

# Analysis of Japanese Sentences by Using Semantic and Contextual Information (I)-Semantic Analysis

Makoto NAGAO\*, Jun-ichi TSUJII\* and Kazutoshi TANAKA\*

## Abstract

We have constructed a powerful parser of Japanese sentences. This parser can resolve ambiguities of input sentences by utilizing various kinds of semantic relationships among words. Especially the 'case' concept introduced by C.J. Fillmore, plays an important role in this parser. Case is the relation which is established between an object and an event. In a sentence, an object is usually expressed by a noun phrase and event by a verb.

The information as to what kinds of cases should be described and what kinds of nouns should be selected as case-elements of a verb of a sentence is called 'case-frame' of the verb and is described in the verb dictionary.

A verb may have more than one case frame corresponding to different usages of the verb. A noun represents a set of objects which has certain same properties. We express properties by attribute-value pairs in the noun dictionary. The dictionary also contains the set-inclusion relationships among nouns. In this paper, we will explain how the parser utilizes these dictionaries in order to analyze a noun phrase and a simple sentence.

## 1. Introduction

In this paper we describe the organization of the natural language parser developed over the last two years. The parser can transform fairly complex sentences into abstract structures marked for case. It utilizes detailed semantic dictionary descriptions and contextual information abstracted from the preceding sentences.

For the present, we have confined the domain of the system to the field of elementary chemistry where we can describe the semantic world in rather concrete terms. At the same time, various complex events occur in this field. For example, substances which participate in particular events may disappear, new substances may emerge or

---

This paper first appeared in Japanese in Joho-Shori (Journal of the Information Processing Society of Japan), Vol. 17, No. 1 (1976), pp. 10~18.

\* Department of Electrical Engineering, Kyoto University

some properties of the substances may be altered. To treat these complex situations, it was necessary to formally represent relationships between events and changes of state and to devise an appropriate scheme of representing context.

The parser consists essentially of four fixed components:

1) The grammar consists of rules written in PLATON. PLATON is a new programming language which is a variant of the system developed by Woods called 'Augmented Transition Network'.

2) In the dictionary are stored words along with their various semantic relationships.

3) The contextual representation is similar in form to the semantic network of R. F. Simmons or the nodespace of D.A. Norman.

4) Semantic and contextual functions are programmed in LISP. These functions are incorporated in the PLATON rules along with rewriting patterns.

In this paper, we are concerned with semantic analysis which utilizes semantic descriptions in the dictionary. The detailed description of contextual analysis is found in another paper.

## 2. Analysis of Noun Phrase

### 2.1 Properties of a Noun Phrase

In Japanese, two or more nouns are often concatenated by the postposition NO to form a noun phrase. A typical example is shown in Figure 2.1.

HANNOU	-NO	ATO	-NO	NATORIUMU	-NO	TAISEKI	-NO	HENKA
reaction		after		sodium		volume		change

changes of the sodium's volume after the reaction

Figure 2.1 Examples of NOUN+NO phrases

The phrase NOUN+NO can modify, in principle, any or all of the succeeding nouns in the extended NO construction so that many different patterns of modification relationships are syntactically permitted. We must decide which one is correct by considering semantic restrictions.

We have classified nouns in the field of elementary chemistry into five categories; Entity Nouns, Attribute Nouns, Value Nouns, Action Nouns and prepositional Nouns. Their meanings are described in the dictionary in different ways. Typical examples of





noun phrase, and an activity by a verb. A sentence instantiates an activity by supplying noun phrases to the cases associated with the activity. The problem is to decide what case a noun phrase holds in relation to a verb in any particular event. In order to determine these relationships we need both syntactic and semantic information. A verb has its own special usage patterns. That is, certain cases are necessary for the activity and certain objects are preferable as fillers for the case. We call these labeled patterns Case Frame for Verbs, and express them as a list of case pairs such as (CASE NOUN). A verb usually has more than one case frame corresponding to different usages.

In the course of analysis we must assign appropriate case labels by considering the case frames of the main verb along with meanings of the head nouns of the noun phrases. But sentences can have several possible interpretations for the following reasons:

(1) A verb may have more than one usage (i.e., a verb may have several case frames).

(2) A postposition can indicate more than one case. Some postpositions can occur with almost any case; WA is an example.

(3) A noun modified by an embedded sentence is usually a case slot filler of the embedded sentence. But we may have no syntactic clues as to what case to assign to the noun.

In the event of multiple interpretations the program derives labeled interpretations showing all possible case relationships between specific nouns and verbs. We choose the interpretation showing the preferable matching of nouns and case by using an evaluation function which has been established empirically.

#### 4. Conclusion

There are many scholars who are interested in using case structures as a representation of natural language utterances. We also believe that the case system is a promising approach to the representation of meanings in natural language.

#### Bibliography

Nagao, M. and Tsujii, J. : 1976, "PLATON - A New Programming Language for Natural Language Analysis", AJCL

Nagao, M. and Tsujii, J. : 1976, "Analysis of Japanese Sentences by Using Semantic and Contextual Information", AJCL