

制限依存文法とその表現

福本文代 佐野 洋

(財) 新世代コンピュータ技術開発機構

1 はじめに

一般に自然言語を解析する文法には、句構造文法に代表される構成文法と、依存文法とがある。論理型文法の記述枠組の中でこれら両者を取り扱う場合、句構造文法は文脈自由文法を基本にしていることから、容易に扱える。一方、依存文法に関しても、依存関係の書き換え規則を宣言的に記述することにより、論理型文法の枠組の中で自然に取り扱うことが可能であることが示されている[1]。またこの形式化と同時に実験実証もなされている[2]。

本稿ではこの実験結果に基づき、従来の書き換え規則の記述能力に制限を加えることで、改良をおこなった制限依存文法の記述枠組について述べる。まず書き換え規則とその規則中で実現されている統語制約について述べる。次に依存関係、等位関係を判断する際に使用する制約のうち、文節の持つ統語上の性質を反映させた意味制約について述べる。

2 書き換え規則と統語制約

従来の書き換え規則は、依存関係の制約を構文情報として式(1)で示される0型の規則と、式(2),(3)で示される1型の規則を基本に宣言的に定義されている。

$$\text{word bar word} \rightarrow \text{word word} \quad (1)$$

$$\text{word bar word} \rightarrow \text{word wba word} \quad (2)$$

$$\text{word bar word} \rightarrow \text{word bar wba word} \quad (3)$$

式(1)は、隣り合う文節同志に依存関係が認められる場合の書き換え規則であり、式(2),(3)は、一つ以上の文節列を挟む場合の書き換え規則である。式中 word は文節を表し、wba は文節列を表す。また bar は、依存関係が成立した際どの文節がどの文節に依存しているかを示す情報が蓄積されている。

この書き換え規則を用いて解析された結果、文節ノードの左から右へ1パスでたどることが可能なリンクが張られた場合について考察する。例として文節数が3の場合を考察する。

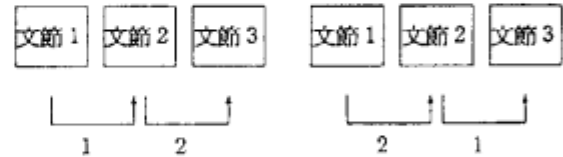


図1: 書き換え規則を用いた解析例

(図1)のように文節と文節の係り方の順序の違いにより2通りの解釈結果が得られることがわかる。一般に n 個文節が存在する場合、 $(n-1)!$ 通りの解釈結果が得られる。したがってこの過剰生成を防ぐために、従来の書き換え規則の持つ統語制約[2]に線形順序依存性という形でより強い制限を加えた。

線形順序依存性: 文節ノードの左から右へ1パスでたどることが可能なリンクがある場合その順序は保持される。

上の統語制約を加えた書き換え規則は以下の通りである。

$$\text{sentence} \rightarrow [\text{begin}] \text{ws} \quad (4)$$

$$\text{word} \rightarrow \text{word} \quad (5)$$

$$\text{ws} \rightarrow \text{wordn ws} \quad (6)$$

$$\text{ws} \rightarrow \text{wordn} \quad (7)$$

$$\text{ws} \rightarrow \text{word} [\text{end}] \quad (8)$$

$$\text{wordn word} \rightarrow \text{word word} \quad (9)$$

$$\text{wordn word} \rightarrow (\text{word; wordn}) \text{ws word} \quad (10)$$

書き換え規則中、線形順序依存性は、ノードカテゴリ wordn により実現されている。つまり書き換え規則適用後 wordn のノードカテゴリは、書き換え規則右辺の右端のノードとして出現しないという制約を記述することにより実現されている。また本規則は同じ記述枠組を用いて依存関係と同様等位関係も扱うことが可能となっている。つまり式(9)は、隣り合う文節同志に依存関係、あるいは等位関係が認められる場合の書き換え規則

則であり式(10)は、離れた文節同志に依存関係あるいは等位関係が認められる場合の書き換え規則である。なお規則中、wsは1つ以上の文節列を表し word は文節を表す。式(9)(10)の右辺の右端のノードはその文節の範疇名が、体言あるいは用言を、左端のノードは連体詞あるいは連用詞を表す。そして左辺の右端のノードは体言、用言、連体詞あるいは連用詞を、左端のノードは連体詞あるいは連用詞を表す。

3 意味制約

一般にある文節がある文節に依存するか否か、またはある語がある句の構成要素になるか否かは、ある程度言語に依存することが[5]により示されている。

本文法では意味制約、つまり文節と文節に依存関係が認められるか否かを判定するための制約条件として、連体修飾、連用修飾、格関係による制約条件の他に、文の構造面を反映させた制約条件を導入した。具体的には各文節に依存関係の充足度という値を持たせることにより実現した。文の構造に関しては[5]、および[3]の述部の階層構造を基本にしている。

依存関係の充足度: 文節同志に意味的依存関係が成立するか否かの指標を示す値。文の構造面を反映させた制約条件である。

以下充足度と、その値を取りうる文節の一部を記述する。(表1)において名詞+格助詞の中に主語を導入したのは、一部の scrambling を処理したいためである。

表 1: 充足度と文節	
充足度	文節
1	名詞+格助詞, 程度副詞, ながら etc
2	場所の修飾語, ないで, ても etc
3	は, おそらく, から, し, etc

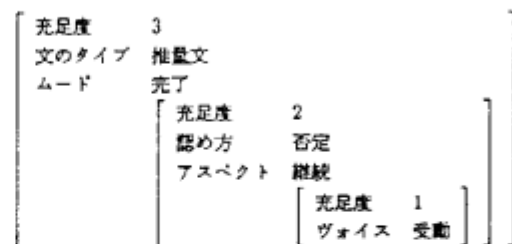
文節の持つ充足度は、その文節が体言あるいは用言の場合は受けの充足度となり、連体詞あるいは連用詞の場合は、係りの充足度となる。そして依存関係が成立するか否かはこの充足度が満足されるか否かにより決定される。つまり文節 A 文節 B の並びが存在する場合、文節 A の充足度を A(Satis)、文節 B の受けの充足度を B(Satis) とすると、

$$A(\text{Satis}) = B(\text{Satis}) \quad (11)$$

の制約条件が満たされるとき依存関係が認められるとした。

また特に受けの文節が用言の場合には、充足度の値と述部の統語範疇とを以下のように対応させた。

食べられて いなかった だろう



以上述べた充足度を用いて、文「おそらく(A) 京都で(B) 会議が(C) 開かれるだろう(D)」を例に解析した結果を図示すると(図2)のようになる。

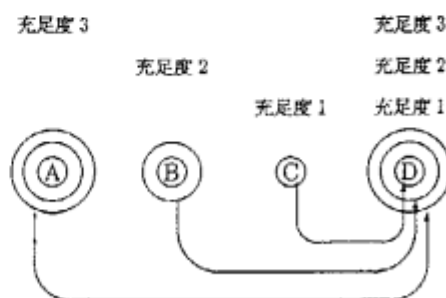


図 2: 充足度を用いた解析例

(図2)において各ノードに付与されている円弧の数が、そのノードの充足度を表す。また、ノードを結ぶアークは依存関係を表す。

4 おわりに

論理型文法の記述枠組の中での依存構造解析手法に、線形順序依存性、依存関係の充足度という制約を加えることで実現した、制限依存文法とその表現について述べた。これにより、これまで直感的解釈によることが多かった依存文法に対し、意味制約を加えることで、より日本語に適した依存文法の形式化が可能となった。今後はこの記述枠組を利用し、特に文脈処理に適した意味表現の抽出を念頭に評価改良を行っていく予定である。

参考文献

- [1] Sugimura, R. *Logical Dependency Grammar and Its Constraint Analysis* ICOT technical memorandum TM-0679, pp. 10, 1989
- [2] 杉村 領一, 福本 文代, 「論理型文法による依存構造解析」 情報処理学会, 談話理解モデルとその応用シンポジウム, 89-100, 1989
- [3] 佐野 洋, 「述部の階層分析と文脈情報」 情報処理学会, 談話理解モデルとその応用シンポジウム, 65-75, 1989
- [4] 杉村 領一, 三言秀夫, 向井国昭, 「コンストレイントに基づく日本語の係り受け解析」 情報処理学会第 95 回全国大会, 1331-1332, 1988
- [5] 南 不二男, 「現代日本語の構造」 大修館書店 1986
- [6] 池田 尚志, 「語法規則方式による日本語文の構文・意味解析」 情報処理学会論文誌, Vol.26, No.6, Nov, 1985