

コア概念を用いた英語句動詞の意味解析

川辺 諭

朝倉慎一

宮崎正弘

新潟大学大学院自然科学研究科

1 はじめに

英語文解析において、前置詞句や句動詞の意味解析や訳し分けは、文脈に応じて着目されている名詞の意味合いを特定せねばならず、難しい問題の1つである。言語学者田中茂範によって提唱されているコア理論[1][2]では、単語は文脈によらない共通の概念を持つものとされている。本稿ではこのコア理論の考え方に基いて、前置詞が表現し得る名詞や動詞のさまざまな関係をあらかじめ計算機に記述しておき、それらの関係を自動抽出することにより、前置詞句、句動詞の持つ意味合いを解析する方法を提案する。実装には宣言型言語である prolog を用いている。

は、上位下位といった関係付けを行なうことにより、木構造を構築している。

表 1: 実体の性質を示すさまざまな属性

具体属性	属性(属性値)の例
視覚情報	a_色(赤)
聴覚情報	a_音色(男性)
味覚情報	a_味(コーヒー)
嗅覚情報	a_匂い(コーヒー)
触覚情報	a_触感(暑い)
有機情報	a_感覚(疲れ)

抽象属性	属性(属性値)の例
抽象属性	性別(女)
空間属性	空間属性(空間)

2 言語データ

本稿で述べる前置詞句、句動詞解析に用いられる言語データについて述べる。

2.1 属性辞書

属性辞書には、文中の名詞が指示する実体の性質を説明する要素の中でも、もっとも原始的な要素であるプリミティブを記述する。このような属性として、感覚器官によって認知される知覚属性などの具体属性、空間のように大脳によって推論される抽象属性、2つ以上の実体の間に存在する関係属性を準備している(表1)。またこれらの属性

関係	関係(2実体)の例
空間関係	内部(c_人, c_建物)
位置関係	上(c_りんご, c_机)
数値関係	大(c_数値, c_数値範囲)
時間関係	内部(script_事象, c_時間範囲)
社会関係	構成員(c_社員, c_組織)

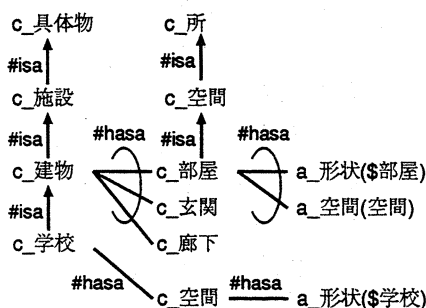


図 1: 多次元シソーラス

Semantic Analysis of English Phrasal Verbs by using Core Concepts

Satoru Kawabe(kawabe@tinlp.info.eng.niigata-u.ac.jp), Shinichi Asakura, Masahiro Miyazaki (miyazaki@info.eng.niigata-u.ac.jp)

Niigata University

2.2 実体辞書

実体は前節で述べた属性や、実体の集合として記述される。さらに、実体の間に上位下位といった概念の関係付けを行なう。その際、概念の多重継承を行なうことにより、実体概念のネットワークである多次元シソーラス [3] を構築している (図 1)。

2.3 動詞辞書

動詞辞書には、英語基本動詞のコア概念が表現する統語的、意味的格要素となる実体のクラスと、それぞれの実体の属性変化をリスト形式で記述する (図 2)。

```
[verb, v_break % 動詞クラス
 [clist, % 格要素リスト
 [主格, c_animate, % 統語格, 実体クラス
 [alist, % 属性変化リスト
 [atrb, 心理, [意志, 満足]]],
 [対格, c_conc,
 [alist,
 [atrb, 機能, [有, 無]]]]].
```

図 2: 動詞コア概念の格文法流記述

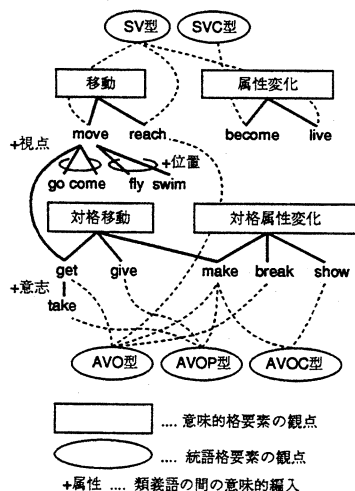


図 3: 動詞類語弁別ネットワーク

またこの形式で記述された動詞を統語的、意味的な観点から分類し、類語弁別ネットワーク [4] を構成する (図 3)。さらに、生活習慣としてある程度定まっている、動作の特定の流れを、スクリプトとして記述している。

2.4 空間詞辞書

空間詞辞書¹には、前置詞、副詞が表現する実体同士の関係を記述している。表 2 に空間の分類と、空間の要素となる実体の焦点となる属性を示す。

表 2: 空間詞の対象となる空間と焦点化属性

空間	焦点化属性	言語表現の例
物理空間	空間関係	in the box.
配置空間	位置関係	near my house.
数値空間	数値関係	over 20 years old.
時間空間	時間関係	at the age of 12.
社会空間	社会関係	with the company.
心理空間	心理属性	in love.
状態空間	状態属性	in a jam.
事象空間	スクリプト	in the office.

前置詞句の解析を行なうために必要な言語データとして、

- (1) 前置詞の種類
 - (2) 先行詞となる実体の焦点化属性
 - (3) 前置詞目的語となる実体の焦点化属性
- を記述する。また、前置詞や副詞を用いた句動詞表現を解析するために必要な言語データとしては、
- (1) 動詞型
 - (2) 補充される意味格要素
 - (3) 目的語となる実体の焦点化属性
- を記述する (図 4)。前置詞辞書に記述されるこれらの概念は、基本的にただ 1 つのコア概念から派生したものであるが、比喩の解析や推論を行なう必要があるので、現段階では派生した意味合いを網羅的に記述している。

3 空間詞解析

空間詞解析部では、表 3 に示される前置詞句、句動詞の解析を行なう。

¹ 田中茂範は前置詞、副詞を実体の空間的関係を捉える語として“空間詞”と呼んでいるので、ここではそれに準じている。

```

prepdic(prepare_at,X):-
  X=[実体関係リスト,
    % 前置詞句の例
    [nprep, % 統語型
      [対象実体, % 先行詞, 前置詞目的語クラス
        [実体,s1,c_具体物],
        [実体,s2,c_所]],
      [付加属性, % 実体に付加される属性
        [関係,s1,[空間関係, 内部,s1,s2]],
        [関係,s2,[空間関係, 内部,s1,s2]]]],
    % 句動詞の例
    [vtprep,
      [付加格要素,
        % [動詞型, 格要素, 実体の条件],
        [移動, 終点格,c_所],
        [存在, 場所格,c_所]]]]].

```

図 4: 空間詞辞書の記述例

表 3: 空間詞解析部での解析内容

表現型	言語表現の例
n1 prep n2	pens on the desk.
vt prep n	get on it.
vt n adv	put it on.
vi adv	wake up.

※表中の略号 n:名詞句,prep:前置詞

vt:他動詞,vi:自動詞,adv:副詞

3.1 前置詞句解析

前置詞句解析では、“先行詞+前置詞+前置詞目的語”といった形式の表現の解析を行う。

(1) “a man at the park.”

という表現の解析に関して、解析の様子を具体的に説明する(表 2の空間詞辞書記述、図 5の解析の様子を参照)。

空間詞辞書には前置詞 “at” に関して、“先行詞としてクラス c_具体物の実体を、前置詞目的語としてクラス c_所の実体をとる時に、それら2つの実体の間に、— c_具体物が c_所に存在している — という位置関係を表現する” という意味内容の記述がある。また、実体辞書を参照すると、実体 “man”

は c_人 というクラスの実体であり、クラス c_人は多次元シソーラス上で c_具体物という上位概念を持つ。また同様に実体 “park” は c_施設というクラスの実体であり、クラス c_施設は多次元シソーラス上で c_所という上位概念を持つことが推論される。これらのことから、先行詞と前置詞目的語の性質(具体的には上位クラス)が、前置詞の要求する条件を満たすので解析は成功し、“a man at the park.” といった前置詞句表現から、“a man” という実体が “the park” という実体の内部に存在するという位置関係属性が抽出される。この解析結果は、「公園の人」「公園にいる人」といった日本語表現に対応する。

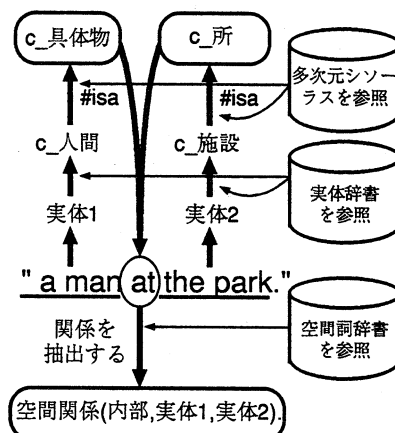


図 5: 前置詞句解析の様子

(2) “a man at the office.”

という表現の場合は、(1) とほぼ同様に c_オフィスが上位概念として c_所を持っているため、人がオフィスに存在しているといった位置関係が抽出され、これは「オフィスにいる人」といった日本語表現に対応する。さらに別解として、c_office は script_仕事といったスクリプトを持っており、それらの属性に着目した解析結果として、c_man に script_仕事という属性が付加される。これは「オフィスで仕事をしている人」といった日本語表現に対応する。

3.2 句動詞解析

句動詞解析では、“動詞+前置詞+名詞句”や、“動詞+名詞句+副詞”といった形式の表現の解析を行

なう。

(1) “get on the bus.”

空間詞辞書では前置詞 on に関して以下の関係リストが記述されている。

[前置詞関係リスト,
..
[vtpprepn,
[付加格要素,
[移動動詞, 場所格, c_平面]]],
..]

動詞 v_get のコアは“対格を主格の内部に取り入れる”といったものであり、この意味合いから派生して“動作主体の位置属性を取り入れる”という、動作主体の移動の意味合いを持つ。そのため、動詞類語弁別ネットワーク上では v_get は移動動詞ノードの動詞としても記述されている。また実体辞書には、クラス c_バスは c_平面という属性を持つと記述されている。これらの属性は前述の関係リストの条件を満たすので、解析の結果、“クラス c_バスの持つ c_平面へ移動する”といった意味合いが抽出される。これは「バスに乗る」といった日本語表現に対応する。

(2) “put the hat on.”

副詞 on に関しては、以下の関係リストが記述されている。

[副詞関係リスト,
..
[vttnadv,
[付加格要素,
[対格変化動詞,
[位置関係, 依存, c_衣服, c_体]]]],
..].

言語データを照合することにより、(1)と同様に動詞 v_put は対格変化動詞(動作主体が目的語の属性を変化させる動詞)であり、クラス c_帽子は c_衣服の下位概念であることが解析される。また、この表現は“on the body”といった表現が省略されており、このような意味内容を補完できるように、熟語知識としてあらかじめ記述してある。解析の結果、“クラス c_衣服を、クラス c_体に付着した状態に変化させる”といった意味合いが抽出される。これは「帽子をかぶる」といった日本語表現に対応する。

4 おわりに

動詞や名詞の持つさまざまな性質から、前置詞、副詞を用いた前置詞句、句動詞表現の持つ意味合いを自動抽出する方法を提案した。この方法を用いることにより、英語文で数多く出現する未知の前置詞句、句動詞表現において、名詞や動詞、前置詞や副詞の持つ意味内容を推論することにより、可能な解釈を抽出できるようになる。今後の課題としては、

1. 文脈情報などを用いての、抽出された解釈の確からしさの推定。
2. 対応する日本語表現生成部の実装。
3. 言語辞書の拡充。

などが挙げられる。

参考文献

- [1] 田中茂範:認知意味論 英語基本動詞の多義構造, 三友社出版 (1990).
- [2] 田中茂範 編著:基本動詞の意味論, 三友社出版 (1987).
- [3] 川村, 片桐, 宮崎:語種々の観点から分類した多次元シソーラス, 信学技法, NLC94-48, pp.33-40(1995).
- [4] 中田, 宮崎:類語弁別ネットワークによる語義記述法, 自然言語処理の基本問題シンポジウム論文集, pp.1-8(1992).