

範疇文法における格選択制限の一手法

3C-2

関洋平 飯島正 木下知貴 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、範疇文法を用いて日本語の構文解析を行う際に、基本範疇に対して素性構造などを利用して意味情報を付加できる枠組において、日本語の述部と名詞の格関係を特定するための選択制限の手法を提案することにある。範疇文法に素性構造の概念を組み合わせた提案の一つとして、Hans Uszkoreitによって提案された**範疇単一化文法** (Categorial Unification Grammar)[4]と呼ばれる枠組がある。本研究ではこの枠組のもとに、格助詞の共演(必須)成分と付加(任意)成分の区別[3]や、提題助詞「は」による主題化、格の交代、二重現象などに対応することを目標としている。

2 範疇文法と範疇単一化文法

範疇文法では、Joachim Lambekの提案によるLambek Calculusの適用規則を基礎としている。適用規則をシーケント・カリキュラスの形式で書く[1]と、次のようになる。

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

構文構造などの文法情報は、2~3個の基本範疇(名詞、文など品詞の基本単位が設定される)を /, \ を用いて組み合わせた**導出範疇**に、語彙のテンプレート(品詞)を対応させて表現する。形態素分割された単語に対応した語彙範疇(基本範疇または**導出範疇**)に対して、上記の除去規則を適用することでボトムアップな構文解析が行われる。範疇文法自体は、実は文脈自由文法の別表記にすぎず、したがって、文の生成能力も文脈自由文法と同等である。

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。具体的な融合の方法は、トップレベルの素性として value(値), di-

rection(方向), argument(引数)を与えることによって、/, \ を素性構造に埋め込むことで実現している。これは、範疇文法を単一化文法として定式化したものであるといえる([4]では PATR 形式の dag 表現を用いて表されている)。基本範疇に対応させて素性構造を用いることによって、格役割の特定や、意味素性情報の伝搬などが取り扱えるようになる。具体例として、名詞「太郎」と、助詞に動詞が接続した動詞句「を愛する」、および名詞に助詞が接続した「3時に」の素性構造を示す。「を愛する」は前方の名詞を引数に取り、「3時に」は後方の動詞を引数に取った結果、文を形成する。

太郎 (n)

範疇	: n(名詞)
分類	: 生物
意味	: 太郎

を愛する (n\s)

方向	: '/'
引数	: [範疇 : n 分類 : 普通 意味 : Y]
値	: [範疇 : s(文) 意味 : 愛する 共演成分 : [(が 生物) : (動作主 X) (を 普通) : (対象 Y)] 付加成分 : [Z1 : Z2]]

3時に (s/s)

方向	: '/'
引数	: [範疇 : s 意味 : X 共演成分 : [Y1 : Y2] 付加成分 : [Z1 : Z2 (に 時制) : (時間 3時)]]
値	: [範疇 : s 意味 : X 共演成分 : [Y1 : Y2] 付加成分 : [Z1 : Z2 (に 時制) : (時間 3時)]]

3 基本的な方針

本研究における提案の概要を説明する。まず、基本範疇としては、s(文) および n(名詞) の二つを設定する。続いて、助詞「が」、「を」、「に」、「で」などいわゆ

A Technique for Selectional Restriction of Case in Applying Categorial Grammar to Japanese.

Yohei SEKI, Tadashi IJIMA, Kazuki KINOSHITA and Ken'ichi HARADA

Faculty of Science and Technology, Keio University
3-14-1 Hiyoshi, Kouhoku-ku, Yokohama 223, JAPAN

る動詞に依存して格役割が特定される助詞に対しては、 $(n\backslash s)/s$ という範疇を与える。すなわち、動詞を引数として取った後、その格フレームにおける助詞に対応する意味情報を用いて、名詞に対して選択制限をかけることで、動詞が求めている格役割かどうかを判定する。

ただし、助詞「に」などは時間格の役割など、動詞というよりは前方の名詞に依存した格役割を与える場合もある。このような場合、「に」の前方の名詞の意味情報は、動詞に依存した「に」の格役割を満たす名詞の意味情報は保持していないと思われる¹。すなわち、範疇 $(n\backslash s)/s$ による構文解析が、意味情報による選択制限によって失敗した場合、助詞の範疇を $n\backslash(s/s)$ に変化させる推論規則を導入しておき、名詞の意味役割を用いて格役割を特定する。この規則は、総記の役割の「が」などについても適用される。また、提題助詞「は」については、はじめから $n\backslash(s/s)$ という範疇を与えておき、動詞が潜在的に保持している共演成分を補完できるかどうか、高階関数に基づいた手法 [2] を用いて判定する。

なお、ここで用いた $n\backslash(s/s)$ という範疇を助詞に与えるという考え方は、[1] で提案されており、「が」「を」「に」などの表層格と、動作主格、対象格などの深層格を高階関数として結びつけるという考え方は [2] による。

4 格助詞のさまざまな役割に対する取扱い

4.1 「が」について

- 私が七曲署の藤堂だ (主格)

「が」は、前方の名詞および後方の名詞 + 「だ」の関係より「主格」となる。「が」の役割は、前方の名詞に依存して特定されていると思われるので、 $n\backslash(s/s)$ という範疇を与える。

- 戸が開く (対象格)

「が」は「開く」という動詞に接続し、前方の名詞は「開く対象」となる。したがって、「が」の範疇は、動詞に接続した後、名詞に選択制限をかける $(n\backslash s)/s$ を与える。

4.2 「を」について

- 手紙を書く (対象格)

「を」が「書く」という動詞に接続すると、その前方の名詞は「書く対象」となる。

- 山に登る (場所格)

「を」が移動性動詞に接続すると、前方の名詞は経

過する場所を意味する。これも後方の動詞に依存している。

4.3 「に」について

- 会社に行く (場所格)

「に」が移動性動詞に接続すると、前方の名詞は場所を意味する。この格役割は、動詞に依存しており、「～に会う」などもこれに準ずる。

- 買物に行く (目的格)

動作名詞に「に」が接続すると、目的を意味する。名詞に依存しているため、範疇 $n\backslash(s/s)$ を与える。

4.4 「で」について

- 塩は海水でつくる (材料格)

「つくる」という動詞および、「塩」という主格名詞によって「海水」に材料という役割が与えられている。したがって、「つくる」の格フレームに、直接補語および主語の意味素性の組合せによって格役割を特定できるような構造を埋め込んでおき、「で」自身は動詞に最初に付く範疇 $(n\backslash s)/s$ によって構文解析が行われる。原因格などもこれに準ずる。

- 学校で調べる (場所格)

場所格の「で」は、名詞が「場所」という意味役割を取れるかどうかによって決定され、動詞には依存しない。しかし、「場所」を取れる名詞というのは一意には分類しにくい²。具格などもこれに準ずる。

5 おわりに

本研究で提案した選択制限の手法は範疇文法の考え方に基づいており、素朴ではあるが、名詞と動詞の間の格役割を特定する際に、どちらが中心となる場合にも対処できるという特徴がある。

参考文献

- [1] 飯島正, 関洋平, 柳原正秀, 木下知貴, 原田賢一: “範疇文法のための関数的な計算機構”, 情報処理学会研究報告 96-NL-111-6(1996), pp.33-40.
- [2] 飯島正, 関洋平, 原田賢一: “高階関数に基づく範疇文法系の日本語への適用 - 表層と深層の語彙主導的な対応づけ -”, 情報処理学会研究報告 96-NL-113-10(1996), pp.59-66.
- [3] 新田義雄: 日本語の格をめぐって, くろしお出版 (1993).
- [4] Hans Uszkoreit: “Categorial Unification Grammars”, Proc. of COLING 1986(1986), pp.187-194.

¹ 「時計を 3 時に合わせる」という例外も考えられる

² 「警察で調べる」という場合、厳密には警察署を指していると思われる