

日英の構文的対応関係に着目した用言意味属性の分類

5P-7

中岩浩巳 横尾昭男 池原悟

NTT 情報通信網研究所

1. はじめに

近年、様々な機械翻訳システムが実用化されているが、今だ十分な訳文品質を得ているシステムはないと言える。特に、省略や照応解析などの文脈解析を必要とする言語現象の解析が不十分である。この原因は、機械翻訳システムが取り扱う言語現象は限定することが困難であるので、文脈などによる制約を導入するためには、文と文との意味的関係の決定等に利用する、語がもつ意味(語義)に関する膨大な知識が必要となる点にある。

知識量の爆発を避けるための手段として、語のもつ意味を分類して、その語のもつ代表的属性値によって、語と語や文と文の関係を決定していく手法が考えられる。特に文脈処理では、文の流れを解析するためにキーとなる用言の意味的属性が重要な知識となる。

用言の分類に関しては従来から様々な研究がなされているが⁽¹⁾⁽²⁾、用言自身の意味だけを分類したものが多く、語義と文中での用法の関係に着目して分類したものは少ない。このため、用言の多義解消などの解析の面で十分な効果が期待できなかった。

本稿では用言のもつ語義と用法の関係に着目して、用言意味属性を分類する。本意味属性は日英構造変換辞書での各用言の用法定義に使用され、文中で使用された複数の用言の意味的関係を追跡する際のキーとなる。

2. 用言意味属性の付与対象

本稿で提案する用言意味属性は、図1に示すような、用言を中心とした2種類の日英構造変換辞書(意味的結合価パターン変換辞書、慣用表現変換辞書)に付与することを前提に考案したものである。これらの辞書は、日本語用言(慣用表現変換辞書においては、「油を売る」等の慣用表現のコア部分)とその格要素への意味的制約からなる日本語単文パターンと、その日本語表現に対応する英語表現である英語パターンの対からなる。統語構造が大きく異なる日本語と英語の二言語間で共通な語義とその用法を示す表現を対で保持することにより、概念的な曖昧性の排除が可能となるため、詳細で正確な属性値の付与が可能となる。また、個々のパターンには日本語格要素への意味的な制約が付与されており、機械翻訳で正確な訳語を得るための詳細な意味的条件が記述されているので、同一の用言に複数の属性が付与されていても、日本語解析の結果、適切な用言意味属性の選択が可能となる。さらに、個々の日英パ

ーン対に用言意味属性を付与する場合には、日本語の用言のもつ意味と共に英語の意味も参考にできるので、用言意味属性の付与が容易になる。

[意味的結合価パターン変換辞書]

N1 (主体) が N2 (文化/人間活動) を 暗記する。
=>N1 learn N2 by heart

N8 (施設) で N2 (動物) を 飼う。

=>N8 raise N2

N1 (主体) が N2 (食料) を 食べる。

=>N1 eat N2

[慣用表現変換辞書]

N1 (主体) は 背が高い =>N1 be tall

図1 日英構造変換辞書

*N1,N2,N8等は結合価のラベルを示し、括弧内は格への意味的制約を示す

3. 用言意味属性の分類基準

本稿で提案する日本語の用言分類体系は、高松らの用言分類⁽³⁾をもとにして、以下の2つの観点から拡張したものである。

・用言がもつ動的特性

用言のもつ概念と談話場面に与える作用による分類：用言単体が発話されることにより、どのような種類の行為がなされ、その発話の結果どのような状況になるかという観点で分類。

E x. 「持つ」 → Possession

「開発する」 → Production

「持つ」という用言は、Possessionの行為を行なうことを文脈の状況の中に提示するのに対し、「開発する」という用言は、何かをProduceする行為を行なうという事実を文脈の状況の中に提示する。

・用言の格に対する関係

用言の支配する格が用言に対してもつ役割による分類：格要素が支配される用言のもとで発話される場合、その格要素が状況の中でどのような役割を果たすかの観点で分類。

E x. 「完成する」

→ SUBJ が 完成する : SUBJ be produced

「開発する」

→ SUBJ が OBJ を 開発する: SUBJ produce OBJ

「完成する」、「開発する」とともにProductionの行為を示す用言であるが、「完成する」はSUBJをProduceするのに対し、「開発する」はSUBJがOBJをProduceすることを示す。

A System of Verbal Semantic Attributes Focused on the Syntactic Correspondence between Japanese and English.

Hiromi NAKAIWA, Akio YOKOO and Satoru IKEHARA

NTT Network Information Systems Laboratories

4. 用言意味属性体系

前節の基準で作成した用言意味属性の体系を図2に示す。図の左部は用言がもつ動的特性という観点(語義の側面)で分類し、右部はそれらについて用言の格に対する関係(用法の側面)で分類したものである。現在、約80種類の用言意味属性を設定している。

ここで提案する用言意味属性は、日英構造変換辞書の各日英パターン対(約15,000対)に付与している。これらのパターンの格要素には、概念を約2,800種類に分類した一般名詞意味属性体系の属性値が意味的制約として付与されており⁶⁾、これによって正確な構文構造の選択ができる。

用言単位でみれば意味的に多義が存在する場合でも、このようにパターン対単位でみれば意味の絞り込みが可能となる。これは、全パターン対の約9割は、付与した用言意味属性の数が1種類に特定できた点からも推測できる。

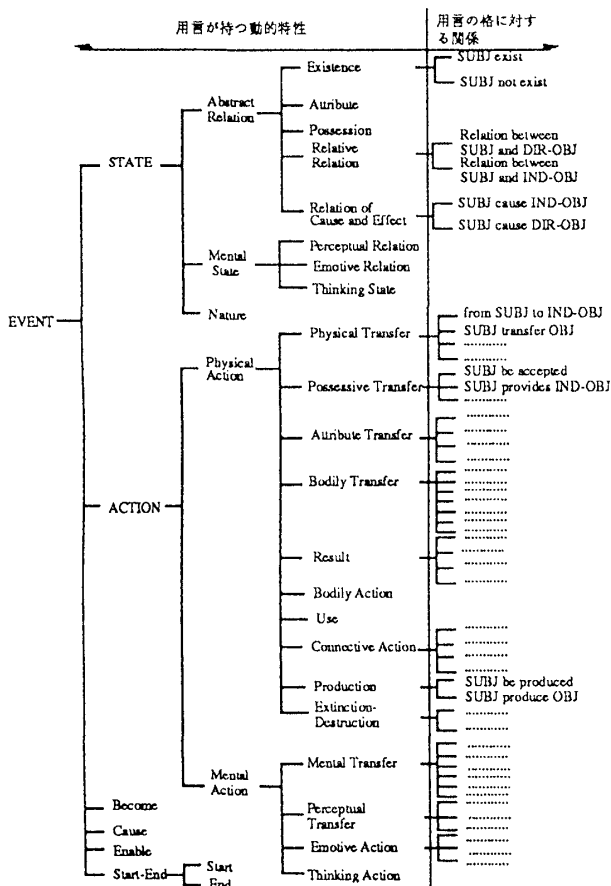


図2 用言意味属性体系

5. 用言意味属性の応用例

本稿で提案する用言意味属性は、以下のような処理に応用できる。

5.1 日本語ゼロ代名詞の照応解析⁷⁾

日本語テキスト中に現れるゼロ代名詞に対する、用言意味情報を用いた照応解析処理への応用が考えられる。本手

法は、用言がもつ意味的属性に着目し、ゼロ代名詞が支配される用言の意味属性と、それと照応する格要素候補が支配される用言の意味属性の関係から、文脈を調べ、照応格要素を決定するものである。この方法は、機械翻訳システムのように対象分野を限定することができない文章でも、用言の意味的属性の対による翻訳対象分野とは独立したルールを記述すればよいので、少ない知識量でゼロ代名詞の照応解析ができる。

5.2 省略格要素への文章外要素の補完処理

日本語の省略格要素に対して、文章中に現われない要素を補完する場合に用言意味属性が利用できる。このような省略現象の解析手法としては、格要素への意味的制約条件を用いて、制約となる意味属性に応じて補完要素を推定する手法が考えられる。例えば、図1の「N8(施設)でN2(動物)を飼う」のN8が省略された場合には、意味的制約である「施設」のカテゴリに応じた格要素が補完される。しかし、この手法では、補完要素の推定に利用する意味情報が格に対する制約条件であるため、その多くは抽象的な意味情報を付与しているため、補完すべき要素が一意に決まらない場合が多い。用言意味属性に応じて省略格要素の補完要素を推定する手法を用いれば、より正確に補完要素を推定できることが予想される。

5.3 その他の文脈解析

用言と用言の意味的な関係を用言意味属性の対によって認定することにより、ある文が示す事象と他の文が示す事象の時間的関係の解析、文章要約等への応用が考えられる。

6. おわりに

本稿では、日本語と英語の構文的な対応関係を利用した用言意味属性体系を提案した。今後は、本体系のより一層の詳細化、厳密化を行なうと共に、本体系の応用について検討していく予定である。

<参考文献>

- [1] 村木, 述語素による動詞の分類, 情報処理学会, 自然言語処理研究会, NL48-5, 1985
- [2] 富浦, 吉田, 動詞の多義性とその記述について, 情報処理学会, 自然言語処理研究会, NL55-2, 1986
- [3] EDR, 概念辞書(第2版)TR-02, EDR, 1989
- [4] 徳永, HARTONO, 田中, 対訳辞書からの中間概念の抽出, 情報処理学会, 第39回情処全大, 7G-5, 1989
- [5] 高松, 西田, 動詞パターンと格構造に基づく英日機械翻訳, 電子通信学会論文誌, Vol. J69-D, pp.815-822, 1981
- [6] 池原, 宮崎, 横尾, 日英機械翻訳のための意味解析辞書, 電子情報通信学会, 言語理解とコミュニケーション研究会, NLC91-19, 1991
- [7] H.Nakaiwa, S.Ikehara, Zero Pronoun Resolution in a Japanese to English Machine Translation System by using Verbal Semantic Attributes, ACL, Proc. of ANLP92, pp. 201-208, 1992