

心像意味論と単一化文法の融合による 意味解析の試み

笠 晃一 横田 将生
福岡工業大学

心像意味論は、人間の心に生ずるイメージに基づいて構築された理論であり、モンタギュー意味論や状況意味論などと比べ、より深い意味の記述が可能である。心像意味論に基づくシステムは、今までにもいくつか作成されてきたが、それらはすべて文法理論として依存文法を採用していた。文法理論には、この他にも単一化文法というのがあり、これを用いれば、統語論的情報、意味論的情報、語用論的情報などを素性構造とその単一化という枠組みによって記述することによって、統一的な形式的体系のもとに理論を展開することが可能となる。ここでは、単一化文法の一つである句構造文法と心像意味論を融合し、日本語文の意味解析を試みた。

An Experiment on Semantic Analysis by Fusion of Mental-image Directed Semantic Theory and Unification Grammar

Koichi RYU Masao YOKOTA
Fukuoka Institute of Technology

Mental-image directed semantic theory (MIDST) can describe meanings deeper than Montague semantics or situation semantics because it is based on the images evoked in human mind. Natural language understanding systems based on MIDST have been constructed before, and all these systems used dependency grammar. Unification grammar, another grammatical theory, describes syntactic, semantic or pragmatical information by the scheme of feature structures and their unification. Making use of it, we can develop theories by uniform formal description. In this paper, we attempt to analyze meanings of Japanese sentences by fusion of MIDST and phrase structure grammar which is a sort of unification grammar.

1. はじめに

心像意味論 [1] は、人間の心の中に生ずるイメージ (心像) に基づいて構築された理論である。つまり、人間の生の知覚経験記憶を座標とするいくつかの属性空間なるものを仮定し、事物概念をそこで形成される軌跡と対応づけている。このため、従来のモンタギュー意味論や状況意味論などと比べ、より深い意味の記述が可能となっている。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おり、それらは、IMAGES - I [2]、IMAGES - II [3]、IMAGES - III [4] として発表されている。しかしながら、これらのシステムでは、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全く独立したものになっていた。

依存文法とは別に、単一化文法というものが知られている。これは、素性構造の単一化に基づく文法理論であり、統語的情報、意味的情報、語用論的情報を、すべて素性構造とその単一化という共通の枠組みを用いて記述することが可能である。このため、統一的な形式的体系に基づいて、これらの情報に関する議論を展開することが可能になる。また、統語解析と意味解析を同時並行的に行なうことによって、統語的な組合せの爆発を抑制できるという利点もある。

単一化文法として、語彙機能文法 (LFG)、一般化句構造文法 (GPSG)、主要部駆動句構造文法 (HPSG) などが知られているが、本研究では、日本語句構造文法 (JPSG) [5] を用いて日本語の文から意味を抽出し、それを心像意味論を用いて表現することを試みた。心像意味論に基づく自然言語理解システムのうち、今までに作成されたものはすべて、単語の意味として、概念部と結合操作部を持たせていた。概念部は論理式により宣言的に記述されるが、結合操作部では、依存関係にある単語の概念部同士を結合させる操作手順が手続的に記述される。単一化文法を用いれば、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、意味を二つの部分に分ける必要性はなくなる。これも、単一化文

法を用いる利点である。以下では、まず、心像意味論における意味記述について簡単に説明した上で、JPSGによる意味処理について述べることにする。

2. 心像意味論における意味表現

人間が持つ概念の中で最も基本的なのは、事物の概念である。事物の概念は、物質あるいは物体である外界物と、それらが形成する関係状態であるところの事象とに区分されるが、ここでは、事象の記述法についてのみ説明することにする。

2. 1 事象概念の定義

心像意味論では、事物の概念を記述するのに一階述語論理を用いる。たとえば、事象概念は次のように定義される。

$$E(y_0) \triangleq E^{++}(y_0, y_1, \dots, y_n) \wedge E^+(y_0) \\ \wedge C_1(y_1) \wedge \dots \wedge C_n(y_n) \quad (1)$$

ただし、各記号は次のような意味である。

E : 事象 y_0 の概念

E^{++} : ある時間帯、事象 y_0 に事物 y_1, \dots, y_n が出現することと、それらの間の関係に関する概念

E^+ : E^{++} と同一時間帯における事象 y_0 自身の状態に関する概念

C_i : 事象 y_0 に出現する事物 y_i の概念

2. 2 軌跡式

属性空間における最も基本的な軌跡と対応づけられる式を原子軌跡式、原子軌跡式の論理的結合列を単に軌跡式と呼んでいる。そして、事象概念の定義において、 E^{++} と E^+ はこの軌跡式によって記述される。原子軌跡式は、次の式 (2) に示される述語式、または、それを略記した式 (3) によって表現される。

$$L(x, y, p_{ti}, p_{tj}, a, g, k) \quad (2)$$

$$(x, y, p_{ti}, p_{tj}, a) \quad (3)$$

これらの式の意味は、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_{ti} から p_{tj} に変化するということであり、 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。また、 $a = (a, g, k)$ である。

原子軌跡式や軌跡式を結びつける結合子には、同時的連言と継時的連言があり、それぞれ記号 \square と記号 \cdot によって表示される。同時的連言 $A \square B$ は、 A と B が同時に生起することを表すし、継時的連言 $A \cdot B$ は、 A に引き続いて B が生起することを表しているが、正確な定義については、別の論文 [1] を参照されたい。

2. 3 事象概念の簡略記法

ここで、「太郎が本を運ぶ」という事象を軌跡式を用いて表現してみると、次のようになる。

$$\begin{aligned} & (\phi, y, x, x, a'_{11}) \square \\ & (\phi, z, x, x, a'_{11}) \square \\ & (\phi, y, p, q, a'_{12}) \square \\ & (\phi, z, p, q, a'_{12}) \square \\ & \wedge \text{運ぶ}^+(x) \wedge \text{太郎}(y) \wedge \text{本}(z) \quad (4) \end{aligned}$$

ただし、 a'_{11} は事物の存在場所を表す属性であるので、式 (4) の一段目と二段目は、それぞれ y と z が、事象 x に出現することを表している。また、 a'_{12} は物の物理空間での位置を表す属性であり、式 (4) の三段目と四段目は、それぞれ y と z が、位置 p から別の位置 q まで移動することを表している。

しかしながら、上記のような意味表現は、計算機に載せるには若干煩雑であり、もう少し簡略化しても、基本的意味処理能力への本質的な影響は少ない [4]。そこで、次のような簡略化を行なうことにする。

(1) 事物が当該事象に出現することを示す軌跡

式を省略する。

- (2) 事象自身の属性に関する概念 (E^+) を省略する。
- (3) 当該事象に出現する事物概念 (C_i) そのものを軌跡式の項とする。
- (4) 属性値は単語そのもので代用する。
- (5) パラメータ g と k を省略する。

たとえば、これらの簡略化を式 (4) に適用すると、次の式 (5) のようになる。

$$\begin{aligned} & (\phi, \text{太郎}, p, q, a_{12}) \square \\ & (\phi, \text{本}, p, q, a_{12}) \quad (5) \end{aligned}$$

以下の説明における事象概念の記述には、すべて簡略記法を用いることにする。

3. J P S G による軌跡式の抽出

この節では、J P S G を用いて日本語の文から軌跡式を抽出する方法について述べる。しかし、その前に、軌跡式や論理的結合子の素性構造による記述について述べておく必要がある。なお、J P S G で用いる素性や構造原理などについては、いちいち説明するようなことはしないので、詳しいことは論文 [5] などを参照して欲しい。

3. 1 素性構造による軌跡式の記述

単一化文法を用いて自然言語の文から軌跡式を抽出するには、もともと述語論理で記述されている軌跡式を素性構造を用いて記述し直してやる必要がある。ここでは、式 (3) に対して、図 1 に

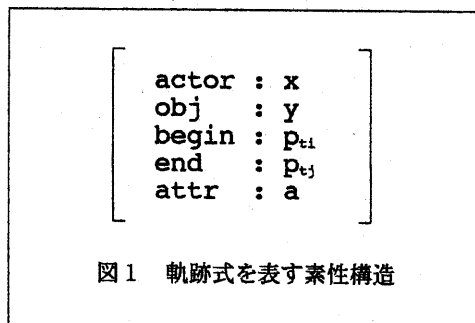
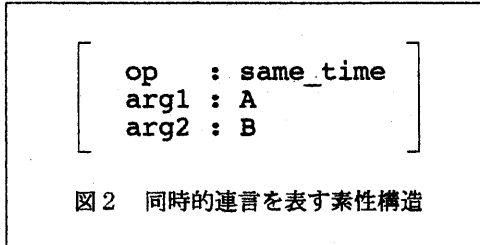


図1 軌跡式を表す素性構造

示すような素性構造を用いることにする。

3. 2 素性構造による結合子の記述

論理的結合子のうち、同時的連言 $A \sqcap B$ を図2に示すような素性構造によって表すことにする。継時的連言も同様の素性構造によって表される。



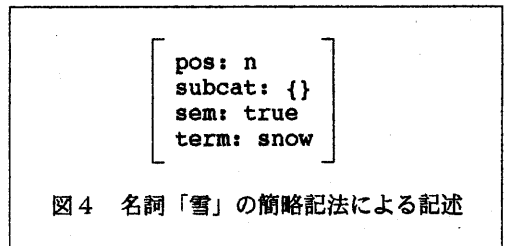
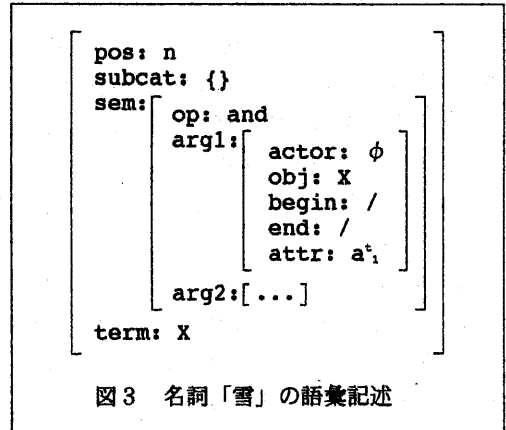
3. 3 名詞の意味と語彙記述

ここでは、「太郎」や「雪」などの外界物を表す名詞のみを考えることにする。心像意味論では、外界物も軌跡式などを用いて詳細に記述される。たとえば、「雪」の記述は次のようになる。

$$\begin{aligned}
 \text{snow}(x): & \\
 & (\phi, x, /, /, a^1_1) \wedge \\
 & (\phi, x, /, /, a^1_2) \wedge \dots \wedge \\
 & (\phi, x, \alpha_{11}, \alpha_{12}, a^1_{23}) \wedge \dots \wedge \\
 & (\phi, x, \alpha_{22}, \alpha_{23}, a^1_{32}) \wedge \dots \wedge \\
 & (y, v, x, z, a^1_{41}) \wedge \\
 & \text{heat}(y) \wedge \text{water}(z) \quad (6)
 \end{aligned}$$

したがって、JPSGにおける「雪」の記述として図3のようなもの考えることができる。ここに、sem 素性が単語の意味を表し、term 素性はこの単語に対する個体変数を表している。op 素性に対する素性値 and は論理積を表している。また、素性構造の記述において、英字の大文字は束縛されていない変数を表すものとする。

しかしながら、事象概念の記述には簡略記法を用いることにしたので、当該事象に出現する外界物は軌跡式の項（個体定数）として扱われることになる。これに伴い、「雪」の語彙記述として新



たに図4のようなもの考えることにする。この場合、名詞自身の意味はほとんど消失し、sem 素性の値は true になっている。これは、「真」を表す定数であり、次の式を満たすものとする。

$$\text{true} \sqcap A = A \sqcap \text{true} = A \quad (7)$$

$$\text{true} \cdot A = A \cdot \text{true} = A \quad (8)$$

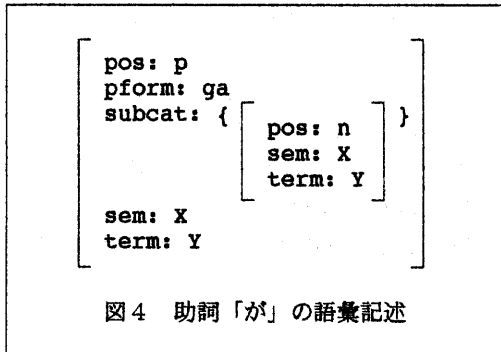
名詞の sem 素性がこのような値を取るの、次のような理由からである。つまり、名詞の意味は、軌跡式の中に個体定数として出現するだけであり、軌跡式そのものではない。他方、統語・意味解析の段階で、二つの句を結合して、より大きな句を作成する場合、新しい句の sem 素性の値は原則として、二つの句の sem 素性の値を同時的連言や継時的連言で結合したものになる。そこで、名詞の sem 素性の値として true を与えておけば、式(7)や式(8)によって、最終的にこの値は消滅する。ただし、名詞の意味を個体定数として出現させる必要があるので、別に term 素性というのを用意して、ここにその値を記述するこ

とにする。

3. 4 後置詞の語彙記述

ここでは、後置詞（助詞）として格助詞のみを考えることにする。格助詞の中でも、特に「が」と「を」は補語となる名詞句に対して選択制限を持たず、IMAGES - Iなどで使用していたPAT操作に相当する記述を行なう必要はない。ここに、PAT操作とは、結合操作部における操作命令の一種であり、事象パターンΓ₂をΓ₁およびΓ₃の両者がともに含むときに限り、Γ₂を共有するようにΓ₁およびΓ₃を結合する操作である。したがって、語彙記述も単純なものになり、たとえば、格助詞「が」に対する記述を考えると、図5のようになる。

ただし、これら以外の後置詞では、選択制限を扱わなければならない場合もあり、PAT操作に準

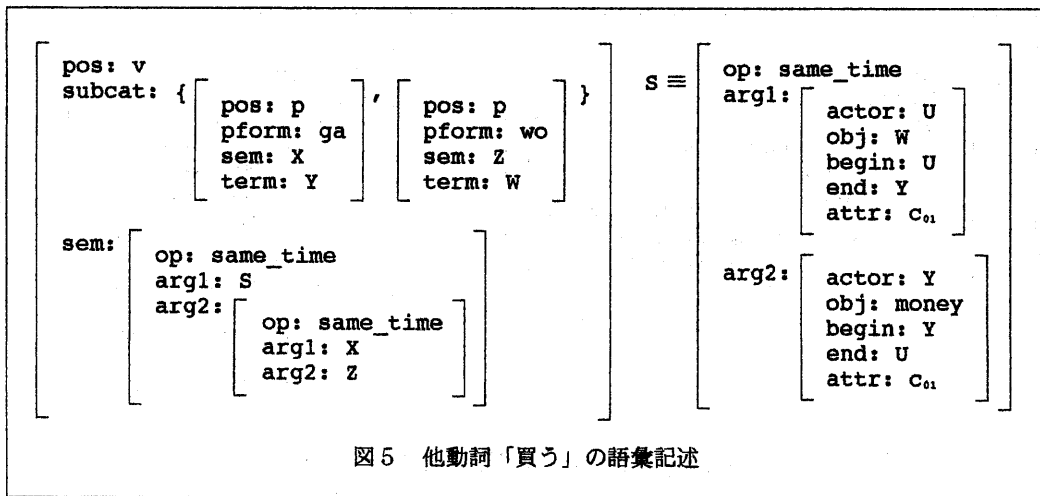


ずることを行なう必要が生ずるが、これは今後の課題とする。

3. 5 動詞の語彙記述

自然言語の文に心像意味論を適用すると、自動詞句の意味は、動詞自身の持つ意味に、主語が持つ意味を同時的連言などで結合したものになる。他動詞句の場合は、さらに目的語の持つ意味も結合される。したがって、たとえば、他動詞「買う」に対する語彙記述を考えると、図5のようになる。ここでは、主語と目的語の意味が同時的連言によって結合されている。

前に、名詞の意味は軌跡式の中の個体定数としてのみ出現するということを述べたが、これは、subcat 素性の中の term 素性の値を、sem 素性の中の軌跡式の対応する部分と共有することで実現している。つまり、素性構造を用いれば、IMAGES - Iなどで使用していたARG操作は自動的に実行され、実質的に不要となる。また、従来のIMAGESでは、動詞に対する選択制限が、LOG操作（PAT操作とほとんど同等であるが、事象パターンではなく事物概念に対する操作である）によって記述されていた。しかし、これも動詞の語彙記述に選択制限のための素性を導入するだけで同等の処理が可能となる。このように、素性構造を用いることにより、従来のIMA



GESで結合操作部として手続きの形で記述されていた部分が不要になり、単語を完全に宣言的に記述できるようになる。

3. 6 統語・意味解析について

単一化文法では、統語解析と意味解析は同時並行的に行なわれ、それはJPSGも例外ではない。JPSGでは、さらに下位範疇化素性の原理や束縛素性の原理などの、いわゆる構造原理を用いて解析が進められる。

図6に、「太郎が本を買う」という日本語文に対する解析結果を示す。sem素性として、文全体の意味が得られているが、前出の式(7)を用いることにより、図中のS自身にまで簡略化することが可能である。つまり、与えられた文の意味はSになる。このSにおいて、taroは「太郎」を、bookは「本」を、また、moneyは「お金」を、それぞれ表している。さらに、変数Uは未束縛であるが、これは「太郎が本を買う相手」を表している。c₀₁は、所有概念を表す属性である。以上より、「太郎が本を買う」という文の意味は、「相手が相手から太郎に本の所有権を移す」という事象と「太郎が太郎から相手にお金の所有権を移す」という事象の同時的連言として得られ、期待どおりの結果となっている。

4. まとめ

本稿では、心像意味論と単一化文法の一つであるJPSGを用いて、日本語文の意味を抽出する方法について説明した。紙面の都合で紹介できなかったが、形容詞の語彙記述や複文の処理もここで述べた方法の拡張によって行なえることが分かっている。今後は、理論をより精密化するとともに計算機上に実働システムを構築してみたいと考えている。

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] 横田, 吉武, 田町: "自然言語理解システムIMAGES-Iの意味解釈過程について", 信学論(D), J69-D, 5, pp.777-784 (1986)
- [3] 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)
- [4] 白石正人: "談話理解システムIMAGES-IIIの試作 — 内界事物概念の意味理解 —", 昭和59年度九州大学大学院修士論文 (1984)
- [5] 郡司隆男: "自然言語の文法理論", 産業図書

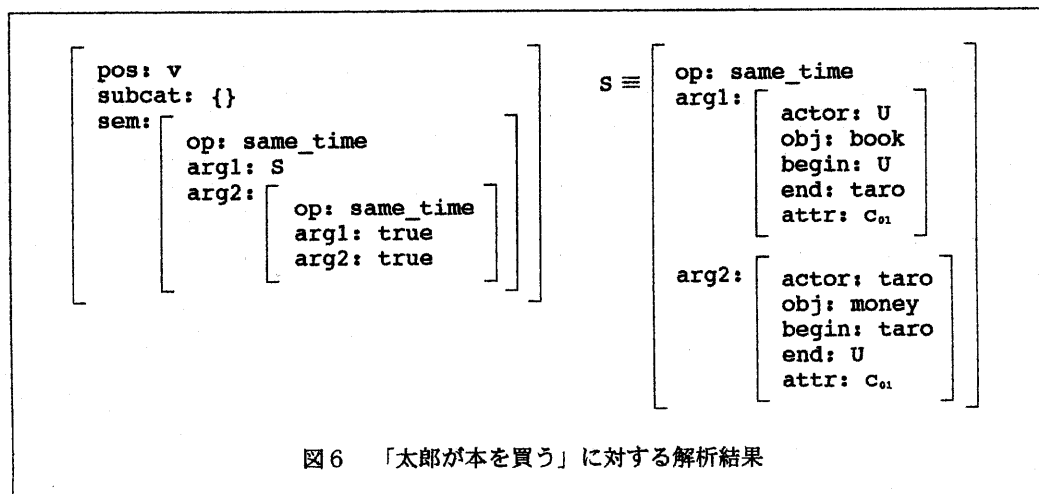


図6 「太郎が本を買う」に対する解析結果