

範疇単一化文法の日本語への適用

5 B-9

関洋平 飯島正 原田賢一

慶應義塾大学理工学部

e-mail: yohei@hara.cs.keio.ac.jp

1 はじめに

本研究の目的は構文・意味レベルにおいて日本語を理解する計算機上のシステムを構築することにある。このような方向で近年取り組まれている文法の枠組の一つとして、単一化文法と呼ばれる一連の文法形式がある。単一化文法は、属性と属性値の組合せをリスト化したデータ構造である素性構造を用いることにより、単語の係り受け関係ならびに意味情報を単語に対応させて表現を行っている。すなわち、文脈自由文法における書換え規則のようなルールを、できるだけ辞書に埋め込む形で取扱うことを目標としている文法枠組である。

本研究では、係り受け関係の取扱いに範疇文法の考え方を取り入れた単一化文法の枠組の一つであり、日本語に対する適用例のあまりない範疇単一化文法 [1] を日本語に対して適用することで、その有効性を検証する。

2 範疇文法

範疇文法では Joachim Lambek の提案による Lambek Calculus の

$$\begin{aligned} / \text{除去規則;} & \frac{\Delta \Rightarrow B/A \quad \Gamma \Rightarrow A}{\Delta, \Gamma \Rightarrow B} \\ \backslash \text{除去規則;} & \frac{\Gamma \Rightarrow A \quad \Delta \Rightarrow A \backslash B}{\Gamma, \Delta \Rightarrow B} \end{aligned}$$

を基礎としている。具体的には、2~3の基本範疇を /, \ を用いて組み合わせることにより、品詞の表現を行い、上記の除去規則を適用することにより、構文解析が行われる。

3 範疇単一化文法

範疇文法と単一化文法を融合させる最初の試みとして 範疇単一化文法 (CUG) は提案された。融合の基本的な考え方は、範疇文法の /, \ を素性構造に埋め込むことで実現しており、単一化文法を基礎にして範疇文法を導入した考え方であるといえる。このような考え方に基づく

An Approach for Applying Categorical Unification Grammar to Japanese.

Yohei SEKI, Tadashi IJIMA and Ken'ichi HARADA  
Faculty of Science and Technology, Keio University  
3-14-1 Hiyoshi, Kouhoku-ku, Yokohama 223, JAPAN

融合の例としては、他に、単一化範疇文法 (UCG)、コンビネータ範疇文法 (CCG) がある。これとは逆に、範疇文法の考えを基礎として、素性構造を基本範疇に対応させた例としては、Type Logical Grammar や Typed Applicative Grammar [3] がある。

4 日本語の範疇

日本語の基本範疇としては、名詞 (N)、文 (S)、後置詞句 (PP) の三つを設けることにする。その三つを /, \ を用いて組み合わせることによって、以下のような品詞を表現することができる。

品詞分類	範疇
文	S
名詞	N
動詞 (格三つ, あげる)	PP \ (PP \ (PP \ S))
動詞 (格二つ, 会う)	PP \ (PP \ S)
動詞 (格一つ, 歩く)	PP \ S
必須格助詞 (が, を, に)	N \ PP
係助詞 (は)	N \ (S / (PP \ S))
任意格助詞 (時間格に, 場所格で)	N \ ((PP \ (PP \ (PP \ S))) / (PP \ (PP \ (PP \ S))))
終助詞 (ね, か)	(PP \ S) \ (PP \ S)
時制助動詞 (た, ている)	(PP \ S) \ (PP \ S)
断定助動詞 (です, だ, である)	N \ (PP \ S)
連体詞 (この, その, あの)	N / N
接尾辞 (時, 君)	N \ N

5 日本語の諸現象に対する取扱い

日本語特有の文法現象に対する取扱いは、範疇の設定とルールの導入により行う。ルールを導入することで、辞書の繁雑さを避けることができる。

● 範疇による取扱いの具体例

1. 助動詞「させる」による使役化

助動詞「させる」には、動詞に接続して文の述部の

意味を「その行為の強制」に変化させる役割がある。例えば、「花子が料理を食べる」という文の動詞「食べる」に接続して「太郎が花子に「花子が料理を食べる」という行為を強制する」というように意味を変化させる。これを範疇によって表すと、以下のようになる。

- 「させる」の範疇

$$(PP_{sem_x} \setminus (PP_{nom, が^s, sem_y} \setminus S_{sem_z})) \\ \setminus (PP_{sem_x} \setminus (PP_{acc, に, sem_y} \setminus (PP_{nom, が^s, sem_w} \\ \setminus S_{force_{agt:w, pa:e:y, obj:z} }))))$$

## 2. 助動詞「られる」による受動化

助動詞「られる」は、「太郎が花子に料理を食べさせる」というような文章に接続して、行為主体者と行為受動者の意味役割を入れ替えるという機能をもつ。範疇によって表すと、以下のようになる。

- 「られる」の範疇

$$(PP_{acc, sem_x} \setminus (PP_{nom, が^s, sem_y} \setminus X)) \\ \setminus (PP_{acc, sem_y} \setminus (PP_{nom, が^s, sem_x} \setminus X))$$

## 3. 助詞「は」による主題化

以下の二つの例文による「は」の役割の違いを扱うために、以下の範疇を「は」に対して与える。

- 例文

太郎はあの子を愛している。  
太郎はあの子が愛している。

- 「は」の範疇

$$N \setminus (S / (PP \setminus S))$$

※ N と PP は 情報を共有

これにより上の例文の「は」が動詞句「あの子を愛している」の欠落格である主格であり、下の例文の「は」は動詞句「あの子が愛している」の欠落格である目的格であることが扱えるようになる。

- ルールによる取扱いの具体例

### 1. 動詞、形容詞による連体修飾

日本語の動詞、形容詞は連体形と終止形が同じ活用をもつという特徴がある。これにより格後置詞句が欠落した文が、後方に名詞を補って連体修飾を行なう、ということが可能になっている。この現象を取扱うために、以下のルールを導入する。

$$(a) PP \setminus S \ N \rightarrow N / N \ N$$

※ PP \ S の PP と N / N の N は 情報を共有

これにより「私は望遠鏡を持つその少年をみる。」のような文章が与えられた際、望遠鏡を持っているのは「私」ではなくて「その少年」であることが取扱えるようになる。

## 2. 語順の自由度

語順の自由度を取扱う方法としては、引数の集合を取ることが出来るように枠組を拡張するという方法がある([2]など)。しかし、これは範疇文法の根本的な利点を損なう感觸も受ける。ここではルールにより取扱う方法について検証を行う。以下の7つのルールを導入することにより、格後置詞句の入れ替えが可能になった。

- 必須格後置詞句(が, に, を格)の入替えルール

$$(a) PP_{が} \ PP \setminus (PP_{が} \setminus X) \rightarrow PP_{が} \ PP_{が} \setminus (PP_{を} \setminus X) \\ (b) PP_{が} \ PP_{に} \setminus (PP_{が} \setminus X) \rightarrow PP_{が} \ PP_{が} \setminus (PP_{に} \setminus X) \\ (c) PP_{に} \ PP_{を} \setminus (PP_{に} \setminus X) \rightarrow PP_{に} \ PP_{に} \setminus (PP_{を} \setminus X) \\ (d) PP_{が} \ PP_{を} \setminus (PP_{に} \setminus (PP_{が} \setminus X)) \\ \rightarrow PP_{が} \ PP_{が} \setminus (PP_{を} \setminus (PP_{に} \setminus X))$$

- 任意格後置詞句(時間格に, 場所格で)の入替えルール

$$(a) (PP \setminus (PP \setminus (PP \setminus S))) / (PP \setminus (PP \setminus (PP \setminus S))) \ PP \setminus (PP \setminus S) \\ \rightarrow (PP \setminus (PP \setminus S)) / (PP \setminus (PP \setminus S)) \ PP \setminus (PP \setminus S) \\ (b) (PP \setminus (PP \setminus (PP \setminus S))) / (PP \setminus (PP \setminus (PP \setminus S))) \ PP \setminus S \\ \rightarrow (PP \setminus S) / (PP \setminus S) \ PP \setminus S \\ (c) (PP \setminus (PP \setminus (PP \setminus S))) / (PP \setminus (PP \setminus (PP \setminus S))) \ S \\ \rightarrow S / S \ S$$

## 6 おわりに

範疇単一化文法の試作システムの実装は Sicstus-Prolog を用いて行った。現在、日本語に対して取扱えている範囲はまだまだ狭いと考えられる。よって、これからも日本語の取り扱える範囲を拡張することを第一に行わなければならない。

## 参考文献

- [1] Hans Uszkoreit: "Categorial Unification Grammars", Proc. of COLING 1986(1986), pp.187-194.
- [2] Beryl Hoffman: "A CCG APPROACH TO FREE WORD ORDER LANGUAGES", Proc. of 31th Annual Meeting of ACL(1993), pp.300-302.
- [3] 飯島正, 関洋平, 柳原正秀他: "範疇文法のための関数的な計算機構", 情報処理学会第111回自然言語処理研究会(1996).