Automated question generation methods for intelligent English learning systems and its evaluation

ARTICLE · JANUARY 2004

CITATIONS
20

4 AUTHORS, INCLUDING:

Hidenobu Kunichika
Kyushu Institute of Technology
18 PUBLICATIONS **48** CITATIONS
SEE PROFILE

Tsukasa Hirashima
Hiroshima University
110 PUBLICATIONS **342** CITATIONS
SEE PROFILE

Automated Question Generation Methods for Intelligent English Learning Systems and its Evaluation

Hidenobu Kunichika, Tomoki Katayama*, Tsukasa Hirashima and Akira Takeuchi
Dept. of AI, Kyushu Institute of Technology, Japan
{kunitika, tsukasa, takeuchi}@ai.kyutech.ac.jp

Abstract: This paper presents methods of automated generation of questions about the contents of an English story. The methods are used for realizing adaptive Question and Answer functions which give suitable questions to learners of different understanding states. To realize adaptive Question and Answer, our methods generate five kinds of questions; to ask about the content of one sentence, to use antonyms and synonyms, to use modifiers appeared in plural sentences, to ask the contents represented by plural sentences by using a relative pronoun, and to ask time and space relationships. Our methods cover 80% of questions in collection books of problems for novices. The result of an experiment shows that the question generation module generates 1977 question sentences from 301 sentences in textbooks for Japanese junior high school students, and 1848 of the generated question sentences are semantically competent.

Keywords: Intelligent CAI, CALL Systems, English Learning, Question Generation, Question and Answer

1. Introduction

QA (Question and Answer) about the contents of a story is widely used in language learning. QA in a target language is effective because the learner must use the plural skills with the language to read a story and question sentences and to compose answer sentences. When a teacher and a learner do such QA, the teacher will give suitable questions for the learner. For example, when the learner could not answer a question, the teacher will paraphrase the question sentence with easy words or give more simple questions.

Some computer assisted language learning systems are equipped with test functions to ask about the contents of a story. Most of them, however, have limited capability to ask questions adaptively depending on learners' understanding states because questions are prepared beforehand (Levin & Evans 95; DynED 97). In preparing questions, teachers are burdened with preparing not only question sentences but also correct answers for every learning material. And it is very difficult to prepare sufficient amount of questions to correspond with various states of learners' understanding.

The target of our study is to realize a QA function which provides suitable questions for each learner and tailored advice according to the answer. To achieve the target, we need the following functions: (1) to understand English sentences, (2) to generate various kinds of question sentences automatically, (3) to select a suitable question for a learner from a set of generated question sentences and (4) to analyze learner's answer sentences and to diagnose errors. In the previous studies, we have proposed a method of representing meanings of stories and a method of diagnoses (Kunichika et al. 95) for the functions (1) and (4). This paper presents methods of question generation for the function (2) and its evaluation.

In the second chapter, we describe the outline of the QA function. Then in the third chapter, we explain information used for question generation and methods of the question generation. After presenting evaluation of the methods, our conclusions and discussions follow.

2. The outline of the QA function

The QA function gives questions about a story in English to a learner who has studied the contents of the

* Currentry NEC Fielding, Ltd.

story by hearing or reading, and the learner answers the questions in English. The methods we propose is used for generation of various kinds of questions automatically from syntactic and semantic information of the story for realizing adaptive QA. In this chapter, we describe aims of the QA, query forms of questions and purposes of questions to achieve the aims.

2.1 Aims of the QA

(1) To master the use of grammatical knowledge

In general, one of important educational targets in language learning is to make learners master the use of grammatical knowledge about words and grammars. In fact, each lesson in textbooks for novice learners has learning targets about such grammatical knowledge, and teachers perform teaching activities for learners to master the targets.

The aims of the QA are to assess a learner's understanding state about the learning targets and to make her master the use of the targets by using the particular knowledge. The particular knowledge includes both the learning targets and the knowledge which the learner used incorrectly in the past. To achieve the aims, the QA function asks questions which require the learner to use the particular grammatical knowledge such as new words or grammars to interpret, or questions which make the learner use the particular knowledge to compose an answer sentence.

(2) To train for conversation

Learning by the QA requires learners to use reading, hearing and writing skills, that is to say, the learners read a story, hear/read question sentences and write answer sentences. The QA is, therefore, effective to train conversation skill for exchanging information each other.

To achieve the aim, the QA function gives the learner questions which have moderate difficulty and ask about the contents of a story. It is desirable to generate question sentences by replacing words and phrases or summarizing sentences, instead of using only words in original sentences of the story. For example, questions which use synonyms of words in the original sentences or ask the contents such as context represented by plural sentences are used.

2.2 Query forms

There are two kinds of skill to use grammatical knowledge: interpretation, which is understanding meaning of a sentence, and composition, which is expressing meaning as an English sentence. In order to distinguish between these two skills in setting up a learning target, the QA function treats two query forms as follows.

(1) Alternative questions

A learner answers alternative questions with sentences which follow "Yes" or "No". This query form is used for confirming whether or not she has understood contents through listening or reading a story. In general, learners will describe correct information when they answer a question with "No", but they may describe the correct information by using the question sentence as hints. Thus we treat this query form as targeting for interpretation.

(2) Questions using interrogative pronouns

To answer this kind of questions, learners translate the answer into an English sentence by using the syntactic knowledge which has already been acquired. We treat this query form as for composition.

2.3 Purposes of questions

As described in the previous section, the QA has two educational aims and there are two kinds of skill. We prepared different types of questions for each combination of aims and skill as shown in Table 1.

(1) To master the use of grammatical knowledge for interpretation

This kind of questions is used for confirming whether or not a learner understands the contents of sentences

Table 1: Purposes of questions

		Skill	
		interpretation	composition
Aims	to master the use of knowledge	(1)	(2)
	to train for conversation	(3)	(4)

which has particular grammatical knowledge by listening and/or reading. Because the aim of the questions is to assess the learner's interpretation skill of particular grammatical knowledge, it is not necessary to ask the learner about extensive contents such as context. So questions for this purpose ask about the content represented by one sentence in a story. For example, question sentences generated by transforming affirmative form of original sentences in a story to interrogative form, or by replacing words in an original sentence with synonyms or antonyms belong to this kind.

(2) To master the use of grammatical knowledge for composition

This kind of questions is used for confirming whether or not a learner can compose sentences which have particular grammatical knowledge. Because the aim of this kind of questions is to assess the learner's composition skill with particular grammatical knowledge, the questions which ask the learner about the content of one sentence are used. The query form of this kind of questions is "questions using interrogative pronouns" and a query part which a question asks is subjects, objects, predicates or their modifiers.

(3) To train for conversation through interpretation

To realize conversation, learners must understand not only the content of each sentence separately but also context of sentences. This kind of questions is used for confirming whether or not a learner grasps context of sentences. In this paper, the context means referring relationship between nouns or pronouns, modifying relationship of words and phrases, and time-space relationship. For example, alternative questions in which a pronoun is replaced with the referent are used.

(4) To train for conversation through composition

This kind of questions makes a learner express information about context of sentences to confirm whether or not she understands the context and composes sentences. For example, questions using interrogative pronouns, which ask the referent of a pronoun or time-space relationship are used.

3. Automated question generation

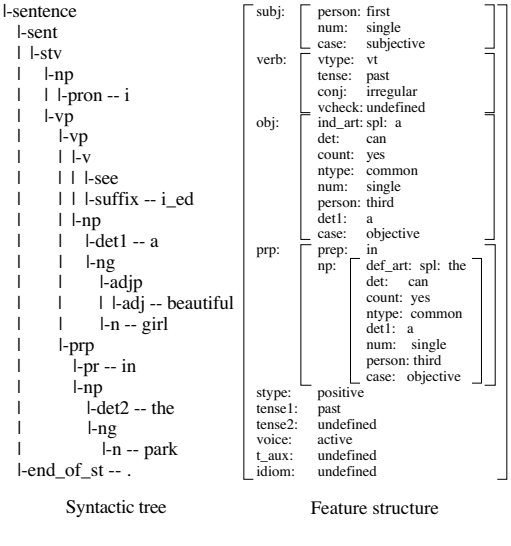
In this section, we explain the information used for automated question generation and methods of question generation.

3.1 Information for question generation

Our methods of question generation use syntactic and semantic information of an original story generated by natural language processing, and dictionaries of synonyms and antonyms. We have proposed a method of extracting syntactic and semantic information of stories based on DCG (Definite Clause Grammar) (Kunichika et al. 95). Semantic information is also important to judge semantic correctness of learners' answers. Here we briefly describe a representation method of syntactic and semantic information.

Syntactic information consists of a syntactic tree which expresses parts of speech and modification relationships of words and phrases, and feature structure which expresses both grammatical functions of words and phrases and grammatical information such as sentence structure and idioms. Fig. 1 shows an example of syntactic information.

Semantic information consists of time and space information and information about verbs, nouns and modifiers. Fig. 2 shows the semantic information of the four sentences placed on the upper left-hand corner. One of features of the expression is that each piece of information is stored separately and relationships of correspondence are expressed by links shown as arrows, so context such as referring relationship by pronouns, or the



Syntactic tree Feature structure

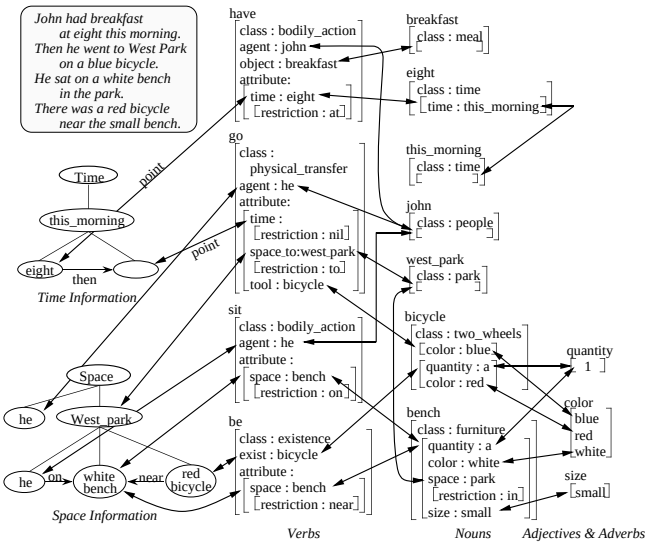


Figure 1: An example of syntactic information Figure 2: An example of semantic information

time order of some events can be used easily.

Our methods of question generation use the dictionaries of synonyms and antonyms about nouns, modifiers and verbs. We could not declare that a word and its synonym have the quite same meaning, so the module must take care to use synonyms. At present, the module treats only the words considered that their meaning are the same and used for training of rewording, e.g, "can" and "be able to".

3.2 Methods of question generation

To generate question sentences to achieve the four purposes described in 2.3, we have prepared five methods of generating question sentences. This section presents each method.

3.2.1 To ask about the content of one sentence

A question that asks the content represented by one sentence in a story is generated by one of the following methods.

(a) To ask a noun or the referent of a pronoun
(b) To ask a modifier
(c) To ask a verb
(d) To ask by transforming an affirmative sentence into an interrogative sentence
(e) To ask an important clause which follows when, because, that, relative pronouns and so on
(f) To ask an infinitive phrase
(g) To transform the passive voice into the active voice and apply one of previous methods (a) - (f)
(h) To select the sentence which has a topic and apply one of previous methods (a) - (g)

Table 2 shows relationship between the purposes of questions described before and the methods of question generation.

We present an example of a flow of question generation to ask a noun by using an interrogative pronoun.
(1) To extract the subject, the predicate, the object, the complement and modifiers of a sentence from its

Table 2: Purposes of questions and methods of question generation from one sentence.

Purposes of questions	Methods of question generation
to master the use of grammatical knowledge for interpretation	d, g
to master the use of grammatical knowledge for composition	a, b, c, e, f, g
to train for conversation through interpretation	a, h
to train for conversation through composition	a, h

Table 3: Semantic categories of nouns and interrogative pronouns

Semantic categories of nouns	interrogative pronouns
things	what
people	who
organization	what, where
time	when
location	where
physical	what, where
abstract	what
etc	what
animate	who
artificial	what, where
natural	what, where
language	what
animal	what
country	where, what
prefecture	where, what
city	where, what
school	where, what

Table 4: Semantic categories of verbs and interrogative pronouns

Semantic categories of verbs	Case	Interrogative pronouns
phenomena	agent	what
	object	what
phisical_movement	agent	what, who
	object	what
transfer	agent	what, who
	object	where, what
information	agent	what, who
	object	what
bodily_action	agent	who
	object	what, where
emotional	agent	what
	object	what
thought	agent	who
	object	what
event	attribute	what, where

Table 5: Roles of prepositional phrases and interrogative pronouns

Preposition	Case of prepositional phrases	Interrogative pronouns
from	origin	who, what
to	space	where
as	role	what
when	time	when
in	space	where
during	during	when

syntactic tree and the feature structure.

(2) To extract attributes of the sentence such as tense, voice, kind of sentence, idioms and auxiliary verbs from the feature structure.

(3) To replace pronouns with their referent by referring to the semantic information.

(4) To get candidates for the interrogative pronoun by referring to both the semantic category of the query part and Table 3. And to determine the interrogative pronoun from the candidates by referring to constraints such as between semantic categories of verbs and interrogative pronouns (see Table 4) and between roles of prepositional phrases and interrogative pronouns (see Table 5).

(5) To replace the query part with the selected interrogative pronoun and translate the sentence into an interrogative sentence

(6) To conjugate the verb in the interrogative sentence by referring to tense and person and number of the subject expressed in the feature structure of the syntactic information.

We give you an example of question generation by the method (a) to ask the subject "John" of the first sentence in Figure 2. Firstly, the method gets the case "agent" of the query part "John" from the semantic information of the verb "have", and the semantic category "people" of "John" from the semantic information of the noun "John". Secondly, the method gets "who" which is a candidate for the interrogative pronoun by referring to both the semantic category of "John" and Table 3. And the method also gets a candidate "who" by referring to both the semantic category "bodily_action" of the verb and Table 4. From these results, the method decides to use "who" as the interrogative pronoun. Finally, the method replaces the query part with the interrogative pronoun, transforms the sentence into the interrogative sentence and gets the question sentence "Who had breakfast at eight this morning?"

3.2.2 To use synonyms or antonyms

This method replaces a word with its synonym or antonym by referring to dictionaries of synonyms or antonyms and generates a question sentence by using one of the methods to ask about the content of one sentence described in 3.2.1. Relationship between the methods of question generation and the purposes of questions is the same as 3.2.1.

For example, the method generates "Is Jane free?" from "Jane is busy." by referring to the dictionary of

antonyms.

3.2.3 To use modifiers appeared in plural sentences

Semantic information of each noun is stored separately. So, even if different objects represented by the same word are stored as different instances in the same entry of semantic information of noun. Modifiers of the same object are stored into the same entry of an instance. Because information of modifiers which are expressed in different sentences is gathered into the same entry of an instance, it is easy to combine the modifiers of an object, which appeared separately, and generates question sentences. Moreover the method generates question sentences which have different contents with an original story easily by using modifiers of different instance or different words. The question sentences generated by such a way are more natural and difficult than to use the words which are not concerned with the original story because the words appeared in sentences in the story are used. This method of question generation is used for the purpose "to train for conversation through interpretation" or "to train for conversation through composition", depending on the query form.

For example, the method generates "Did John sit on a small white bench?" from the two sentences "John sat on a white bench." and "There was a red bicycle near the small bench." which have the same object "bench". The method also generates "Did John sit on a red bench?" which makes learners answer "No" by using modifiers of the different word.

3.2.4 To ask the contents represented by plural sentences by using a relative pronoun

This method combines the contents represented by two sentences by using a relative pronoun and generates a question sentence in the following way. Firstly, the method selects a main sentence and a subordinate sentence which have an antecedent, and selects a relative pronoun by referring to the case and the semantic category of the antecedent in the subordinate sentence. Secondly, the method takes the antecedent away from the subordinate sentence. Finally, the method combines two sentences and the relative pronoun and generates a question sentence. This method of question generation is used for the purpose "to train for conversation through interpretation" or "to train for conversation through composition", depending on the query form.

For example, "Was there a red bicycle near the small bench which John sat on?" is generated from the two sentences "John sat on a white bench." and "There was a red bicycle near the small bench."

3.2.5 To ask relationship of time and space

Time and space information are expressed as hierarchies as shown in Figure 2. Inclusion relationship information is expressed, vertically, and partial ordering in time information and positional relationships in space information, horizontally. This method generates, for instance, question sentences which ask when an event occurs or what event occurs in a place by using inclusion relationship. And the method generates such question sentences which ask what event occurs after an event by referring to partial ordering in time information. This method of question generation is used for the purpose "to train for conversation through interpretation" or "to train for conversation through composition", depending on the query form.

For example, the method generates a question sentence "What did happen this morning?" by referring to inclusion relationship in time information in Figure 2.

4. Evaluation

We have implemented a question generation module which adopts the methods of question generation described in the previous chapter. This chapter presents evaluation of the module from the following two kinds of views. One is to investigate how our methods of question generation cover collection books of problems on the market. The other is to investigate how the question sentences generated by the module are correct semantically.

4.1 Coverage of the kinds of questions

To assess our methods of question generation, we have investigated whether or not the methods correspond to kinds of questions in collection books of problems: "THRESHOLD" (Ferguson & O'Reilly 87) and "A self

learning class for English learning" (Horiguchi 98) for novices and "Chigasaki Method The Latest NEWS in English bimonthly " (Chigasaki 98) and "Special listening training for TOEFL" (Iwamura 95) for advanced learners. The questions investigated are the same form with the QA function, that is to say, questions are written in English and learners must write answers in English.

In the investigation, we have confirmed that our methods cover 160 of 201 question sentences (80%) in the collection books for novices. Table 6 shows the kinds of 160 question sentences. The table shows that 64% of whole question sentences is the kind "to ask about content of one sentence". So we can say that the simple questions which ask narrow range of the content of a story are generally used for novices. The 41 question sentences which can not be generated are the question sentences which require common sense to generate, which ask about implicit information and which use replaced words such as other wording than original sentences in a story.

We also found that our methods cover 42 of 99 question sentences (42%) in the collection books for advanced learners. The reason of the result is that such collection books have many questions which require to understand deeply to answer them, such as the question sentences which have summarized content or ask the implicit emotion of a character.

Table 6: Kinds of questions should be generated

Methods of question generation	Number of questions
to ask about content of one sentence	103
to use synonyms or antonyms	17
to use modifiers appeared in plural sentences	7
to ask the contents represented by plural sentences by using a relative pronoun	13
to ask relationship of time and space	20
Total	160

4.2 Semantic correctness of generated question sentences

To assess how the question generation module generates correct question sentences semantically, we verified the correctness of the question sentences generated automatically from sentences in textbooks "New Horizon" (Ota et al. 89) for second and third grade of Japanese junior high school. For question generation, the module excepts interrogative sentences, exclamatory sentences, imperative sentences, meaningless sentences such as "I see." or "You are welcome." and the words placed in subordinate clauses or prepositional phrases.

The sentences used for question generation is 176 sentences in the textbook for second grade and 125 sentences in the textbook for third grade. We excepted the method of questions "to ask relationship of time and space" because the natural language understanding module has not implemented completely yet and the module could not prepare complete information of time and space. The question generation module generated 1263 question sentences from the sentences in the textbook for second grade and 714 questions from the

Table 7: Methods of question generation and semantic correctness of generated question sentences

Methods of question generation		Number of generated questions	Number of correct sentences
to ask about content of one sentence	to ask a noun or the referent of a pronoun	443	408
	to ask a modifier	52	49
	to ask a verb	104	100
	to ask by transforming an affirmative into an interrogative	22	22
	to ask an important clause	12	12
	to ask an infinitive phrase	2	2
	to transform the passive into active	233	231
	to use a topic	129	118
to use synonyms or antonyms		309	294
to use modifiers appeared in plural sentences		473	417
to ask the contents represented by plural sentences by using a relative pronoun		198	195
Total		1977	1848

sentences for third grade. 1179 (93%) and 669 (94%) question sentences are correct semantically respectively. We present methods of question generation and semantic correctness of generated question sentences in Table 7.

The main reason of generation of incorrect question sentences is that the classification of semantic categories is little rough for question generation. For example, "Is Mike many years old?" was generated by replacing "thirteen" with "many" which is in the same semantic category "quantity".

5. Conclusions and Discussions

In this paper, we described necessity, implementation methods and evaluation of automated question generation for QA. We have employed the five methods for question generation: to ask about the content of one sentence, to use synonyms or antonyms, to use modifiers appeared in plural sentences, to ask the contents represented by plural sentences by using a relative pronoun, and to ask relationship of time and space. As an evaluation, we investigated questions in collection books of problems on the market. We have found that our methods cover about 80% of the questions for novices. Thus we can say that our methods are useful for novices. The methods can be used for not only QA but also supporting teachers to prepare paper tests and so on. As the other evaluation, we verified semantic correctness of the generated question sentences. As the result, 1848 (93%) of 1997 question sentences generated from sentences in textbooks for Japanese junior high school students are correct semantically.

Our methods do not use the knowledge which depends on particular stories, so the methods can generate question sentences as correctly as the result of our experiment if sentences which are level with the sentences in the textbooks we used and syntactic and semantic information of the sentences are prepared.

In our experiment, 7% of generated question sentences are incorrect semantically. The QA function should not give learners such questions. Thus to avoid the situation, we must subdivide the classification of semantic categories of words. The other short-term target of our research is to investigate, realize and evaluate a dialog strategy decision method to decide how and what to ask in what order for the adaptive QA function. To achieve the target, first, we need a method of calculating difficulty of a question for a learner by referring to syntax and semantics of the question and the learner's understanding state. Then we investigate a method to give learners suitable advice and questions by considering context of both a story and dialog in the past, educational targets and the calculated difficulty.

Acknowledgments

This work is supported in part by Grant-in-Aid for Scientific Research on Priority Areas (A) (1) No.12040104 from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

Chigasaki Method English Society (Ed.). (1998). *Chigasaki Method The Latest NEWS in English bimonthly.* Tokyo: Chigasaki Press.

DynED. (1997). *Dynamic English.* DynED Japan Co., Ltd.

Ferguson, N., Ferguson, C., & O'Reilly, M. (1987). *THRESHOLD unit 1-6.* Castle Publications SA. Hermance, Switzerland.

Horiguchi, T. (1998). *A self learning class for English learning.* Tokyo: Seishinsha.

Iwamura, K. (1995). *Special listening training for TOEFL.* Tokyo: ALC Press Inc.

Kunichika, H., Takeuchi, A., & Otsuki, S. (1995). An Multimedia Language Learning Environment with Intelligent Tutor. In Chan, T., & Self, J. (Eds.). *Emerging Computer Technologies in Education.* VA: Association for the Advancement of Computing in Education. Ch.10.

Levin, L., & Evans, D. (1995). ALICE-chan: A Case Study in ICALL Theory and Practice. In Holland, V., Kaplan, J., & Sams, M. (Eds.). *Intelligent Language Tutors: Theory Shaping Technology.* NJ: Lawrence Erlbaum Associates Inc. Ch.5.

Ota, A. et al. (1989). *New Horizon English Course 2, 3.* Tokyo: Tokyo Shoseki.