

語義文における動詞間の上位一下位関係

5G-5

富浦 洋一¹, 日高 達¹, 吉田 将²

九州大学

九州工業大学

1. まえがき

国語辞典の語義文から動詞間の上位一下位関係を抽出する方法と抽出結果について述べる。動詞の多義性を考慮して、同一の動詞でも語義が異なれば異なる述語に対応させ、上位一下位関係を第一階の述語言語で記述する。

まず、語義文の論理的性質、および見出し動詞とその語義文の関係から、語義文がどのような統語構造のとき上位一下位関係が抽出できるか、その場合、上位一下位関係を記述している部分は語義文の統語構造のどこかについて述べる。さらに、抽出の際問題となる、定義動詞の語義の選択、見出し動詞と定義動詞の変数の対応のための必要条件とヒューリスティックについて述べ、最後に、抽出結果について述べる。

2. 見出し動詞とその語義文の関係

本稿では、内包的動詞、補文を要求する動詞は取扱わない。また、文の述語が動詞である語義文のみ対象とする。

動詞は個体間の関係を示し、動詞の語義文はその動詞が示す個体間の関係を他の単語で記述したものである。したがって、一般に、見出し語に対応する論理式の変数はすべて語義文に対応する論理式の変数に対応する。語義文の変数のうち、見出し語の変数に対応しない変数がある場合も考慮して、見出し語とその語義文の関係を次の様に定義する。

定義1：見出し語とその語義文の関係

$$\forall x [V^*(x) \equiv \exists z S(x, z)] .$$

ただし、 $V^*(x)$, $S(x)$ は、それぞれ、見出し動詞に対応する論理式、語義文に対応する論理式である。

3. 語義文の論理的性質

本稿では、主節の述部が一つの動詞である文を単文と呼ぶことにする。したがって、連体修飾節を含む文、および、2つの文が主従の関係にある複文も単文と呼ぶ。単文から修飾部を取り除いた文を骨格文と呼ぶ。

性質1：単文とその骨格文の関係

$$\forall x \forall y [S(x, y) \supset S^*(x)] .$$

ただし、 $S(x, y)$, $S^*(x)$ は、それぞれ、

単文、その骨格文に対応する論理式である。

$S^*(x)$ は、

$$N_1(x_1) \wedge \cdots \wedge N_N(x_N) \wedge V^d(x) \quad (1)$$

の形式の論理式であり、一階の述語言語で表現できる。ただし、 $N_i(x_i)$ は骨格文の名詞に対応する論理式、 $V^d(x)$ は骨格文の動詞(定義動詞)に対応する論理式である。

複数の文から構成され、それらが対等の関係にある文を重文と呼ぶ。「かつ」的意味で結ばれる重文を連言的重文、「または」的意味で結ばれる重文を選言的重文と呼ぶ。文 s' が重文 s を直接構成する場合、 s を親の文、 s' を子の文と呼ぶ。単文 s の親の文、子の文は、 s 自身であるとする。親-子の関係の推移的閉包の関係を、祖先-子孫の関係と呼ぶ。

性質2：文 s とその子孫の単文 s' の関係

(a) s' の祖先が単文、連言的重文のみの場合

$$\forall x \forall y [S(x, y) \supset S'(x)] . \quad (2)$$

(b) s' の祖先が単文、選言的重文のみの場合

$$\forall x \forall y [S'(x) \supset S(x, y)] . \quad (3)$$

ただし、 $S(x, y)$, $S'(x)$ は、 s , s' に対応する論理式である。

4. 動詞間の上位一下位関係の抽出

4. 1. 上位一下位関係記述部の抽出

定義1, 性質1, (2)より次式を得る。

$$\exists z [V^*(x, y) \supset S^*(x, z)] . \quad (4)$$

(1)と上式から、

$$\exists z [V^*(x, y) \supset \prod_i N_{zi}(z_i) \wedge V^d(x, z)] . \quad (5)$$

を得る。上位一下位関係は、同時的必然的に成り立つ関係であり、一階の述語論理では含意に相当する。(5)は見出し動詞と定義動詞間に上位一下位関係(一般化上位一下位関係)が成立することを示している。

また、定義1, (3)より、

$$S'(x, z) \supset V^*(x) . \quad (6)$$

を得る。

$$S'(x, z) \triangleq \prod_i N_{xi}(x_i) \wedge \prod_i N_{zi}(z_i) \wedge V^d(x, z)$$

の場合、制限付ではあるが、(6)も見出し動詞と定義動詞間の上位一下位関係（特殊化上位一下位関係）を示している。

抽出規則

s' を語義文の子孫の単文とする。

(a) s' の祖先が、単文、連言的重文のみの場合、 s' の骨格文が上位一下位関係記述部であり、このとき見出し動詞との間に一般化上位一下位関係が抽出できる。

(b) s' の祖先が、単文、選言的重文のみで、修飾部を含まない場合、 s' が上位一下位関係記述部である。このとき見出し動詞との間に特殊化上位一下位関係が抽出できる。

次節で述べる必要条件とヒューリスティックを用いて、定義動詞の語義の選択、見出し動詞と定義動詞の変数の対応をとると、(5)(6)に対応する上位一下位関係が抽出できる。

4. 2. 必要条件とヒューリスティック

定義2：述語の有意味定義域

述語 $P(x)$ とその x_i に関する有意味定義域 $D_{Pi}(x_i)$ は次の関係にある。

$$\forall x [P(x) \supset D_{Pi}(x_i)]$$

有意味定義域は、名詞句の意味素性、意味マーカと呼ばれて来たものに相当する。

まず、有意味定義域を用いて、構文解析で用いられる選択制限により、定義動詞の語義の候補をさぼる。(4)、(6)、定義2より、次の必要条件を得る。

必要条件

定義動詞のある語義を選択し、見出し動詞との変数の対応を次の様にとったとする。

$$V^e(\dots x \dots) : \dots \wedge N(x) \wedge \dots \wedge V^d(\dots x \dots)$$

このとき、この語義の選択、変数の対応が適切であるための必要条件は、

$$\exists x [D_{ex}(x) \wedge D_N(x) \wedge D_{dx}(x)]$$

である。ただし、 $D_{ex}(x)$ 、 $D_N(x)$ 、および $D_{dx}(x)$ はそれぞれ、 $V^e(x)$ の x に関する有意味定義域、 $N(x)$ の有意味定義域、 $V^d(x)$ の x に関する有意味定義域である。

ヒューリスティック

見出し動詞、骨格文間で対応しない変数が少ない定義動詞の語義の選択、変数の対応が適切なものである。

5. 抽出実験

語義文、有意味定義域の記述がある計算機可読の辞書として、情報処理振興事業協会技術セ

ンターから出された計算機用日本語基本動詞辞書 I P A L¹⁾ を用いて、抽出実験を行なった。

抽出結果は以下のとおりである。

対象：見出し動詞と上位一下位関係記述部の組 1 4 6 7

得られた上位一下位関係：1 0 3 0

(すなはち、真の解が必要条件を満たす)

残り 4 3 7 は、辞書中に定義動詞の適切な語義がないか、有意味定義域の不備か、今回のアルゴリズムが任意格に対処していないために抽出に失敗したものである。

1 組当りの 1 位解の平均の個数：4. 1

1 位解に真の解が含まれる割合：7 0. 1 %
(内、語義文が単文の場合：8 1 %)

1 0 3 0 のうち単文は 5 3 3 あり、上位一下位関係抽出の支援システムとしては有効と言える。得られた上位一下位関係の性質は

最上位語義：3 5 3

上位一下位関係の最大の深さ：6

LOOP：2 (長さ共に 2)

補う II $\Leftrightarrow$ 埋める IV

応じる II $\Leftrightarrow$ 受ける VI

下位語義の多い最上位語義：

変える VII $\dots$ 5 5

及ぼす I $\dots$ 3 1

持つ V $\dots$ 2 5

等である。最上位語義は全体数に比べて多いが、今回の実験で行なわなかった「なる」「する」を定義動詞とするものに対して抽出実験を行ない、さらに、I P A L に登録されていないサ変動詞のうち特に基本的と思える動詞（移動する、存在する、発生する、など）を登録して、これらを定義動詞とするものに対しても抽出実験を行なえば、最上位語義は 2 0 0 程度になると思われる。

6. あとがき

見出し動詞と語義文の関係を定義し、語義文の論理的性質を考察して、見出し動詞と定義動詞間の上位一下位関係を抽出する方法を述べた。I P A L を対象に抽出実験を行ない 1 0 3 0 の関係を得た。

今後、一部のサ変動詞、「なる」「する」を定義動詞とする語義に対しても抽出実験を行ない、関係を強化する予定である。

謝辞

本研究は、国語辞典として計算機用日本語基本動詞辞書 I P A L を用いた。I P A L 編集者ならびに関係各位に厚く謝意を表します。

1) 計算機用日本語基本動詞辞書 I P A L (Basic Verbs) 説明書、I P A 技術センター

2) 富浦, 日高, 吉田：国語辞典の語義文からの動詞の上位一下位関係の抽出, NL73-3, 1989