

素性伝播による構文生成*

2D-5

徳梅喜啓, 柴田昇吾, 柵木孝一†
キャノン(株) 情報システム研究所‡

1 はじめに

我々は複写機マニュアル等の日英機械翻訳システムを構築中である[1]。現在、英文生成系では、CFGベースの構文生成を作成している。

我々の構文生成は、(1)英語の意味ネットワークを変換して得られる素性構造(feature structure)から句構造を生成するものである。(2)文法規則の拡張部には素性の伝播方法が宣言的に記述されており、これに従って素性伝播を行いながら句構造を生成する。また、(3)文法規則には条件部とメッセージ部があり、これによって文法規則を適用していくときの制御を行っている。

本稿では構文生成における素性伝播と文法規則の制御について述べる。

2 文法規則の記述形式

文法規則は、次の形式を持っている。

[句構造部 意味部 条件部 メッセージ部]

ここで、句構造部には句構造規則が、意味部には親カテゴリから子カテゴリへの素性の伝播方法が記述されている。条件部はその文法の適用条件を記述するためのもので、メッセージ部はその文法の子カテゴリを親カテゴリとする文法規則を適用するときの制約を与えるためのものである。つまり、素性伝播は意味部によって行われ、文法規則の制御は条件部とメッセージ部によって行われる。

文法規則の例を以下に示す。ここで、@sem は意味部を、@con は条件部を、@mes はメッセージ部を示す。↑は親カテゴリの素性構造を表す。= は素性構造が等しいことを、- は素性構造の削除を表す。== および ≠ は値の一致、不一致を表す。++ はメッセージの付加を表す。

1. bun → SDEC END
@sem ↑ = SDEC
2. SDEC → NP BEP PRED
@sem (↑ subcat SUB) = NP
(↑ head) = BEP
(↑ subcat PRE) = PRED
@con (↑ head syn) == Wv1
@mes (NP rel) → NP ++ SREL
3. SDEC → NP VP
@sem (↑ subcat SUB) = NP
↑ - (↑ subcat SUB) = VP
@con (↑ head syn) ≠ Wv1
@mes (NP rel) → NP ++ SREL

4. NP → DDET NOMHD SREL/NP
@sem (↑ head) = NOMHD
(↑ rel) = SREL
@con mes == SREL
5. SREL/NP → RELPRO SDEC
@sem (↑ relpro) = RELPRO
↑ - (↑ relpro) = SDEC
6. VP → V NP
@sem (↑ head) = V
(↑ subcat OBJ) = NP
@con (↑ head syn) == T1

3 素性伝播

本節では、図1の "The boy who I saw is John." に対応する素性構造を用いて説明する。なお、図1において [1] は素性構造のリンクを表す。図1の素性構造の処理に関する文法規則は前節の如くである。

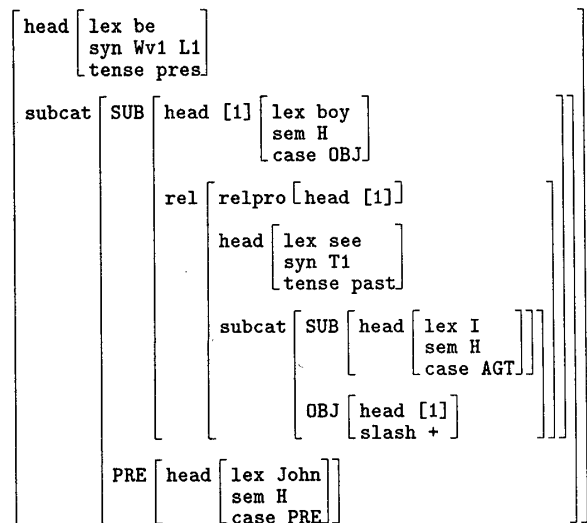


図1: 素性構造

さて、素性伝播による構文生成について説明する。まず、図1の素性構造が bun にセットされ、構文生成が始まる。最初に、文法 1. が適用され、bun の素性構造はそのまま SDEC に伝播される。次に、SDEC の主動詞の文法コードが Wv1 (LDOCE の文法コード) であるので文法 2. の条件部とマッチし、文法 2. が適用される。文法 2. では親である SDEC の subcat の SUB の素性構造が NP に伝播される。文法 2. ではメッセージ部の記述がある。これは文法 2. で生成される NP が関係節(rel)を伴っていれば SREL というメッセージを NP に付加することを示している。今の場合、関係節なので SREL が付加される。これは

* Syntactic Generation Based on Feature Propagation

† Yoshihiro TOKUUME, Shogo SHIBATA, Koichi MASEGI

‡ CANON Inc. Information Systems Research Center

文法 4. の条件部とマッチし、文法 4. が適用され、さらに、文法 5. へとつながる。文法 5. では親である SREL/NP の素性構造から relpro の素性構造を削除したものが SDEC に伝播される。SDEC は文法 3. を経て文法 6. につながる。なお、欠けの情報は素性構造の中に slash 素性として持たせてあり、素性伝播に従って自動的に送られる。そして、NP にこの情報が送られると ϕ を出力するようになっている。

最終的に得られる句構造を図2に示す。

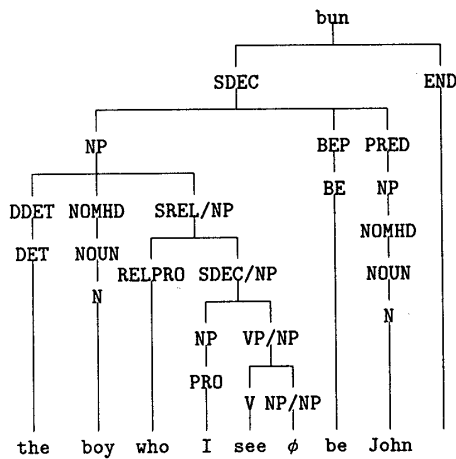


図 2: 句構造

4 文法ルールの制御

本節では文法ルールの制御について任意格の処理を例に述べる。素性構造としては図3の "I stop at Kobe to rest." に対応するものを用いる。

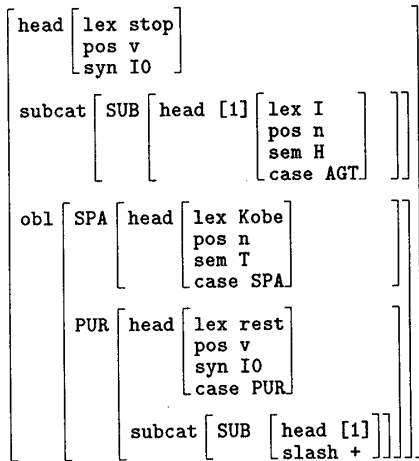


図 3: 素性構造

図3の素性構造の処理に関する文法ルールを以下に示す。

1. SDEC \rightarrow NP VP
 $@sem (\uparrow subcat SUB) = NP$
 $\uparrow - (\uparrow subcat SUB) = VP$
 $@con (\uparrow head syn) \neq Wv1$
 $@mes (NP rel) \rightarrow NP ++ SREL$
 $(VP obl PUR head pos) == v \rightarrow VP ++ INF$
2. VP \rightarrow VP INF
 $@sem \uparrow - (\uparrow obl PUR) = VP$

- ($\uparrow obl PUR$) = INF
 $@con mes == INF$
 $@mes (VP obl) \rightarrow VP ++ PP$
3. VP \rightarrow VP PP
 $@sem \uparrow - (\uparrow obl *x) = VP$
 $(\uparrow obl *x) = PP$
 $@con mes == PP$
 $@mes (VP obl) \rightarrow VP ++ PP$

さて、任意格の処理の説明を行う。本例では stop の任意格要素として目的格と場所格がある。任意格の中では目的格、原因格、条件格などの処理が優先されるので、目的格の処理が最初に行われる。これは、文法 1. のメッセージ部で、目的格要素が動詞のときは VP にメッセージとして INF を付加することによってなされる。つまり、文法 3. の条件部はメッセージ PP を要求しているため適用されず、INF を要求している文法 2. が適用される。文法 2. のメッセージ部では任意格があれば、子の VP にメッセージとして PP を付加する。よって次には文法 3. が適用される。文法 3. の意味部では任意格があれば、そのうちの1つの素性構造を PP の素性構造とすることが記述されている。よって、今度は場所格が前置詞句として生成される。

最終的に得られる句構造を図4に示す。

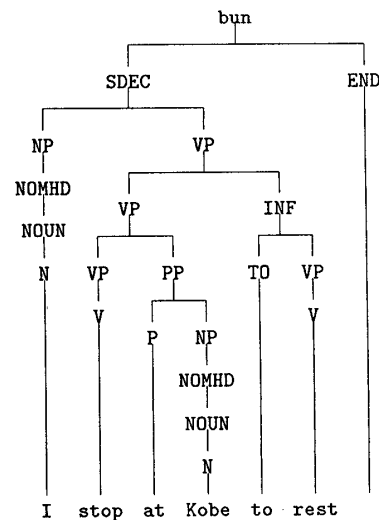


図 4: 句構造

5 おわりに

本稿では、宣言的な記述によって素性伝播と文法ルールの制御を行う方式を示した。今後は、幅広い言語現象に対応できるように本方式の記述能力を高めていきたい。

参考文献

- [1] 根本 他：機械翻訳におけるドメイン知識の活用，情報処理学会第34回全国大会 7W-3 (1987)
- [2] Schieber：An Introduction to Unification-Based Approaches to Grammar, CSLI Lecture Notes No. 4, 1986
- [3] Sells：Lectures on Contemporary Syntactic Theories, CSLI Lecture Notes Series No. 3, 1986
- [4] Bresnan (Ed.)：The Mental Representation of Grammatical Relations, MIT Press, 1982