Proceedings of the IASTED International Conference
APPLIED INFORMATICS
February 14-17, 2000, Innsbruck, Austria

ACQUISITION SYSTEM OF DEEP CASE USING QUERIES AND REPLIES

YOSHIYUKI MASATOMI, KENJI ARAKI and KOJI TOCHINAI
Graduate School of Engineering, Hokkaido University
N13 W8, Kita-ku, Sapporo, 060-8628 Japan
TEL: +81-011-706-6823 FAX: +81-011-709-6277
E-mail: {tome, araki, tochinai}@media.eng.hokudai.ac.jp

ABSTRACT

In natural language processing, case frames are treated as essential in expressing the meaning of sentences. A deep case expressing semantic relation between a verb and its case is important for the analysis of sentences. In the case of use of deep cases for semantic analysis, it is an important problem that they have ambiguity. Each user has various classifications of deep cases. In the case that the system has not acquired a deep case, the system asks the user what kind of the deep case. Then, the system acquires deep cases using queries and replies to adapt to a user. In this paper, we propose the acquisition system of deep case which adapts to a user using queries and replies. Then, we carried out the experiment for evaluating the system. We got the result which was effective to decide deep case which adapts to a user.

Key Words: Deep Case, Ambiguity, Queries and Replies, Semantic Relation

1 INTRODUCTION

It is used two classes of cases to express the structure of sentences, that are surface cases and deep cases. For the analysis of sentences in natural language processing, the deep cases are important, because the deep cases can express semantic relations which the surface cases can not express. Therefore, it is necessary to determine deep cases of words in the semantic analysis of a natural language processing system.

A deep case expresses the semantic relation between a verb and the other words in a sentence. However, the definition of deep cases has ambiguities, that there are no outstanding opinions about the kinds of deep cases, the standards to determine a deep case in a sentence, and etc. Accordingly, when the deep cases are fixed in a natural language processing system, a user has to use the given deep cases even if they differ from the user's definition. Therefore the user's satisfaction to the system becomes low.

We consider that there are two types of case relations. One is the typical case relation that the judgments of almost all users correspond. The other is the case relation that has ambiguities. In the case, the judgments of them are different each user. They are not always same in the specific user. These ambiguities are called "the user's ambiguity" in this paper. In the past, the deep case decision methods do not allow the user to have the user's ambiguity.

To increase the user's satisfaction, we propose a deep case decision method in this paper. The system based on our proposed method uses queries and replies to adapt case decisions to a specific user. And the system allows the user to have the user's ambiguity. Because we consider that the allowance of the user's ambiguity is necessary, for example, in the semantic comprehension system adapting to a specific user. This proposed system analyzes the Japanese sentences and acquires the deep cases between words of them using information of morphological analysis and parsing. The 11 kinds of deep cases are used in our proposed method.

First, the system acquires information of the deep case using queries and replies. Next, the system determines a deep case for a relation with a particle between a noun and a verb in a new sentence by using information of the deep case obtained. Determining all deep cases using queries and replies needs heavy labor because there are a lot of relations with particles between nouns and verbs. Therefore, in the case that there are lack of information of the deep cases about the noun and the verb, this system predicts and determines the deep cases by the usage of the particle. And the system uses the thesaurus to the deep case decision. Therefore, this system is able to acquire deep cases efficiently.

We also describe the experimental results of the evaluation of this system. In this experiment, the system acquires information of the deep cases in 1,600 Japanese sentences. We evaluate this system by using the last 400 Japanese sentences because the system adapts to a specific user from 1,201 to 1,600 Japanese sentences.

Table 1: Kinds of deep cases

Deep Cases	Meaning
AGENT	Thing that causes action
OBJECT	Object of action
PLACE	Place where action is taken
INSTRUMENT	Instrument for action
CAUSE	Cause of action
TIME	Time for action
BENEFICIARY	Thing that action is taken for
MANNER	Manner of action
MATERIAL	Material of action
SOURCE	Source of action
GOAL	Goal of action

2 OUTLINE

Figure 1 shows the outline of the processing process. First, the input Japanese sentence is analyzed in the morphological analysis system and the parsing system. Information on the noun and the verb with a particle is obtained by both analyses of the morphological analysis and parsing. Next, this system processes it in the deep case decision part. Deep cases are decided by using the deep case decision rule in this part. The one that a deep case can be decided in the deep case decision part outputs a deep case and the one that it is not possible to decide processes it in the deep case acquisition part. The user judges the output deep case. In the following, details of each processing are described.

2.1 MORPHOLOGICAL ANALYSIS AND PARSING

Japanese morphological analysis system JUMAN [5] is used about the processing of morphological analysis of this system.

The parsing processing in this system uses Japanese parsing system KNP [6]. As a result, information on the noun and the verb with a particle is obtained.

2.2 DEEP CASE DECISION RULES

"Deep case decision rules" are rules to decide deep cases of the verb and the noun with a particle obtained by this research. In this research, the system treat with only relation that can decide a deep case from the relation between a noun and a verb with a particle in classification of deep cases. The acquired rules are accumulated by the combination of [noun]-[particle]-[deep case] and [verb]-[particle]-[deep case]. The reason why the deep case decision rules are described separately for the noun and the verb is to have considered that the system can decide deep cases when one of the noun and the verb does not exist in the deep case decision rules. Table 2 and table 3 show the example of deep case decision rules.

2.3 DEEP CASE DECISION PROCESS

It is very difficult to decide deep cases using queries and replies with a user about the relations between all nouns and verbs. When there is one or no deep case decision rule

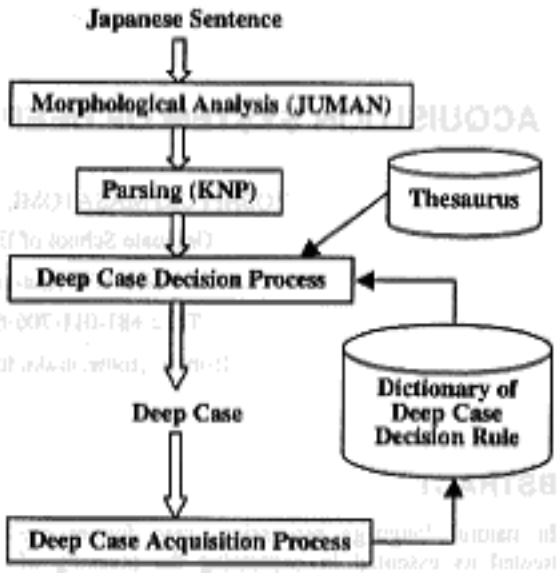

Figure 1: Process

about the noun and the verb in new input sentence, the system decides deep cases by using deep case decision rules and thesaurus [3].

In the case that the deep case decision rules of the verb and the noun exist, the deep case can be decided by these rules.

Next, the procedure in the case that the deep case decision rules of the verb only exist is explained. First, the deep case which can be obtained from the relation between the particle and the verb in the deep case decision rules is chosen. At this time, the system judges whether the kind of the deep case which can be extracted is singular or plural. In the case that the kind of the deep case is singular, the system decides its deep case. In the case that the kind of the deep case is plural, the system searches the word in the deep case decision rules that is similar in the meaning to the input noun using the thesaurus. If there is the similar word, the deep case is decided by using the deep case decision rules of the word. If there is not the similar word, the system processes it in the deep case acquisition process.

In the case that the deep case decision rules of both the verb and the noun do not exist, the procedure of the deep case decision is explained as follows. First, the system searches the input verb and noun in the thesaurus. When there are the input words in the thesaurus, the system searches the words in the deep case decision rules that are similar in the meaning to the input words using the thesaurus. If there are the similar words, the deep case is decided by using the deep case decision rules of the words. If the similar words do not exist, the system processes them in the deep case acquisition process.

Table 2: Examples of deep case decision rules (noun)

Noun (English)	Particle:Deep Case
watashi (I)	*ga*:AGENT *ni*:BENEFICIARY *wa*:AGENT *wo*:OBJECT
kare (he)	*ga*:AGENT *ni*:BENEFICIARY *wa*:AGENT *wo*:OBJECT
keisanki (computer)	*de*:INSTRUMENT *wo*:OBJECT
ginko (bank)	*ni*:OBJECT *ni*:PLACE
konchu (insect)	*wa*:AGENT *wo*:OBJECT
denwa (telephone)	*ga*:OBJECT *de*:INSTRUMENT *wo*:OBJECT *wa*:OBJECT
douro (road)	*wo*:OBJECT *wo*:SOURCE *de*:SOURCE
sensou (war)	*de*:CAUSE *wo*:OBJECT
eki (station)	*de*:SOURCE *ni*:GOAL *wo*:SOURCE

Table 3: Examples of deep case decision rules (verb)

Verb (English)	Particle:Deep Case
ataeru (give)	*ga*:AGENT *ni*:BENEFICIARY *wa*:AGENT *wo*:OBJECT
wataru (cross)	*de*:INSTRUMENT *wa*:AGENT *wo*:PLACE
yaku (burn)	*de*:INSTRUMENT *wo*:OBJECT *wa*:AGENT
taberu (eat)	*ga*:AGENT *de*:PLACE *ni*:TIME *de*:INSTRUMENT *ni*:PLACE
soujisuru (clean)	*de*:INSTRUMENT *wa*:AGENT *wo*:PLACE
hairu (enter)	*ga*:AGENT *de*:INSTRUMENT *ni*:GOAL *heGOAL*:
ueru (plant)	*wo*:OBJECT *ni*:PLACE *wa*:AGENT
deru (go out)	*kara*:SOURCE *wo*:SOURCE *ga*:AGENT

2.4 DEEP CASE ACQUISITION PROCESS

Processing in the deep case acquisition process is as follows:

- In the case that it is not possible to decide deep cases in the deep case decision process,
- In the case that the user judged that it was error about the output deep case.

In the deep case acquisition process, deep cases are acquired using queries and replies with a user. In this process, the system asks the user which deep case is correct. The user answers it. Therefore, the deep case is decided and the system adds new rules to the dictionary of deep case decision rule. Figure 2 shows the example of this process. In this example, the input sentence is a Japanese one. First, the morphological analysis is done. Second, the parsing is done. Then, if it is not possible to decide deep cases in the deep case decision process, the system asks the user about the kind of deep case. The user answers it, and so the system adds deep case decision rules showing in figure 2.

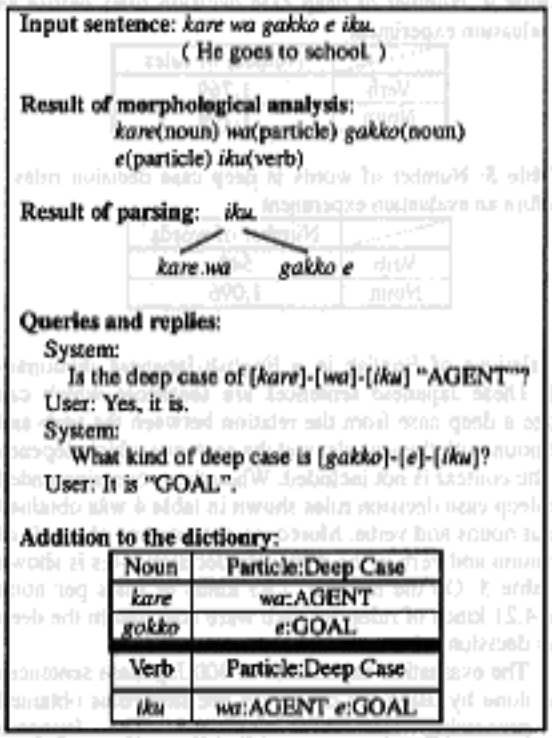

Figure 2: Example of deep case acquisition process

3 EXPERIMENT

The deep case decision rule is empty in the initial state. The input sentence is given, and then the system acquires the deep case decision rules by using queries and replies with a user. First, the system learned about the input 1,200 Japanese sentences and obtained the deep case decision rules. Next, the evaluation experiment with 400 Japanese sentences was done, and then the effectiveness of the system was verified. These Japanese sentence are the active voice.

3.1 EXPERIMENT METHOD

The examinee of this experiment is a university student of the science course and a male. First of all, 1,200 Japanese sentences were input and the morphological analysis and parsing were done. The errors of the morphological analysis result and the parsing result was corrected by a user. The results of both analyses after correcting was input to the system and it acquired the deep case decision rules. These Japanese sentences are

Table 4: Number of deep case decision rules before an evaluation experiment

	Number of rules
Verb	1,769
Noun	1,708

Table 5: Number of words in deep case decision rules before an evaluation experiment

	Number of words
Verb	548
Noun	1,096

translations of English in a English-Japanese dictionary [7]. These Japanese sentences are sentences which can judge a deep case from the relation between the verb and the noun with the particle, and the sentence which depends on the context is not included. When the processing ended, the deep case decision rules shown in table 4 was obtained about nouns and verbs. Moreover, the number of words of the noun and verb in the deep case decision rules is shown in table 5. On the average, 1.95 kinds of rules per noun and 4.21 kinds of rules per verb were acquired in the deep case decision rules.

The evaluation experiment of 400 Japanese sentences was done by using the deep case decision rules obtained by processing the above-mentioned. The Japanese sentences used in this experiment are translations of English in English-Japanese dictionary [7]. First, the morphological analysis and parsing were done and the errors of both analysis results were corrected by the user. The evaluation experiment by this system was done by using the correct result of the morphological analysis and the parsing. As for the correctness of the deep case decision, it is the one by the judgment of the user.

3.2 EXPERIMENT RESULTS

Table 6 and table 7 show the evaluation experiment results. Table 6 shows the correct answer rates by the existence of the noun and the verb in the deep case decision rules. As for the result as a whole, the correct answer rates became 72.8%. Table 7 shows the correct answer rates in the case that at least one of the noun or verb existed in the deep case decision rules. Table 8 shows the number of deep case decision rules after the evaluation experiment with 400 Japanese sentences finished. Compared with the state before the evaluation experiment, the number of rules of verbs increased by 540 kinds and the number of rules of nouns increased by 428 kinds. Moreover, the numbers of nouns and verbs in the deep case decision rules are shown in table 9. Compared with the state before evaluation experiment, the verbs increased by 252 words and the nouns increased by 320 words. On the average, 1.51 kinds of rules per noun and 2.89 kinds of rules per verb were acquired in the deep case decision rules after the evaluation experiment.

4 DISCUSSION

In this section, we consider the results of evaluation experiment about table 6 which is classified by the existence of the noun and the verb in the deep case decision rules.

In the case that the deep case decision rules of both the verb and the noun exist, the system outputs the result according to the deep case decision rules because both the relation between input verb and particle and the relation between input noun and particle exist in them. Then, the user judges the output deep case correct because the relation between input verb and noun with particle was once acquired by the user. However, the user judges that the output deep case is erroneous by user's ambiguity about deep cases. In this experiment, the output result of one kind was assumed by user's ambiguity in the combinations of 250 kinds of nouns and verbs and it was corrected. According to the result of this experiment, this System decided deep case which adapts to the user. Figure 3 shows the example of user's ambiguity. In this example, the relation between the verb "*neru*" and the noun "*beddo*" with the particle "*de*" was decided the deep case "INSTRUMENT" by processing the sentence of (a) in figure 3. Therefore, the deep case result of the output is "INSTRUMENT" for the input sentence of (b) in figure 3. However, the user judged that a deep case at this time was not "INSTRUMENT" and it was "PLACE".

In the case that the deep case decision rules of the noun only exists, there were 258 kinds of the combinations (the verb and the noun with the particle) in this case.Their output deep cases included two errors. These errors were by user's ambiguity. There were 15 kinds of the

Table 6: Correct decision rates by existence of deep case decision rules

Deep Case Decision Rules	Correctness	Error	Impossible to Decide	Total	Correct Decision Rates(%)
Verb and Noun Exist	249	1	0	250	99.6
Only Noun Exists	241	2	15	258	93.4
Only Verb Exists	140	10	25	175	80.0
Verb and Noun Never Exist	43	15	183	241	17.8
Total	673	28	223	924	72.8

Table 7: Correct decision rates in the case that deep case decision rules of verb or noun existed

Deep Case Decision Rules	Correctness	Error	Impossible to Decide	Total	Correct Decision Rates(%)
Verb or Noun Exists	630	13	40	683	92.2

Table 8: Number of deep case decision rules after an evaluation experiment

	Number of rules
Verb	2,309
Noun	2,136

Table 9: Number of words in deep case decision rules after an evaluation experiment

	Number of rules
Verb	800
Noun	1,416

(a) *kanojo wa itsumo beddo de neru.*
(She always sleeps in the bed.)
(b) *kare wa asa made beddo de neru.*
(He sleeps in the bed until morning.)

Noun	Particle	Verb
beddo	*de*	*neru*
INSTRUMENT ⟶		PLACE

Figure 3: Example of user's ambiguity

combinations that two or more deep cases were able to be taken. Because the system could not decide a deep case by using thesaurus, the queries and replies with the user in the processing of the above-mentioned deep case acquisition process was performed and deep cases were decided.

In the case that the deep case decision rules of the verb only exists, there were 175 kinds of the combinations in this case. Their output deep cases included ten errors. These errors were occurred by lack of deep case decision rules. Moreover, there were 25 kinds of the combinations that two or more deep cases were able to be taken. Because the system could not decide a deep case by using thesaurus, the queries and replies with the user in the processing of the above-mentioned deep case acquisition process was performed and deep cases were decided.

In the case that the deep case decision rules of both the verb and the noun never exist, there were 241 kinds of the combinations. 58 kinds of output deep cases included fifteen errors. These errors were occurred by lack of deep case decision rules. There were 183 kinds of the combinations that the system were not able to decide a deep case using thesaurus. These were also processed in deep case acquisition process.

5 CONCLUSION

The acquisition system of deep case which adapts to a specific user using queries and replies was proposed in this paper. We confirmed that this system was effective to decide deep case which adapts to a specific user and deep cases of verbs and nouns with particles from the results of the evaluation experiment. Especially, deep cases were decided without using queries and replies with a user in the case that one of the noun or the verb was existed in the deep case decision rules by high accuracy.

The principal cause which the system can not decide a deep case is what there is no word in deep case decision rules that is similar to the input word. However, increasing the input sentences, decreasing the number of combinations not to be able to decide. Moreover, when the word that wants to judge in the thesaurus not registered, this system cannot decide deep cases even if the input sentences are increased. To solve this problem, the automatic expansion method of the thesaurus by the automatic addition registration of the word from acquired deep cases information to the thesaurus is examined.

Whether ambiguity between users can be absorbed with this system is scheduled to be verified as a research problem in the future. The expressions of passive voice and so on are examined.

ACKNOWLEDGEMENTS

This work is partially supported by the Grant-in Aid for Scientific Research (C) (No.10680367) of Japan Society for the Promotion of Science.

REFERENCES

[1] A.Ohishi and Y.Matsumoto, Lexical knowledge acquisition for Japanese verbs based on surface case pattern analysis, *Transactions of Information Processing Society of Japan*, 36(11), 1995, 2597-2610.

[2] Y.Namioka, N.Uratani and T.Aizawa, The method of deep case extraction in news sentences, *Information Processing Society of Japan SIG Notes*, 77(3), 1990.

[3] National Language Research Institute Publications, *Word List by Semantic Principles* (Tokyo: Shuei Shuppan, 1964).

[4] Y.Masatomi, K.Araki and K.Tochinai, Deep case acquisition system by dialogue with users, *IEICE Technical Report*, NLC98-35, 1998.

[5] S.Kurohashi and M.Nagao, Japanese morphological analysis system JUMAN version 3.5, Graduate School of Engineering, Kyoto University, 1998.

[6] S.Kurohashi, Japanese parsing system KNP version 2.0 b6, Graduate School of Information Science, Kyoto University, 1998.

[7] Y.Ogawa et al, *SUNRISE English-Japanese Dictionary* (Tokyo: Obun Sha, 1986).

[8] R.L.Liu and V.W.Soo, An empirical study on thematic knowledge acquisition based on syntactic clues and heuristics, *Proceedings of the 31th Annual Meeting of ACL*, 1993, 243-250

[9] M.Nagao, *Shizen Gengo Shori - The Iwanami Software Science Series 15* (Tokyo: Iwanami Shoten, 1996)