

語彙機能文法の実現方式

4F-5

加藤正明 +
甲斐郷子 ++大里真理子 ++
椎野努 ++

+(株)沖テクノシステムズラボラトリ ++ 沖電気工業(株)

1.はじめに

近年、構文解析処理と意味解析処理をつなぐ文法理論として Bresnan らによる語彙機能文法(LFG)[1]が注目されその応用例も報告されている[2][3][4]。しかし、LFG を応用した実用的なシステムを構築しようとする際には、辞書の記述に冗長性が生じ辞書の構築が困難になること、f構造と呼ばれる表示形式を生成する処理に多くの計算を必要とすること、という問題がある。

本稿ではこの2つの問題を解決する方式を提案する。

2. LFGについて

LFG は、文脈自由文法の句構造規則によって構文木(c構造)を作るという解析に加えて、c構造から機能構造(f構造)と呼ばれる表示形式を生成するという解析を行う文法理論である。

f構造は機能(文法機能)という属性とその値の組の集合として表される。機能とは構文的に決まる要素の総称であり SUBJ、PREDなどを指す。f構造に含まれる各機能は何らかの意味役割を持つ。例えば英語の paint という動詞に対して SUBJ は動作主、OBJ は動作対象という意味役割を持つ。SUBJ、OBJ、OBJ2 の各機能は f構造内の関係によって様々な意味役割を持つことができ、意味役割の割り当てについて制限がないために、意味的な制限のない機能と呼ばれる。これに対して、意味的な制限のある機能、例えば COMP(補文) は、動作主という意味役割を持つことはできない。

f構造の生成には機能注釈と呼ばれる一種の等式を用いる。機能注釈で表現された手続きを実行すると、c構造の各節点に対応した局所f構造や文全体のf構造が生成される。

f構造を使うことにより、非文の判定など多くの重要な文法現象を説明できる。英語の動詞 paint は平叙文では、主語と目的語をもつ文中に出現する。このとき述語 paint は、SUBJ と OBJ の機能を下位範疇化するといい、述語に下位範疇化される機能を項と呼ぶ。述語の下位範疇化に関する情報は語彙形式と呼ばれ、paint の例では `paint<((↑SUBJ)(↑OBJ))>' と表される。ある文を非文でないと判定するためには、その文を表すf構造が次の条件を満たさなければならない。

- ・そのf構造に含まれる述語の項が全て含まれていること(完全性条件)
- ・その述語の項以外の機能が含まれていないこと(一貫性条件)
- ・そのf構造内の各機能が一つ以上の値を持たないこと(唯一性条件)

A Study for an Implementation of
Lexical-Functional Grammar.

Masaaki Kato +, Mariko Osato ++,
Kyoko Kai ++, Tsutomu Shiino ++
+Oki Technosystems Laboratory, Inc.
++Oki Elec. Ind. Co., Ltd.

3.辞書の冗長性に関する改良

3.1 語彙形式における冗長性

述語の項には省略可能なものがある。図1(1)で述語である fly は SUBJ、OBLgoal(目標格)、OBLsource(源泉格)の三つを下位範疇化する。(2)では OBLsource が省略され、(3)では OBLsource と OBLgoal が省略されている。fly にとって、OBLgoal と OBLsource は省略可能な項である。

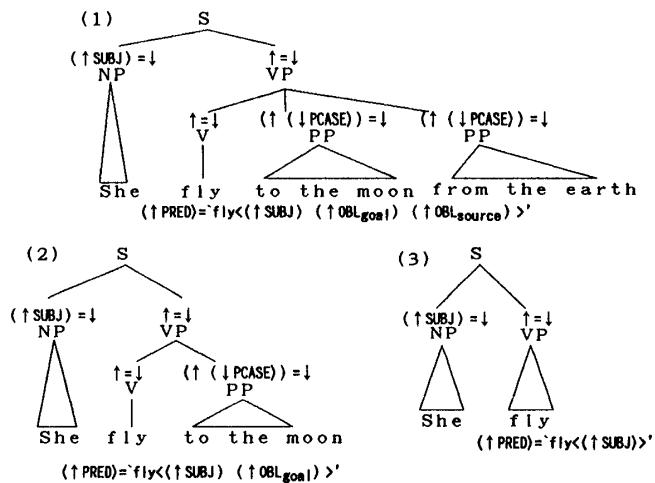


図1 fly の語彙形式の曖昧性

しかし、従来の語彙形式には省略可能な項とそうでない項の区別がない。そのため、図1の三つの文の述語 fly にそれぞれ次の語彙形式が必要になる。

- (1) `fly<((↑SUBJ)(↑OBLgoal)(↑OBLsource))>'
- (2) `fly<((↑SUBJ)(↑OBLgoal))>'
- (3) `fly<((↑SUBJ))>'

さらに、述語 fly に対する省略可能な項としては by the rocket のような OBLinst(道具格)がある。その3つを全て下位範疇化する場合から、一つも下位範疇化しない場合まで合わせて8通りの語彙形式を辞書に記述しなければならない。

この fly のように省略可能な項を下位範疇化する述語には数多くの語彙形式を記述する必要があるので、語彙数の多い実用的な辞書の構築が困難になる。

3.2 語彙形式の記述の改良

上記の問題は、本来は省略可能である項を語彙形式の中では省略不可能な項として表現しているところに原因がある。そこで、下位範疇化に関して辞書に記述する情報を意味的に制限のない項と意味的に制限のある項とに区分し、語彙形式を図2のように記述する。図2の A の部分は意味的に制限のない項を、B の部分は意味的に制限のある項を表している。ここで意味的に制限のある項は省略可能であり、意味的に制限のない項は省略不可能であるとみなす。この仮定に立って、A の部分に現れた項を省略不可能な項、B に現れた項を省略

可能な項として扱う。こうすることで、一つの述語について語彙形式を数多く記述する必要がなくなり、辞書の記述量を減少させることができる。

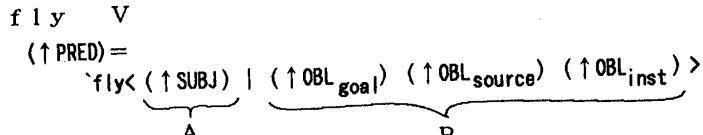


図2 改良した語彙形式

4. f構造生成手続きの改良

4.1 機能注釈の曖昧性によって発生する処理

図3に、 $\text{VP} \rightarrow \text{V NP PP}$ という句構造規則の右辺に含まれる PP に対して、(1) $(\uparrow \text{XCOMP}) = \downarrow$ と (2) $(\uparrow (\downarrow \text{PCASE})) = \downarrow$ の二通りの機能注釈が付与されている例を示す。(1)は、I kept him in the house. を解析するのに必要であり、(2)は、I cut the cake with a knife. を解析するのに必要である。これは、述語である keep と cut の語彙形式の違いから生じる相違である。

- (1) $\text{VP} \rightarrow \text{V NP PP}$
 $\uparrow = \downarrow (\uparrow \text{OBJ}) = \downarrow (\uparrow \text{XCOMP}) = \downarrow$
(2) $\text{VP} \rightarrow \text{V NP PP}$
 $\uparrow = \downarrow (\uparrow \text{OBJ}) = \downarrow (\uparrow (\downarrow \text{PCASE})) = \downarrow$

図3 機能注釈の曖昧性の例

句構造規則の機能注釈に曖昧性がある場合、従来のf構造生成の方式ではその曖昧性の数だけ文全体のf構造が生成される。例えば、図3の句構造規則を用いて I kept him in the house. を解析すると、図4に示すc構造が生成される。図4の(1)と(2)は全く同じ形の構文木で表されているが、(1)のc構造は図3(1)の機能注釈を含み、(2)のc構造は図3(2)の機能注釈を含んでいる。この2つのc構造からそれぞれ文全体のf構造が生成される。

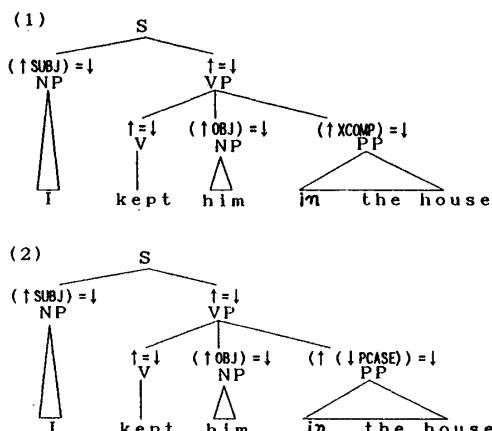


図4 図3の二つの機能注釈に対応するc構造

生成された複数のf構造から、与えられた文の構造を正しく表すf構造を選び出すには、非文の判定条件を用いる。図4(1)から生成されたf構造は三つの判定条件を満たす。しかし、図4(2)から生成されたf構造は完全性条件及び一貫性条件を満たさない。このf構造はVPに対応する局所f構造が不適格である為に、文全体のf構造が不適格と判定される。

4.2 f構造生成手続きの改良

不適格な局所f構造を生成するようなc構造に従って文全体のf構造を生成する必要はない。そこで、f構造生成手続きを次のように改良する。

各c構造の末端からSラベルをもつ最上位節点へ向けて各部分木ごとにf構造を生成してゆく。そして、ある部分木に対応する局所f構造が生成されたとき、その局所f構造に局所適格性条件を適用する。もし、適格な局所f構造であればf構造生成の処理を続行し、不適格な局所f構造であればそのc構造に関するf構造生成の処理を打ち切る。

局所適格性条件とは、f構造生成手続きの中でそれ以上の解析を行っても適格性条件を満たす可能性のない局所f構造を発見するための条件であり、前述の非文の判定条件のうちの一貫性条件と唯一性条件から成る。図5に、図4(2)の中のVPに対応する局所f構造を示す。図5の局所f構造の中の機能OBLloc は述語keep の語彙形式 $\text{'keep} < (\uparrow \text{SUBJ}) (\uparrow \text{OBJ}) | (\uparrow \text{XCOMP}) >$ の中に現れない。従ってこの局所f構造は局所適格性条件の一貫性条件に違反しており不適格である。

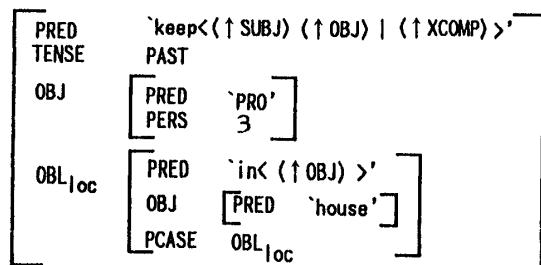


図5 不適格な局所f構造

以上の生成手続きを採用することによって計算量を抑えることができる。

5. おわりに

本稿では、LFGを応用した実用システムの構築時に発生する、辞書の冗長性の問題と、計算量の問題を解決する方式を提案した。

辞書の冗長性については、語彙形式の中で意味的に制限のある項とない項とを分離して記述する形式を採用し、さらに意味的な制限のある項は全て省略可能であるものと仮定することにより解決する。この仮定の適用可能範囲については、言語間の相違や、文脈依存性の有無と関係があり、その一般的な正当性の検証は今後の課題である。

構文規則の機能注釈の曖昧性による計算量の問題については、局所f構造の不適格性を早期に発見しこれを排除することで曖昧性を減らし、計算量を抑える。

参考文献

- [1]Bresnan, J. (ed): The Mental Representation of Grammatical Relations, pp.183-390, MIT Press, (1982).
- [2]向井他: 談話理解システムDUALS の概要, ICOT 研究速報 TM-0118, (1985).
- [3]新田: LFGと意味解釈の融合に向けて, 情処研報 88-NL-68 (1988).
- [4]二口, 寺下: LFGに基づく並列型バージング法, 情処研報 89-NL-72 (1989).