

話者の対象認識過程に着目した英語名詞句解析法

7H-4

川辺 諭 宮崎正弘

新潟大学大学院自然科学研究科

1 はじめに

自然言語処理において英語文の本質的な意味理解を実現するために、時枝誠記の言語過程説 [1] を発展的に継承した、三浦つとむの文法 [2] を適用する。本研究では、三浦の文法で提案されている“対象 ⇄ 認識 ⇄ 表現”という人間の過程的な言語運用に着目し、英語名詞句から言語表現の持つ統語的構造である“表現構造”と、認識された対象世界の構造である“認識構造”を抽出する方法を提案する。

2 核名詞

本研究で名詞句の核名詞として取り扱われる品詞を表1に示す。これらの名詞はほとんどが単独で名詞句となり得る。ただし、可算名詞の単数形は通常限定詞を伴い、また、可算名詞の複数形は語尾が屈折変化する。

表1：核名詞となる語

品詞	語例		
名詞	可算名詞	普通名詞	apple
		集合名詞	class
	不加算名詞	固有名詞	John
		物質名詞	coffee
		抽象名詞	heart
所有代名詞	mine, yours, ours		
不定代名詞	one, none, some		
再帰代名詞	myself, himself, themselves		

English Noun Phrase Parsing Algorithm Based On The Speaker's World Recognition Process
Satoru Kawabe, Masahiro Miyazaki
Niigata University

3 名詞の静的属性

静的属性とは、実体に関する時間変化の概念を持たない属性であり、形容詞、名詞などによって表現される。核名詞の静的属性を表現する品詞とその意味情報を表2に示す。

表2：名詞の静的属性

品詞	意味情報
形容詞	実体の静的属性
冠詞	個別性の捉え方（特殊、普遍）
代名詞	実体間の人称、空間関係
数詞	数、順序
名詞	実体、実体として捉えた属性
前置詞	空間関係

4 名詞の動的属性

動的属性とは、実体に関する時間変化の概念を扱った属性であり、名詞句においては分詞（現在分詞、過去分詞）により表現される。分詞の伝える核名詞の動的概念を表3に示す。

表3：分詞が伝える動的概念

概念	統語要素	意味情報
相	現在分詞	動的概念の性質、様態
態	過去分詞	動作主と受動者の関係

また動詞の取る格要素を表4に示す。

表4：動詞の格要素

格	統語要素	意味情報
主格	主語	動作主体
対格	直接目的語	目的物
与格	間接目的語	受益者
属格	補語	主格、対格の属性

5 解析例

一般化 LR 文法 (富田法) を Prolog で記述した SGLR パーザ [3] を用いて、英語名詞句解析部のプロトタイプを作成した。パーザによる解析結果例を図 1 に示す。

```
input sentence :
>>> a present given him by the pretty girl .
```

【表現構造】

```
| -名詞句
  | -名詞句
    | -冠詞 -- a
    | -名詞 -- present
  | -過去分詞句
    | -過去分詞 -- given
  | -直接目的語
    | -名詞句
      | -名詞
      | -人称代名詞 -- him
  | -主格前置詞句
    | -主格前置詞 -- by
  | -名詞句
    | -冠詞 -- the
  | -名詞句
    | -形容詞 -- pretty
    | -名詞 -- girl
```

【認識構造】

```
| -複合名詞句構造
  | -実体
    | -物体 -- present
    | -数 -- 単
    | -class -- a
  | -動的属性
    | -動詞
      | -コアフレーム -- give
      | -態 -- 受動
    | -視点 -- 対格
  | -主格
    | -人 -- girl
    | -数 -- 単
    | -instance -- the
    | -静的属性
      | -pretty
      | -属性 -- 美醜
      | -属性値 -- 美
  | -対格
    | -物体 -- present
    | -数 -- 単
    | -class -- a
  | -与格
    | -関係フレーム -- him
    | -人称 -- 3
    | -性 -- 男
    | -数 -- 単
```

図 1 : パーザによる名詞句解析の例

表現構造では、入力された英語文名詞句が限定詞（冠詞）a と可算名詞present からなる核名詞（先行詞）と、過去分詞given、間接目的語him、主格前置詞句by the pretty girl からなる過去分詞句により構成されていることが解析されている。

認識構造では、入力された名詞句から、話者の視点がおかれている実体と、その動的属性の構造を抽出している。ここではまずpresent が単数の実体であり、話者にclass として一般的に把握されていることが解析されている。また動的属性として、動詞give が主格the pretty girl と対格a present、与格himを取り、先行詞となっている対格に話者の視点があるという構造が抽出されている。さらに実体girl に関して、話者にinstance として特殊的に把握されており、静的属性：美醜として属性値：美を持つといった意味的な内容も、解析されている。

6 おわりに

本研究では計算機上における本質的な自然言語意味理解の実現を目指し、三浦つとむの文法理論に基づいて、英語文名詞句からその統語的構造である“表現構造”と、認識された対象世界の構造である“認識構造”を抽出する方法を提案した。今後、単語辞書、関係代名詞など文法の拡充を行い、本手法の有効性を検討する予定である。

7 謝辞

本研究を進めるにあたり、解析部の実装を行なった後藤友美（現三井海上システム開発）、田中陽子（現コロナ）の両嬢に深謝する。

参考文献

- [1] 時枝誠記：日本文法口語篇，岩波書店(1950).
- [2] 三浦つとむ：日本語とはどういう言語か，講談社学術文庫(1976).
- [3] 沼崎、田中：SGLR:逐次型一般化 LR パーザの Prolog による実現，情報処理学会論文誌, Vol32, No3, PP396-403(1991).