範疇文法へ単一化文法形式を導入する新たな試み

4H－6

飯島 正, 関 洋平, 柳原 正秀, 木下 知貴, 原田 賢一
(iijima@ae.keio.ac.jp)
慶應義塾大学 理工学部

1 はじめに

範疇文法(Categorial Grammar)[1]と単一化文法(Unification Grammar)[2]は，いずれも計算的な文法記述として極めて興味深い枠組みである．計算的側面から眺めた場合，互いに相補的な役割を果たしていながら，共通して項書き換え(term rewriting)に基づく計算モデルを持っており並列化の可能性も高いといった特徴を有している．そこでそうした関係を踏まえ，両文法記述を融合する試みが従来よりなされている．本報告では，そのためのアプローチとして，CUG[3]で与えられているような複合範疇に対応する素性構造を(高階)関数の定義とみなして，それを直接計算する素朴なモデルを提案する．

範疇文法は，構文的な構造規則を範疇と呼ばれる単語単位の要素（ほぼ，いわゆる品詞に対応）に分解し，辞書中の各語彙記述に埋め込むものである．このとき，範疇として，原子的な基本範疇だけでなく，(型付き)関数定義に似た表記をつかって基本範疇から組み合わせてつくる複合範疇の表記法を与えることにより，構造的な表現を部分的に与えることを可能にしている．それにより，範疇文法では，ボトムアップに範疇の列の項書換えを行っていくことで文の構文的な解析ができる．

一方，単一化文法形式は，構文規則を論理式として表現する論理文法の枠組みに，単一化操作を定義したレコード表現（「属性名ラベル＋属性値」対のリスト）である素性構造を取り入れたものである．その素性構造をつかって各単語のもつ性質を，同じ構文規則に現れる他の語の性質と制約として結び付けることにより，構文解析と同時に制約に基づく意味解析を行うことができる．ここでも，構文解析は，構文規則の項書換えとして行うことができる．

この両者を融合する研究アプローチの一つにCUG (Categorial Unification Grammar) [3] がある．CUGにおける両者の融合は，単一化文法の枠組みの中に範疇文法をコーディングして埋め込むことで行われている．例えば，関数子を用いて基本範疇から作られる複合範疇は，その関数に対応する素性構造に変換されて，単一化文法の枠組に取り入れられる．しかし，この対応づけだけでは，単一化操作によって，意味制約を伴った構文構造を作り上げることはできても，範疇文法の特徴的な関数的な項書換えの動作を規定してはいない．

本報告では，範疇文法と単一化文法を融合するという共通の目的のもとで，むしろ逆に範疇文法の枠組の中に素性構造とその単一化操作を埋め込むことを試みる．そこで原子的な基本範疇に素性構造を割り当てておき，それに加えて素性構造を対象とした単一化に基づく関数結合演算を与えることで，複合範疇のための素性構造を組み立てていく方法を与える．同演算には複合範疇に固有の素性を追加的に与えることができる．これによって，従来よりもより範疇文法の考え方に適した形で，単一化文法形式（意味情報の制約記述）を取り入れることができる．

以下では，２章で範疇文法を概説し，３章で単一化文法の紹介とCUGにおける両文法記述の融合方法について解説した後，４章で本報告のアプローチを簡潔に紹介する．

2 範疇文法

範疇文法では，文法規則は範疇と呼ばれる要素に分解されて，文法規則では終端記号に対応する各単語の辞書中の語彙項目に埋め込まれる．範疇は，基本範疇と複合範疇に分けられる．ここでは，簡単な英語を想定して，基本範疇として，S(文)，NP(名詞句)，N(名詞)の三つを与える．すると，複合範疇は，/と\および括弧を使って基本範疇を組み合わせたものである．例えば，

NP/N	冠詞
$NP\backslash S$	自動詞
$(NP\backslash S)/NP$	他動詞

となる．ここで，「値/引数」は「引数」を右側に与えると「値」を返す関数，「引数\値」は「引数」左側に与えると「値」を返す関数を意味する．実は，この表記はLambeckによるものであり，\の両辺を入れ替えた順序で表記する方法(Steedmanによるもの)も広く使われているので注意されたい(本報告では，筆者の好みによりLambeckによる表記を採用している)．

上記のような範疇は，各語に対し，例えば，以下のように辞書中の語彙項目に範疇が与えられる．

"Another Approach to Incorporate a Unification Grammar Formarism into a Categorial Grammar,"
o Tadashi Iijima, Yohei Seki, Masahide Yanagihara, Kazuki Kinosita, and Ken'ichi Harada,
Faculty of Science and Technology, Keio University
3-14-1 Hiyoshi, Kohoku-ku YOKOHAMA 221, Japan

(単語)	→(範疇)
Hanako	→NP
like	→$(NP\backslash S)/NP$
Taro	→NP

すると，文

$$\dfrac{\underset{NP}{\text{Hanako}} \quad \dfrac{\underset{(NP\backslash S)/NP}{\text{likes}} \quad \underset{NP}{\text{Taro.}}}{NP\backslash S}}{S}$$

というようにボトムアップに書き換えていくことで文を構文的に解析できる．

3 単一化文法とCUG

ここでは，単一化文法についての解説は省き，素性構造の表記だけを紹介する．素性構造は，「属性ラベル＋属性値」対の集合であり．以下のように記述する．

$$\begin{bmatrix} \text{ラベル}_1 & \text{値}_1 \\ \text{ラベル}_2 & \text{値}_2 \\ \vdots & \vdots \\ \text{ラベル}_n & \text{値}_n \end{bmatrix}$$

単一化文法では，文法規則に現れる非終端記号や終端記号(および，それに対応する各単語の辞書中の語彙項目)に，こうした素性構造を用いることで，単語の字面表現，意味項目，音韻表現などに加えて，文法的な制約(性・数の一致など)の検査に必要な属性を取り扱うことができる．

1.「はじめに」で述べたようにCUGでは，複合範疇を素性構造へマッピングすることで範疇文法を単一化文法に取り入れている．例えば，他動詞に対応する範疇は，

$$(NP\backslash S)/NP$$

というように与えられるが，それは以下のような素性構造となる(当初の論文[3]ではPATR形式のdag(directed acyclic graph)表現で与えられている)．

$$\begin{bmatrix} \text{結果:} & \begin{bmatrix} \text{結果:} & S \\ \text{方向:} & \backslash \\ \text{引数:} & NP \end{bmatrix} \\ \text{方向:} & / \\ \text{引数:} & NP \end{bmatrix}$$

もちろん，本来は他にも多くの素性が含まれており，性・数の一致といった制約を伝播するためにも使われる．

4 素性構造を扱う関数による融合

前章で示した素性構造自身は，ラベル名称に「引数」とか「結果」とあるにもかかわらず,「計算」とは独立した中立的な表現といえる．しかし，この素性構造を関数の定義とみなして直接実行するような計算セマンティックスを与えることは可能である．実引数を「引数」素性値と単一化させて，そこで入力されたデータの一部が加工されて更に「結果」素性値へ伝播された後，その「関数定義」を「結果」素性値に置き換えてしまえばよい．

本報告は，このモデルの有用性を示すことを目的としているが，未だ十分な形式化にも実装による確認にも至っていない．紙数の制限から事例は割愛するが，従来の範疇文法の計算の自然な拡張になっていることは明らかである．この「計算」の結果は，文をあらわす範疇Sの素性構造であり，その素性値に，必要な情報や構造は伝播されて格納されることになる．

ここで重要な点は「結果」として素性値をとることができるので，関数定義を「結果」に返すことができる，すなわち，必然的に高階関数を与えることができる点である．

5 おわりに

本報告のアプローチ自体は極めて素朴であり従来なされてきた研究の自然な延長上にあるが(範疇文法自体の有用性が必ずしも明確ではないというのは別にしても)，実際にはまだ多くの課題が残されている．たとえば，関数定義のために素性構造を使うためには，素性構造にどのようなコンストラクトを追加すべきであるかも明確ではない．少なくとも，帰納的な定義のためには，選択関数(disjunction)が必要である．また，各種の文法的制約を評価するのに計算順序は，ボトムアップ（最内優先）だけで十分なのだろうか? 少なくとも，範疇文法において興味深い演算の一つであるタイプレージング[1] 等を扱うためには代数的等価性にしたがった変換を取り入れる必要がある．こうした問題は，一般的な計算能力の解析とは別に，実際に文法記述を試みながら，ニーズ指向で解消していくことが必要であるとも考えられる．

参考文献

[1] Mary McGee Wood: *"Categorial Grammars,"* Routledge (1993).

[2] Peter Sells: *"An Introduction to Unification-Based Approaches to Grammar,"* CSLI Lecture Notes No.3(19).

[3] Hans Uszkoreit: *"Categorial Unification Grammars,"* Proc. of COLING'88,pp.187-194.