

語義文からの動詞の階層関係の抽出

富浦 洋一* 日高 達* 吉田 将**

*九州大学工学部 **九州工業大学情報工学部

我々は、国語辞典からの動詞の階層関係の抽出に関する研究を進めている。制限付きではあるが、見出し語とその語義文の間には、ある階層関係が成り立っていると考えられる。そこでこの仮定に基づき、以下のことを議論する。

- (1) 述語間の階層関係の定義。
- (2) 語義文に対応する論理式の諸性質。
- (3) どの様にして、語義文に対応する論理式から動詞間の階層関係を求めるか。

EXTRACTION OF HIERARCHICAL RELATION AMONG VERBS FROM DEFINITION SENTENCES IN JAPANESE DICTIONARY

Yoichi TOMIURA* , Toru HITAKA* and Sho YOSHIDA**

*Kyushu Univ, 6-10-1, Hakozaki, Higashi, Fukuoka, 812, Japan.
**Kyushu Institute of Technology.

We are studying how to extract hierarchical relation among verbs from definition sentences in Japanese dictionary.

We assume that there exists an equivalent relation under some restrictions, between an entry word and its definition sentence. Based on this assumption following points are discussed,

- (1) the definition of the hierarchical relation between predicates,
- (2) characteristics of the logical formula corresponding to a definition sentence,
- (3) how to extract hierarchical relations among verbs from the corresponding logical formulas.

1. まえがき

自然言語の意味処理を行なう際、様々な知識が必要不可欠となる。この膨大な知識を一つ一つ整理し、入力するのは大変な作業になり、従って計算機により半自動的に知識の収集ができることが望ましい。我々は、その試みとして、国語辞典からの知識の抽出に関する研究を進めている。

動詞の意味の定義として、同義語、反義語、上位語、下位語などを示したり、その動詞の原因、結果などの因果関係を示したりするが、国語辞典では普通、見出し語とある意味で等価な文をその語義文として記述してある。すなわち、【見出し語】が成立すれば、同時に、【語義文】が成立するという関係があり、逆に制限付きではあるが、【語義文】が成立すれば、同時に、【見出し語】が成立するという関係がある。これから、動詞間の上位-下位関係を抽出し、non-logical axiom とすることができる。

本稿では、動詞間の上位-下位関係とは何かを定義し、この関係をどのようにして抽出するかについて説明する。そしてこの際問題となる語義文中の動詞の語義番号の選択、上位-下位関係にある動詞間の格スロットの対応について述べる。

2. 見出し語とその語義文の関係

見出し語とその語義文の関係を考える前に、まず文の論理式への翻訳について少し触れておく。簡単のために次のような単文を考える。

建物ができる。

これは次のような論理式に翻訳できる。

$\$ \text{建物}(x) \wedge \$ \text{できる}(x)$ 。

$\$ \text{建物}(x)$, $\$ \text{できる}(x)$ は、それぞれ、「建物」、「できる」の論理式への翻訳である。上式は「建物」という集合の要素 x が「できる」ことを示している。このように名詞句は集合を示す単項述語（特性関数）と考えられ、動詞に係る場合、動詞の変数の領域を制限するのとして扱える。一般に単文は次のように論理式に翻訳できる。

$\text{NP}(x) \wedge \text{P}(x)$ 。

$\text{P}(x)$ は文の述語とそれに係る連用修飾句を合わせたものの論理式への翻訳、 $\text{NP}(x)$ は文の述語に係る名詞句の論理式への翻訳である。本稿では次のような略記法を用いる。

$\text{P}(x) : \text{P}(x_1, \dots, x_n)$

$\text{P}(x) : \text{P}_1(x_1) \wedge \dots \wedge \text{P}_n(x_n)$

$\exists x : \exists x_1 \dots \exists x_n$

$\exists(x-x_i) : \exists x_1 \dots \exists x_{i-1} \exists x_{i+1} \dots \exists x_n$

単純に考えると、見出し語とその語義文の間には同値関係が成立しているように思われる。これは次式で表現できる。

$$\begin{cases} \forall x [V^n(x) \supset S(x)], \\ \forall x [S(x) \supset V^n(x)]. \end{cases} \quad (1)$$

ただし、 $V^n(x)$, $S(x)$ は、それぞれ、見出し語、語義文の論理式への翻訳である。

次の例を考える。

建つ VIII 建物ができる。

「駅前にビルが建った。」

語義文中の「できる」は次の意味で用いられていると考えられる。

できる IV それまでは存在しなかった何かが完成・成立する。

「最近新しい法律ができた。」

「Aが建つ」と言うとき、Aは「建物」であり、そのAが「できる」ことであるから、「建物Aができる」と言える。また逆に、「建物Aができる」と言うとき、「Aが建つ」と言える。つまり、

$$\begin{cases} \forall x [\$ \text{建つ}^3(x) \supset \$ \text{建物}(x) \wedge \$ \text{できる}^4(x)], \\ \forall x [\$ \text{建物}(x) \wedge \$ \text{できる}^4(x) \supset \$ \text{建つ}^3(x)]. \end{cases}$$

が成立する。では、次の例を考えてみよう。

愛する I 誰かや何かに対してとても好きだという気持ちを持つ。

「彼は妻子を愛している。」

語義文中の「持つ」は次の意味で用いられていると考えられる。

持つ V 何かに対してある感情を抱く。

「私は将来に希望を持っている。」

「愛する I」, 「持つ V」はそれぞれ次のように論理式に翻訳でき、

x_1 が x_2 を愛する $\rightarrow \$ \text{愛する}^1(x_1, x_2)$

x_1 が x_2 に x_3 を持つ $\rightarrow \$ \text{持つ}^5(x_1, x_2, x_3)$

「愛する I」の語義文は次のように論理式に翻訳できる。

$\$ \text{誰かや何か}(x_2)$

$\wedge \$ \text{とても好きだという気持ち}(x_3)$

$\wedge \$ \text{持つ}^5(x_1, x_2, x_3)$ 。

よって、(1)の第一式に当てはめて、

$\forall x_1 \forall x_2 \forall x_3 [\$ \text{愛する}^1(x_1, x_2) \supset$

$\$ \text{誰かや何か}(x_2) \wedge$

$\$ \text{とても好きだという気持ち}(x_3) \wedge$

$\$ \text{持つ}^5(x_1, x_2, x_3)]$ 。

を得る。この論理式は、 x_1, x_2 が

$\$ \text{愛する}^1(x_1, x_2)$

を満たす場合、任意の x_3 に対して

$\$ \text{とても好きだという気持ち}(x_3)$

が成立することを主張しているが、これは明らかにおかしい。しかし、見出し語の意味をその語義文で説明しているからには、 $\$ \text{愛する}^1(x_1, x_2)$ が成立していれば、同じ世

界、同じ時間で成立することが「愛する I」の語義文にかかれてはいるはずである。そこで、語義文に現われて見出し語に現われない変数 z に関して、限量子を全称から存在に変えて、

$$\forall x_1 \forall x_2 \exists z [\$ \text{愛する}^1(x_1, x_2) \supset \\ \$ \text{誰かや何か}(x_2) \wedge \\ \$ \text{とても好きだという気持ち}(z) \wedge \\ \$ \text{持つ}^5(x_1, x_2, z)] .$$

を考えてみる。確かに、 x_1, x_2 が

$$\$ \text{愛する}^1(x_1, x_2)$$

を満たすとき、 x_2 は、

$$\$ \text{誰かや何か}(x_2)$$

を満たし、ある z が存在して、

$$\$ \text{とても好きだという気持ち}(z)$$

$$\wedge \$ \text{持つ}^5(x_1, x_2, z) .$$

を満たす。よって上式は成立する。

一般に、述語 $P^1(x, y)$ が成立していれば同じ世界、同じ時間に、述語 $P^2(x, z)$ が成立するという関係（これを $P^2(x, y), P^1(x, z)$ 間の上位-下位関係と呼ぶ）は次のように表わせる。

[定義 1] 述語間の上位-下位関係

述語 $P^2(x, z), P^1(x, z)$ 間に上位-下位関係がある。
if and only if

$$\forall x \forall y \exists z [P^1(x, y) \supset P^2(x, z)] .$$

定義 1 に基づき (1) を次のように修正する。

$$\left\{ \begin{array}{l} \forall x \forall y \exists z [V^n(x, y) \supset S(x, z)] , \\ \forall x \forall z \exists y [S(x, z) \supset V^n(x, y)] . \end{array} \right. \quad (2)$$

次の例を考える。

上がる IX ある現象が生じる。

「会場に歓声が上がった。」

語義文中の「生じる」は次の意味で用いられていると考えられる。

生じる I ある物や現象が発生する。

「壁にひびが生じる。」

「上がる IX」、「生じる I」の論理式への翻訳をそれぞれ次のように表わす。

$$x_1 \text{ が } x_2 \text{ に上がる} \rightarrow \$ \text{上がる}^9(x_1, x_2)$$

$$x_1 \text{ が } x_2 \text{ に生じる} \rightarrow \$ \text{生じる}^1(x_1, x_2)$$

ここで注意を要するのは語義文中の「ある」の取扱である。「ある」を存在限量子と解釈すると、「上がる IX」の語義文は、

$$\exists x_1 [\text{現象}(x_1) \wedge \$ \text{生じる}^1(x_1, x_2)] .$$

と論理式に翻訳される。しかし x_1 は $\$ \text{上がる}^9(x_1, x_2)$ の自由変数として現われるので、この場合、「ある」を存在限量子として解釈するのはおかしい。「ある現象」は「現

象」の部分集合を示していると考えるのが妥当である。このように、語義文には曖昧な表現があるのであるが、我々はそのに対してたいい何等不都合を感じない。人は普通、恒偽となるような文は話さない。従って辞書も見出し語が恒偽になるような領域まで説明しているわけではない。我々は語義文を、次に示す有意味定義域に制限して解釈していると考えられる。

[定義 2] 有意味定義域

述語 $P(x)$ の x_i に関する有意味定義域は、

$$\{ x \mid X_i(x) \}$$

である。ただし、

$$X_i(x_i) = \diamond \exists (x-x_i) P(x)$$

\diamond は可能を示す様相演算子である。

以後、有意味定義域のことを単に定義域と呼び、混同の恐れがない限り、定義域を表わす述語 $X_i(x_i)$ も単に定義域と呼ぶ。

定義域について、次の補題 1 が成り立つ。

[補題 1]

$$\forall x \forall y \exists z [P^1(x, y) \supset P^2(x, z)]$$

$$\supset \forall x_i [X^1_i(x_i) \supset X^2_i(x_i)] , \quad (i=1, 2, \dots)$$

ただし、 $X^1_i(x_i), X^2_i(x_i)$ はそれぞれ述語 $P^1(x, y), P^2(x, z)$ の x_i に関する定義域である。

定義域は、見出し語よりもその語義文の方が一般に広い。したがって、補題 1 により (2) の第二式は一般に成立しない。(2) の $S(x, z)$ を $X^n(x) \wedge S(x, z)$ で置き換え修正する。ただし、 $X^n(x)$ は見出し語の x に関する定義域である。また、見出し語の変数と対応する変数の領域を制限している「ある」付きの名詞句の論理式への翻訳は、 $X(x) \wedge NP_z(x)$ と等価と考えられる。ただし、 $NP_z(x)$ は、その名詞句から「ある」を取り除いたものの論理式への翻訳である。したがって、上記のように $S(x, z)$ を置き換えることにより、このような名詞句の「ある」は取り除いて考えられる。定義域の定義より、

$$\forall x \forall z [V^n(x, y) \supset X^n(x)] .$$

ゆえ、(2) は次のように修正される。

[定義 3] 見出し語と語義文の関係

$$\left\{ \begin{array}{l} \forall x \forall y \exists z [V^n(x, y) \supset S(x, z)] , \\ \forall x \forall z \exists y [X^n(x) \wedge S(x, z) \supset V^n(x, y)] . \end{array} \right.$$

語義文を論理式に翻訳するときに、語義文中の動詞の語義番号の選択、および見出し語と語義文の変数の対応をどう決めるかが問題となる。定義 3 の関係は適切な語義番号の選択、見出し語との語義文の適切な変数の対応が成されて、初めて成り立つものである。この問題に関しては第 4

節で述べる。

3. 動詞間の上位-下位関係の抽出

見出し語とその語義文の関係から次に示す動詞間の上位-下位関係を抽出することができる。

[定義4] 動詞間の上位-下位関係

動詞 v^1, v^2 の論理式への翻訳をそれぞれ, $V^1(x, y), V^2(x, z)$ とする。

$X(x) \wedge Y(y)$ が成立するとき, v^2 と v^1 が上位-下位関係にある。

if and only if

$$\forall x \forall y \exists z$$

$$[X(x) \wedge Y(y) \wedge V^1(x, y) \supset V^2(x, z)].$$

定義3の第二式から得られる関係は, 上記のような関係になる。このような関係も, 実際の利用を考えた場合使えることもあると考えられるので, 定義1を拡張して, 制限付きの上位-下位関係も含めて定義4のように定義する。

本節では, 語義文がどのような統語的構造のとき, どのような上位-下位関係が得られるかについて述べる。

まず, 最も単純な場合から考える。文が単文で, 連用修飾句, 連体修飾句部を含まない場合, その論理式への翻訳は,

$$N_x(x) \wedge N_z(z) \wedge V^0(x, z)$$

である。ただし, $N_x(x), N_z(z)$ は名詞の論理式への翻訳であり, $V^0(x, z)$ は述語である動詞 v^0 の論理式への翻訳である。 x は見出し語との変数の対応が取れるもの, z は取れないものを示す。

これから, 次の規則が導出できる。

[規則1]

文が単文で, 連用修飾句, 連体修飾句部を含まず, その論理式への翻訳が,

$$S(x, z) = N_x(x) \wedge N_z(z) \wedge V^0(x, z) \quad (3)$$

であるとする。

このとき,

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

が成立しているならば, 上位-下位関係,

$$\forall x \forall y \exists z [V^n(x, y) \supset V^0(x, z)].$$

が得られる。

また,

$$\forall x \forall z \exists y$$

$$[X^n(x) \wedge S(x, z) \supset V^n(x, y)]. \quad (5)$$

が成立しているならば, 上位-下位関係,

$$\forall x \forall z \exists y$$

$$[X^n(x) \wedge N_x(x) \wedge N_z(z) \wedge V^0(x, z)$$

$$\supset V^n(x, y)].$$

が得られる。

特に, (3), (4), (5) がともに成立する場合を考える。

(3), (4) より,

$$\forall x \forall y [V^n(x, y) \supset N_x(x)].$$

が成り立つ。補題1より,

$$\forall x_i [X^n(x_i) \supset \diamond N_{x_i}(x_i)]. \quad (5)$$

が成り立つ。ここで次の仮定を導入する。

[仮定1]

語義文中の名詞の解釈は, 世界, 時間によらず一定である。

語義文は, 見出し語の意味を説明するものであるから, 理想的には万人に同じ意味を伝えられるものでなければならぬ。したがって, 少なくとも, 語義文を構成する名詞には, 万人が共通に解釈する名詞を選ぶような注意を辞書編集者はしていると考えられる。

仮定1より, 次式が成り立つ。

$$\forall x [\diamond N(x) = N(x)]. \quad (6)$$

上式と(5)より次式が成り立つ。

$$\forall x_i [X^n(x_i) \supset N_{x_i}(x_i)].$$

これより次の規則を得る。

[規則2]

文が単文で, 連用修飾句, 連体修飾句部を含まず, その論理式への翻訳が,

$$S(x, z) = N_x(x) \wedge N_z(z) \wedge V^0(x, z)$$

であるとする。このとき,

$$\left\{ \begin{array}{l} \forall x \forall y \exists z [V^n(x, y) \supset S(x, z)], \\ \forall x \forall z \exists y [X^n(x) \wedge S(x, z) \supset V^n(x, y)]. \end{array} \right.$$

が成立しているならば, 上位-下位関係,

$$\left\{ \begin{array}{l} \forall x \forall y \exists z [V^n(x, y) \supset V^0(x, z)], \\ \forall x \forall z \exists y \end{array} \right.$$

$$[X^n(x) \wedge N_x(z) \wedge V^0(x, z) \supset V^n(x, y)].$$

が得られる。

次に, (複文の場合はその主節に) 重文を含まない文を考える。語義文について以下の仮定を導入する。

[仮定2]

(a) 複文の論理式への翻訳を $S(x)$, その主節のみの論理式への翻訳を $S'(x)$ とすると,

$$\forall x [S(x) \supset S'(x)].$$

(b) 語義文の論理式への翻訳を $S(x)$, それから文の動詞を修飾する副詞などの連用修飾句を取り除いた文の論理式への翻訳を $S'(x)$ とすると,

$$\forall x [S(x) \supset S'(x)].$$

(c) 名詞句の論理式への翻訳を $NP(x)$, それから連体修飾

部を取り除いたものの論理式への翻訳を $N(x)$ とすると、
 $\forall x [NP(x) \supset N(x)]$.

語義文中の複文の従属節は原因、理由、手段、目的などを示してあり、どれも主節を修飾する節である。たとえばある辞書によると、「捜す」の語義文は、「自分の求めるものを見つけるために、ある場所を調べる。」であるが、

(自分の求めるものを見つけるために、ある場所を調べる) \supset (ある場所を調べる)

が成り立つ。また、連用修飾句、連体修飾部は、それぞれ、動詞、名詞の論理式への翻訳を真にする変数の領域を制限しているものである。したがって、それらを含んだものの論理式への翻訳が真ならば、それらを除いたものの論理式への翻訳も真である。たとえば、「しっかりとつかむ」、

$\forall x \forall y [\$ \text{しっかりとつかむ}(x, y)$
 $\supset \$ \text{つかむ}(x, y)]$.
 $\forall x [\$ \text{とても好きだという気持ち}(x)$
 $\supset \$ \text{気持ち}(x)]$.

が成り立つ。

仮定2より次のことが結論される。

【規則3】

文 s の論理式への翻訳を $S(x, z)$ とし、
 $\forall x \forall y \exists z [V^n(x, y) \supset S(x, z)]$.

が成立しているとする。このとき、 s から従属節、連用修飾句、連体修飾部を取り除いた文の論理式への翻訳を $S'(x, z)$ とすると、

$\forall x \forall y \exists z [V^n(x, y) \supset S'(x, z)]$.

が成立する。

最後に重文を含む文について考える。語義文中の重文について、次の仮定を導入する。

【仮定3】

(a) 連言的重文は次のように論理式に翻訳できる。

$S^1(x) \wedge \dots \wedge S^n(x)$

(b) 選言的重文は次のように論理式に翻訳できる。

$S^1(x) \vee \dots \vee S^n(x)$

ただし、 $S^1(x)$ は重文を分解してできる個々の文の論理式への翻訳である。

仮定3より、次のことが結論される。

【規則4】

文 s の論理式への翻訳を $S(x, z)$ とする。いま、

$\forall x \forall y \exists z [V^n(x, y) \supset S(x, z)]$.

が成立しているかつ、 s が連言的重文のとき、

$\left\{ \begin{array}{l} \forall x \forall y \exists z [V^n(x, y) \supset S^1(x, z)] . \\ \vdots \\ \forall x \forall y \exists z [V^n(x, y) \supset S^n(x, z)] . \end{array} \right.$

が成立する。また、

$\forall x \forall z \exists y [X^n(x) \wedge S(x, z) \supset V^n(x, y)]$.

が成立している、かつ、 s が選言的重文のとき、

$\left\{ \begin{array}{l} \forall x \forall z \exists y [X^n(x) \wedge S^1(x, z) \supset V^n(x, y)] \\ \vdots \\ \forall x \forall z \exists y [X^n(x) \wedge S^n(x, z) \supset V^n(x, y)] \end{array} \right.$

が成立する。ただし、 $S^i(x, z)$ は重文を分解して得られる個々の文の論理式への翻訳である。

定義3, 規則1, 規則2, 規則3, 規則4より動詞間の上位-下位関係は次のようにして得られることが解る。

【アルゴリズム】 動詞間の上位-下位関係の抽出(概略)

ある動詞の見出し語 v^n とその語義文に対して以下の操作をする。ただし、 v^n の論理式への翻訳を $V^n(x, y)$ 、 x に関する定義域を $X(x)$ とする。

Step1 語義文から従属節、連用修飾句、連体修飾部をすべて取り除いた文(骨格文)を求める。

Step2 骨格文が単文ならば、その述語である動詞の論理式への翻訳を $V^g(x, z)$ として、

$\forall x \forall y \exists z [V^n(x, y) \supset V^g(x, z)]$.

を出力する。

Step3 骨格文が連言的重文ならば、それを分解して得られる個々の文 s に対して、

s が単文ならば、その述語である動詞の論理式への翻訳を $V^g(x, z)$ として、

$\forall x \forall y \exists z [V^n(x, y) \supset V^g(x, z)]$.

を出力する。

Step4 語義文が単文であり、かつ、連用修飾句、連体修飾部を含まないならば、その述語である動詞の論理式への翻訳を $V^g(x, z)$ 、述語の動詞に係る名詞のうち変数 z を制限するもの論理式への翻訳を $N_z(z)$ として、

$\forall x \forall z \exists y [X(x) \wedge N_z(z) \wedge V^g(x, z) \supset V^n(x, y)]$.

を出力する。

Step5 語義文が選言的重文ならば、それを分解して得られる個々の文 s に対して、

s が単文であり、かつ、連用修飾句、連体修飾部を含まないならば、その述語である動詞の論理式への翻訳を $V^g(x, z)$ 、述語の動詞に係る名詞の論理式への翻訳を $N_z(x)$ 、 $N_z(z)$ として、

$$\forall x \forall z \exists y \\ [X(x) \wedge N_z(x) \wedge N_z(z) \wedge V^0(x, z) \\ \supset V^n(x, y)].$$

を出力する。

4. 語義番号の選択と変数の対応

動詞間の上位-下位関係を抽出するもとなる定義3の関係は、語義文中の動詞の語義番号の選択、見出し語との変数の対応が適切でなければ成り立たない。本節では、まず語義文が単文の場合について、どのようにすれば、適切な語義番号の選択、見出し語との変数の適切な対応ができるかを述べ、次に複文、重文を含む一般の場合について考える。

まず、我々が語義文を読む場合のことを考える。我々は語義文中の動詞の意味を、格パターンとそれに係る名詞句を頼りに選択している。辞書編集者は語義文の意味ができるだけ明確となるように注意を払っているであろうから、語義文中の動詞の語義番号の選択の曖昧さが少なくなるように、次式を満たすような名詞句を使っていると考えられる。

$$\forall x_i [\diamond N P_i(x_i) \supset X^0_i(x_i)]. \quad (i=1, 2, \dots)$$

ところが、名詞句の定義域というのは考えにくい。そこで、我々は、名詞句の定義域の代わりに、単にその名詞句を代表する名詞の定義域で語義番号の選択を行っていると予想される。補題1、仮定2(c)より、

$$\forall x [\diamond N P(x) \supset \diamond N(x)].$$

ではあるが、語義文においては、

$$\forall x_i [\diamond N_i(x_i) \supset X^0_i(x_i)]. \quad (i=1, 2, \dots)$$

が成立していると仮定する。ただし、 $X^0_i(x_i)$ は語義文の動詞の x_i に関する定義域である。これと(6)より次式を得る。

$$\forall x_i [N_i(x_i) \supset X^0_i(x_i)]. \quad (i=1, 2, \dots) \quad (7)$$

従って、語義文中の動詞のある語義番号の選択が妥当であるための必要条件の一つは、その語義番号における格パターンが語義文中でそれに係る格を含み、その格に対応する変数の定義域とその動詞に係る名詞が(7)の関係を満たすことである。

また、定義3が成立するためには、補題1より、

$$\forall x_i [X^m_i(x_i) \supset X^s_i(x_i)]. \quad (i=1, 2, \dots) \quad (8)$$

が成立しなければならない。ただし、 $X^s_i(x_i)$ は語義文の x_i に関する定義域である。補題1、仮定2(b)(c)より、語義文から、連用修飾句、連体修飾部を取り除いた文の x_i に関する定義域 $X^{s'}_i(x_i)$ について次式が成り立つ。

$$\forall x_i [X^s_i(x_i) \supset X^{s'}_i(x_i)]. \quad (9)$$

また、(3)より次式が成り立つ。

$$X^{s'}_i(x_i) \equiv \diamond \exists (x-x_i) \exists z \\ [N_z(x) \wedge N_z(z) \wedge V^0(x, z)] \\ \supset \diamond \exists (x-x_i) N_z(x) \wedge \diamond \exists z N_z(z) \\ \wedge \diamond \exists (x-x_i) \exists z V^0(x, z) \\ \supset \diamond N_{x_i}(x_i) \wedge \diamond \exists (x-x_i) \exists z V^0(x, z) \\ \equiv N_{x_i}(x_i) \wedge X^0_i(x_i)$$

ただし、 $X^0_i(x_i)$ は、語義文中の述語の動詞の x_i に関する定義域である。よって、上式と(8),(9)より、語義文中の動詞の選択と、見出し語との変数の対応が妥当であるための必要条件の一つは、次式を満たすことである。

$$\forall x_i [X^m_i(x_i) \supset N_{x_i}(x_i) \wedge X^0_i(x_i)]. \\ (i=1, 2, \dots) \quad (10)$$

この二つの必要条件だけではまだ、語義番号の選択、見出し語との変数の対応に関して、曖昧さがかなり残るので、さらに次の仮定を導入する。

【仮定4】

見出し語と語義文の変数の対応について、

- (a)見出し語にあって語義文にない変数(定義3では y に相当)の数が少ない方が妥当な対応である。
- (b)語義文にあって見出し語にない変数(定義3では z に相当)の数が少ない方が妥当な対応である。

以上より、語義文中の動詞の妥当な語義番号の選択、見出し語との変数の妥当な対応を求めるには次のように行なう。

【アルゴリズム】 語義番号の選択、変数の対応の決定

まず、語義文の述語である動詞について、次のような可能な語義番号をすべて求める。すなわち、その語義番号における格パターンが語義文中でそれに係る格を含み、その格に対応する変数の定義域とその動詞に係る名詞が(7)の関係を満たす。

次に、上述の可能な語義番号について、(10)を満たすすべての語義番号の選択、変数の対応を求め、そのうち、見出し語にあって、語義文にない変数の数が最少の語義番号の選択と変数の対応を求める。

曖昧さがあれば、さらにそのうちから、語義文にあって、見出し語にない変数の数が最小の語義番号の選択と変数の対応を求める。

つぎに文 s が、複文、重文を含む一般的な場合を考える。この場合も基本的には単文の場合と同じである。異なるのは、定義域に関する条件で、(10)の代わりに以下で述べる(14)を用いることである。

語義文が複文を含む場合、補題1、仮定2(a)より次の関係が成り立つ。

$$\forall x_i: [X^{s_i}(x_i) \supset X^{s'_i}(x_i)]. \quad (i=1,2,\dots) \quad (11)$$

ただし、 $X^{s_i}(x_i)$ 、 $X^{s'_i}(x_i)$ はそれぞれ、語義文の x_i に関する定義域と、その主節のみからなる文の x_i に関する定義域である。

語義文が連言的重文または選言的重文の場合、その定義域について、それぞれ次式が成り立つ。

$$X^{s_i}(x_i) = X^{s^1_i}(x_i) \wedge \dots \wedge X^{s^n_i}(x_i)]. \quad (i=1,2,\dots) \quad (12)$$

$$X^{s_i}(x_i) = X^{s^1_i}(x_i) \vee \dots \vee X^{s^n_i}(x_i)]. \quad (i=1,2,\dots) \quad (13)$$

ただし、 $X^{s^j_i}(x_i)$ は語義文を個々の文に分解して得られる文の x_i に関する定義域である。(11)、(12)、(13)を用いて、

$$\forall x_i: [X^{s_i}(x_i) \supset X^{s''_i}(x_i)]. \quad (i=1,2,\dots)$$

なる $X^{s''_i}(x_i)$ を求めることができる。

ただし、 $X^{s''_i}(x_i)$ は語義文を構成する名詞、動詞の論理式への翻訳の定義域で表わされている。よって、(10)の代わりに、

$$\forall x_i: [X^m_i(x_i) \supset X^{s''_i}(x_i)]. \quad (14)$$

を定義3が成り立つための必要条件の一つとする。

5. IPALについて

情報処理振興事業協会技術センターから計算機用日本語基本動詞辞書(IPAL)⁽¹⁾が公開された。これには、前節で述べた定義域、格パターンに相当する記述がある。本研究では、国語辞典としてIPALを用いる。この節では、IPALについて簡単に説明する。

IPALでは日本語の動詞のうちで語彙体系上ならびに使用頻度上重要であると考えられる基本的な和語動詞861語を取り上げ、それを意味及び統語的特徴に基づいて下位区分し(サブエントリと呼ぶ)、それを一つの単位として、意味、形態、統語、文法的カテゴリ、慣用表現などに関する情報を記述してある。記述例を表1に示す。

《見出し》あいする
 《通番》001
 《No》001
 《枚数》002
 《意味》誰かや何かに対して、とても好きだという気持ちを持つ。
 《上位》思う
 《類義》愛おしむ、恋する
 《反義》嫌う
 《分名1》対人感情
 《文番1》2.302

《文名2》愛, 恋愛
 《文番2》481a, 482
 《意味分類》動作(動き), 生理・心理
 《表記》愛(あい)する
 《活用》サ変
 《語幹》ais
 《異音同語》あいす
 《自他》他
 《派可》あいせる
 《派使》あいさす
 《転名》あい
 《文型》N1ガ N2ヲ
 《文1》彼は 妻と子を 愛している。
 《述語素》02a, A
 《格1》ガ
 《素1》HUM
 《名1》彼
 《格2》ヲ
 《素2》HUM/ANI/ORG
 《名2》妻, 子, 主人/犬/家族, 日本, 会社
 《ヴ1》二使役
 《ヴ2》直受, 間受, 尊敬
 《ヴ3》ヲ
 《ヴ4》ニ
 《ヴ7》能動
 《テ1》未来
 《テ2》進行
 《テ3》テイク, テクル, テシマウ, カケル, ハジメル, ダス, ツヅケル
 《ム1》命令
 《ム2》意志・勧誘
 《ム3》タイ・タガル, ナサイ, ナ(禁止)
 《ム4》3a
 《マナ》マス, ナイ

表1. 辞書内容の例

この内で、《見出し》、《通番》(同音異義語を区別するための3桁の番号)、《No》(当該見出しの何番目のサブエントリかを示す3桁の番号)、《枚数》(当該見出しのサブエントリ数を示す3桁の番号)、《意味》(人間向けにサブエントリの意味を分かりやすく説明したもの)、《格i》(《文型》の中でi番目に出現する格形式)、《素i》(《文型》の中でi番目に出現する名詞の意味素性)を用いる。名詞の意味素性がその格スロットに対応する変数の定義域に対応していると考えられる。意味素性を表2に示す。

略号	素性名	例
CON	具体名詞	
ANI	動物	犬・馬・鳥・猿・コアラ
HUM	人間	姉・先生・男性・学生
ORG	組織・機関	国・企業・警察・研究所
PLA	植物	花・桜・松・バラ
PAR	生物の部分	頭・足・腕・腰・根・羽
NAT	自然物	山・空・石・川・丘
PRO	生産物・道具	紙・車・パン・布・鉄
PHE	現象名詞 (自然/生理)	光・音・火・風・雨・涙 声・匂い・病気・しわ
ABS	抽象名詞	
ACT	動作・作用	勉強・練習・見学・散歩
MEN	精神	心・意識・思い出・悩み
LIN	言語作品	名前・ニュース・説教
CHA	性質	美・欠点・みかけ・寛容
REL	関係	縁・原因・条件・根拠
LOC	空間・方向	外・公園・東・右
TIM	時間	昨日・日曜日・夕方
QUA	数量	3日・5kg・4m・全部
DIV	(制限緩やか)	

表2. 意味素性 (文献(1)より)

但し、次の包含関係が成り立つ。

$CON \subseteq DIV$, $PHE \subseteq DIV$, $ABS \subseteq DIV$,
 $ANI \subseteq CON$, ..., $PRO \subseteq CON$,
 $ACT \subseteq ABS$, ..., $QUA \subseteq ABS$

6. 抽出例

IPALによると、「愛するI」の辞書記述は次のようになっている。

愛する I 誰かや何かに対して、とても好きだという気持ちを持つ。

「彼は妻子を愛している。」

ガ：HUM

格パターン ヲ：HUM/ANI/ORG

語義文中の名詞、「誰か」、「何か」、「気持ち」の示す集合はそれぞれ、HUM, CON U ABS, MENとする。

語義文中の動詞「持つ」の語義番号のうち、格パターンの中に「二格」と「ヲ格」を含む語義番号とその記述は、

持つ I 手に何かを取る。

「彼は右手に筆を持った。」

ガ：HUM/ANI

格パターン ニ：PAR

ヲ：CON

III 何かを所有する。

「彼は田舎に田畑を持っている。」

ガ：HUM

格パターン ニ：LOC

ヲ：CON/ABS

V 何かに対して、ある感情を抱く。

「彼は相手のやり方に反感を持った。」

ガ：HUM/ORG

格パターン ニ：CON/ABS

ヲ：MEN

このうち、「二格」の定義域がCON U ABSを含み、「ヲ格」の定義域がMENを含む語義番号はVのみである。これに対して、「愛する」にあって、「持つ」にない変数がない対応は、

①{(1 1),(2 2)} ②{(1 1),(2 3)}

③{(1 2),(2 1)} ④{(1 2),(2 3)}

⑤{(1 3),(2 1)} ⑥{(1 3),(2 2)}

である。ただし、例えば⑥は、「愛するI」の「ガ格」と「持つV」の「ヲ格」、「愛するI」の「ヲ格」と「持つV」の「二格」が対応することを示す。このうち定義域に関する必要条件を満たしているものは①のみである。

「愛するI」の語義文は連体修飾部を持つ単文なので、動詞の上位-下位関係抽出アルゴリズムにしたがって、次の上位-下位関係を得る。

$\forall x_1 \forall x_2 \exists z$

[\$愛する^I(x_1, x_2) \supset \$持つ^V(x_1, x_2, z)] .

6. あとがき

述語間の上位-下位関係を定義し、これに基づいて国語辞典の見出し語(動詞)とその語義文の関係を定義した。語義文に対応する論理式を持つ諸性質を仮定して、見出し語とその語義文から動詞間の上位-下位関係が求まることを示した。その際に問題となる語義文中の動詞の語義番号の選択、見出し語との変数の対応をどの様にして求めるかを示した。今後、これに基づき、計算機による実験を進めていく予定である。

参考文献

- (1) 計算機用日本語基本動詞辞書IPAL (Basic Verbs) 説明書 情報処理進行事業協会技術センター
- (2) 富浦, 日高, 吉田: 「語義文からの動詞の階層関係の抽出」 第40回電気関係学会九州支部大会連合大会講演論文集