ISCA Archive
http://www.isca-speech.org/archive

3rd International Conference on
Spoken Language Processing (ICSLP 94)
Yokohama, Japan
September 18-22, 1994

A SEMANTIC INTERPRETATION BASED ON DETECTING CONCEPTS FOR SPONTANEOUS SPEECH UNDERSTANDING

Akito Nagai Yasushi Ishikawa and Kunio Nakajima

MITSUBISHI Electric Corp. Computer and Information Systems Lab.
5–1–1, Ofuna, Kamakura, Kanagawa 247, Japan

ABSTRACT

This paper describes a two-stage semantic interpretation method where an intention is associated with an input sentence and it integrates concept hypotheses detected from the sentence. In this approach, a *concept* represented by several phrases is an unit of semantic interpretation, and a spontaneously spoken sentence is regarded as a sequence of *concepts*. Syntactic constraint is limited within a phrase to robustly understand various sentential expressions. For disambiguation of concept hypotheses, they are evaluated with scoring criteria based on linguistic knowledge. The method has been evaluated on test sets collected as text data including various type of sentences which subjects made with no limitation. The experimental results of semantic understanding using 22 concept frames with approximately 1000-word vocabulary shows that the scoring method contributes to disambiguation of meaning candidates, and that 75 % of test sentences are correctly understood.

1 INTRODUCTION

Many intensive researches have been done towards the development of a spoken language system which realize successful communication between a naive user and an intelligent dialogue system [1][2].

To realize linguistic robustness is one of the most important issues of understanding spontaneously spoken language. When a naive user communicates with a dialogue system, he will talk about items of task information freely with a large variety of ill-formed expressions. How does a language model cover such a variety of sentences with sufficient constraints for continuous speech recognition ?

Syntactic constraint which defines sentential patterns grammatically is contribute to improve performance of speech recognition. But this approach is fragile and no longer robust because it limits sentential expressions and hardly has a sufficient coverage of various utterances of a naive user.

Several recent works have tried to solve this linguistic robustness problem by taking an approach of relaxing grammatical constraints and partially parsing technique [3][4][5]. This approach exploits syntactic constraint on an ordinal relation between phrases which occupy a part of a sentence. In Japanese, however, partial parsing technique may not be successful because the order of Japanese phrases (*bunsetsu*) is considered to be free, although expressions of *bunsetsu* itself are well represented by *bunsetsu* grammar. Thus, a key issue for attaining linguistic robustness in Japanese spontaneous speech is how to exploit semantic knowledge to represent relations between *bunsetsu*'s by semantic-driven processing.

One of the methods of representing a meaning of a sentence is to use semantic frames which define case relations between *bunsetsu*'s from a viewpoint of the predicative usage. In this approach, it is difficult to define suitable semantic frames to a meaning of the sentence, since various combinations of concepts for task information can be included in a sentence of a naive user. Such a frame that covers all combination of concepts reduces semantic constraints, and leads to the explosion owing to both word sense ambiguity and many candidates of speech recognition. Therefore, a framework to evaluate growing meaning hypotheses is indispensable in the process of semantic interpretation from a sequences of phrases to a meaning representation.

Our approach is based on an idea that a task item represented by a partial expression can be an unit of semantic interpretation. We call this unit a concept. The followings are bases of this method; (1) a concept is represented by several words or phrases continuously uttered in a part of a sentence, (2) a spontaneously spoken sentence is regarded as a sequence of concepts, and (3) a sequence of concepts are produced to represent an intention of the sentence. We believe that syntactic constraint limited only inside a phrase will be helpful for realizing robustness, and that the explosion of a meaning hypothesis will be avoided by treating a concept as a target of semantic constraints.

2 SEMANTIC INTERPRETATION METHOD

The proposed method has two stage of semantic processing (Fig. 1). At the first stage, a system hypothesizes an intention of the input as an intention frame, considering predicates and semantic relations between phrases in the early stage of understanding process. The intention frame constraints a set of concept frames. At the second stage, the system detects concepts from a phrase lattice using concept frames. Slots of the concept frame will be filled with words or phrases which can be concatenated in the lattice, specifying the phrases if semantic values of them are suitable as constituents of the concept. Meaning hypotheses of an entire sentence is finally formed by the intention frame filled with concept frames.

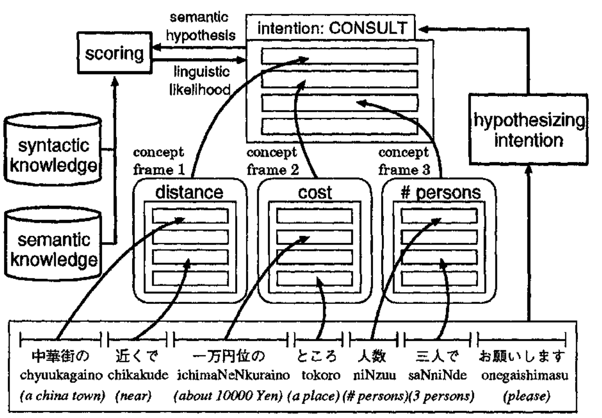

Figure 1: Principle of the semantic interpretation method based on detecting concepts

To reduce ambiguity of the meaning hypotheses, some conditions of existence of a concept and scoring based on syntactic/semantic knowledge are used. The conditions are required, for example, that a concept should occupy a continuous part of a sentence. Concerning the scoring knowledge, linguistic likelihood of a set of concepts is estimated using scoring rules considering both syntactic and semantic dependency between case slots.

The advantageous features of the method are considered to be followings: (1) a better coverage of sentential expressions than syntactic rules of a sentence, (2) less increase in meaning hypotheses by filling concepts to the intention frames than by filling directly *bunsetsu*'s to general case frames, and (3) portability of common concepts which are defined to be shared for different tasks.

2.1 Hypothesizing Intention Types

The decision of an intention of user's utterance includes two main purposes;

1. to transfer an information of user's request to a system as an instruction like, for example, consulting a database with some conditions, inputting a value of a certain item, changing a value of the item.
2. to integrate concepts using a relation between an intention and concepts. If sorts of concepts are defined so as to depend on the intention, the relation can be represented by a structure of the intention frame.

The intention is defined as a semantic frame which treats concept hypotheses as slot fillers. The types of intentions are defined by analyzing a target task. For example, intention types in "Hotel Reservation" task are defined as *reservation*, *change*, *cancel*, *WH-inquiry (when, where, how, etc.)*, *Y/N-inquiry (yes or no question)* and *consultation*. A type of them is hypothesized for each sequence of phrase candidates.

Fig. 2 shows an example of the procedure. To hypothesize the intention type, morphological features are used; key predicates (verbs and adjectives) which relate semantically to each intention, a particle "ka" standing for an inquiry sentence and interrogative adverbs. If a key predicate is omitted in an input sentence, the intention type is guessed using task knowledge of semantic relation between concepts. Fig. 3 shows the conceptual relation as a hierarchy of indispensable information of reservation (a date, a room, the number of persons) and attributions of a hotel.

1. a key predicate is found in a sequence of phrase candidates.
   - (a) if the predicate is related to *reservation* or *change* or *cancel*, then the intention type is decided corresponding to the predicate.
   - (b) else (this case is about a consultation of a task database.)
     - i. if interrogative adverbs are found, then intention type is *WH-inquiry*.
     - ii. else
       - A. if a particle "ka" is found, then intention type is *YN-inquiry*
       - B. else *consultation*.
2. if a key predicate is not found, (omission of predicate)
   - (a) if indispensable information of reservation is found, then intention type is *reservation*.
   - (b) else
     - i. if a name of hotel is found,
       - A. if attributions of hotel are found, then go to 1.(b)
       - B. else intention type is *reservation*.
     - ii. else
       - A. if attributions of hotel are found, then go to 1.(b)
       - B. else (rejection).

Figure 2: An example of a procedure for hypothesizing user's intention types in *Hotel Reservation* task.

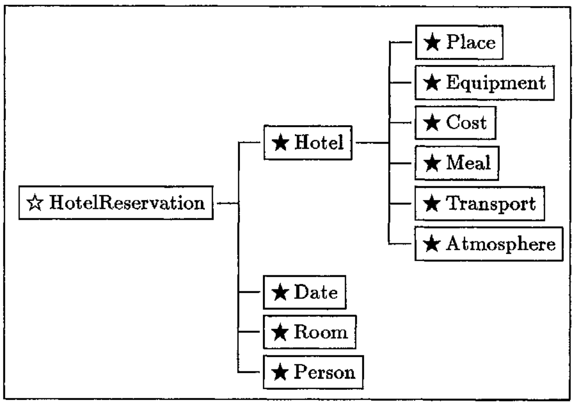

Figure 3: A semantic relation of *Hotel Reservation*, ☆: intention, ★: concept

2.2 Detecting Concepts

Concepts as an unit of semantic understanding are defined to represent semantic contents of a target task. Tab. 1 shows an example of concepts for representing *Hotel Reservation* task. A concept representation is based on a semantic frame which has several case slots with semantic values and particles as shown in Tab. 2. Only case slots which are concern semantically with a concept are defined.

In detecting concepts from a phrase lattice, filling slots based on examining semantic value and a particle is done with phrase candidates which can be concatenated at each phrase segment. A phrase candidate which has no particle is examined using only semantic value. Because this phrase candidate has case-level ambiguity, each case is hypothesized. If a phrase candidate has possibility to be filled to plural slots, each concept hypothesis is produced. To realize robust understanding, a concept hypothesis where all slots are not filled is allowed to exist.

Finally, concept hypotheses are combined using intention frames which are associated with sequences of phrase candidates. All meaning hypothesis for an entire sentence are generated as the intention frames which have slots filled with concept hypotheses.

This is the principle of detecting concepts. Many senseless hypotheses of a meaning, however, remain owing to ambiguities of word sense, case slots of a phrase and concept segmentation. Therefore, meaning hypotheses are evaluated by conceptual disambiguation techniques as described the next section.

Table 1: An example of concepts for representing *Hotel Reservation* task.

Distance, Place, TakeTime, UseTraffic, Equipment, UseEquipment, Cost, Meal, View, Atmosphere, ReserveDate, ReserveStay, ReserveHotelname, ReserveRoom, ReservePerson, ChangeDate, ChangeStay, ChangePerson, ChangeHotelname, ChangeRoomtype, ChangeRoomNum, Cancel

Table 2: An example of concept representation for *Distance*

Concept: Distance		
case slot	semantic value	particle
agent	*place	wa, ga
modifier	*place	no
source	*place	kara, yori
goal	*place	made, ni
attribute	*distance	-

note: "-" means that all particle can be supposed.

3 CONCEPTUAL DISAMBIGUATION

To reduce ambiguity of meaning hypotheses, conditions of existence of a concept and linguistic knowledge of scoring it are used. This methods are to filter out senseless concept hypotheses from both syntactic and semantic viewpoints.

3.1 Conditions of Existence of a Concept

Two conditions of existence of a concept are supposed. One is that a concept should have filled slots which are indispensable to forming a gist of the concept. These *indispensable slots* are defined in each concept frame. For example, a concept representing *Distance* should have an attributive case slot filled with a phrase which means *near, close, around, etc.* as the *indispensable slot* (Tab. 2).

Another condition is that a concept should occupy a continuous part of a sentence as shown Fig. 4. Such a concept hypothesis that has case slots scattered in a sentence is abandoned. This assumes that an user speaks an item information as a chunk of phrases and only once talks about the item in one utterance which includes various other concepts.

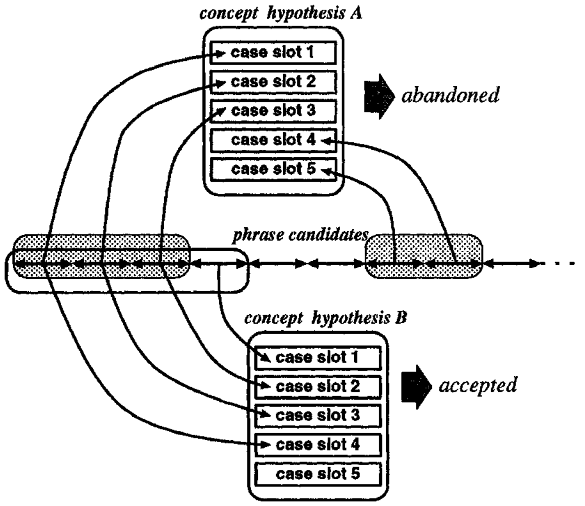

Figure 4: A condition of existence of a concept: assumption that a concept should occupy a continuous part of a sentence.

3.2 Estimating Linguistic likelihood

Syntactic and semantic likelihood of a concept hypothesis is evaluated by a scoring method which considers linguistic dependency between phrases. In the scoring method, penalty rules for a concept hypothesis are used. If a new hypothesis of a concept is produced, the hypothesis is examined based on all penalty rules. Total score of concept hypotheses are evaluated as linguistic likelyhood of a meaning hypothesis.

In defining penalty rules, pragmatic knowledge is not used because the proposed method of semantic interpretation should be as independent of specific task domain as possible. Therefore, penalty rules are defined using general syntactic knowledge of, mainly, particles and modifiers, and shallow semantic knowledge about relation between cases.

Table 3: Types of case slots

slots to modify	**attributive* case, **adverbial* case
slots to be modified	agent case, source case, goal case, object case, instrument case
slots for predicates	predicative case

* note: *attributive* case stands for a case which modifies the substantive case (agent, source, etc.), *adverbial* case for a case which modifies the predicative case.

Types of case slots are classified as shown in Tab. 3 to evaluate the dependency between case slots in a concept. In this classification, three types of cases based on modifying relation and eight semantic roles of cases are defined. All slots of a concept frame are assigned to these cases as features of linguistic dependency between phrases.

Several principles of defining penalty rules based on general syntactic knowledge and shallow semantic knowledge are mainly as follows:

< *syntactic features* >

- If a particle is omitted in input utterance or missed in speech recognition, a case of the phrase is ambiguous. In this case, a concept hypothesis including the phrase is penalized according to the priority of the relation between a case and a particle in each case slot. For example, a particle "kara"(from) is key information for a source case. If a phrase candidate omitted a particle is filled in the source case, the concept hypothesis is penalized.
- If a phrase candidate is filled in *attributive* case, the substantive case should follow the *attributive* case. A concept hypothesis where the both cases are not in this order is penalized.
- If a phrase candidate is filled in *adverbial* case, the predicative case should coexist in same concept hypothesis. A concept hypothesis where the both cases do not coexist is penalized.
- A heuristic method where a sentence hypothesis with minimum number of phrases is regarded as most likely output is well known for improving a performance of Japanese morphological analysis. A meaning hypothesis where total number of phrases is not minimum among all meaning hypotheses is penalized.
- Verb conjugations of phrase in the predicative cases are also penalized from a viewpoint of phrase position in a sentence.

< *semantic features* >

- Only predicative case by itself has insufficient information for existence of a concept. Similarly, a concept hypothesis without the predicative case has insufficient information. Therefore, a concept hypothesis where both the predicative case and other cases do not coexist is penalized.
- The order of the predicative case and other cases is considered. A hypothesis where an inversion of the predicative case and other case occurs is penalized.
- Semantic mismatch between phrase candidates is also considered. For example, Japanese adjective "yoi (good)" which means *permission* should succeed a phrase which means *assumption.*
- Abstract noun like *place* or *neighborhood* without modifiers has insufficient information. Therefore, a concept hypothesis where an object case filled with the abstract noun is not modified by other cases is penalized.

4 EXPERIMENTS

Semantic understanding experiments has been carried out for evaluating conceptual disambiguation by the scoring method and linguistic robustness of the proposed semantic interpretation method to open test sentences.

The input sentences on the *Hotel Reservation* task were made by 10 subjects as character string. As shown in Tab. 4, subjects were instructed to make test set A with as plain expressions as possible, and test set B with no limitation of sentential expressions.

Table 4: Two test sets of sentences (*Hotel Reservation*)

set name	#sent.	conditions to subjects
set A	51	to make as plain sentences as possible
set B	64	to make various sentences freely

4.1 Procedure

Test sentences were described as sequences of phonetic symbols. As a pre-processing for semantic interpreting process, the lattice generator analyzes the test sentence by driving chart parser using phrase grammar of approximately 1000-word vocabulary and outputs all possible phrase candidates as a lattice. Each phrase candidate are given a semantic value looking up a semantic dictionary.

Semantic interpreting process decides an intention and selects the intention frame for each phrase sequence (same structure of intention frame has been used in the experiments). The process detects concepts using 22 concept frames, filling slots with phrase candidates and examining conditions of concept existence. Then, all meaning hypotheses are penalized to find the best hypothesis from a viewpoint of linguistic likelyhood. Penalty of one is added to a meaning hypothesis every time it has a linguistic feature of penalty rules.

4.2 Results and Discussion

Two kinds of experiments has been performed according to the two test sets A and B.

The contribution of scoring methods has been evaluated for test set A. Tab. 5 reports the results. Ambiguity of meaning hypotheses was 50.8 a sentence using only conditions of existence of a concept with semantic understanding rate of 100 %. This shows that the conditions are proper for detecting correct concepts but they are still insufficient to reject senseless hypotheses.

The use of penalty rules based on syntactic relations reduced the ambiguity of 50.8 to 4.1 without missing a correct hypothesis. The rules worked well in disambiguation on syntactic modification like ***attributive*** case. A heuristic rule which gives a priority to a meaning hypothesis with minimum number of phrases was also good for restrainting such undesirable separation that *Yokohama* and *eki* (a station) are separated from a compound word *Yokohama-eki.*

Using all rules including semantic penalty rules achieved almost complete disambiguation to 1.1. But a correct hypothesis was missed in the case of conceptual utterance where predicative case and other cases did not coexist (case omission). For example, Japanese utterance "*Yokohama hukiN, tasho tookutemo iinodesuga*" (around Yokohama, I don't mind if [a hotel is] a long way [from the station]) occurs the omission of both agent case *a hotel* and source case *from the station.* Then, *Yokohama* which has case ambiguity without particle is mistaken for agent case instead of *a hotel* . This is because semantic penalty rule gives priority to such a conceptual hypothesis that agent case and predicative case coexists. To solve this problem, a process to compensate the omitted case using context information within an utterance should be required.

Table 5: test set A (48 sentence). Conditions are; (1) conditions of existence of a concept, (2) syntactic scoring rules, (3) semantic scoring rules.

Conditions	#Candidates	Understanding Rate(%)
(1)	50.8	100.0
(1)+(2)	4.1	100.0
(1)+(2)+(3)	1.1	98.0

Linguistic robustness of the method has been evaluated for test set B including various sentential expressions. 75.0 % (48 sentences) of sentences were correctly understood as shown in Tab. 6. 20.3 % (13 sentences) were rejected including linguistic phenomena as follows:

- A complex or compound sentence including a relative clause and plural predicatives: a mechanism for precisely analyzing this type of sentence should be implemented to the method, although the method can handle parallel clauses which modify a predicative. As for a relative clause which modifies a noun, one of approaches could be that concept of a relative clause is registered as ***attributive*** case slot of a subjective concept including the noun.
- Repetition or restart: this occurred in a sentence like "Yokohama-eki-no chikaku-de narubeku yasukute eki-kara chikai (in neighborhood of Yokohama-station, as inexpensive as possible, in neighborhood of the station)". The utterance including repetition may be understood based on conceptual search method which merges a repeated concept into a concept already hypothesized. Clue words like interjections of repairing a concept are helpful for detecting the repeated or restarted concept.
- A chain of same case element: this means a succession of adjectives or nouns in parallel, and can be analyzed by administering plural phrase candidates at a same case.
- Unknown concepts: detecting boundaries of unknown concepts and exploiting partial information already attained by other concepts are important to realize robust interpretation.

These phenomena were observed by examining 13 sentences of *No Answer.* The proposed method, however, showed a performance that 94.1% (48 / 51 sentences) were correctly understood with ambiguity of 1.5 excluding the *No Answer.*

Table 6: test set B (64 sentences). Conditions are; (1) conditions of existence of a concept, (2) syntactic scoring rules, (3) semantic scoring rules.

Conditions	#Candidates	Correct	Error	No Answer
(1)+(2)+(3)	1.5	75.0%	4.7%	20.3%

5 CONCLUDING REMARKS

We proposed a two-stage semantic interpretation method for robustly understanding spontaneous speech. In this approach, concepts are detected from an input lattice to form a meaning hypothesis with an intention decided from the input. Conceptual disambiguation was done by estimating linguistic likelihood of concept hypotheses using syntactic and semantic penalty rules. The experimental results showed that scoring method greatly contributed to reduce ambiguity of meaning hypotheses, and 75% of test sentences including various sentential expressions were correctly understood with ambiguity of 1.5. Future enhancements will include; (1) interpretation of more various expressions like a complex or compound sentence, (2) conceptual search method, (3) exploiting top-down knowledge for speech recognition, (4) data collection of naive speaker and (5) integration to speech recognition as semantic-driven method.

REFERENCES

[1] Victor Zue, Stephanie Seneff, Joseph Polifroni, Michael Phillips, Christine Pao, David Goodine, David Goddeau and James Glass, "PEGASUS: A Spoken Dialogue Interface for On-Line Air Travel Planning," Proc. ISSD-93 (Waseda Univ., Tokyo, Japan), pp. 157–160, Nov. 1993.

[2] Wayne Ward and Sheryl R. Young, "Flexible Use of Semantic Constraints in Speech Recognition," Proc. ICASSP 93, pp. II-49–50, Apr. 1993.

[3] David Stallard and Robert Bobrow, "Fragment Processing in the DELPHI System," Proc. DARPA Speech and Natural Language Workshop, pp. 305–310, Feb. 1992

[4] Stephanie Seneff, "A Relaxation Method for Understanding Spontaneous Speech Utterances," Proc. DARPA Speech and Natural Language Workshop, pp. 299–304, Feb. 1992

[5] Paolo Baggia and Claudio Rullent, "Partial Parsing as Robust Parsing Strategy," Proc. ICASSP 93, pp. II-123–126, Apr. 1993.