

## 構文意味解析システム SAX における係り受け解析

## - 構文意味情報の構成について -

福本文代

(財) 新世代コンピュータ技術開発機構

杉村領一

松下電器産業株式会社

## 1 はじめに

ICOTでは、自然言語処理の研究過程で得られた要素技術を汎用日本語処理系(Language Tool Box 以下 LTB)としてまとめ上げている。SAXは、このLTBの一モジュールであり、論理型言語により高速に構文意味解析を行なうシステムであるが、[1]、今回この上に新たに文を文節と文節の係り受け関係で表現する係り受け文法を開発した。この文法は、係り受けの判定処理の細分化のみで文が解析でき、かつその処理は文法規則中、局所化しているため、効率よく日本語解析が行なえる。本文法では、文の構成単位である文節の意味を物または関係として定義し、出力する文の意味情報を、物と物との間の関係がネットワーク構造で表現されるようにする。また解析過程で得られた意味情報を蓄積するために、その文節の持つ構文情報を保持しながら、文節の意味情報を後文節へ伝播する。本稿では、この係り受け文法で意味を構築する手法、すなわち意味の蓄積伝播について述べる。

## 2 論理型言語による係り受け解析

本文法は DCG(definite clause grammar)と呼ばれる文法記述形式に 0 型文法を導入することによって実現された拡張記述形式を用い、文節と文節の係り受けを判定する係り受け文法である [2]。文を解析する際用いる文法規則は、

- 「A と B と」のような等位構造を認識し一文節にまとめるルール群 [(9),(11)~(17)]
- まとめた等位構造を含む文節の依存関係(文節の係り受け)を調べるルール群 [上記以外]

より成る。

|          |   |              |     |
|----------|---|--------------|-----|
| sentence | → | wordit wba   | (1) |
| wordit   | → | {begin}      | (2) |
| wordit   | → | {begin}      | (3) |
| bar      | → | {end}        | (4) |
| wba      | → | word bar wba | (5) |

Dependency Structure Analysis in SAX

-Composition of Syntactic and Semantic Information-

Fumiyo FUKUMOTO ICOT

Ryoichi SUGIMURA Matsushita Electric Industrial Co.,Ltd.

|                |   |                                |      |
|----------------|---|--------------------------------|------|
| wba            | → | word bar                       | (6)  |
| word bar       | → | gapH bar                       | (7)  |
| word bar word  | → | word word                      | (8)  |
| word bar pword | → | word pword                     | (9)  |
| word bar word  | → | word (bar; []) wba word        | (10) |
| wordit pword   | → | wordit wordg                   | (11) |
| wordit gapH    | → | wordit wordg                   | (12) |
| word pword     | → | gapH wordg                     | (13) |
| word gapH      | → | gapH wordg                     | (14) |
| pword pword    | → | pword wordg                    | (15) |
| pword gapH     | → | pword wordg                    | (16) |
| pword          | → | pword (bar; []) (wba; []) word | (17) |
| word           | → | pword                          | (18) |

文節の持つ意味情報の蓄積は、この文法の補強項部分で行われている。ここでは、補強項でのこの処理に係り受け決定関数と呼ぶ。係り受け決定関数では、文節が等位を示し、かつ文節と文節が同種の等位であるかどうかの判定処理、あるいは文節と文節が係り受け可能か否かの判定処理、そしてこれらを判定した結果、同種の等位であり、また係り受け可能と判明した時に、文節の持つ意味情報を生成する処理を行なっている。

## 3 データ表現

本文法は、文の構成単位である文節の意味を、物あるいは関係として定義する。また扱う文節の構文意味情報は、形態素解析システム LAX の出力が与えられる。文法はこれを入力し、解析を行なった結果、その意味情報を物と物との間の関係がネットワーク構造で表現されるよう蓄積する。以下に文節「食べる」の構文意味情報の一例を示す。ここで、意味情報は、soas/ で記述し、構文情報はそれ以外とする。

|         |                                  |     |
|---------|----------------------------------|-----|
| sur/    | [食, べ, る]                        | (1) |
| moded/  | cat/ 用言                          | (2) |
| frame/  | [(agt, が, 46), (obj, を, 47)]     | (3) |
|         | :                                |     |
|         | :                                |     |
| moding/ | cat/ 連体詞                         | (4) |
| soas/   | [(46, heap: {食べる, 46, 47, yes})] | (5) |

図 1 : 係り受け文法で扱う文節の構文意味情報

