

タイプ付き素性構造による生成過程の制御

1 E - 2

上田良寛 小暮 潔

ATR自動翻訳電話研究所

1. はじめに

対話文翻訳システムの生成部として、双方向文法を用いた生成システムを開発した。生成システムでは、同じ規則が何度も適用されて停止に至らなかったり[2]、最終的に書換えの失敗に至る無駄な規則が適用される場合がある。これを避けるために、**Assertions** という制約条件を文法規則に付加し、宣言的記述で制御を行う方法を提案した[1]。

ここでは、このような宣言的な制御を、タイプ付き素性構造[3]の導入によって行う方法について述べる。さらに、素性構造に選言を含める[4]ことにより、派生木全体を複製することを避ける方法について述べる。

また、発話意図を生成するためにHPSG[5]に基づいて作成された文法を、生成例とともに示す。

2. タイプ付き素性構造の導入

2.1 文法規則の選択

[1]で提示した問題は、次の2つの文法規則で左辺のVP1として得られる素性構造(図1)の意味構造の違いが通常の素性構造では記述できず、結果的にこの2つの規則を選択することができないということであった。

- ```

VP1 -> VP2 XP (1)
 <VP1 SEM CONT> == <VP2 SEM CONT>
 <VP2 SUBCAT FIRST> == <XP>
 <VP1 SUBCAT> == <VP2 SUBCAT REST>

VP1 -> VP2 PP (2)
 <VP1 SEM CONT> == <PP SEM CONT>
 <PP SUBCAT FIRST> == <VP2>
 <PP SUBCAT REST> == NONE

```

|                                               |                                                           |
|-----------------------------------------------|-----------------------------------------------------------|
| (1) [RELN PLAY<br>ARG1 JOHN'<br>ARG2 SOCCER'] | (2) [RELN IN<br>ARG1 JOHN-PLAY-SOCCER'<br>ARG2 THE-PARK'] |
|-----------------------------------------------|-----------------------------------------------------------|

図1 素性構造の例

これを避けるために導入したのが親の素性構造に対して制約条件を与える**Assertions**という記述である。(1)、(2)のそれぞれの規則に対して、次のような記述を与える。ここで**TYPE**はキーとなる素性(ここでは<SEM CONT RELN>素性)に対して辞書項目中で与えられている。

- ```

(1) (:ASSERTIONS (:EQ (:TYPE <SEM CONT RELN>) V))
(2) (:ASSERTIONS (:EQ (:TYPE <SEM CONT RELN>) P))
    
```

Assertionsを導入した動機は、意味素性中のキーとなる素性のタイプを制約として扱うことにあつ

た。これは通常の素性構造では扱うことができないが、Ait-Kaci [3]による階層的タイプ付き素性構造はこの目的にうまく合致する。

規則(1)は、タイプ付き素性構造を用いて次のように記述することができる。1

- ```

VP =HC*=> (VP XP) (1')
 (<lm sem cont> == <head-dtr sem cont>)
 (<lhead-dtr lsubcat first> == <lcomp-dtr-1>)
 (<lm lsubcat> == <lhead-dtr lsubcat rest>)
 (<lhead-dtr sem cont rein> == [V])

```

ここで下線部は、<SEM CONT RELN>素性がタイプVと単一化されることを示す。この単一化が成功する、すなわち、<SEM CONT RELN>素性の値がタイプ階層の中でタイプVの下にある場合に、この文法規則は適用可能である。

#### 2.2 カテゴリとのインタラクション

タイプの利用方法のひとつに、カテゴリ(ノンターミナルシンボル)との関連づけがある。カテゴリはタイプを用いて表現され、階層化がなされている。例えば、“**must**”の辞書項目中では、下位範疇化フレームに動詞句が入るということが、次のように記述されている。

- ```

(deflex "must" dyadic
  (<lsubcat first> == [vp])
  ...)
    
```

VPをXPのサブタイプと定義しておけば、“**must**”が規則(1)と単一化されるときに、[VP]は[XP]と単一化され[VP]となる。このような未指定カテゴリに対して、例えばD-PATR [5]では、特別なシンボルX, YなどをCFGに導入し、CFG規則中でこのようなシンボルにあたったときに特別な処理を行うようにしている。未指定カテゴリをタイプ階層の中に組み入れることにより、未指定カテゴリに無理なく対処することができる。

2.3 停止性問題への対処

規則(1)は、解析の規則として見ると、子ノードのsubcatリストから要素を1個取り除くという意味である。この規則はsubcatリストが消費されてしまうまで適用される。一方、生成の規則として見ると、子ノードのsubcatリストに1個要素を加えるという意味になり、無限に適用可能である。

この問題は、subcatリストの長さ上限を与えることによって解決可能である。英語の場合では主語を除けばsubcatリストの長さはせいぜい2であ

1. この規則で、=HC*=>リンクは、右辺の最初の要素がhead daughter、その他がcomplementになるということを表している。S=CH=>NP VPのようなcomplementが先にくる構成のために、=CH=>リンクも用意されている。!マークがつけられたシンボルは別に定義されるテンプレートを意味している。

る。しかし、全ての動詞に対して規則(1)を2回適用させるのは無駄である。タイプ階層の中で動詞を3つのサブタイプ(Monadic, Dyadic, Triadic)に分類し、次のような制約を規則(1)に加える。

```
(:or ((<sem cont reIn> == [dyadic])
      (<lhead-dtr lsubcat rest> == [list-end]))
      ((<sem cont reIn> == [triadic])
      (:or
        (<lhead-dtr lsubcat rest> == [list-end])
        (<lhead-dtr lsubcat rest rest> == [list-end]))))
```

3. 選言の導入

HPSGでは、主語、目的語などは、主動詞の下位範疇化フレームで指定されるため、主動詞が決定されるまでは決定されない。一方、主動詞は、主語・動詞の一致により、主語が決定されるまでは完全には決定されない。このため、主動詞を派生する際にその候補の数だけ派生木全体を複製しなければならなかった。主語が決定されればこのうちの一つしか生き残らないので、無駄である。

素性構造に選言を導入する[4]ことにより、動詞の表層形の違いを一つの辞書項目にまとめることができる。“be”動詞の場合の例を示す。

```
(deflex-unit lbe-Unit| dyadic
  (:or (lfinite-form lpresent-tense
        (:or ((<word> == "am") l1sg-subj-agr)
              ((<word> == "are")
                (:or (l2sg-subj-agr) (lpl-subj-agr)))
              ((<word> == "is") l3sg-subj-agr))
        .....)) ;過去形、過去分詞など
```

主語が派生され、単一化されると同時に、この3つの表層形の候補のうちから1つが決定される。

4. 文法と生成例

この生成システム用に開発した英語文法は、BorsleyによるHPSGの修正版[7]を基礎にしている。HPSGでは、ほとんどの言語情報は辞書項目中に置かれる。同様に発語内的力関係(illocutionary force)も辞書項目中に記述されることが多い。例えば、発語内的力タイプPROMISE(行為拘束型)は、“will”の辞書項目中に次のように記述される。

```
(deflex-unit |will-Unit| DYADIC
  (:or ((<sem cont reIn> == [*promise*])
        (<sem cont obje> == <lsubcat-1 sem cont>)
        (<lsubcat-1 sem cont reIn> == [*action*])
          ;主動詞は行為型
        l1st-subj-agr) ;主語は1人称(肯定文の場合)
        ... ;;「未来」の意味の記述
```

発語内的力のうちのいくつかは、文法規則中に記載されている。例えば、発語内的力REQUEST(行為指導型)は、命令文の規則に記述されている。また、“Would you give ...?”のような疑問文でもREQUESTを表現できる。

例として、発語内的力タイプREQUESTを含む図2のような素性構造表現を考える。この素性構造(fs-1)から、次のような生成結果が得られる。

```
> (gen3 fs-1)
("would you send me a registration form" "could you send
me a registration form" "send me a registration form")
```

```
[[SEM [[CONT [CIRC[RELN [*REQUEST*]]
            [AGEN ?X03[IND-OBJ[LABEL *SPEAKER*]]]
            [RECP ?X02[IND-OBJ[LABEL *HEARER*]]]
            [OBJE [CIRC[RELN [*SEND-1*]]
                  [AGEN ?X02]
                  [RECP ?X03]
                  [OBJE IA-REG-FORM' ]]]]]]]]
      [PRAG [[HEARER ?X02]
             [SPEAKER ?X03]]]]]
```

図2 入力素性構造

これらの文はいずれも同じ発語内的力を表現するが、それぞれの表現に込められた丁寧さのレベルが異なっている。丁寧さのレベルを表す素性を導入し、対応する文法規則に丁寧さのレベルとの対応づけを記述する。先の素性構造に丁寧さのレベルを最高に設定したもの(fs-1-3)を入力すると、次のような結果が得られる。

```
> (gen3 fs-1-3)
("could you send me a registration form")
```

このように、双方向文法生成システムは、過生成の傾向があるため、表層発話に含まれる意図の違いを、文法作成者に気がつかせるという働きがある。このため、発話意図の研究の道具としても有用である。

5. おわりに

生成過程をタイプ付き素性構造と選言を含む素性構造を用いて制御する方法について述べた。また、この生成システムが発話意図を扱う文法を研究するのにも適切であることを示した。

現在の文法は発話意図と表層発話の対応づけがまだ十分ではない。今後はこの文法をさらに拡充する必要がある。

本研究の機会を与えて下さいましたATR自動翻訳電話研究所 樽松明 社長、および、有益な助言をくださる言語処理研究室の諸氏に感謝します。

参考文献

- [1] 上田、小暮:「素性構造を入力とする英文生成」、情報処理学会第38回全国大会、2D-4、1989。
- [2] S. M. Shieber et al.: "A Semantic-Head-Driven Generation Algorithm for Unification-Based Formalisms", in 27th ACL, pp. 7-17, 1989.
- [3] Ait-Kaci, H.: "An Algebraic Semantics Approach to the Effective Resolution of Type Equations", Theoretical Computer Science 45, pp. 293-351, North-Holland, 1986.
- [4] Kasper, R. T.: "A Unification Method for Disjunctive Feature Descriptions", in 25th ACL, pp. 235-242, 1987.
- [5] Pollard, C. and I. A. Sag: "An Information-Based Syntax and Semantics, Volume 1, Fundamentals", CSLI Lecture Notes Number 13, CSLI, 1987.
- [6] Karttunen, L.: "D-PATR: A Development Environment for Unification-Based Grammars", Report No. CSLI-86-61, CSLI, 1986.
- [7] Borsley, R. D.: "Subjects and Complements in HPSG", Report No. CSLI-87-107, CSLI, 1987.