

心像意味論のための HPSG の意味原理について

笠 晃一 弘中 大介 横田 将生
福岡工業大学

〒811-0295 福岡市東区和白東3丁目30-1

あらまし 我々は数年前より、心像意味論に基づく自然言語理解システムを HPSG を用いて記述する試みを行なっている。HPSG では、従来より意味論として状況意味論が使用されることが多かったが、心像意味論を使用することにより、より深い意味の記述が可能になる。また、逆に、心像意味論に基づくシステムは、以前は依存文法を用いて記述されていたが、HPSG を使用することにより統一的な意味記述が可能になる。さらに、統語解析と意味解析を同時並行的に行なうことが可能なので、統語的な組合せ的爆発を抑制できるという利点もある。今回、意味記述形式と意味原理を文生成にも利用できるように改良したので報告する。

キーワード 自然言語理解、意味解析、心像意味論、HPSG

Semantic Principles of HPSG for Mental-image Directed Semantic Theory

Koichi RYU Daisuke HIRONAKA Masao Yokota
Fukuoka Institute of Technology

3-30-1 Wajirohigashi, Higashi-ku, Fukuoka, 811-0295 Japan

Abstract We have been attempting describing a natural language understanding system based on mental-image directed semantic theory (MIDST) with HPSG for some years. Situation semantics has been used for HPSG in most cases. If MIDST is used, deeper meaning can be described for sentences. On the other hand, the systems based on MIDST have been described with dependency grammar. If HPSG is used, meaning of words can be described uniformly and combinatorial explosion in parsing can be restrained. In this paper we will report on recent improvement of semantic description form and semantic principles aiming at sentence generation.

Keywords natural language understanding, semantic analysis, mental-image directed semantic theory, HPSG

1. はじめに

心像意味論 [1] は、自然言語の意味を人間の心像現象と対応づけて説明する理論であり、述語論理の基盤に立ち、計算にも適している。従来のモンタギュー意味論や状況意味論 [2] などと比べ、より深い意味の記述が可能であり、(1) 意味的曖昧さを指摘できる、(2) 意味的異常性を検出できる、(3) 同義性を指摘できるなどの特徴を持っている。HPSG [3] では、従来より意味論として状況意味論が使用されることが多かったが、心像意味論を使用すれば、自然言語の意味のより深い理解が可能となる。

一方、心像意味論に基づく自然言語理解システムはすでにいくつか作成されている [4][5] が、それらは統語解析時に依存文法を使用しており、統語解析と意味解析が独立していた。また、単語の意味として概念部と結合操作部を持たせており、宣言的記述と手続き的記述を併用したものになっていた。心像意味論を HPSG を用いて処理すれば、統語解析と意味解析を同時並行的に行なえるので、統語的な組合せ的爆発を抑制可能である。また、従来の概念部と結合操作部は、ともに素性構造を用いて宣言的に記述できるので、意味を二つの部分に分ける必要性もなくなる。さらに、当然のことながら、HPSG の研究において得られているさまざまな知見を利用することも可能になる。

以上のような理由により、我々は数年前より、心像意味論を HPSG に組み込む試み [6] を行なっている。しかしながら、最初に作成したシステムの意味記述形式は文生成に向けたものではなかった。文生成においては、論理形式の「平坦な (flat)」表現 [7] が使用されることが多いが、このような表現になっていなかったのである。我々は、将来的に多言語間翻訳も視野に入れているので、今回これを修正することにした。以下、意味記述形式の修正とそれに伴う意味原理の修正について述べることにする。

2. 素性構造による概念部の表現

心像意味論に基づく従来のシステムは、単語の意味を概念部 [4] と結合操作部 [4] に分けて記述していた。概念部はその単語の表す概念を記述するものである。まずは、概念部を素性構造で表現する。

2. 1 素性による原子軌跡式の記述

心像意味論において、意味記述は軌跡式 [1] とい

うものを用いて行なわれるが、この軌跡式の最もプリミティブな要素が原子軌跡式と呼ばれるものである。これは次のような形をした述語論理式である。

$$L(x, y, p_{t_j}, p_{t_j}, a, g, k) \quad (1)$$

この式 (1) の意味は、時間帯 $[t_i, t_j]$ において、事物 x の作用により事物 y の属性値が p_{t_i} から p_{t_j} に変化するということである。また、 a は属性の種類、 g は事象が時間的か空間的かを表すパラメータ、 k は属性の基準値である。ここでは、この原子軌跡式を図 1 のような素性構造で記述することにする。

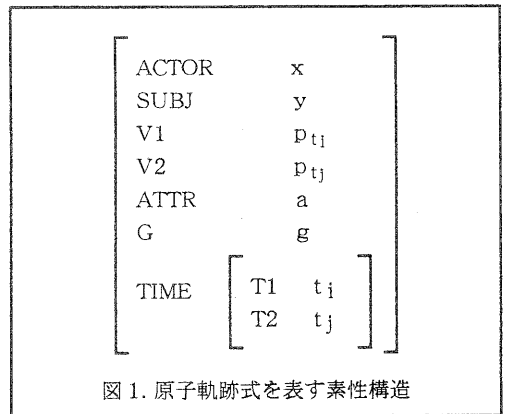


図 1. 原子軌跡式を表す素性構造

ここに、属性の基準値 k は省略した。また、TIME 素性に対応するものは式 (1) には現れていないが、時間的結合子を表現するために導入した。これについては、2. 3 節で詳述する。

2. 2 集合を用いた概念の記述

物や事概念を記述するのに積標準形を使用し、さらに、これを集合記述を用いて表すことにする。この様子を図 2 に示す。ただし、 \wedge は連言を表す記号であり、 $\Psi_i (i = 1, 2, \dots, N)$ は述語論理式 Φ_i の素性構造による表現を表すものとする。

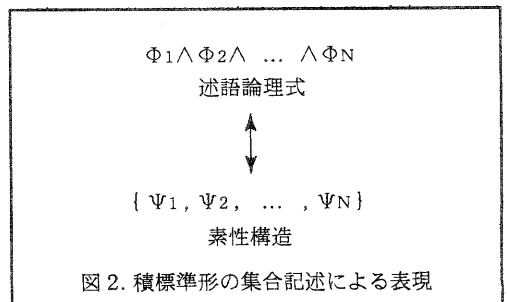


図 2. 積標準形の集合記述による表現

集合記述を素性構造に導入したので、集合記述の単一化について定義しておく必要がある。ここでは、定義として次のようなものを与えることにする。

定義 1 AとBを集合記述であるとする。このとき、AからBへの一対一写像fが存在して、Aの任意の要素aに対してaとf(a)が単一化するとき、そしてこのときに限り、AとBは単一化するという。

$A = \{ a_1, a_2, \dots, a_N \}$ とすると、定義1の条件を満たす写像fが存在すれば、単一化の結果は直観的に次の式(2)のようになる。ここに、 \wedge は単一化を表す記号である。

$$\{ a_1 \wedge f(a_1), a_2 \wedge f(a_2), \dots, a_N \wedge f(a_N) \} \quad (2)$$

ただし、定義1の条件を満たす写像は一つとは限らないので、一般には式(2)のようなものを選言 \vee によって連結したものになる。集合記述の単一化の例を図3に示す。ここでは、縦棒を使用した記法も示されているが、これは Prolog のリスト記述の場合とよく似た意味で使用されている。すなわち、 $A = \{ a | X \}$ とBの単一化を考えると、aがBのどれかの要素bと単一化するときAとBは単一化に成功し、XはBから要素bを除いた集合と単一化する。

(1)

$$\left\{ \left[A \ a \right], \left[B \ b \right] \right\} \wedge \left\{ \left[B \ b \right], \left[B \ c \right] \right\}$$

$$= \left\{ \left[A \ a \right], \left[B \ b \right] \right\}$$

(2)

$$\left\{ \left[A \ a \right], \left[B \ b \right] \right\} \wedge \left\{ \left[C \ c \right], \left[D \ d \right] \right\}$$

$$= \left\{ \left[A \ a \right], \left[B \ b \right], \left[C \ c \right], \left[D \ d \right] \right\} \vee \left\{ \left[A \ a \right], \left[B \ b \right], \left[D \ d \right], \left[C \ c \right] \right\}$$

(3)

$$\left\{ \left[A \ a \right], \left[A \ b \right] \right\} \wedge \left\{ \left[A \ b \right] | X \right\}$$

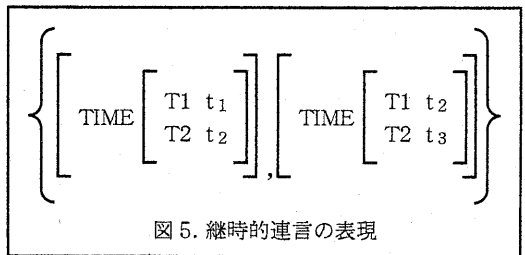
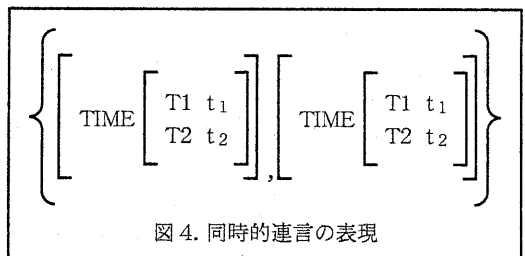
$$= \left\{ \left[A \ a \right], \left[A \ b \right] \right\}, \quad X = \left\{ \left[A \ a \right] \right\}$$

図3. 集合記述の単一化の例

2. 3 時間的結合子の表現

文生成のためには論理式の平坦な表現が必要になるが、心像意味論ではもともと平坦に近い表現がとられているので、大幅な書き換えは必要でない。ただし、時間的結合子[1]のために平坦でない表現になってしまっているため、ここは修正が必要である。

時間的結合子は原子軌跡式を結合して軌跡式を作るもので、同時的連言(\sqcap)と継時的連言(\cdot)がある。同時的連言は、二つの事象が同じ時間帯に起こっていることを表し、継時的連言は、一つの事象が起こった直後に、別の事象が起こることを表す。ここでは、これらの演算子を2.1節で導入したTIME素性を使用して表現することにする。この様子を図4と図5に示している。



3. 素性構造による結合操作部の表現

従来のシステムの結合操作部は、概念部に対する処理を手続き的に記述していた。ここでは、新たに素性を導入することにより、結合操作部が担っていた処理を、素性構造の単一化によって行なうことを考える。

3. 1 結合操作部の概要

新たな素性を導入する前に、従来から用いられてきた結合操作部について簡単に説明しておく。結合操作部は、操作命令と条件命令より構成される。

操作命令には次の3種類がある。

(1) ARG(Γ_1, Γ_2)

当該単語の概念部の変項 Γ_1 に、その単語と依存関係にある単語の概念部の変項 Γ_2 を代入する。

(2) PAT($\Gamma_1, \Gamma_2, \Gamma_3$)

事象パターン Γ_2 を Γ_1 と Γ_3 の両方が共に含むときに限り、 Γ_2 を共有するように Γ_1 と Γ_3 を結合する。

(3) LOG($\Gamma_1, \Gamma_2, \Gamma_3$)

概念 Γ_2 を Γ_1 と Γ_3 の両方が共に含むときに限り、 Γ_2 を共有するように Γ_1 と Γ_3 を結合する。

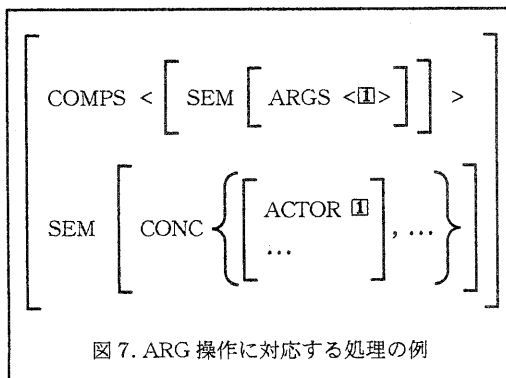
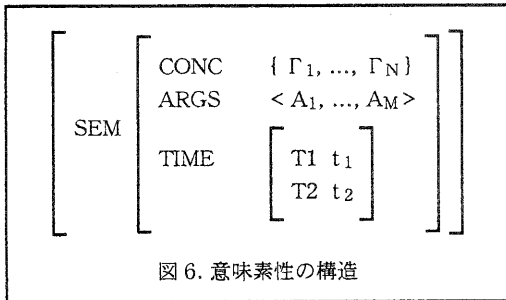
ただし、PAT 操作と LOG 操作の両方において、 Γ_2 が演算子の時はこれを用いて Γ_1 と Γ_3 を結合することになる。操作命令は次の条件命令との組合せにより、一つの結合操作を形成する。

(4) COND(Γ_1, Γ_2) $\rightarrow O_1, O_2, \dots, O_n$;

単語 Γ_1 の品詞が Γ_2 であるときに、 O_1 から O_n までを実行する。ただし、 O_1 は一つの操作命令を表す。

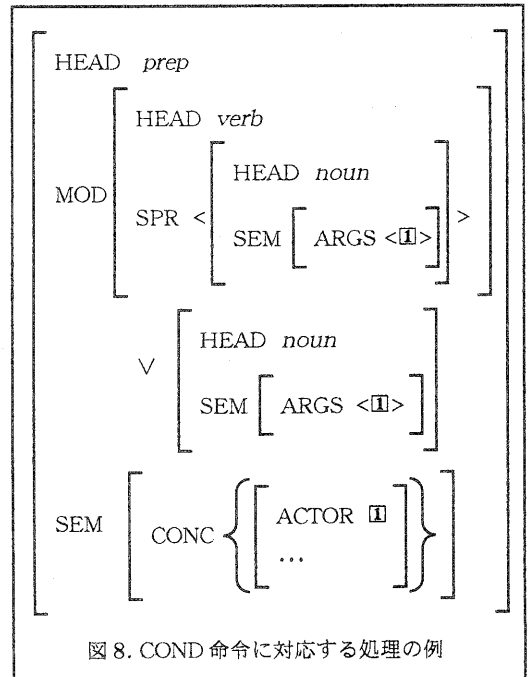
3. 2 概念部の処理に必要な素性

図6に示すように、概念部に対する素性として、CONC 素性を導入する。さらに、概念部の処理に必要な素性として ARG 素性と TIME 素性を導入し、これら三つの素性を合わせて SEM 素性（意味素性）とする。このうち、ARG 素性はリストであり、当



該句の概念部の主要な変項を表している。この素性は、ARG 操作に相当する処理を単一化によって行なうために使用する。たとえば、図7の例においては、当該句の概念部の変項を補助部の変項と単一化している。

一方、TIME 素性は概念部が表す事象が起きている時間帯を表す。PAT 操作において Γ_2 が時間結合子の場合に行なわれる処理と同様な処理を単一化によって行なうために、この素性が使用される。しかし、PAT 操作や LOG 操作と同等の処理は主として意味原理を用いて行なわれるので、詳細は次章で議論することにする。なお、COND 命令と同様な処理は、選言によって結合された素性構造によって可能である。たとえば、図8に前置詞 with に対する記述を示しているが、被修飾句の品詞が動詞であるか名詞であるかによって異なった処理がなされるのが分かるだろう。



4. 意味原理

前章までに心像意味論に基づく意味解析に必要な素性はすべて揃ったので、ここで意味原理について述べることにする。さらに、PAT 操作や LOG 操作と同等の処理を意味原理を使って実行できることも示す。

4. 1 意味合成原理と意味継承原理

心像意味論に対する意味原理は、次の二つにより

構成される。

(1) 意味合成原理

適格な句構造においては、母親の CONC 値はすべての娘の CONC 値の和集合になる。

(2) 意味継承原理

主辞を持った句において、母親の ARGS と TIME の値は、主辞となる娘のそれぞれの値と同一である。ただし、主辞でない娘が修飾句であるときは、母親の TIME 値はその修飾句の TIME 値と同一である。

意味合成原理において、すべての娘の CONC 値を寄せ集めたものが母親の CONC 値になるが、これは集合の直和を意味しない。心像意味論の場合は、異なる娘の CONC 値の中に全く同一の要素が出現するこ

ともあり得るので、そのような場合、一方だけを母親の CONC 値に含ませる必要がある。すなわち、母親の CONC 値はすべての娘の CONC 値の正確に和集合でなければならないのである。なお、図 9 の例において、上の二つの意味原理の効果を示している。二つの娘の CONC 値の和集合が母親の CONC 値になっているが、左側の娘の CONC 値の第 1 要素と右側の娘の CONC 値の第 1 要素が一致しているため、一方だけが母親の CONC 値に含まれている。また、主辞の ARGS 値が母親の ARGS 値に、修飾句の TIME 値が母親の TIME 値になっている点にも注意されたい。

4. 2 PAT 操作と LOG 操作に対応する処理

意味原理についての議論が終わったので、最後に、

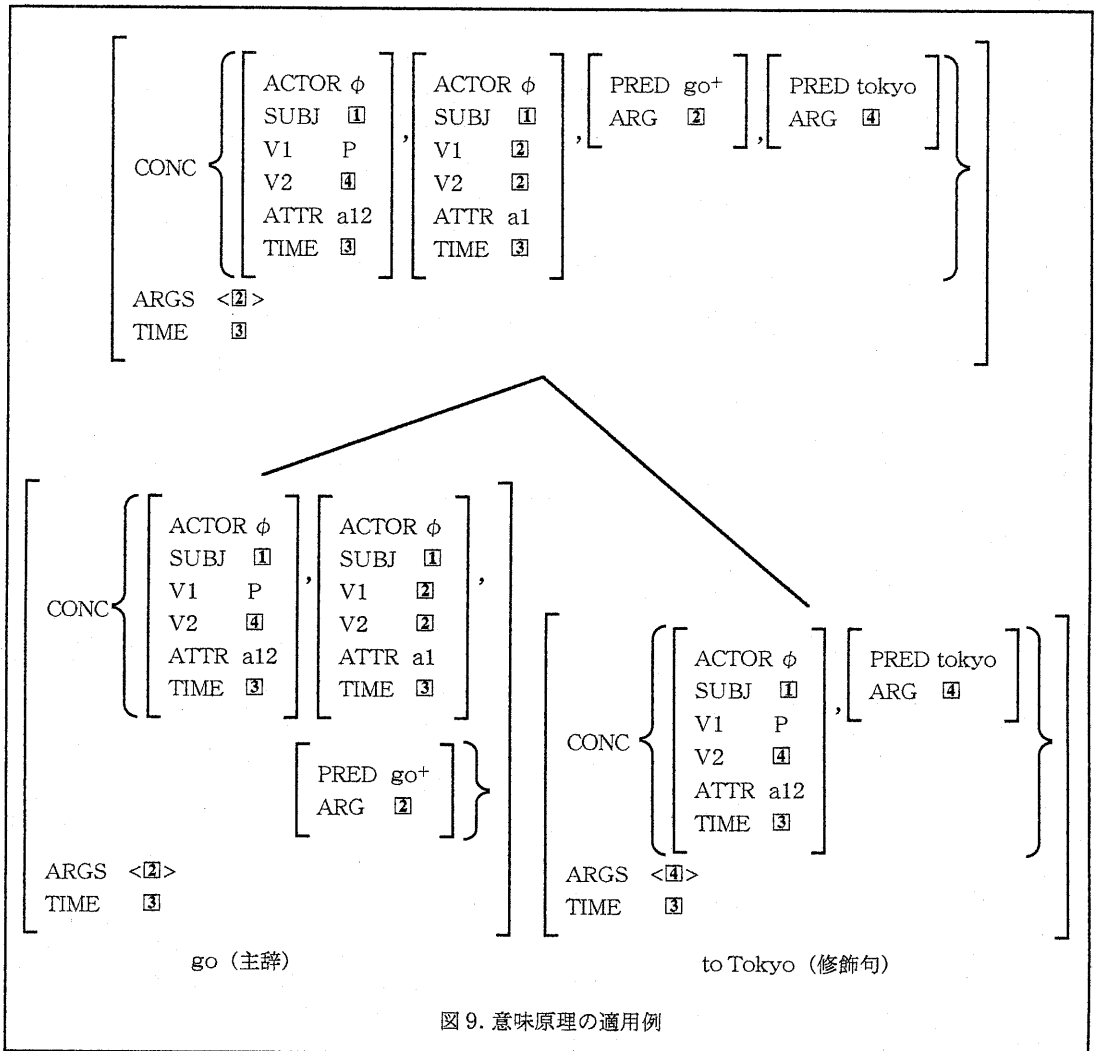


図 9. 意味原理の適用例

PAT 操作や LOG 操作に対応する処理について説明する。まず、PAT($\Gamma_1, \Gamma_2, \Gamma_3$)において Γ_2 が時間的結合子の場合を考えてみよう。このときは、TIME 素性を用いて、同時的連言や継時的連言を表現すればよい。たとえば、図 10 において COMPS|FIRST|SEM|TIME|T2 は、SEM|CONC|E1|TIME|T1 と同一の値を共有しているので、意味合成原理によって COMPS|FIRST|SEM|CONC の値と SEM|CONC の値の和集合が求められるとき、補助部の概念と主辞の概念が継時的連言によって結合されることになる。なお、意味継承原理により、SEM|TIME の値が母親に引き継がれるが、これは補助部の概念の時間帯と主辞の概念の時間帯を合わせたものになっている。

次に、PAT($\Gamma_1, \Gamma_2, \Gamma_3$)において Γ_2 が事象パターンの場合を考えてみよう。これと同等の処理は、集合の単一化と意味合成原理を用いて行なうことができる。たとえば、図 11 は前置詞 to に対する記述で

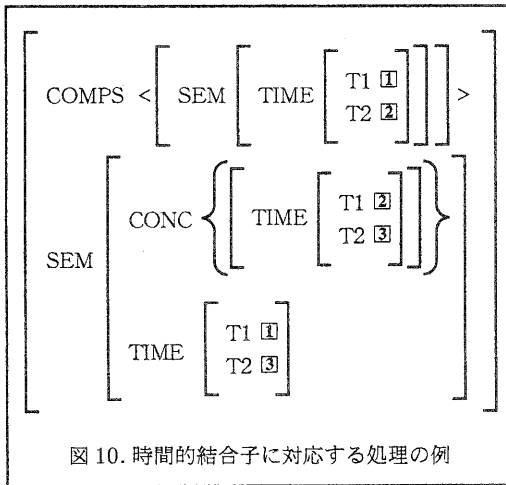


図 10. 時間的結合子に対応する処理の例

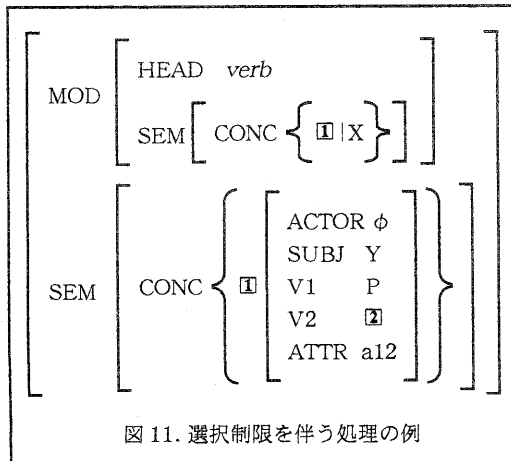


図 11. 選択制限を伴う処理の例

あるが、MOD|SEM|CONC と SEM|CONC が共通の要素を含まなければならないことが記述されている。したがって、意味合成原理によって CONC 値の和集合が求められるとき、修飾部の概念と主辞の概念が共通の要素を共有するように結合されることになる。なお、LOG 操作に対応する処理も同様なので省略する。

5. まとめ

本稿では、心像意味論に基づく意味解析を HPSG を用いて行なうために必要な素性と意味原理についての考察を行なった。その結果、従来、結合操作部で行っていた処理はすべて、HPSG によっても実現可能であることが分かった。特に、PAT 操作や LOG 操作に対応する処理は、意味合成原理を用いて行なうことが可能である。なお、ここで紹介した意味解析法は集合の単一化も含めて、すでにコンピュータ上に実現されており、自然言語文からの空間的概念事象の抽出などに利用されている。

今後は、心像意味論に基づく文生成を HPSG を用いてやってみたいと考えており、本稿で平坦な意味記述を採用したのは、このためでもある。我々は、生成を統語解析の逆の過程とみなす立場に立って、研究を行なう予定である。

参考文献

- [1] 横田将生: "人間の心像現象に基づく自然言語の意味記述について", 人工知能学会研究会資料, SIG-FAI-8802, 1988.
- [2] Barwise, J. and Perry, J.: "Situations and Attitudes", MIT Press, 1983.
- [3] Sag, I.A. and Wasow, T.: "Syntactic Theory: A Formal Introduction", CSLI Publications, Stanford University, 1999.
- [4] 横田, 吉武, 田町: "自然言語理解システム IMAGE S-I の意味解釈過程について", 信学論 (D), J69-D, 5, pp.777-784, 1986.
- [5] 横田, 白石, 笠, 西村, 田町, 寺田: "自然言語理解システム IMAGES-II", 信学論 (D-II), J74-D-II, 9, pp.1243-1254, 1991.
- [6] 笠兎一, 横田将生: "心像意味論と単一化文法の融合による意味解析の試み", 情報処理学会研究報告, Vol.97, No.109, pp.127-132, 1997.
- [7] Kay, M.: "Chart Generation", Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics (ACL-96), Santa Cruz, CA, pp.200-204, 1996.