

# 黒板モデルの導入により後戻り処理を伴わない べた書き日本語文の解析

一本木 真史 浅野 晃 上原 邦昭 豊田 順一  
大阪大学産業科学研究所

本報告では、格文法をベースにした係り受け解析による意味解析システムについて述べる。一般意味解析は階層的な処理と仮説の生成および検証を伴う。これは問題解決の典型的なパターンである。意味解析を問題解決の一種と見なすと、問題解決器はパーザに相当するものと考えられる。我々が実現したシステムは黒板モデルをパーザの基本機能として採用し、解析過程は文の先頭から解析を進め後戻り処理を伴わないように構成している。また、意味解析の階層的な処理内容に従つ必要な知識を細分化し、それぞれ独立した知識源としてまとめることが可能となっている。さら細分化された知識の管理が容易となっており、自然言語特有の例外的な表現にも対処しやすくなっている。動詞の文型・述語素などを抽出するための辞書には「日本語基本動詞辞書 IPAL」を採用しており、解析システムとしての汎用性が高くなっている。

## Japanese Language Understanding System Using Blackboard Model

Masashi IPPONGI, Akira ASANO, Kuniaki UEHARA and Jun'ichi TOYODA

The Institute of Scientific and Industrial Research,

Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka 567, Japan

This paper presents a Japanese language understanding system using case grammar. In general, parsing process can be seen as a hierarchical procedure which generates hypotheses and validates them. This mechanism is also a typical pattern in problem solving. We, thus, regard parsing as problem solving and consider that a problem solving model corresponds to a parser. In our system, the blackboard model is used as a basic mechanism for parser and knowledges for parsing is partitioned into some couples of knowledge modules, each of which knowledge transforms partial solutions into more larger one during the course of parsing.

# 1 はじめに

従来より日本語文の解析システムが多数試作されているが<sup>[13]</sup>、日本語文の構文的な特徴による数多くの問題が残されたままである。日本語の文節は文節内部の語順に関して強い制約があるが、文節どうしの順序に関しては非交叉の原理などの緩い制約しかない。したがって、日本語文解析においては、英文解析のように語順の制約情報を利用することができず、名詞の意味と助詞から文節間の係り受け関係を抽出して文全体としての意味構造を得なければならない<sup>[16]</sup>。

文節間の係り受け関係を解析するための手掛かりは、動詞からその動詞に係るべき文節の構造を予測するという形式で得られる場合が多い<sup>[1][12]</sup>。このような方針で設計されたシステムでは、まず最初に動詞に注目し、その動詞と共に起しうる名詞句を予測して係り受け関係が抽出される。そのため、動詞が最終文節に存在する日本語文では、解析文を文末まで認識した後でなければ係り受け関係の抽出を開始できず、既に解析された文節をプッシュダウンスタックなどの短期記憶に貯めて、動詞が現れた時点で初めて動詞との係り受け関係が決定される。しかしながら、個々の係り受け関係を逐次的に決定していくプッシュダウンスタックなどでは、複文などの解析において後戻り処理を伴うという問題がある<sup>[3][5]</sup>。

一方、日本語文の解析に通常の句構造文法を用いた場合、文法規則の数は膨大なものとなり、すべての規則の適用可能性を解析途中のあらゆる段階で調べることは解析効率を非常に悪くする。また、文節を解析するための規則、文の構造を解析するための規則などは適用される対象が異なっているものである。このような適用対象の異なる知識は、分類されて、各々が必要とされる段階で適用されることが望ましい。つまり、処理過程の種々の段階で用いられる文法規則をすべて同等に扱うのではなく、それぞれの「モジュール化」して適用すべきである。合文性を調べると同時に、各自立語に言語情報（品詞・活用形等）を添える。本システムで使用した「モジュール化」は「日文語基本文法-複文編-（電総研）」で提案された辞結合モジュールをもとにして、助詞・助動詞などの品詞を同定できるように修正を加えたものである。

我々が実現している解析システムは、あらゆる係り受け関係の候補を並列的に考慮し、文の先頭から後戻り処理を伴わずに意味

解析を実行する。また、語の意味・用法・機能などの語そのものに属する知識を文法規則として利用して解析を進めるものである。さらに、意味解析の過程を明確に階層化し、黒板モデル<sup>[2][10]</sup>を導入して個々の処理内容にあわせて知識を細分化し、それぞれ独立した知識源としてまとめている。

本システムは、日本 DG 社の MV-8000-II 上で KCL (Kyoto Common Lisp) を用いて実現されている。本システムで採用した「日本語基本動詞辞書 IPAL」<sup>[9]</sup> は関係データベース DG/SQL に記憶されており、両者は互いに結合されている。また、関係データベースに記憶されている情報は、Lisp 関数による検索が可能となっている。処理の対象としている日本語文の形式は、システムの汎用性を考慮して、一般的な日本語の記述方式である漢字仮名混じりべた書き文としている。

## 2 意味解析に必要な知識

従来の係り受け解析システムでは、まず動詞に注目して、その後係り受け関係を決定していることが多い。しかしながら、このような処理の流れでは「文節に対する処理」、「文節どうしの関係に対する処理」、「文全体に対する処理」が混在して実行され、それぞれの処理に必要な知識の管理が困難となっていた。

我々は、システムのメンテナンスの向上および知識管理の容易さを重視して、相互依存がないように知識源を細分化している。本システム意味解析を行う場合、その処理過程は階層的に分割されたものとなる。個々の処理で適用される知識は互いに独立であり、他の処理で適用される知識に影響を及ぼさない。以下に本システムで使用する知識を分類して説明する<sup>[7]</sup>。

### 2.1 チェッカ<sup>[14]</sup>

自立語と付属語、付属語と付属語の結合の合文性を調べるために、結合の正当性を有限状態オートマトンとして表現した知識を用いられる文法規則をすべて同等に扱うのではなく、それぞれの「モジュール化」して適用すべきである。合文性を調べると同時に、各自立語に言語情報（品詞・活用形等）を添える。本システムで使用した「モジュール化」は「日本語基本文法-複文編-（電総研）」で提案された辞結合モジュールをもとにして、助詞・助動詞などの品詞を同定できるように修正を加えたものである。

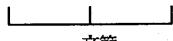
### 2.2 文節知識

「文節知識」は、形態素解析されている文より自立語と付属語

の順序を調べ、一文節を抽出するための知識である。「文節知識」では、基本的に一文節内には自立語が1個だけ存在し、かつその自立語は文節の先頭に現れることを利用している。現在、「文節知識」には約30個の規則がある。

#### 文節知識の例)

前提条件：名詞 + 助動詞 + 助詞 + 名詞

結果      :       一文節

#### 2.3 意味素性テーブル

意味解析においては、文中に最も多く存在する自立語である名詞の意味を正確に把握する必要がある。本システムでは動詞辞書「IPAL」にもとづき、名詞の意味素性として「人間」、「組織」、「動作」などの19個を設定している(表1)。各名詞に対して意味素性を与えるための知識を「意味素性テーブル」と呼ぶ。

表1) 名詞の意味素性一覧<sup>[9]</sup>

記号	素性名
CON	--- 具体名詞
ANI (+CON)	--- 動物
HUM (+CON, +ANI)	--- 人間
ORG (+CON)	--- 組織・機関
PLA (+CON)	--- 植物
PAR (+CON)	--- 生物の部分
NAT (+CON)	--- 自然物
PRO (+CON)	--- 生産物・道具
PHE	--- 現象名詞
ABS	--- 抽象名詞
ACT (+ABS)	--- 動作・作用
MEN (+ABS)	--- 精神
LIN (+ABS)	--- 言語作品
CHA (+ABS)	--- 性質
REL (+ABS)	--- 関係
LOC (+ABS)	--- 空間・方角
TIM (+ABS)	--- 時間
QUA (+ABS)	--- 数量
DIV	--- (制限なし)

#### 2.4 格構造知識

文の先頭から意味解析を進めるためには、各文節の格構造候補を抽出する必要がある。我々は、日本語の文を意味内容により12種類に分類し、それぞれの文のタイプごとに共起しうる格構造を設定している(表2)<sup>[4]</sup>。表中の「往来に関して述べる文の出発点」は、「場所を示す名詞+から」、「場所を示す名詞+より」の

いずれかの文節であることがわかる。このように名詞の意味素性と助詞の結合状態から各文節のとりうる文のタイプと格構造候補を予測する規則を「格構造知識」と呼ぶ。

#### 格構造知識の例)

前提条件：場所を表す意味素性を持つ名詞

+ から(助詞)

結果    : 文のタイプ→往来に関して述べる文

格構造 →出発点

表2) 文のタイプと格構造

文のタイプ	設定格構造
精神的行為に関して述べる文	主体、内容、対象
観念的行為に関して述べる文	主体、対象、目標、相手
	手段/道具
往来に関して述べる文	主体、場所、出発点、到着点
もののやりとりなどの共同的行為に関して述べる文	主体、対象、相手
動作・労働などの行為に関して述べる文	主体、対象、目標、相手 対象の始状態、 対象の終状態 目標物の材料 手段/道具、方法
現象に関して述べる文	対象、原因、変化の始点 変化の終点、所属の始点 所属の終点
感情・感覚などに関して述べる文	主体、源
ものの属性に関して述べる文	対象、属性、比較対象
可能性・必要性などに関して述べる文	主対象、副対象
存在に関して述べる文	対象、場所
存在の状態を述べる文	対象
対象間の関係などの認識に関して述べる文	対象、相手

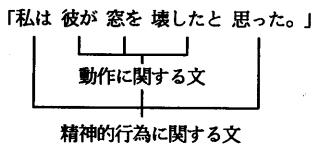
#### 2.5 係り受け範囲決定知識

解析文が複文の場合、各文節がいずれの動詞に係るかを決定することが必要となる。しかし、日本語は語順に対する制約が弱いため、名詞の意味的な情報や助詞の類から各文節が係る動詞を決定しなければならない。このような手法に基づいて、各文節の格構造候補より係り受け関係を決定する知識を「係り受け範囲決定知識」と呼ぶ。「係り受け範囲決定知識」の基本方針は以下の三点である。

- a) 非交叉の原理
- b) 一文一格の原理
- c) 一つの係り受け範囲内に、複数のタイプの文の格構造は存在しない

c) は、2.4節で述べたように12種類に分類された文のタイプの係り受け関係が、それぞれ交叉することができないということを示している。

#### 係り受け範囲の例)



#### 2.6 変形規則

本システムは原則的に後戻り処理を行わない方針で設計している。しかしながら、受け身や使役などの格構造候補を文の先頭から考慮すると、各文節の格構造候補の数が爆発する可能性がある。したがって、文末まで解析が進み、助動詞から受け身・使役の文であることが発見された段階で、「変形規則」を適用し、各文節にすでに与えられている格構造候補を書き換えるようにしている。たとえば、受け身文の場合は、主格と目的格が入れ替わっているので元に戻すという「変形規則」を用意している。また、使役文の場合は、使役格を追加し、目的格となっているものを主格に戻すという「変形規則」を用意している。これらの「変形規則」は、2.4節で述べた文のタイプごとに設定されており約20個ある。

#### 2.7 動詞辞書

本システムでは、「計算機用日本語基本動詞辞書 IPAL」を利用し、動詞の文型・テンス・ヴォイス・アスペクト・述語素などを抽出している。「IPAL」は、情報処理振興事業協会より提供された汎用的動詞辞書で、動詞が約900語収録されている。「IPAL」の各動詞には、述語素が組み込まれている。述語素とは、名詞句と動詞、二つの名詞句と動詞、名詞句どうしの結びつきに基づいて、名詞の意味と助詞の類より文の格関係を表したもので、文の意味構造として利用されるものである。「IPAL」で用いられている述語素は、約60種に分類されている。たとえば、名詞句と動詞の結びつきを表す述語素には、動作の主体や対象といったものがあり、二つの名詞句と動詞の結びつきを表す述語素には、起点や着点などがある。また、名詞句どうしの結びつきには、全体-部分の関

係などがある。このような動詞辞書「IPAL」を、黒板上の係り受け範囲を動詞の文型と正しく対応づけられている候補のみに絞り込むために使用する。

### 3 動作原理

#### 3.1 意味解析の手法

日本語文では、文の意味構造の中心を成す動詞が最終文節まで出現しない。したがって、文の先頭から意味解析を開始しようとする場合には、各文節に対するあらゆる格構造の可能性(仮説)を設定し、すでに解析されている文節までの係り受け関係(仮説)と組み合わせながら新たな係り受け関係を設定(生成)したり、既に設定されている係り受け関係を棄却(検証)したりしなければならない。このような仮説の生成と検証の繰り返しは典型的な問題解決のパターンであるため、日本語文の意味解析は一種の問題解決と見なすことができる。我々は日本語文の意味解析を「単語の意味を得る」、「文節の意味構造を得る」、「文節間の関係を得る」、「文全体の意味を得る」などをサブ・ゴールとしても階層的な問題解決と見なし、個々のサブ・ゴールを達成するために多数の仮説を立て、その仮説を検証していく形で解析を進めるようしている。以上のような手法では、各単語の意味から格構造を抽出したり、既に抽出されているいくつかの格構造を組み合わせて文体を予測する必要がある。このような処理を実現するには、処理内容にあわせた階層的な知識をもつシステムを構築しなければならない。このようなシステム構成は、一般に黒板モデルと呼ばれる問題解決システムがもっているものである。我々は日本語文の意味解析を一種の問題解決とみなし、パーザの基本機能として黒板モデルを利用して本システムを構築している。

#### 3.2 格構造候補の生成および絞り込み

2.4節で述べたように、本システムでは日本語の文を意味内容により12種類に分類し、それぞれの文のタイプごとに共起しうる格構造を設定している。このように、文および格構造を分類することにより、文の先頭から解析を進めても解析文のタイプを予測することが可能となる。たとえば、システムが文節「～は」を読み込むと ①精神的行為に関する文、②観念的行為に関する文、③往来に関する文、④やりとりに関する文、⑤一般的行為に関する文、⑥現象に関する文 などが予測できる。統いて「～と」という文節を読み込むと、既に設定された予測は ①精神的行為に

に関する文、②概念的行為に関する文、③やりとりに関する文、  
④一般的行為に関する文 に絞り込むことができ、さらに「～について」という文節を読み込むと、最終的に ①やりとりに関する文 であると文のタイプを決定することができる。このように、文の先頭から解析を行う場合でも、動詞を認識せずに解析文のタイプをある程度決定できる。このような方法を採用すると、あらゆる可能性を常に考慮しながら解析しているためにプッシュダウンスタックなどの短期記憶を用いる必要はなく、後戻り処理も最小限になるため解析効率に大きな損失はない。また、文節や文節どうしに関する部分的な知識のみで解析を進めることができるとなるために、知識が簡潔となり、かつ自然言語特有の例外的な表現を処理する場合に対処しやすくなる。以上のような手法では、助詞の類のみから文体を予測するために、必ずしも一意に文のタイプが決定されるとは限らない。たとえば、「～が～を～に～」という文の場合、①往来に関する文、②やりとりに関する文、③一般的行為に関する文 という予測はできるが、これらのうちの一つに絞ることはできない。このような場合には、以後の章で述べるような各単語の意味的な情報を利用する必要がある。

また、以上のような手法は複文などの意味解析における文体の抽出や係り受け範囲の決定に非常に有効なものとなっている。たとえば、

「私は彼が窓を壊すのを見た。」

という文が与えられたとき、システムは「私は彼が」という部分まで読み込んだ時点で、「私は」と「彼が」がどちらも文の主体となる文節であることを抽出し、解析文が複文であると予測する。このように、解析の初期の段階で複文などの文体を予測できることは、以後の解析における無駄な処理を省くことができ、効率的な解析が実行される。

## 4 システムの基本構成

### 4.1 黒板モデル

一般に、黒板モデルは以下の三要素から構成されている。

#### a) 知識源

知識源は、問題解決に必要な知識を相互依存なしに分割したものである。個々の知識源は、黒板上のデータを取り入れ、知識源中の知識を適用してそれらを更新するようになっている。知識は黒板上のデータのみを書き換えることができ、逆に黒板上のデータは知識のみによって書き換えられる。個々の知識源には、その中の知識が発火するために必要な前提条件として、黒板上のデータ状態が指定されている。

#### b) 黒板データ

黒板データは、黒板上に書かれたデータのこと、知識源が変更しようとするデータや知識源の変更によって作り出された状態を保持するものである。黒板上のデータは、入力データ、部分解候補、最終解などからなり、それらは解析のレベルに応じて階層化されている。

#### c) 制御

制御部は黒板上のデータの変化を常に監視し、個々の知識源のもつている前提条件を調べながら次の行動を決定する。

我々は、以上のような構成をもつ黒板モデルをバーザの基本機能として利用している。具体的には、意味解析に必要な知識を相互独立な知識源にまとめ、黒板を解析処理のためのワーキングメモリとして使用し、黒板上に処理過程を階層的に書き込んでいる。

### 4.2 解析手順

本システムは、7個の知識源を持っており、システムの制御部が黒板上のデータを常に監視し、機会主義的に知識源中の各知識を発火させながら解析を進行する。解析手順は以下のようになる。

- 1) 入力文の形態素解析を「チェック」を適用しながら黒板上で開始する。
- 2) システムの制御部は常に黒板上の形態素解析の変化を把握し、「文節知識」を適用して一個の文節を取り出す。
- 3) 2)で抽出された文節に「意味素性テーブル」を適用して文節中の各名詞に意味素性を与える。
- 4) 「格構造知識」を文節中の意味素性および助詞に適用し、文節の格構造候補を取り出す。
- 5) 文頭から現在解析中の文節までの各文節に与えられた格構造候補に、「係り受け範囲決定知識」を適用し、可能な係り受け関係を黒板上に生成する。
- 6) 文節中の助動詞から、解析文が受け身文あるいは使役文であることが判明した場合は、「変形規則」を用いて黒板上に生成された格構造候補を変形し、5)に戻る。
- 7) 文末まで 1)～6) の解析を終了すると、各文節の係り受け関係がいくつかの候補に絞られる。動詞辞書 IPAL を用いて、これらの候補に適合する文型を選択し、両者の整合性を調べる。最終

的に黒板上に残った述語素が文の意味構造になる。以上の解析過程の詳細については次章で説明する。

## 5 解析過程の詳細

### 5.1 文字解析部

入力文に対しては、まず形態素解析が行われる。形態素解析を行うには、自立語辞書を用いるのが一般的である。しかしながら、一般の辞書は項目数が数万から数十万と非常に大きいため、主記憶上に全体の内容をもつ場合、二次記憶装置の使用が不可避免となる。しかも、ポインタを使ったデータ構造では、新しいポインタをとるたびに主記憶装置と二次記憶装置との間に転送をひき起こすことになり、主記憶上での項目どうしのつき合わせに要する時間よりも、主記憶装置と二次記憶装置との間でのデータ転送に要する時間の方が大きくなってしまう。このような時間の無駄を無くすために本システムでは形態素解析に際して、「通常の日本語文では、漢字列・片仮名列は一つの自立語を構成し、ひらがな列は助詞・助動詞列となる」という字種情報を用いている。これにより、完全な名詞辞書を整備するという困難な作業が必要なくなっている。字種情報より自立語が抽出されると、その後の付属語との結合の正当性を「チェック」を利用して調べる。形態素解析における多義性は、「文節知識」、「意味素性テーブル」、「格構造知識」の適用により解消される。

「チェック」は大別すると次の四種類に分類できる。

- ・名詞など無活用語とその直後の付属語との結合をチェックするもの
- ・動詞など活用語の語幹とその直後の活用語尾との結合をチェックするもの
- ・動詞など活用語の活用語尾とその直後の付属語との結合をチェックするもの
- ・付属語どうしの結合をチェックするもの

以上のような「チェック」および字種情報を用いて、黒板上で形態素解析が進められる。

### 5.2 単語解析部

システムの制御部は常に黒板上の形態素解析の状態を監視しており、黒板上のデータが変化することに「文節知識」を適用し一文節を抽出する。さらに、抽出された一文節中に名詞が含まれて

いる場合、「意味素性テーブル」を用いて個々の名詞に対して、機能別に分類した意味素性を与える。ただし、自然言語では一つの名詞が複数の意味をもつ場合が多いが、意味素性の多義性は「係り受け範囲決定知識」の適用により解消される。

「文節知識」および「意味素性テーブル」が適用されると、意味素性が付与された名詞を含む一文節が黒板上に生成される。

### 5.3 文節解析部

システムの制御部は、黒板上に一文節が生成されると、その文節に対して「格構造知識」を適用する。「格構造知識」が適用された文節には黒板上に格構造候補が与えられる。

また、解析中の文節内の助動詞から、解析文が受け身・使役文であることが発見された段階で「変形規則」を適用し、黒板上ですぐに与えられている格構造候補を書き換える。ただし、使役文であることは助動詞より一意に決定できるので黒板上の格構造候補をすべて書き換えるが、受け身文の場合には、受け身の他に自発・可能・尊敬の場合が考えられるので、黒板上のデータに追加するという方法を採用している。

変形例)

本が (主体) 彼から (与える者) 彼女に (受ける者) 渡された。

↓ (変形規則)

本が (対象) 彼から (主体) 彼女に (受ける者) 渡された。

この例は、やりとりに関する文の場合の変形例であり、「渡された」という文節を読み込んだ時点で、解析文が受け身である可能性を予測して、格構造候補を書き換えている。この例では“主体”と“与える者”という格構造を“対象”と“主体”に変えている。

### 5.4 文節間解析部

黒板上に格構造候補が生成されると「係り受け範囲決定知識」が適用されて、文のタイプが絞り込まれ、係り受け関係が決定されていく。また、一文一格の原理を満たさないような係り受け関係が発生する場合には、解析文が埋め込み文であると予測し、その可能性を黒板上に記述する。さらに、「係り受け範囲決定知識」により名詞の意味素性の多義性が解消される場合もある。たとえば、“会社”という名詞は組織を表す場合と場所を表す場合がある。もし、「会社に」という文節を含んだ解析文がものやりとりに関して述べている文であることが予測されれば、“会社”は組織を表している。また、解析文が往来に関して述べる文であると予測されると、“会社”は場所を表していることが決定できる。

## 5.5 文解析部

### 特殊埋め込み文の例)

解析文中に動詞が現れると動詞辞書「IPAL」から動詞の文型を取り出し、黒板上の係り受け範囲を動詞の文型と正しく対応づけられている候補のみに絞り込んでいく。係り受け関係が一意に決定されている場合についても、正当性の検証の意味も兼ねて、動詞の文型との対応づけを行う。ただし、解析文が埋め込み文の場合には、埋め込み文の種類に応じて処理を以下のように分類している<sup>[8]</sup>。

#### a) 単純埋め込み文の場合

この種の埋め込み文は、埋め込み文の動詞(副動詞)を含む文節が主動詞の必須格の一つとなる。全体の構造としては单文の構造が二つあると見なして個々の動詞の文型と係り受け範囲を対応づけている。

#### 単純埋め込み文の例)

解析文 「私は 彼が 窓を 壊したと 考えた。」

文型 「壊す」：(人間 が) (生産物 を)

「考える」：(人間 が) (\_\_\_\_ と)

↓

「壊す」：(彼 が) (窓 を)

「考える」：(私 が) (彼が窓を壊した と)

#### b) 埋め込み文の動詞直後の名詞(被修飾名詞)が、埋め込み文中の格要素に対応する一般的連体形埋め込み文の場合

被修飾名詞を文型中の格要素の一つにあてはめて、動詞の文型と係り受け範囲を対応づける。

#### 一般的連体形埋め込み文の例)

解析文 「私は 彼が 弟に 渡した 資料を 読む。」

文型 「渡す」：(人間 が) (人間 に) (生産物 を)

「読む」：(人間 が) (生産物 を)

↓

「渡す」：(彼 が) (弟 に) (資料 を)

「読む」：(私 が) (資料 を)

#### c) 埋め込み文による被修飾名詞が動詞の文型に影響を及ぼさない特殊連体形埋め込み文の場合

この種の埋め込み文の被修飾名詞には「こと」、「とき」など、その語自身が明確な意味をもたない特殊な語が多いので、被修飾名詞より特殊連体形埋め込み文であることを認識している。処理は単純埋め込み文の場合と同様である。

解析文 「私は 彼が 弟に 資料を 渡したことを 知らない。」

文型 「渡す」：(人間 が) (人間 に) (生産物 を)

「知る」：(人間 が) (\_\_\_\_ を)

↓

「渡す」：(彼 が) (弟 に) (資料 を)

「知る」：(私 が) (こと を)

最後に「IPAL」からテンス・ヴォイス・アスペクトを取り出し、文の深層構造、すなわち述語素の各要素が決定される。「IPAL」を用いても複数の候補がある場合には、曖昧性があるものとして、複数の解析結果を残している。

2.7節で述べたように述語素は大別して、名詞句と動詞、二つの名詞句と動詞、名詞句どうしの関係を示すものの三つになる。本システムでは約60個の述語素を用いて最終的な文の意味構造を表わしている。

#### 解析結果例)

入力文：三原山が噴火口から煙を吐く

動詞 「吐く」の述語素：

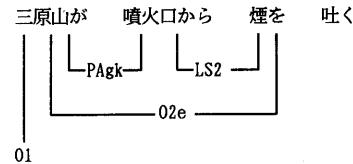
LS2 -- [を、から] 空間的起点

02e -- [が、を] 対象(出現)

PAgk -- [が、から] 部分

01 -- [か] 対象

#### 解析結果：



## 6 まとめ

本稿では、各語の意味・用法・機能などを利用して解析を行う日本語解析システムについて述べた。本システムの特徴は以下の4点である。

- 文の先頭より解析を行い、後戻り処理を伴わない。
- 解析過程の処理内容に従って必要な知識を細分化し、それぞれ独立した知識源にまとめており、知識管理が容易になっている。
- 動詞の文型・テンス・ヴォイス・アスペクト・述語素等を抽出

- するために、「計算機用日本語基本動詞辞書 IPAL」を利用してい 参考文献
- る。
- ・日本語文の意味解析を一種の問題解決とみなし、バーザの基本機能として、黒板モデルを導入している。
- 本システムは、プッシュダウンスタックなどの短期記憶を用いた場合と異なり、すべての格構造の可能性を考慮するため、係り受け関係を一意に決定する必要がなくなっており、埋め込み文の解析においても可能性を棄却するだけで係り受け関係を修正することが可能である。つまり、後戻り処理を行わずに解析を進める手法としては、Connectionist Model を用いた連想記憶に基づく [15]。これは従来の逐次型計算機ではなく、並列型計算機をもとに開発されているシステムである。
- 並列的な処理を行うことは本システムの処理過程と類似している。しかしながら、並列型計算機は我々が即座に手にすることができるものではないため、システムの実用性という点においては劣っていると思われる。また、基本メカニズムの設計が行われているのみで、比較的容易な文しか解析できず、システムの拡張にはかなりの時間がかかると思われる。
- 我々は後戻り処理を伴わないという方針でシステムを構築しているが、本システムでは受け身・使役の文に関して予め可能性を考慮しておくことができず、受け身・使役の文であることが判明した後に仮説を再生成するという後戻り処理を伴う。人間が受け身・使役の文を解析するといった場合には、受け身や使役の文であることを文脈や解析途中の特定の語句から予測して、予め可能性を考慮しているものと考えられるが、詳細については明らかでない。このような人間の予測方法を解明して本システムに融合させることは今後の課題である。
- また、本システムは各動詞に対するほとんどの必須格を解析可能なが、任意格や各文節に対する修飾語句については無視している。今後は、任意格用の格構造知識を追加するなどしてシステムを拡張し、任意格についての解析機能を実現する予定である。
- 謝辞
- 「日本語基本動詞辞書 IPAL」を御提供頂いた、情報処理振興事業協会の皆様に感謝致します。
- [1] 日高 他: 格文法による日本語の構文解析, 情報処理学会, 自然言語処理技術シンポジウム, pp. 41-46 (1983)
- [2] H.P. Nii: Blackboard Systems: The Blackboard Model of Problem Solving and the Evolution of Blackboard Architectures, The AI Magazine, Vol. 7, No. 2, Summer, pp. 38-53 (1986)
- [3] 市山 他: 日本語処理の観点から見たプッシュダウンオートマタの再評価, 情報処理学会, 自然言語処理技術シンポジウム, pp. 23-27 (1983)
- [4] 池田尚志: 日本語文における格の種類についての考察, 情報処理システムが提案されている [15]。これは従来の逐次型計算機 处理学会, 自然言語処理研究会, 41-2 (1984)
- ではなく、並列型計算機をもとに開発されているシステムである。[5] 池田尚志: 語法規則方式による日本語文の構文・意味解析, 並列的な処理を行うことは本システムの処理過程と類似している。情報処理学会論文誌, Vol. 26, No. 6, pp. 1079-1088 (1985)
- しかしながら、並列型計算機は我々が即座に手にすることができるものではないため、システムの実用性という点においては劣っていると思われる。また、基本メカニズムの設計が行われているのみで、比較的容易な文しか解析できず、システムの拡張にはかなりの時間がかかると思われる。
- 我々は後戻り処理を伴わないという方針でシステムを構築しているが、本システムでは受け身・使役の文に関して予め可能性を考慮しておくことができず、受け身・使役の文であることが判明した後に仮説を再生成するという後戻り処理を伴う。人間が受け身・使役の文を解析するといった場合には、受け身や使役の文であることを文脈や解析途中の特定の語句から予測して、予め可能性を考慮しているものと考えられるが、詳細については明らかでない。このような人間の予測方法を解明して本システムに融合させることは今後の課題である。
- また、本システムは各動詞に対するほとんどの必須格を解析可能なが、任意格や各文節に対する修飾語句については無視している。今後は、任意格用の格構造知識を追加するなどしてシステムを拡張し、任意格についての解析機能を実現する予定である。
- [15] 田村 他: Connectionist Model を用いた自然言語処理システム, 情報処理学会論文誌, Vol. 28, No. 2, pp. 202-210 (1987)
- [16] 辻井潤一: 日本語構文解析, 情報処理学会, 自然言語処理研究会, 38-5 (1983)