English Grammar Learning System Based on Knowledge Network of Fill-in-the-Blank Exercises

Takuya Goto[1], Tomoko Kojiri[1], Toyohide Watanabe[1], Takeshi Yamada[2], and Tomoharu Iwata[2]

[1] Graduate School of Information Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464-0803, Japan
{tgoto,kojiri,watanabe}@watanabe.ss.is.nagoya-u.ac.jp
[2] NTT Communication Science Laboratories, Japan

Abstract. To understand English grammar is essential to write/speak/read English appropriately. Fill-in-the-blank exercise of English grammar is one of the popular types of exercises, which is introduced to check acquired/in-acquired grammatical knowledge by evaluating the word selected by a learner for each sentence. Since such exercise-based learning is effective to acquire practical knowledge, our objective is to construct English grammar learning system using fill-in-the-blank exercises. In order to make learners study with the exercises effectively, the system should grasp the grammatical knowledge that learners need to obtain and give series of appropriate exercises. In our system, the knowledge network, which represents relations among exercises according to the differences in their knowledge, is introduced. In the knowledge network, exercises are linked by the inclusive relation of their knowledge. Based on the acquired/in-acquired grammatical knowledge determined by the learner's answer, the system traverses knowledge network and poses the exercise which contains the knowledge that the learner needs to study.

Keywords: Fill-in-the-blank exercise, English grammar, individual learning support, knowledge network.

1 Introduction

To understand English grammar is essential to write/speak/read English. The appropriate words are different for each situation, so it is necessary to study English grammar with various sentences. Fill-in-the-blank exercise of English grammar is one of the popular types of exercises, which is adopted in many tests for evaluating grammatical knowledge, such as TOEIC and TOEFL. In the exercises, learners select appropriate words/phrases with the correct grammar for the blank in the sentence from several choices. By tackling many exercises, learners acquire grammatical knowledge. Therefore, the objective of our research is to construct the English grammar learning system using fill-in-the-blank exercises. In order to make learners acquire grammatical knowledge effectively, the

I. Lovrek, R.J. Howlett, and L.C. Jain (Eds.): KES 2008, Part III, LNAI 5179, pp. 588–595, 2008.

system should grasp the knowledge which learners do not obtain and give a series of exercises corresponding to such knowledge. The effective learning is that the learner is able to notice his/her dis-understanding knowledge automatically. So, it is important to grasp learner's dis-understanding knowledge and give exercises based on it. However, acquired/in-acquired knowledge is different even when learners chose the same incorrect answer for the same exercise, since there is various grammatical knowledge to grasp in the sentence. So, if learners cannot answer a exercise, the system is not able to judge which grammatical knowledge is understood incorrectly. In our research, understanding degree is set for individual knowledge which represents ratio of correctness of exercises that include the grammatical knowledge. By defining the understanding degree, grammatical knowledge whose understanding degree is low is considered as in-acquired knowledge.

In learning support systems based on exercises, exercises are classified by their knowledge used to derive the answer. Kojiri, et al. proposed a mechanism for automatically generating answers to mathematical exercises and constructed network structure which represented the relation between mathematical exercises based on the steps of learners' solving process[1]. Ishima, et al. developed an intelligent tutoring system for high school chemistry and classified chemistry exercises into five types based on solving algorithms of exercsies[2]. In the solving process of English fill-in-the-blank exercises, since learners analyze grammatical structure of the whole sentence, it is necessary to classify exercises by grammatical knowledge that is acquired to solve the exercise.

In our system, exercises are categorized based on all grammatical knowledge which are included in the sentence of the exercise. Based on the inclusive relations of such grammatical knowledge of exercises, the knowledge network is introduced. The knowledge network represents relations between exercises according to the difference of grammatical knowledge. By traversing the knowledge network toward exercises that contain in-acquired grammatical knowledge, the system can give exercises that contain similar grammatical knowledge to the current exercise which holds in-acquired knowledge.

2 Learning by Fill-in-the-Blank Exercises

Fill-in-the-blank exercises of English consist of three parts: sentence, blank, and choices. When the learner solves the exercise, he/she firstly decides the main structure of the whole sentence by analyzing types of every word in it, such as verb or noun. The grammatical knowledge used to determine the structure of the whole sentence, i.e. sentence patterns, is called a *Sentence Structure Knowledge(SS-Knowledge)*. If choices consist of different types of words, the correct choice can be selected successfully by the SS-Knowledge. If there are choices of the same type of words, the learner has to consider not only the sentence pattern but also the appropriate declension or conjugation of particular words or phrases. The grammatical knowledge which decides such declension or conjugation is defined as *Partial Structure Knowledge(PS-Knowledge)*. Not only

PS-Knowledge of the blank in a sentence, but also that of related words to the blank are necessary.

In adaptive learning using exercises, the method of posing exercises is very important and various posing algorithms based on the understanding degrees of individual learners were developed [3,4,5]. Those algorithms change the difficulty of exercises from basic ones to practical ones step by step. Basic exercises consist of simple SS-Knowledge and contain a little PS-Knowledge. On the other hand, practical exercises are composed of complicated SS-Knowledge with a lot of different PS-Knowledge. Since it is necessary for learners to grasp all grammatical knowledge included in the exercise globally, such as PS-Knowledge and SS-Knowledge, difficulties of exercises are different for individual learners according to the combination of grammatical knowledge. So, for the purpose of making learners understand the new grammatical knowledge, exercises by which learners can focus on the target grammatical knowledge should be provided appropriately. That is, exercises that hold a lot of acquired grammatical knowledge and a little in-acquired grammatical knowledge may be effective to notify learners the new grammatical knowledge. Therefore, exercises to provide learners should be managed based on the acquired/in-acquired grammatical knowledge which they contain. For example, if a learner answers the exercise correctly, another exercise with one more PS-Knowledge is effective to acquire new knowledge. When a learner cannot answer the exercise, another exercise with one grammatical knowledge eliminated may help the learner to understand the in-acquired grammatical knowledge.

For the purpose of giving exercises along grammatical knowledge, the database of fill-in-the-blank exercises needs to be structured based on grammatical knowledge. In our approach, knowledge network is introduced, which indicates all combinations of grammatical knowledge existing in the database and their inclusive relations. Figure 1 shows the conceptual imagination of the knowledge network. In the knowledge network, the nodes that are composed of the smallest number of the least knowledge correspond to the most basic exercises. Paths from the root nodes to the goal nodes represent series of exercises for particular grammatical knowledge from basic ones to practical ones. By selecting appropriate nodes on the knowledge network based on the learner's answer and providing exercises that correspond the nodes, it is possible to support individual learners to study specific grammatical knowledge efficiently.

3 Knowledge Network

Knowledge network represents the relations between grammatical knowledge included in the database. Knowledge network is composed of nodes and links. Nodes correspond to the set of grammatical knowledge which consists of fill-in-the-blank exercises. Links indicate inclusive relations between nodes.

By analyzing the English grammar textbook[6], 5 types of SS-Knowledge and 24 types of PS-Knowledge are defined. Figure 2 is the example of attaching grammatical knowledge. "Subject(S) + Verb(V) + Direct Object(O)" is added to this

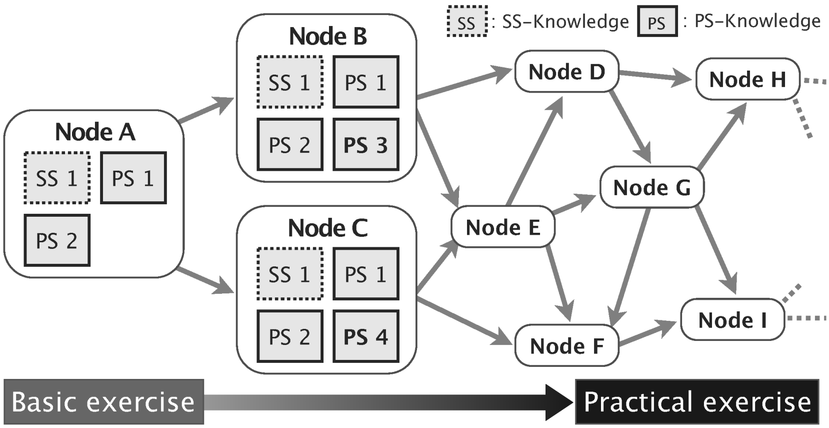

Fig. 1. Conceptual imagination of knowledge network

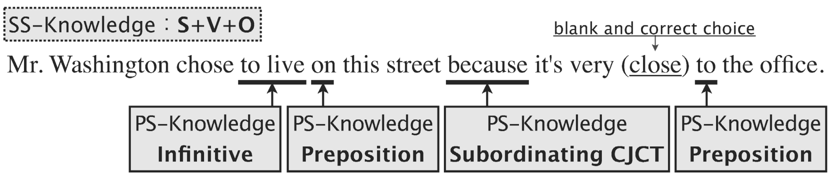

Fig. 2. Example of attaching grammatical knowledge

sentence as SS-Knowledge and "Infinitive", "Preposition", and "Subordinating conjunction(Subordinating CJCT)" are attached as PS-Knowledge.

Links between nodes indicate inclusive relations of grammatical knowledge. There are two types of inclusive relations: one is the inclusive relation of sentences and the other is the inclusive relation of words. The inclusive relation of words is defined by the embracement of existing PS-Knowledge. Since words included in exercises are different according to the given sentences, the amount and the types of PS-Knowledge differ according to them. If the amount of PS-Knowledge increases, more grammatical knowledge is needed to be understood. Therefore, links are attached to the nodes that have one more PS-Knowledge than the current node. On the other hand, the inclusive relation of sentences is derived from the relations of SS-Knowledge. SS-Knowledge represents the complicacies of sentences and some of one type of SS-Knowledge are extended by another one. For example, "S+V" needs to be understood before studying "S+V+O". Thus, links are also attached to nodes that contain the same PS-Knowledge and different SS-Knowledge of inclusive relations.

Figure 3 is the example of knowledge network. This network consists of four nodes. Nodes A and B, and nodes C and D have the same PS-Knowledge but different SS-Knowledge, such as "S+V+O" and "S+V+O+O". Since their SS-Knowledge have inclusive relations of sentence, nodes A and B, and nodes C and

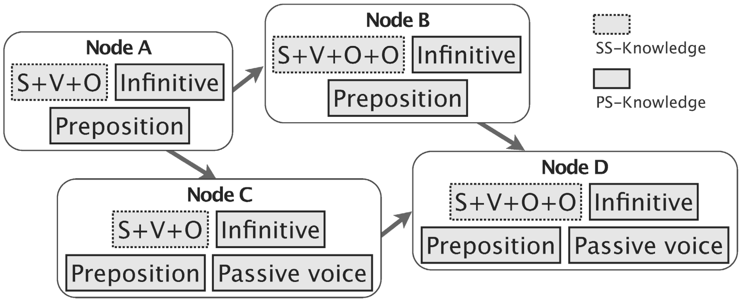

Fig. 3. Example of knowledge network

D are connected by links, respectively. On the contrary, nodes A and C contain the same SS-Knowledge, and PS-Knowledge of node C consist of that of node A and "passive voice". So, the link is attached from node A to node C. Similarly, another link is added from node B to node D.

4 Mechanism for Selecting Exercises

Based on the learner's answer, exercises which help the learner acquire dis-understanding knowledge are estimated. When the learner selects a correct answer, he may understand all grammatical knowledge included in the exercise. However, understanding of it can be inferred according to the word class of the choice that the learner selected, when he could not answer correctly. Table 1 shows the understanding/dis-understanding of knowledge based on the difference in word class of the learner's choice with the correct one. A learner tends to grasp the SS-Knowledge by observing the structure of the whole sentence and to find appropriate declension or conjugation from choices. Therefore, if the learner selected the wrong choices which have the different word class from the correct one, he/she may not understand the SS-Knowledge. On the other hand, when the learner selected the wrong choices with the same word class, he/she is expected to understand the SS-Knowledge, but may not be able to decide appropriate declension or conjugation. So, the PS-Knowledge in the blank is estimated as dis-understanding knowledge.

Table 1. Understanding/dis-understanding knowledge based on learner's choice

Learner's choice	Understanding knowledge	Dis-understanding knowledge
Different word class with right choice	-	SS-Knowledge
Same word class with right choice	SS-Knowledge	PS-Knowledge

According to the understanding/dis-understanding knowledge estimated by the learner's answer, understanding degree for each grammatical knowledge is determined. Understanding degree for the grammatical knowledge i, which is represented as c_i, is calculated as follows:

$$c_i = (s_i - f_i)/(s_i + f_i) \quad (-1 \leq c_i \leq 1) \tag{1}$$

c_i is defined so as to take the value between -1 to 1. When c_i is 1, the learner may understand the knowledge i, but if c_i takes -1, he/she does not understand the knowledge i at all. c_i is calculated based on s_i and f_i. s_i indicates the number of exercises by which the learner is estimated to understand the knowledge i. f_i corresponds to the number of exercises by which the learner is estimated no to understand it. For example, if the learner answered eight exercises that contain "Infinitive" correctly, $s_{\text{Infinitive}}$ becomes 8. When the learner selected incorrect choices of the PS-Knowledge for two exercises that also have "Infinitive", $f_{\text{Infinitive}}$ is set to 2. In this example, $c_{\text{Infinitive}}$ is calculated as 0.6.

The exercise to give to the learner is determined based on the understanding degrees for individual grammatical knowledge. If a learner selected a right choice and the understanding degree of all grammatical knowledge in the current exercise is higher, the system grasps that the learner understands all the knowledge successfully and tries to teach new grammatical knowledge. So, it follows the link from the corresponding node to the node in which one new type of grammatical knowledge whose understanding degree is the lowest is added to the current one. If the understanding degree of grammatical knowledge of the current exercise is lower, the system regards that the learner does not understand the knowledge and tries to teach it again by giving another exercise of the dis-understanding knowledge in the other node. On the other hand, when the learner answers wrongly and he/she has not answered all exercises of the current node, the system tries to teach the knowledge again and gives another exercise of the current node. If he/she has already answered all exercises, the system follows the link from the corresponding node to the node in which one type of grammatical knowledge whose understanding degree is the highest is removed. By traversing the knowledge network according to the learner's answer and understanding degrees, the learner can tackle exercises that contain grammatical knowledge that he/she does not understand deeply.

5 Evaluation

355 fill-in-the-blank exercises were prepared in the database and 29 grammatical knowledge is applied. Since links based on inclusive relation of PS-Knowledge are attached from nodes to nodes which contain one additional type of PS-Knowledge, 104 knowledge networks were generated. Of all knowledge networks, 103 were smaller-scale networks that have less than 10 nodes. The last one is a larger-scale network which consists of 54 nodes. In order to make learners study acquired/in-acquired grammatical knowledge, it is necessary to construct knowledge network using all grammatical knowledge. By connecting nodes that

have inclusive relation of more than two PS-Knowledge, all exercises can be structured as one knowledge network.

We implemented an web-based English grammar learning support system based on our mechanism for selecting exercises. The knowledge network of larger-scale was applied. After a learner selects a choice through the web page, the system displays the result of learner's answer and the explanation to derive the correct answer. The system, then, determines the next exercise. If the system cannot select appropriate exercise because current node is the root node, the leaf nodes or the node that does not contain enough exercises, it finds the new node that contains the knowledge whose understanding degree is the lowest.

The experiment to evaluate the effectiveness of our selection mechanism is executed. Sixteen students in our laboratory were asked to participate in the experiment. The experiment consisted of three phases: pretest phase, learning phase, and posttest phase. In the pretest phase, participants were asked to answer twenty exercises. During the pretest, neither the results of their answers nor the right answer were displayed. In the learning phase, participants were divided into two groups: namely eight participants for each group, and participants of each group were given twenty exercises. Participants in one group studied using our system (*group 1*), and those of the other group used the system which randomly selects the exercise from the database (*group 2*). For the group 1, the system grasped the understanding degrees of participants based on results of their pretests and selected the start node for individual participants. In the learning phase, both results of their answers and the explanation to solve each exercise were shown. In the posttest phase, all participants tackled the same exercises provided with the pretest again, but the orders of posing exercises were different.

Table 2. Average scores in pretest and posttest

	Group 1	Group 2
Pretest	15.9	15
Posttest	16.1	15.9

Table 3. Average values of increasing understanding degrees for knowledge that answered incorrectly in pretest

	Group 1	Group 2
Increase of understanding degrees	42.3%	21.8%

Table 2 indicates the average scores in the pretest and the posttest. The changes of the total scores did not show significant differences between both groups. On the other hand, Table 3 indicates the average values of increasing understanding degrees for all grammatical knowledge that participants were not answered correctly in the pretest phase. For the group 1, understanding degrees for dis-understanding knowledge were increased about 42.3%, while 21.8% for the group 2. For both groups, almost all grammatical knowledge which were incorrectly answered were provided in the learning phases. Therefore, from the result, our system can give appropriate fill-in-the-blank exercises, which helps learners to study in-acquired grammatical knowledge effectively.

During the learning, the system could not select exercises from the corresponding node for five times because of the lack of the number of exercises and links. In order to provide appropriate exercises, more exercises should be prepared in our database.

6 Conclusion

In this paper, we proposed a framework of learning English grammar using fill-in-the-blank exercises. We classified exercises based on their grammatical knowledge and represented knowledge of the database as the knowledge network. Based on the experimental result, our selection mechanism and knowledge network was able to give appropriate exercises and was effective for studying dis-understanding knowledge step by step.

Currently, the selection methods of exercises are common to all learners. However, the progresses of understanding the knowledge are different from each learner. If learners can understand the knowledge with a few explanations, they may need a few exercises for one grammatical knowledge. On the other hand, if learners tend to understand the knowledge slowly, they may prefer to tackle the exercises of the same knowledge repeatedly. Thus, the mechanism to navigate the knowledge network according to the learner's learning tendencies should be developed.

References

1. Kojiri, T., Hosono, S., Watanabe, T.: Automatic Generation of Answers for Complex Mathematical Exercise. In: Proceedings of ICCE 2006, Workshop, pp. 25–32 (2006)
2. Ishima, N., Ueda, T., Konishi, T., Itoh, Y.: Developing Problem Representation, Knowledge Representation and Problem Solving System for Intelligent Educational System of High School Chemistry. In: Proceedings of ICCE 2006, Workshop, pp. 41–48 (2006)
3. Fischer, S.: Course and Exercise Sequencing Using Metadata in Adaptive Hypermedia Learning Systems. ACM Journal of Educational Resources in Computing 1(1), 5 (2001)
4. Scheiter, K., Gerjets, P.: The Impact of Problem Order: Sequencing Problems as a Strategy for Improving One s Performance. In: Proc. of 24th Annual Conference of the Cognitive Science Society, pp. 798–803 (2002)
5. Cooley, R.E., Abdullah, S.C.: Controlling Problem Progression in Adaptive Testing. In: Proc. of ICCE 2000, vol. 1, pp. 635–642 (2000)
6. Quackenbush, E.: New Direction of English. Translation by Okada, N., Biseisya, Japan (in Japanese) (2000)