

名詞句「 N の N 」の解析におけるシソーラス知識の利用

1 E-3

森 重樹, 中村 貞吾, 日高 達
九州大学

1 はじめに

日本語文においては, 2つの名詞「 N_1 」, 「 N_2 」が連体助詞「の」で結合された「 N_1 の N_2 」という形の名詞句が頻繁に使われる。この名詞句における2つの名詞の意味関係は多様であるが, 「 N_1 の N_2 」という表現の中には意味関係を決定する情報は明示的には現れていない。本論文では, 名詞をいくつかのタイプに分類して「 N_1 の N_2 」の意味関係を捉え, 更に, 一般知識として名詞のシソーラスを用いて, 「 N_1 の N_2 」という表現の意味的妥当性を論理的に検証する。

2 名詞句「 N_1 の N_2 」の意味関係

通常の名詞は1引数述語に対応し, その引数を指示する(例. 「本」: 「本(x)」)。そして, 一般に「 N_1 の N_2 」に対応する論理表現は,

$$N_1(y) \wedge N_2(x) \wedge P(x, y)$$

となる。ここで, 「 P 」は「 N_1 の N_2 」の明示されない意味関係であり, 名詞句全体は引数「 x 」を指示している。この意味関係を決定するために, 名詞を次の3つのタイプに分類する。(1) 事象名詞(動詞から派生したものであり, 事象を指示対象とする), (2) 関係名詞(2項述語によって記述される関係や属性に対応し, 2つの引数のうちの1つを指示する(例. 「母」: 「母(x, y)」)), (3) 通常名詞(1項述語に対応し, その引数を指示する)。

これらの分類により, 「 N_1 の N_2 」の明示されない意味関係は次のように考えられる。

1. 「 N_1 」が事象名詞で「 N_2 」が「 N_1 」に対応する動詞の引数, または「 N_2 」が事象名詞で「 N_1 」が引数。

例. 「英語の勉強」

$$\text{英語}(y) \wedge \text{勉強}(x) \wedge \text{object}(x, y)$$

「 $\text{object}(x, y)$ 」は事象「 x 」の対象が「 y 」であることを示す。

この場合, 事象名詞に対応する動詞の格フレーム知識を用いて, 引数がどの格役割に対応するかを決定できる。

2. 「 N_2 」が関係名詞。

この場合, 「 N_1 の N_2 」の論理表現は,

$$N_1(y) \wedge N_2(x, z) \wedge P(\dots)$$

となる。ここで, 「 P 」が「 $y = z$ 」, すなわち, 「 N_1 」が関係名詞「 N_2 」の補引数となることがある。そのとき「 N_1 」は個体を指示しており, その論理表現は,

$$N_1(y) \wedge N_2(x, y)$$

と表される(例. 「太郎の母」: 太郎(y) $\wedge$ 母(x, y)。しかし, 「 N_1 」が個体指示である場合がすべてこれにあたるわけではなく, 「故郷の母」などは,

$$\text{故郷}(y) \wedge \text{母}(x, z) \wedge P(x, y)$$

となり, 意味関係「 P 」は次の3.の場合と同様に決定される。このように「 N_2 」が関係名詞の場合, 「 N_1 」は「 N_2 」の補引数に対する選択制約を満足するかどうか吟味する必要がある。

また, 「 N_1 」が個体指示であるか否かによっても曖昧性を持つ。

例. 「日本人の妻」

- (a) 「日本人」がある特定の人を示す場合。

$$\text{日本人}(y) \wedge \text{妻}(x, y)$$

(日本人である男性の妻を指示する)

- (b) 「日本人」が「妻」の属性を示す場合。

$$\text{日本人}(x) \wedge \text{妻}(x, y)$$

(日本人である妻を指示する)

この場合, 文脈情報等を利用して, 「 N_1 」が個体指示であるかどうかを決定する必要がある。

3. 「 N_1 」, 「 N_2 」が通常名詞。

この場合, 様々の可能な意味関係(所有, 全体部分, 場所, ...)があり, これらを決定するためには, 名詞のシソーラスや共起関係等の様々な知識が必要となる。

Utilization of a Thesaurus on Analyzing Noun Phrases
「 N no N 」

Shigeki MORI, Teigo NAKAMURA, Toru HITAKA
Kyushu University

3 「 N_1 の N_2 」の意味的妥当性

「 N_1 の N_2 」の2つの名詞間には様々な意味関係が考えられるが、2つの名詞を常に「の」で結合できるという訳ではない。そこで、「 N_1 の N_2 」の意味的妥当性を一般知識を用いて論理的に判定する2つの原則を提案する。

1. 「 N_1 の」が「 N_2 」に矛盾しない。
(「 N_1 の N_2 」の指示対象が空でない)
2. 「 N_1 の」が「 N_2 」を真に制限する。
(「 N_1 の」が冗長でない)

この原則を論理的に記述すると次のようになる。

$$\Gamma \nVdash \neg \exists x \alpha$$

$$\Gamma \nVdash \alpha \equiv \beta$$

ここで、「 α 」は「 N_1 の N_2 」に対応する論理式、「 β 」は「 N_2 」に対応する論理式、「 x 」は「 N_1 の N_2 」によって指示される変数、「 Γ 」は一般知識である。

これらの原則は「 N_1 の N_2 」の意味的妥当性に対する必要条件である。

4 シソーラスを利用した「 N_1 の N_2 」の意味的妥当性の判定

ここで、一般知識 Γ として筑波大学の荻野先生が作成した「現代日本語名詞シソーラス」を用いる。このシソーラス中には、名詞間の上位下位関係、全体部分関係が、それぞれ以下のような2つの単語間の2項関係として記述されている。

$$\begin{array}{ll} \text{上位-下位 } (N_1, N_2) & N_1: \text{上位}, N_2: \text{下位} \\ \text{全体-部分 } (N_1, N_2) & N_1: \text{全体}, N_2: \text{部分} \end{array}$$

これらを論理式で表現すると、上位下位関係は、

$$\forall x [N_2(x) \supset N_1(x)]$$

と表される。全体部分関係には2つの異なる関係があり、

$$\begin{array}{l} \text{PART-OF}(N_2, N_1) \stackrel{\text{def}}{=} \\ \forall x [N_2(x) \supset \exists y [N_1(y) \wedge \text{part-of}(x, y)]] \\ (\text{全ての } 'N_2' \text{ は } 'N_1' \text{ の部分である}) \end{array}$$

$$\begin{array}{l} \text{HAS-PART}(N_1, N_2) \stackrel{\text{def}}{=} \\ \forall x [N_1(x) \supset \exists y [N_2(y) \wedge \text{part-of}(y, x)]] \\ (\text{全ての } 'N_1' \text{ は } 'N_2' \text{ を部分として持つ}) \end{array}$$

なる論理式に対応する。ここで、 $\text{part-of}(x, y)$ は個体間の全体部分関係を表し、「 x 」が「 y 」の部分であることを意味する。

以下では、シソーラスを用いて「 N_1 の N_2 」の妥当性の判定を行う方法について述べる。もし、 $\text{PART-OF}(N_2, N_1)$ が知識中にあれば、

$$\begin{array}{ll} N_1 \text{ の } N_2 & : \alpha = N_1(y) \wedge N_2(x) \wedge \text{part-of}(x, y) \\ N_2 & : \beta = N_2(x) \end{array}$$

であるので、

$$\Gamma \vdash \alpha \equiv \beta$$

となり、「 N_1 の N_2 」は原則2.により意味的に妥当な表現ではない。 $\text{HAS-PART}(N_1, N_2)$ についても同様のことがいえる。例えば、いま $\text{HAS-PART}(\text{少女}, \text{髪})$ が知識中にあるとする。名詞句「少女の髪」は「少女の」が「髪」の指示対象を制限しているの、原則2.により妥当な表現である。しかし、「髪の少女」は「髪の」が冗長な表現となり、原則2.を満足しない。

ところが、「細胞の核」のように表層的にはこれらの原則に当てはまらないながらも、意味的に妥当な表現として成り立つものがある。「核」には、

- 核₁: 細胞の核
- 核₂: 物事を中心
- 核₃: 原子核
- 核₄: 核兵器

などの意味があり、

$$\text{PART-OF}(\text{核}_1, \text{細胞}) \in \Gamma$$

であるから、「核」の意味が「核₁」であれば、原則2.により「細胞の核₍₁₎」は妥当な表現ではない。しかし、この場合「細胞の」は「核」に対する多義の制約と考えられ、原則2.は「細胞の核」においても成立することになる。

5 おわりに

日本語における「 N_1 の N_2 」という名詞句の意味的妥当性を、一般知識として名詞のシソーラスを用いて判定するための2つの原則を提案した。今後、文脈情報を用いた意味関係の決定、及びその妥当性についても検証して行く。

参考文献

- [1] 池田 博之, 名詞句「 N の N 」の解析におけるシソーラス知識の利用, 九州大学大学院総合理工学研究科修士論文, 1992
- [2] 草薙 裕, 南 不二男, 中野 洋, 吉田 夏彦, 朝倉 日本語新講座4「文法と意味II」, 朝倉書店, 1985