunity for quality feedback hamper using the target language[2]. The language environment and few qualified English teachers in China cannot supply enough chances of authentic talking. So school teachers often complain of working burdens, and do not have adequate time to converse with students in English.

A potential solution to this problem is to apply computer spoken dialogue systems to role play a conversational partner. If we could design an interactive web - based system, which could chat with the English learners anytime anywhere, the great demand for learning partners could be fulfilled. Such a system should aim at helping the learners to improve their skills of using English through frequent chatting with them in English, as well as motivating them through playing and scoring mechanism. Inspired by the great demand for English instruction, we in 2002 began to design such a system[3]. Our design principle is application and evaluation oriented. As soon as the system is applicable, we put it into free usage in the Internet and get the user feedback. Through the application and evaluation, we get more suggestions and critiques, which can direct our research more effectively.

After six years' continuous development, the system's current pedagogical functions include automatic scoring of gap - filling exercises without predefined answers, listening training, talk show of two robots, multimodal user interface and selectable chatting pattern, free chatting adaptive to user preference and topic, guided chatting in given scenarios, and scoring mechanism. So it is not only an intelligent web - based human - computer dialogue system with a natural language for English instruction, but also a learning assessment system for learners and teachers. Therefore, it is on the one side freely utilized by the Internet users, most of which are students at different levels, on the other side integrated into university students English instruction[4] and middle school English classes[5][6][7].

In this paper we only explore its core and specific function, i. e. free chatting based on textual knowledge and reasoning. In Section 2 we review the related works. In Section 3 we explain the free chatting based on textual knowledge and reasoning. In Section 4 we analyze the user dialogue data of free Internet usage within 6 months and the system's integration into English class. In Section 5 we discuss our approach and suggest the further work.

2. Related Work: ELIZA Mechanism

Brennan defined a chatbot as "an artificial construct that is designed to converse with human beings using natural language as input and output"[8].

ELIZA[9] was the first chatbot, and used the keyword or pattern matching mechanism to find a pattern and the corresponding response for a given user text, which we call ELIZA mechanism.

Since 1990s, it has been further inherited in some popular chatbot programs such as ALICEBOT (http://www.alicebot.org), an influential open source project and a prototype of many online chatbots. The ELIZA mechanism can be explained with an example. If the user inputs "*I am a student*", and the robot responds with "*How do you like being a student*?" the keyword pattern and robot output template can be defined as:

"*I am a * student.*" → "*How do you like being a student*?"

The left-hand side of the arrow is user input pattern, and the right-hand side robot output template (in the following paragraphs we still use this style of writing). The star in the user input can represent any words or characters. The user's input matches this pattern, only if it contains the words "*I am a*" and "*student*". Therefore, all the following texts match this pattern: "*I am a teacher and you look like a student*", "*I am a person who likes all students*", etc. Obviously the robot response mentioned above, "*How do you like being a student*?", is not appropriate to these inputs. A so-defined pattern will match plenty of input texts which are semantically different with the original meaning, i.e. "*I am a (any kind of) student*". In order to express this meaning more exactly, the old mechanism has to include many patterns, such as "*I am a student*", "*I am a middle school student*", "*I am a Chinese student*". However, it is very difficult to enumerate all these patterns.

Similarly, in order to express the responses to the inputs such as "*I am a teacher*", "*I was a salesman two years ago*", "*My classmate will be an officer*", the ELIZA mechanism should contain more and more patterns and the corresponding responses.

The ELIZA mechanism can't avoid the discrepancy between the exactness and infiniteness of input patterns, because it only seizes the keywords in the input, but not the syntax and semantics of the whole input text. In order to describe the whole input text and to achieve the apparently syntactic and semantic analysis, it has to contain the whole input text as an input pattern. However, it is impossible to write all the multitudinous expressions. A compromise is just writing the so-called keywords in the input and ignoring the other parts of the input, but it causes the expressive inexactness of input pattern, as illustrated by the examples above.

Moreover, the ELIZA mechanism holds no memory of the conversation, and so it cannot achieve any form of targeted collaboration or negotiation.

The ELIZA mechanism's advantage is its fast response speed, as it does not run deeper syntactical or semantic analysis of the input text.

Besides ELIZA-like chatbots, there are also many other kinds of chatbots. Some are commercial products, and some have been applied in education[10-13]. But we cannot find more references that introduce their underlying technologies. Through actual chatting with some online free chatbots, we can only feel that they seem still ELIZA-like, and the conversation does not hold memory of the dialogue context.

3. CSIEC System Architecture and Underlying Mechanism

In the system design, we attempt the fully syntactical and semantic analysis of the user inputs, as the logician G. Frege pointed out: "*The meaning of a sentence exists in the meanings of all words within the sentence and their conjunction method*"[14]. After parsing the user input we obtain the user information in the form of XML, i. e. NLML and call them the user facts. The facts are retrieved from natural language expressions, and also represented with the annotation of natural language in the sentence ontology. These facts function as the main contextual source of the robot dialogue reasoning. This thought originates from L. Wittgenstein's theory (1918/21) about the world, facts, objects and human language: "*The world consists of facts, the facts consist of objects. The facts are reflected in the language. A logical picture of facts is a thought*"[15].

The current CSIEC system is version10. Corresponding to its multiple functions the whole system has a more complicated architecture, which is shown in Fig. 1. The system components are related with one another. Here, we focus on the components that are crucial to the free chatting function using the textual knowledge and reasoning.

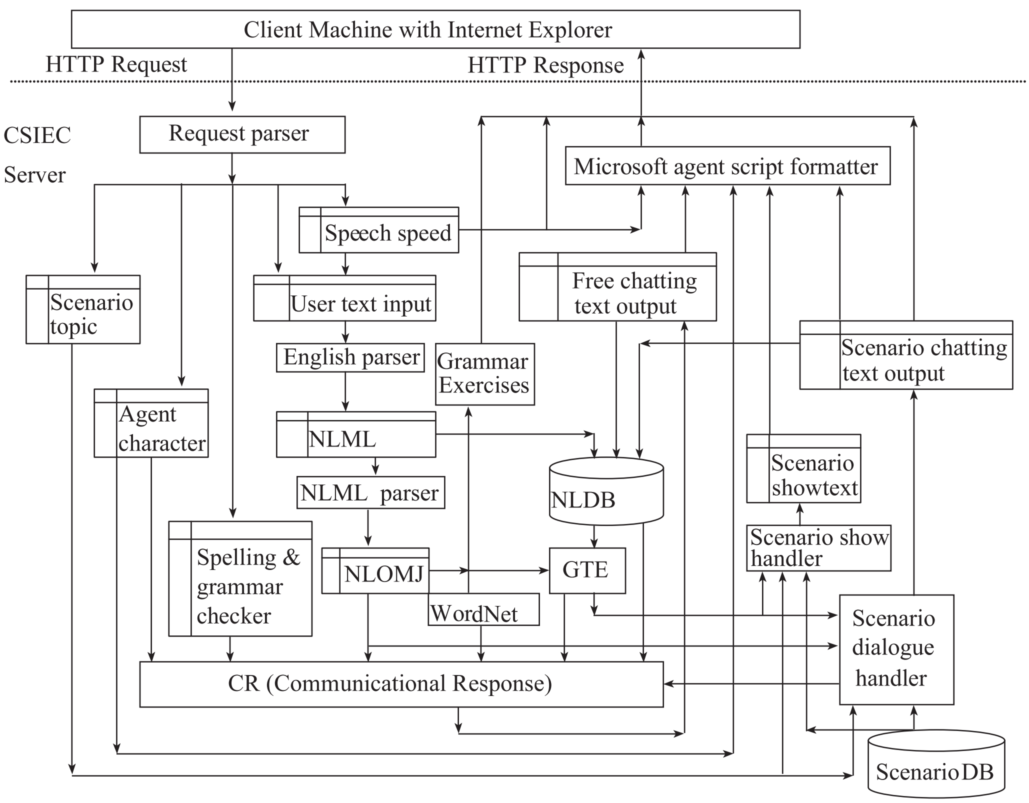

Fig. 1. The Architecture of CSIEC system

3.1 Browser/Server Interface

The Browser/Server interface includes HTTP request parser and Microsoft agent script formatter.

The HTTP request parser resolves the user request from HTTP connection and gets some parameter values: input text, spelling and grammar checker, agent character, speech speed, scenario topic, etc. The input text is needed in all the functions. Spelling and grammar checker is selected if the user wants the robot to check the spelling and grammar error in the input. Anyone of the three female Microsoft agent avatars (Christine, Ingrid and Emina) and two males (Stephan and Christopher) with different figures and voices can be selected to represent the user or the robot. The speech speed defines the synthesized voice's speed of the selected avatar. The scenario topic is selected when the user chats with the robot on a given topic.

Microsoft agent script formatter transforms the output text into VB scripts, considering the selected agent character and speaking speed.

3.2 English Parser

The English parser parses the user input text into NLML (Natural Language Markup Language). NLML is a dependency tree in an XML form, and structurally labels the grammar elements (phrases), their relations and other linguistic information (words, part-of-speech, entity type, chunk tag, grammatical function tag, head word path, etc.). More detailed description about NLML can be seen in[16].

For example, the NLML for the sentence "*What is your name?*" is:

```
<mood>question</mood> <complexity>simple</complexity>
<subject> <prem> <type>possessive</type> <word>your</word> </prem>
<noun> <pers>third</pers> <word>name</word>
<numb>sing</numb> <type>noun</type>
</noun>
</subject>
<verb_phrase> <verb_type>be</verb_type> <tense>present</tense>
<numb>sing</numb> <pers>third</pers> <verb_word>is</verb_word>
<predicate> <predicate_type>np</predicate_type>
<noun> <type>relpron</type> <word>what</word> </noun>
</predicate>
</verb_phrase>
```

3.3 NLML parser

The NLML parser parses the NLML obtained from the user input or loaded from database into NLOMJ (Natural Language Object Model in Java), which represents the grammatical elements and their dependency with the Sentence or Phrase ontology in the working memory[17]. Through NLOMJ we can easily acquire any element from an expression and manipulate the sentence transformation, for example, negation, statement to question, active to passive, and vice versa.

Through NLOMJ the declarative sentence can be retrieved and decomposed into atomic facts consisting of only one subject and one verb phrase, which we call a basic sentence. For example, from the sentence "*I am Peter and live in Europe*" two basic sentences can be extracted: "*I am Peter*" and "*I live in Europe*".

3.4 NLDB (Natural Language Database)

The NLDB stores the historical discourse, the user atomic facts in the form NLML, the robot atomic facts that are also expressed in NLML, the common sense knowledge, the inference knowledge, the discourse history, the dialogue context including users' usage of different functions and obtained scores, and other data.

The robot's atomic facts are manually predefined in the form of natural language, and stored with the form of NLML in a table in the NLDB. For example, an atomic fact about the avatar "*Stephan*" is "*My name is Stephan*". Each of the five avatars has its own personality, which are expressed in its own facts table.

We call the user and robots facts that are stored in the NLDB personality knowledge.

A common sense table in the NLDB stores the textual common sense knowledge with the form NLML, like "*China is in Asia*". The items in this table are predefined by the system maintainer, and can be extended through users' talking, i. e. the facts retrieved from the user inputs, which are neither related with the user itself nor with the robot, are added into the common sense knowledge table. However, the system maintainer should check the correctness of these automatically acquired "knowledge".

3.5 Common Sense Knowledge

The common sense knowledge is the basis for response generation and logical inference. It is now represented by WordNet[18] and the common sense table in NLDB. WordNet is a famous common sense ontology. From it the relation between two nouns, verb, adjectives or adverbs can be retrieved. For example, "*student*" is a kind of "*person*", or formally expressed, "*person*" is one hypernym of "*student*", "*student*" is one hyponym of "*person*". "S*tupid*" is one antonym of "*clever*", and the synonym of "*foolish*". "*Run*" is one troponym of "*rush*", and one of its hypernyms is "*travel*".

3.6 GTE (Generation of Textual Entailment) Mechanism

The GTE mechanism can generate the textual entailments or make inferences for a given text. For example, from "*my name is Stephan*" it can get the following entailments: "*Stephan is my name*", "*I am Stephan*", and "*I am called Stephan*". This mechanism uses a table in NLDB to store the inference rules, runs a pattern recognition technique to look up a matching rule in the table for a given text, and gets the appropriate entailment form based on the rule and input text[19]. The inference table now has 250 rules, and is still being enriched. We call the inference rule table inference knowledge.

3.7 CR (Communicational Response) Mechanism

The CR mechanism generates responses by free talking with the user. It comprehensively takes into account the user input, the user facts and selected robot facts stored in NLDB, the common sense knowledge, the generation of textual entailment, and the chatting pattern represented by the agent avatar selected by the user. The three kinds of knowledge database, i. e., personality knowledge, common sense knowledge and inference knowledge make up the foundation of communicational response.

In free chatting with Microsoft agent characters, the avatar selected by the user defines the chatting pattern. Christine always tells the user stories, jokes and world news. Stephan prefers to listen quietly when the users share with him their own experiences. Emina is a curious girl, and is fond of asking users all kinds of questions related with the user input. Christopher supplies comments, suggestions and advices on the user input. Ingrid behaves as a comprehensive virtual chatting partner, who gives users responses corresponding to both the input text and the discourse context. If the user selects the chatting in pure text, the robot behaves just like the comprehensive avatar Ingrid. So we introduce the response mechanism of the comprehensive avatar.

For a user input, the comprehensive avatar looks for a response in the sequence of personality knowledge considering the discourse context, direct response, inference knowledge, and common sense knowledge.

3.7.1 Direct Response

At first we introduce the direct response without considering the discourse context. With the syntactical and semantic analysis capability the CSIEC system can overcome the shortcoming of ELIZA mechanism. It can define a more general pattern to include all those sentences with a specific syntactic and semantic feature. The pattern is described with the form NLML including some pseudo variables, which stand for the keywords in the ELIZA mechanism.

We illustrate the mechanism still with the example used in Section 2, i. e. how to express theresponse to: "*I am a student*" and so on inputted by the user for the first time. The input pattern can be expressed as the following (due to the space limitation only the core parts are listed, the other parts are represented by ellipsis. The rest NLMLs are described by this way, too.):

```
<mood>statement</mood>
<subject> <pseudo>pseudo variable 1</pseudo> </subject>
<verb_ phrase> <voice>active</voice> <verb_ type>be</verb_ type>……
<predicate> <predicate_ type>np</predicate_ type>
<noun> <word>pseudo variable 2</word> <type>noun</type> </noun>
</predicate>
</verb_ phrase>
```

The response template NLML is:

```
<mood>question</mood> <subject>pseudo variable 1</subject>
<verb_ phrase> <voice>active</voice> <verb_ type>likewishinf</verb_ type>…
<verb_ word>do</verb_ word> <verb_ word>like</verb_ word> <verb_ change/>
<direct_ object> <noun_ clause> <clause_ type>gerund</clause_ type>
<verb_ phrase> <voice>active</voice> <verb_ type>be</verb_ type>
…… <verb_ word>being</verb_ word>
<predicate> <predicate_ type>np</predicate_ type>pseudo variable 2</predicate>
</verb_ phrase>
</noun_ clause> </direct_ object>
</verb_ phrase>
<circum> <circum_ type>adv</circum_ type>
<type>question</type> <word>how</word> <attribute>way</attribute>
</circum>
```

The following input texts can be found matching the input pattern, and can get the corresponding responses with the help of output template:

"I am a university student." → "How do you like being a university student?"
"I was a salesman two years ago" → "How did you like being a salesman two years ago?"
"My classmate will be an officer" → "How will your classmate like being an officer?"
"I am a teacher and you look like a student" → "How do you like being a teacher?"
"I am a person who likes all students" → "How do you like being a person?"

But the following inputs do not match this pattern:

"*I am Peter*", "*I am happy today*", "*I am in the room*", "*he helps this student*", etc.

The pairs of an input pattern and an output template are written in a table "direct – response" in NLDB. The algorithm for this direct response is:

Parse the input text into NLML and get its NLOMJ. Pattern recognition: compare the structure of this NLOMJ with the patterns in direct response table. If there is such a pattern matching this input text, transform the template NLML and calculate the response text. end.

So the direct response mechanism consists of three procedures: input pattern and output template annotation with NLML, pattern recognition (matching), and output template transformation. The pattern and output annotation in NLML must be done by an author who is good at English grammar and dialogue generation, but does not need to learn the complicated description of text with NLML, because we have designed a GUI pattern – template editor to assist the swift and convenient annotation.

There can be multiple responses to a given input text. So, for an input pattern several output templates can be generated. The pair of pattern and template can be labelled by an index integer so that the robot can generate different responses for a given text of a user to avoid the tedious sole output.

The direct response generation only considers the input text, but not the dialogue context. The user input can be declarative, investigative, imperative, exclamative, or a phrase.

As for the example above, because the user input is declarative, the user fact "*I am a student*" is added into the user facts table.

3.7.2 Direct Response Considering Dialogue Context and Personality Knowledge

In many occasions the user facts and robot facts, and the user discourse history must be considered, for example if the user talks about the personality of the user or of the robot, or repeats an input again. The response is generated by searching the user personality or robot personality database, and considering the dialogue context.

For example, if the user inputs "*I am a student*" for the second time, the same response "*How do you like being a student?*" may sound tedious. However, from the user dialogue history table the robot knows that this response template has been used once, so it will look for another response template, for example "*What are you studying?*" Moreover, the robot finds the same fact in the user facts table (personality knowledge), thus it will respond with this knowledge: "*Yes, I know. What are you studying?*"

3.7.3 Searching in Inference Knowledge and Common Sense Knowledge

If no response can be generated based on the direct response for a given input, or on personality facts, the inference (GTE) table and the common sense knowledge table will be searched. If there is still no corresponding response, the WordNet will be looked for to retrieve the definition of a noun in the input (or its symposium, hypernym, hyponym) as the response output.

3.7.4 Communicational Response for Different Mood: Cases

We explain the response generation algorithm with the following user input and robot output pairs in details.

A) If the input is a statement, the response to this statement can be a statement, or a question, or their combination. For example:

"*Hello, I am Peter.*" → "*Hi, Peter. How are you?*"

This output is adirect response to "*I am <name>*". The user's new fact "*I am Peter*" will be added into the user fact table. By this user's next login, the robot will greet with: "*hello, nice to meet you, Peter.*"

"*I am a student.*" → "*How do you like being a student?*"

This output is a direct response to the user input, and associated with the user input.

"*I feel very happy to be a student.*" → "*I'm a college student and my major is math. What is your major?*"

The robot output is still another direct response to "*be a student*", but consists of two sentences. One is a statement about the robot fact, and another is a question about the user.

"*Literature.*" → "*Alas. You study literature. Do you like it?*"

This output isone response according to inference knowledge, i. e. "*Your major is literature*" → "*You study literature.*", plus one direct response to it.

"*Surely.*" → "*Okay, you like literature.*"

This robot analyzes the user input and obtains a new user fact according to the dialogue context, and repeats this new user fact again.

"*You are boring.*" → "*No, I am not boring. Are you boring?*"

This input is a comment on the robot, and the robot can't find such a statement "*I am boring*" from its own facts table, so it denies the user comment. Moreover, the robot can't find such a fact from the user facts table, thus it inquires the user again.

B) If the input is a question, it tries to answer the question.

"*Do you like literature?*" → "*Certainly, I like literature, because reading literature enriches my life. I know you also like literature, because you have told me.*"

The user asks a question about the robot itself, so the robot looks for the fact "*I like literature*" in its own facts table. It finds not only such a fact, but also the explanation to this fact. In addition it finds also such a fact from the user facts table, so it repeats this fact to demonstrate its memory capability.

"*What is a UFO?*" → "*UFO is a phantom.*"

The user asks a concept definition, and the robot finds the definition from WordNet.

User: "*Who am I?*" Robot: "*Peter.*"

The user asks a question about himself/herself, and the robot finds the answer from the user facts table.

"*Do you speak Chinese?*" → "*No, I do not speak Chinese because I can't speak Chinese. Do you speak Chinese?*"

The user asks a question about the robot itself, so the robot looks for the fact "*I speak Chinese*" in its own fact table. It finds not only a negative answer to this fact, but also its explanation. However, it can't find such a fact from the user facts table, thus it asks the user about it.

"*What is twenty plus thirty-one?*" → "51."

The robot can calculate a mathematic expression with the direct response.

"*Where is Australia?*" → "*Sorry I do not know. Do you know the answer? Where is Australia?*"

For the question whose answer can't be retrieved from the personality knowledge and common sense knowledge, the robot acknowledges its insufficiency and asks the user the answer.

"*Australia is in the Pacific.*" → "*Oh, Australia is there.*"

The robotadds the fact uttered by the user into the common sense knowledge table, as it has nothing to do with the user and the robot, and responds to the user input.

"*Where is Australia?*" → "*In the Pacific.*"

The robot finds the answer to thequestion from the updated common sense table.

C) If the input is an imperative, it responds and acts.

"*Please tell me what UFO is.*" → "*UFO is an (apparently) flying object whose nature is unknown.*"

The robot reduces the user request into a question about the concept definition and gets it from WordNet.

"*Tell me a joke.*" → "*This story is a joke. Which Search Engine does Arnold Schwarzenegger use?*"

The robot gets a joke from the table "*story*" and narrates it sentence by sentence.

"*Sing a love song please.*" → "*I Will Always Love You—Whitney Houston* 1993…"

The robot retrieves a song from the table "*song*" and recites it.

D) If the input is an exclamation, the robot responds with feeling the surprise.

"*How simple you are!*" → "*You are surprised, aren't you? Yeah, I am simple. Are you simple?*"

The robots transfers the user exclamation into a declaration about itself: "*I am simple*", and gets this entry also from its robot facts table, so it acknowledges the declaration, and asks the user the same question since no such an entry is found in the user facts table.

4. Implementation, Pilot Application and Evaluation

The CSIEC system is now implemented in JDK1.6.0, and uses MySQL Server 5.0 as a database management system. It can be freely accessed in Internet. The internet users get to the CSIEC website (www.csiec.com) mainly through search engines, because our website has become one of the top 5s in the searching results of famous search engines such as google.com, yahoo.com and baidu.com by related keywords such as "chatbot", "English chatbot", "Online English learning" in Chinese or in English, although we have not made any advertisement. The effectiveness and attractiveness of the system's adaption to English learning has been somewhat demonstrated by this practical achievement.

With the human - computer dialogues recorded in the database, we can make a summarization of the system's chatting function within a given period, for example from January 20th, 2007 to June 20th, 2007. The different users who accessed the CSIEC during this period counted 1783. The analysis of the users' demographic distribution shows that more than half of the users were undergraduate students. The second large user population was middle school students. Totally more than 80% of the users were various students.

4.1 Dialogue Duration

The chatting quality can be measured by the chatting duration between the user and the robot. To calculate the chatting duration we define two terms: round and number of the rounds. A round means a user input and a corresponding robot output. Therefore, the total rounds of a given user cover all dialogues between the user and the chatbot, and can be used to measure the duration of the user's chatting with this chatbot. We divide the number of the rounds into four classes, as Table 1 shows.

Table 1: The relation between the duration of dialogues and number of users

Dialogue duration	Range of the rounds numbers	Number of users	Number of users/ Total user number	Number of users/ Total user number in[20]
Short	(0, 10]	871	48.85%	62.34%
Long	(10, 50]	685	38.42%	30.10%
Longer	(50, 100]	136	7.63%	4.78%
Very long	(100, 580]	91	5.10%	2.79%
Total user number		1783	100.00%	100.00%

The average rounds number is 27.4. The number of the rounds of each user varies from 1 to 580. Table 1 shows that c. a. 49% of the users chatted with the robot briefly (< =10 rounds); c. a. 46% (38.42% + 7.63%) chatted with it long or longer; and only few, c. a. 5%, chatted with it very long (> 100 rounds). Compared with our previous finding about the ALICEBOT's chatting duration in[20] which is listed in the last column of Table 1, the percentage of the brief chatting with the robot decreased by 21.78%. Proportionally, the percentage of the long and longer chatting increased.

4.2 The Distribution of User Chatting Patterns

The CSIEC system supplies multimodal user interface and selectable chatting patterns. Thus we investigate the distribution of chatting patterns. 84.7% of the chatting was held with the free chatting pattern, and only 15.3% used the chatting in a given scenario. The reason may be that the free users do not understand the chatting in a given scenario very well.

Among the free chatting, the chatting without spelling and grammar check (c. a. 66%) was much more used than with check (c. a. 18%). This result reflects that most free users treated the system as a chatting partner, so they would like to chat with it more fluently instead of worrying about grammar and spelling errors. Human – computer chatting is the most unique function of the CSIEC system, thus the users like to use it frequently.

4.3 User Feedbacks

In the foot of almost every webpage of the CSIEC system we leave a feedback text area so that the users can straightforwardly enter their comments, critiques and suggestions. Through scrutinizing of the user feedbacks we find as many critiques as praises. For example, there were the following positive comments:

The robot is more advanced than before, and also personalized.
The dialogue is fluent. I hope the master to enrich the robot's language.
The kind of communication can improve our English.

The negative comments pointed out either technical problems or content shortcoming. Some complained that they could not use the agent version or the agent voice sounded curious. Other problems included the website access speed was too low, and the robot response was too slow, etc. These problems should be tackled in the further improvement.

4.4 Pilot Educational Evaluation

The CSIEC system has also been applied in English education in universities and middle schools. The collected survey data from one university and one high school show the students felt the CSIEC – assisted English learning could help with course unit review, make them more confident, improve their listening ability, and enhance the interest in language learning. One item in the questionnaire for high school students shows 60.5% of the students "liked" or "liked very much" such a form of English learning, whereas only 2.3% disliked it. 60.5% of them would continue using the system after class, even without the teacher's request[5-6].

The system was also integrated into an English class in Grade one of a junior school[7]. The comparison of two examination results before and after the integration class shows great improvement of students' performance. The average exam score of the whole class was improved from 64.39 to 90.81, whereas the standard deviation decreased from 20.129 to 9.572. The survey data indicate the students' favor to this system, too. 76.9% of them hoped to use the CSIEC in the entire English instruction very much, 23.1% hoped to use it in the whole English class.

Although the main functions used in these education applications mentioned above were talk show and chatting on given topics, the free chatting was also often utilized by the students, and contributed to the performance improvement.

5. Discussion and Conclusion

The original goal of the CSIEC system is providing English learners with a virtual chatting partner. So the chatting should be the most fundamental function. The statistical analysis about the users' behavior indicates that the users have a preference for free chatting and for chatting without spelling and grammar checking. This fact proves that the users prefer the unique chatting function, which is lacked in other systems. So we must continue to reinforce this primary utility.

The chattingquality can be somewhat demonstrated through the chatting length. The increased percentage of the long and longer chatting shows that the free chatting quality of CSIEC is becoming better. The underlying design principle, i.e. fully syntactical and semantic analysis of the user input, and communicative response mechanism, as well as the effort of chatting considering the dialogue context all contribute to that quality progress. Certainly, the content analysis of the dialogues should be further conducted in order to investigate the chatting quality more exactly.

The famous logical programming systems such as Prolog and LISP can makereasoning and inference according to the logical rules. However, the laborious transformation from natural language into exact logical language and vice versa seems to be only done by the logician experts. The giant common sense ontology OpenCYC project's logical reasoning interface with natural language is also limited (http://www.opencyc.org/doc/#NL_OpenCyc). In our project, we attempt to directly teach the computer how to understand and extract the syntactical features of texts and how to make responses and inferences according to the direct response rules with the notation of NLML and NLOMJ, just as the English teacher teaches the students how to learn the sentence patterns and transformation rules, and how to apply them in the actual language expression. This naïve approach is consistent with the human language acquirement process, and has great promise.

Through thepilot application and evaluation, we find currently there are still some user requirements which have not been fulfilled well, first of all, the system's stronger ability of natural language understating and generation, which is the fatal factor influencing the human-computer communication. Solely in NLP many problems are still hard to be solved, such as the textual ambiguity and entailment[21]. How to overcome these problems with current available technologies is still a great challenge to us.

The contradiction between the high response speed and complex, deep, syntactic and semantic analysis should also be paid great attention, especially in the web environment. Among the user feedbacks, many are complaints against the slow Internet response. To solve this problem the concrete implementation techniques of programming and database management should be improved furthermore.

Acknowledgments

We are grateful for the support to our projects from Konrad – Adenauer – Foundation Germany, Ministry of Education China, Peking University, Education Committee of Capital Beijing, and Korea Foundation for Advanced Studies.

References

[1] Ministry of EducationChina. Annual Educational Statistics 2006 (People's Education Press, Beijing, 2006).

[2] L. Fryer & R. Carpenter. Emerging Technologies Bots as Language Learning Tools. Language Learning & Technology. 10 (3) (2006).

[3] J. Jia. CSIEC (Computer Simulator in Educational Communication): A Virtual Context – Adaptive Chatting Partner for Foreign Language Learners. In: Kinshuk, et al. ed., Proc. of International Conference of Advanced Learning Technology 2004 (IEEE Computer Society, CA, 2004).

[4] W. Chen, J. Jia, Y. Cheng. Exploring CSIEC's Application in Teaching and Learning of Oral English: Design – Based Research Approach. In: J. Lee et al. ed.: Research in IT and Education: a Multi – Disciplinary Perspective (Normal University Press, Beijing, 2007).

[5] W. Chen, J. Jia, D. Xiang. The Application of CSIEC in English Instruction in High School: Design Based Research. China Educational Technology. 20 (2) (2008).

[6] W. Chen, J. Jia. Design – Based Research and Its application in CSIEC project. Modern Educational Technology. 15 (1) (2008).

[7] J. Jia, M. Ruan, Use Chatbot CSIEC to Facilitate the Individual Learning in English Instruction: a Case Study, In: B. Woolf et al. (Eds.), Proceedings of International Conference on Intelligent Tutoring System 2008, Lecture Notes in Computer Science 5091, Springer, 2008.

[8] K. Brennan. The Managed Teacher: Emotional Labour, Education, and Technology.

Educational Insights. 10 (2) (2006).

[9] J. Weizenbaum. ELIZA - a Computer Program for the Study of Natural Language Communications between Man & Machine. Communications of Association for Computing Machinery. 9 (1) (1966).

[10] A. C. Graesser, P. Chipman, B. C. Haynes, A. Olney. AutoTutor: An Intelligent Tutoring System with Mixed - initiative Dialogue. IEEE Transactions on Education. 48 (4) (2005).

[11] S. Seneff. Interactive Computer Aids for Acquiring Proficiency in Mandarin. In: Huo et al. ed.: Proceeding of the 5th International Symposium on Chinese Spoken Language Processing. 2006, Lecture Notes in Artificial Intelligence 4274 (Springer, Berlin, 2006).

[12] P. Kerfoot, et al. A Multi - institutional Randomized Controlled Trial of Web - based Teaching to Medical Students. Academic Medicine. 81 (3) (2006).

[13] A. Kerly, P. Hall, S. Bull. Bringing Chatbots into Education: Towards Natural Language Negotiation of Open Learner Models. Knowledge Based Systems. 20 (2) (2007).

[14] Frege, G. Begriffsschrift, Eine der Arithmetischen Nachgebildete Formalsprache des Reinen Denkens. (Wissenschaftliche Buchgesellschaft, Darmstadt, 1879).

[15] L. Wittgenstein. Tractatus logico - philosophicus (Suhrkamp, Frankfurt am Main, 1918).

[16] J. Jia. NLML - a Markup Language to Describe the Unlimited English Grammar. http://arxiv. org/abs/cs/0404018 (2004).

[17] J. Jia, Y. Ye, K. Mainzer. NLOMJ—Natural Language Object Model in Java. In Z. Shi, et al. ed.: Intelligent Information Processing II (Springer, Berlin, 2004).

[18] C. Fellbaum. WordNet: An Electronic Lexical Database (MIT Press, MA, 1998).

[19] J. Jia, The generation of textual entailment with NLML in an intelligent dialogue system for language learning CSIEC, in: Proceedings of IEEE International Conference on Natural Language Processing and Knowledge Engineering 2008, IEEE Press, 2008.

[20] J. Jia. The Study of the Application of a Web - Based Chatbot System on the Teaching of Foreign Languages. In: P. Kommers et al. ed.: Proc. of SITE 2004 (AACE, VA, 2004).

[21] I. Dagan, et al. ed.: Proc. of the PASCAL Challenges Workshop on Recognising Textual Entailment. LNCS, vol. 3944 (Springer, Berlin, 2005).