

計算機用日本語生成辞書IPAL (SURFACE/DEEP)の枠組み

村田賢一 岡部了也 井口厚夫 後藤恒男
帝京平成大学 上智大学 獨協大学 ティー・エフ・シー

計算機による日本語生成のための辞書IPAL (SURFACE/DEEP)の枠組みを述べている。

The framework for the lexicon IPAL (SURFACE/DEEP)

Ken'ichi MURATA Ryoya OKABE Atsuo IGUCHI
Teikyo-Heisei University Sophia University Dokkyo University
Tsuneo GOTO
TFC

The framework for the lexicon IPAL (SURFACE/DEEP) is described.

1 はじめに

何らかの人工知能ソフトウェアで作られ出した「意味」を、なるべく自然な日本語文にすること、特に、そのための辞書をいかにして作るかをとりあげる。

「意味」の表現形式は古典的なフレーム形式に近いもの、また意味論的体系としては述語論理のようなもの(正確にはモンタギュー文法や状況意味論のようなフォーマルなものではなく、むしろインフォーマル・ロジックとでもいうべきもの)を考える。

「良い辞書をつくる」という目的からすると、すべての意味がアトミックだというのでは困るので、適当な構造、特に、異表現間の等価性と半順序関係を導入したい。

述語論理の基本的構成要素は述語と項であり、項は基本的には個体であるけれども、実際の個体表現の多くは、量化された普通名詞(ときには、量化詞が省略された裸の名詞)であるので、固有名詞だけでなく普通名詞も視野に入れた枠組みをつくる必要がある。

モデル理論として想定するのはナイーブな集合論で、集合の要素はすべて個体だ、と思うことにする。

一つの部分集合は一つの種(kind)という個体をあらわす。実際には物質名詞も扱うので、問題はいろいろあるが、そのところには立ち入らないことに

する。

述語の構造化にはKarttunen and Peters (1979)の分析にヒントを得て、基本的意味と付加的意思の分解を行った。付加的意思は一般にはpresuppositionと云われるタイプのものであるが、実は、それより項の変域の制限という形で現れることの方が多い。

普通名詞の構造化には古典的な成分分析の手法を用いた。

次に、一つの「意味」に対し、複数の(ほとんど無限といってもいい数の)日本語表現が対応し、人間が種々の修辭学的動機でそれらの表現を(多分に苦勞して)使い分けている現象は、文生成に対する制御機能としてとらえる。

具体的には文型や文体(いわゆる待遇表現も含めて)を指定する種々の制御情報を文生成に介在させることにする。

なお、埋め込み文や並列文のような複文を生成することも制御機能として扱う。

この制御機能は辞書においては、パラメータとして扱われる。というのは、生成辞書は関数形式で与えられるからである。

実際の文生成は、意味ネットワークの形で与えられた深層意味構造が与えられると、まずIPAL (SURFACE/DEEP)で、文生成に用いる単位表現を引き、次に従来型のIPAL (IPAL (BASIC VERBS))を引きその表現の文法情報を求め(このあたりは和文英訳

本研究は、情報処理振興事業協会(IPA)「独創的技術育成事業」の一環として行われたものである。

に際し、まず和英辞書を引いて英語表現を選び、次に和英辞書で英語の文法情報を調べる人に似ている)、最後にテンスや接続詞注入といった整形作業によって文を完成する。

なお、人間が種々な修辞学的な要素を操作する過程は、まさにAIの課題そのものであり、それはこの枠組の外側に位置づけられるものではあるが、全体として深い関係があるので、あわせて研究しており、今回は別タイトルの論文(「IPAL (SURFACE/DEEP)による自然な日本語文生成」)で論じている。なお、辞書試作の経験については、同じく別タイトルの論文(「計算機用日本語生成辞書IPAL (SURFACE/DEEP)の試作」)で述べている。

2 述語フレームと述語生成辞書

2.1 深層意味構造1: 述語フレーム

計算機用日本語生成辞書IPAL (SURFACE/DEEP)の目的は、言いたい内容、つまり深層意味構造を表層表現に変換するための辞書を提供することであるが、この「深層意味構造」には(インフォーマルな)ロジックが用いられ、フレームと呼ぶもので表示される。フレームについては以下で詳説するが、フレームを表層文に変換する操作は、概ね下図の様なものである。この例では単純化のため、固有

| | |
|----------------|-----------|
| 0001 | 貸借行為 |
| 貸し手 | “次郎” |
| 借り手 | “太郎” |
| 貸借物件 | “「否定の意味」” |
| 副詞的意味成分: “昨日” | |
| 夕形,x2ガ格能動,x2ハ格 | |

→ 昨日、次郎は太郎に「否定の意味」を貸した

図1: フレームの変換例

名をリテラルで使用する場合を示した。以下、リテラルは引用符“ ”で囲んで示す。

この節では、まず文に対応する述語フレーム、およびそれを表層表現に変換するための辞書である述語生成辞書の仕組みを述べる。述語フレームは、フレーム同定子、基本部、修飾部、制御部の四つの部分に分かれており、これらを順をおって説明する。

2.2 フレーム同定子と基本部

フレーム同定子と基本部は、論理学の用語で言えば、述語とその項に相当し、これがその述語フレーム

の中核となる命題を表わす。例えば、貸す(次郎, 太郎, 「否定の意味」)では「貸す」がフレーム同定子、(次郎, 太郎, 「否定の意味」)が基本部に対応する。述語フレームと論理式の関係は2.5.1節で述べる。

論理学では通常、述語として対応する自然言語の表現をそのまま用いることが多く、上では「貸す」を述語に用いたが、我々の辞書の目的が一つの深層意味構造に対し、様々な表層表現を与えることから、対応する表層表現とは異なった中立的、あるいはより一般的な述語名を用いる方が好ましい。例えば、次の二文は同一の事象を表現したものであり、貸す(次郎, 太郎, 「否定の意味」)は(1a)に写像されたり、(1b)に写像されたりする。

- (1) a. 太郎が次郎に「否定の意味」を借りた。
b. 次郎が太郎に「否定の意味」を貸した。

そこで、深層意味構造においてはより中立的な「貸借行為」という名前を用いる。従って、我々の辞書においては世の中の種々の事象に対しそれを表わす述語名を付けることとなる。例えば、「授受行為」「飲食行為」「消失行為」などである。これらの述語名を整理することや、コンピュータで処理する際の便宜性を考え、番号をふっておく方が都合がよい。以下では、述語名に対し適当に番号がふられているものとして説明する。

貸借行為(次郎, 太郎, 「否定の意味」)という表示は誰が貸した側で、誰が借りた側であるのかが明示的ではない。そこで、状況意味論、特にDevlin (1991)の方法を用い、役割名を明示するものとする。「貸借行為」に参加するものが「貸し手」「借り手」「貸借物件」の三者であると定義されていたとすると、この命題の我々の枠組みにおける表示は次のようになる。

(2)

| | |
|------|-----------|
| 0001 | 貸借行為 |
| 貸し手 | “次郎” |
| 借り手 | “太郎” |
| 貸借物件 | “「否定の意味」” |

第一行目は、述語名とそれに対応する番号である。以下、この番号のことをフレーム番号、述語名を(述語)フレーム名と呼ぶこととする。フレーム番号とフレーム名の対をフレーム同定子と呼ぶ。下の三行は、各々役割名とその役割を果たす参与者で構成される。以下、役割名を引数同定子、対応する参与者を引数値と呼ぶ。引数同定子と対応する引数値の対の集合が述語フレームの基本部である。つまり、この例では下三行が基本部となる。

引数同定子により役割名が明示されているので、理論上は基本部における各対はどの順序で書いても良いこととなるが、辞書執筆上の便宜等からこの順序は固定されてるものとする。

どの引数同定子が可能であるか、及びその引数値がどのような種類であるか、そしてそれらの間の順序はそれぞれの述語に対して定義されていないならぬ。例えば、「飲食行為」に対しては辞書において次の様な定義が記載される。

(3) 0103 飲食行為

第一項 飲食者 (変域: 人)

第二項 飲食物 (変域: 固体・液体)

これは飲食行為のフレームの一つ目の引数同定子が「飲食者」であり、その引数値の変域が人であることなどが表わされている。ここで、第二引数値の変域で用いた・は和集合を表わすものとする。

2.3 修飾部

修飾部は、基本部で表わされる中心的命題にさらに実質的な意味を付け加えるための部分である。これは、付加的意味成分と副詞的意味成分とに分けられる。付加的意味成分と副詞的意味成分の意味論的な違いはないが、前者は語彙化され一つの語として表わされるのに対し、後者は付け加えられる内容が副詞句で表わされる。図示する際は基本部の下に書く。

2.3.1 付加的意味成分

引数値の変域の制限: 付加的意味成分の一つ目の例は、引数値の変域の制限である。「飲食行為」の定義は(3)であるが、この定義では、飲食物は固体でも液体でも良い。これに対し、「飲む」は飲食行為の一種であるが、第二引数値の飲食物は液体に限られる。この情報は、付加的意味に液体(飲食物)の様に書かれる。これは「飲食物は液体である」という命題を表わす。この情報により図2のフレームは「太郎がコーヒーを飲む」に変換される。この変域制限がない場合は「食べる」ではなく「摂取する」の様に固体に対しても液体に対しても使える述語に変換される。

様態等: 引数値の変域の制限以外にも、述語に意味が付加されることがある。例えば、「掻き込む」は飲食行為であるが、さらに「急いで」という意味合いが加わる。この意味は「掻き込む」という一単語の中に含まれているので、付加的意味という形で書かれる。

| | |
|---------------|------|
| 0103 | 飲食行為 |
| 飲食者 | “太郎” |
| 飲食物 | コーヒー |
| 付加的意味:液体(飲食物) | |

→ 太郎がコーヒーを飲む

図 2: 項の変域制限のある例

| | |
|---------------|------|
| 0103 | 飲食行為 |
| 飲食者 | “太郎” |
| 飲食物 | ご飯 |
| 付加的意味:固体(飲食物) | |
| 付加的意味:急いで | |

→ 太郎がご飯を掻き込む

2.3.2 副詞的意味成分

「掻き込む」を上のように考えると、「掻き込む」とは「急いで食べる」とほぼ同義となる。ただし、「急いで」という意味合いは、前者では単語内に内包されるのに対し、後者では独立した副詞句として明示されている。この様に独立した副詞句として表現したい場合に用いられるのが副詞的意味成分である。

| | |
|---------------|------|
| 0103 | 飲食行為 |
| 飲食者 | “太郎” |
| 飲食物 | ご飯 |
| 付加的意味:固体(飲食物) | |
| 副詞的意味:急いで | |

→ 太郎がご飯を急いで食べる

図 3: 副詞的意味成分

2.4 制御部

出力される表層表現を文体・文脈に沿った語形・語順に整えるための情報が制御部である。この情報にはテンス・アスペクト情報を含むので、「食べる」に対し「食べた」、「食べている」など厳密には意味が変わる場合がある。制御部に書かれるパラメータには次の様なものがある。

文体論的情報: 「食べる」「食う」「召し上がる」など基本的意味は同じで使用される文脈の異なる表現を区別するためのパラメータである。これには、「ぞんざい」「口語的」「事務的」などの文体論的なものと、「尊敬」「謙譲」などの待遇に関わるものがある。
文法的情報: 「貸す」と「借りる」は表わしている事態は同一であるが、何を主語とするかで異なっている。「貸す」は貸し手が主語であるが、「借りる」は借り手が主語となる。この違いを表わすためのパラメータが文法情報である。第一項を主語にする場合

は「 x_1 ガ格」、第二項を主語にする場合は「 x_2 ガ格」の様に書く。ある項を主語にして述べる場合、「借りる」「買う」の様に単一語の能動形を使う場合と「貸される」「売られる」の様に受動形にする場合とがある。これらを区別するために、能動形か受動形かという情報も付記する必要がある。能動形は「ル形」、受動形は「ラレ形」とし、「 x_2 ガ格ル形」「 x_3 ガ格ラレ形」の様に書く。

テンス情報とアスペクト情報: ここでのテンス及びアスペクト情報は、主に述語の語形を指定するために用いられ、過去の場合は「タ形」、進行相などは「テイル形」という形で指定する。「タ形」「テイル形」は、実際には上の文体的情報や文法的情報と違い、意味を付け加えるが、「タ形」は必ずしも過去とは限らないし、「テイル形」も進行相の場合や完了相の場合など述語により様々である。この様な意味的な違いはとりあえず扱わない。

ムード情報: ムードは、中核的な命題内容に対する話者の態度を表し、「肯定」「否定」に関する情報と「断定」「疑問」「推量」「命令」に関する情報、および丁寧さをあらわす「マス形」等の組み合わせで表す。最初の二つに関してはそれぞれ「肯定」と「断定」をデフォルトとする。

談話情報: 文は一般にトピックの部分とコメントの部分に分けられるが、この様な談話に関わる情報を与えるのが談話パラメータである。ここでは特に何がトピックであるかについての情報を取り扱う。トピックはよく係り助詞「は」を用いて表わされるので、「 x_2 ハ格」の様に何をハ格で表わすかという形で表わす。

重文・複文情報: 複雑な文を生成しようとする、いわゆる重文や複文を扱う必要が生ずる。重文や複文では述語が複数現れるが、これは述語フレーム内での別の述語フレームを参照するという形で実現する。

2.5 述語生成辞書

ここまでで文に対する深層意味構造である述語フレームについて説明してきた。述語生成辞書には、この述語フレームを表層化に変換する手助けとなる情報が記載されるが、これは表層化関数 f を用いて書かれる。表層化関数 f は述語フレームを表層表現に写像する関数である。辞書に記載する内容は述語フレームの具体的な引数値の値とは無関係なので、引数値を x_1, x_2, x_3 の様に変項で表わすと、次の様に

なる。この例は、飲食行為でありかつ飲食物が固体

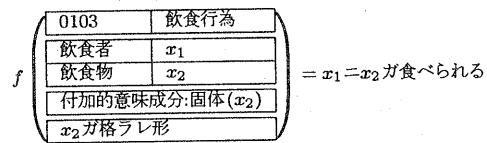


図 4: 表層化関数

であることを飲食物を主格にした受動形で述べる時には「食べられる」が用いられ、飲食物には格助詞「が」が、飲食者には格助詞「に」が用いられることを示している。実際の辞書記述の際には次の様に一行に縮めた表記を用いる。

- (4) $f_{0103,0}(x_1, x_2 : \text{固体}(x_2), : x_2 \text{ガ格ラレ形})$
 $= \text{ニ, ガ食べられる}$

f の添え字のうちの 0103 は対象としてのフレームが飲食行為であることを表わし、0 は出力される表現が述語であることを示す。 f の引数は「:」により三つに仕切られているが、これらは順番に基本部、修飾部、制御部に相当する。修飾部はさらに「;」で区切られるが、「;」の左が付加的意味成分、右が副詞の意味成分を書くための場所である。各引数値に対し用いられる格助詞は、引数値と同じ順序でコンマ「;」で区切って書くこととする。

2.5.1 述語シソーラス

述語フレームの論理式への変換: ここで述語フレームを論理式に変換する方法を説明する。述語フレーム名を π 、引数値を x_1, \dots, x_n 、付加的意味成分を $\alpha_1, \dots, \alpha_m$ 、副詞の意味成分を μ_1, \dots, μ_l とすると、この述語フレームは以下の様なこれらの命題の連言として解釈する。制御部のテンス・アスペクト情報、重文・複文情報も論理的意味に関与するが、ここでは単純化のため制御部の情報は無視することとする。

$$(5) \pi(x_1, \dots, x_n) \wedge \alpha_1 \wedge \dots \wedge \alpha_m \wedge \dots \wedge \mu_1 \wedge \mu_l$$

$$(6) \text{飲食行為}(x_1, x_2) \wedge \text{固体}(x_2) \wedge \text{急いで}$$

例えば、「掻き込む」に対する述語フレームは(6)の様解釈される。

述語シソーラス: 述語辞書にはテンス・アスペクト情報、重文・複文情報は記載されないし、他の制御情報は論理的な意味を変えないので、上で定義した述語フレームの論理式としての解釈を用いて表現間の

等価性および含意関係を定義できる。例えば、「飲食する」「召し上がる」「頂く」は尊敬、謙譲の意味合いにおいては異なっているが、論理的意味は(5)により飲食行為(x_1, x_2)であり、同義表現と見なすことができる。また、先の例を使えば「掻き込む」と「急いで食べる」に対応する述語フレームの論理式は共に(6)であり、これらも同義となる。

また、論理式の含意関係を用いて、この様に定義された同義表現の集合間の上下関係も規定できる。例えば、

(7) 飲食行為(x_1, x_2) \wedge 固体(x_2) \wedge 急いで \rightarrow 飲食行為(x_1, x_2) \wedge 固体(x_2)

の含意関係より、「食べる」は「掻き込む」「急いで食べる」の上位表現であると判定される。同様に、「飲食する」「召し上がる」「頂く」は「食べる」や「掻き込む」「急いで食べる」の上位表現である。

2.5.2 表層化関数及び述語フレームの異表記

表層化関数の異表記: (4)の例では、固体(x_2)という引数値の変域の制限情報を付加的意味のところに書いたが、変域の制限が何番目の引数値に対するものであるかを視覚的に見やすくすることや、変域の制限が複数ある場合の書き易さを考えると、この情報は引数値のところに直接書く方が便利であり、辞書の記載にはこの便法を用いる。(4)で x_1, x_2 と書いていた引数値のところには、代わりに対応する変域を書くこととする。今の例の場合、 x_1 の変域は(3)で「人」と定義されているので、この便法を用いると(4)は次の様に書き換えられる。

(8) $f_{0103,0}(\text{人}, \text{固体}; : x_2 \text{ガ格ラ形}) = \text{ニ}, \text{ガ食べられる}$

別フレームの設定: 「食べる」「掻き込む」など変域の制限に限らず付加的意味成分がある場合は、付加的意味のないより一般的な語(「摂取する」とは異なる一語で表わされる。したがって、これらを元の述語フレームと(関連はあるが)異なるフレームとみなすこともできる。例えば、固形物飲食行為、急速飲食行為の様なフレームを設定し、「食べる」は前者に対応する表現、「掻き込む」は後者に対応する表現とする。この様に新しいフレームを設定することで、コンピュータ内での処理などの場合に付加的意味成分を無視することができる。ただし、新しく設

定したフレームと元のフレームの間の同値関係を辞書内で定義する必要がある。

3 名詞句フレームと名詞生成辞書

3.1 深層意味構造 2:名詞句フレーム

これまでの例では、述語フレームの引数値には表現そのものを入れて来た。固有名詞に関してはこの方法である程度扱えるかも知れないが、普通名詞の場合は形容詞句や関係節などの修飾表現を付け加えたり、文体に応じてぞんざいな表現、丁寧な表現などに変える必要があり、述語と同様に名詞句の深層構造として名詞句フレームを用いる。名詞句フレームも述語フレームと同様にフレーム同定子、基本部、修飾部、制御部の四つの部分に分けられる。

名詞句フレームは述語フレームから参照されるという形で用いられる。例えば以下のフレームは述語フレームの第一引数値 x_1 が名詞句フレーム「人」を参照することを表わしている。これにより、貸し手

| | | | |
|----------------|---------|------|---|
| 0001 | 貸借行為 | 0011 | 人 |
| 貸し手 | x_1 | 性別 | — |
| 借り手 | “太郎” | 年齢 | 老 |
| 貸借物件 | “否定の意味” | 量化:中 | |
| タ形, x_2 ガ格ル形 | | | |

→ 太郎がその老人に「否定の意味を」借りた

x_1 は「その老人に」という表現に変換される。

以下に名詞句フレームの仕組みを説明する。

3.2 フレーム同定子と基本部

フレーム同定子は述語フレームの時と同様にフレーム番号とフレーム名の対とする。フレーム名には「人」「器」などある種類を表わす名前のうち一般的なものをを用いる。

述語フレームは事態を表わすので、基本部にはその事態の参与者に関する情報が記述されるが、名詞句フレームは個体(の集合)を指すので述語フレームとは異なり、いわゆる成分分析的な手法を取り、その個体を区分するための基本的な特徴を記述する。これは基本属性同定子とその属性値の対で書かれる。例えば、「人」を区分するための基本属性としては「性別」や「年齢」があり、「器」を区分するための基本属性には「対象物」「形状」「サイズ」「材質」などがある。例えば、「人」であり「女」で「老」いたもの

は「おばあさん」と表現され、「ご飯」を盛るための「深さ」のある「大」きい「陶器製」の「器」は「どんぶり」と表現される。

| | |
|------|----|
| 0024 | 器 |
| 対象物 | 飯 |
| 形状 | 深 |
| サイズ | 大 |
| 材質 | 陶器 |
| | |
| | |

～どんぶり

基本属性の無指定: 基本属性は常に指定される必要があるとは限らない。例えば、「男」という表現は、典型的には大人であるが、子供を指してもかまわないので、年齢を指定する必要がない。この様にある基本属性に対し値を指定しない場合、___を用いることとする。また、すべての基本属性値を無指定にすることも可能で、例えば「性別」「年齢」をともに無指定にすればより一般的な意味の「人」が得られる。

基本属性の種類及び順序は、述語辞書の引数の場合と同様に次の様に辞書で定義される。

(9) 0011 器

基本属性 1 年齢

基本属性 2 性別

基本属性は成分分析における成分に相当するもので、それらの組み合わせで一つのグループ内の語彙分布が特徴づけられるという考えに基づいている。したがって、どの様な基本属性を設定すべきかの一応の基準は、その意味が語彙化されるか否かである。ただし、どの範囲の語を記述するかによって設定される基本属性の種類が変わる可能性もあり、また指定されることの少ない基本属性があったりと、どの基本属性が有効なのかを判断するのはきわめて難しい。

3.3 修飾部

修飾部に記載される内容には、付加属性、名詞句修飾情報、量化情報の三つがある。

3.3.1 付加属性

名詞句フレームの修飾部に記述される情報の一つ目が、付加属性である。付加属性は、述語フレームでいえば副詞的意味成分に相当するもので、付加属性の表わす意味は、形容詞句として表現される。

付加属性も基本属性と同様に成分分析の手法を取り、付加属性同定子と属性値の対で記述される。

(10)

| | |
|------|----|
| 0024 | 器 |
| 対象 | 飯 |
| 形状 | 深 |
| サイズ | 大 |
| 材質 | 陶器 |
| サイズ | 小 |
| | |
| | |

～小さなどんぶり

基本属性と付加属性の違いは、基本属性が一つの語彙グループを分類するための属性であり、語彙化されるのに対し、付加属性は独立した形容詞句として表現されることである。さらに、基本属性は段階性の無い「分類」属性であるのに対し、付加属性は段階性のある属性と考える。例えば、上の例では基本属性と付加属性の両方にサイズの指定があるが、基本属性における指定の「大」は「飯を盛るための深型の器」の中で大きい部類のものの意で、ある特定の大きさを指しているのに対し、付加属性における「小」は他のものと比較した場合の大きさを述べている。

また、付加属性の種類・数は基本属性と違い不定とする。

3.3.2 名詞句修飾情報

名詞句を修飾する方法は形容詞句だけではなくいわゆる関係節がある。このための情報が名詞句修飾情報である。関係節は節、つまり文であるので、名詞句修飾情報はこの文を表わす述語フレームへの参照という形で表わされる。また、関係節には意味的に欠けた部分があり、修飾されるものがその欠けた部分の役割を果たすの普通であるので、この意味的に欠けた役割を表わすために名詞句修飾情報で参照される述語フレームは引数値の一つが*で埋められているものとする。(図5参照)

3.3.3 量化情報

量化情報は、英語のa(n), the, this, that, every等に相当するいわゆる限定辞表現を生成するための情報である。

日本語では、限定辞なしの「裸の名詞句」で用いられた場合に、定名詞句と解釈されることもあるし、不定名詞句と解釈されることもある。この意味的な区別を表わすために、「量化情報」に「定」あるいは

| | |
|----------|----|
| 0024 | 器 |
| 対象物 | 飯 |
| 形状 | 深 |
| サイズ | 大 |
| 材質 | 陶器 |
| 名詞句修飾情報: | |
| | |

| | |
|------|-------------------------|
| 0001 | 貸借行為 |
| 貸し手 | “次郎” |
| 借り手 | “太郎” |
| 貸借物 | * |
| | |
| | タ形, x ₂ ガ格ル形 |

→ 太郎が次郎に借りたどんぶり

図 5: 名詞句修飾情報

「不定」と書くこととする。ただし、この二つの情報は表層表現には影響を与えない。

指示形容詞「この」「その」「あの」などの表現を生成するためには、これらが文法用語でそれぞれ近称、中称、遠称と呼ばれることより、「近」「中」「遠」と書く。指示形容詞がついた表現が実際に何を指示するかを決定するのは難しい問題であるので、ここでは意味的には上の「定」名詞句と同じ扱いとする。

量化情報に「全」と書かれてある時は、「全ての」という表現を生成するものとする。同じ全称量化の意味の表現として「どの男も」の様なものも考えられるが、これは形的には疑問形であることと、助詞「も」の指定が必要であることなどの理由で、現在の枠組みでは扱えない。

3.4 制御部

制御部では主に文体論的情報を扱う。これは、述語フレームの場合と同様に文脈や状況に即した表現を得るための情報で、「ぞんざい」「事務的」など文体に関する指定をする。例えば、「男」を表すフレームに「ぞんざい」という指定が加わると「野郎」に、「事務的」が加わると「男性」へと変換される。

3.5 名詞句生成辞書

名詞句生成辞書も表層化関数 f を用いて記述され、述語辞書の場合と同様に次の様な一行に縮小した形で記載される。

$$(11) f_{0024}(\text{飯, 深, 大, 陶器} : \text{小}) = \text{小さいどんぶり}$$

f の添え字は、名詞句フレーム番号で、これにより「器」のフレームをベースにしていることが分かる。また、 f の引数は述語辞書と同様に「:」により基本部、修飾部、制御部に分けられている。

基本部は基本属性が辞書に定義された順序で書かれ、指定する必要のないところは $_$ を書く。辞書で

は名詞修飾情報、量化情報を記載する必要がないので、修飾部には付加属性のみが書かれる。

3.6 名詞句シソーラス

名詞句フレームの述語としての解釈: 論理学では普通名詞は一項述語として解釈される。ここでは、名詞句フレームの量化情報を除いた部分を述語として解釈する方法を定義する。

名詞句フレーム名、基本属性値、付加属性値ともに一項述語と見ることができるので、名詞句フレームの論理式としての解釈はこれらの述語の連言と解釈し、 λ 表記すると次の様になる。

$$(12) \lambda x[\text{器}(x) \wedge \text{飯}(x) \wedge \text{深}(x) \wedge \text{大}(x) \wedge \text{陶器}(x) \wedge \text{小}(x)]$$

ただし、ここで $\text{飯}(x)$ は「 x は飯である」ではなく、「 x の対象物は飯である」という命題であるとする。しかし、この様に定義すると $\text{大}(x)$ と $\text{小}(x)$ という一見相反することが同時に述べられてしまう。

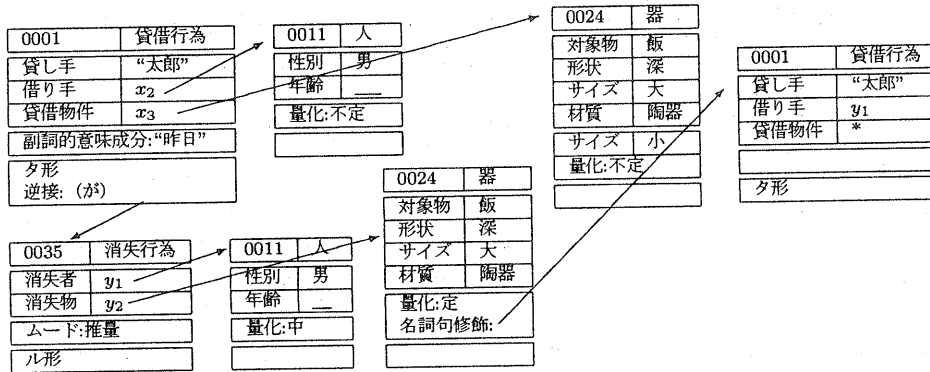
これは、これらの述語は何を基準にして大小を判断しているかが異なっているからであり、この基準を明確にするために添え字を付け、 $\text{大}_{c_1}(x)$ 、 $\text{小}_{c_2}(x)$ の様に記述し、 $c_1 \neq c_2$ であると仮定することで矛盾を回避できる。基本属性に対する基準は語彙的に固定されているのに対し、付加属性に対する基準は文脈で決まるものとする。例えば、今の例では c_1 は「飯を盛るための深型の器」に固定されるのに対し、 c_2 は文脈に依存する。

この様に基準となる集合が必要となる述語に同様の添え字を加えると、上の名詞句フレームの解釈は次の様になる。

$$(13) \lambda x[\text{器}(x) \wedge \text{飯}(x) \wedge \text{深}_{c_1}(x) \wedge \text{大}_{c_2}(x) \wedge \text{陶器}(x) \wedge \text{小}_{c_3}(x)]$$

c_1 、 c_2 は基本属性に対する基準で語彙的に固定されると述べたが、具体的にどの様な値になるかは辞書において記述される。

名詞句修飾情報も、参照する述語フレームから “*” でマークされた項を λ 抽象することにより一項述語と解釈することができる。 $\lambda x[\text{貸借行為}(x, \text{太郎}, \text{「否定の意味」})]$ 一般に、名詞句フレームの (述語的) 解釈はフレーム名、基本属性値、付加属性値、名詞句修飾情報が述語 π_0, \dots, π_n と解されるとき、 $\lambda x[\pi_0(x) \wedge \dots \wedge \pi_n(x)]$ であるとする。したがって、「太郎に本を貸した老人」は $\lambda x[\text{人}(x) \wedge \text{老}_{c_1}(x) \wedge \text{貸借行為}(x, \text{太郎}, \text{「否定の意味」})]$ ような解釈を得る。



→昨日、太郎は男に小さいどんぶりを貸したが、その男は太郎に借りたどんぶりをなくした。

図 6: フレームを用いた深層意味構造

名詞句シソーラス: 上の説明では、量化情報は無視された。これは、量化情報が文に対するオペレータとして解釈されるため、当該の名詞句フレームを参照する述語フレームまで考慮する必要があったからである。これは次の節で説明する。

名詞句辞書には量化情報及び名詞句修飾情報は記載されないので、上で述べた論理的解釈方を用いて同義関係、含意関係が定義できる。例えば、「人」「人間」はともに $\lambda x[\text{人}(x)]$ 、「男」「野郎」「男性」はいずれも $\lambda x[\text{人}(x) \wedge \text{男}(x)]$ と解釈され、それぞれが同義表現のグループをなすことが分かる。

また、一項述語は個体の集合として解釈できるので、同義表現類の間の上下関係に対応する解釈(集合)の包摂関係で捉えると、「人」「人間」は「男」「野郎」「男性」の上位表現、また、同様にして「老人」は「人」「人間」並びに「男」「野郎」「男性」の下位表現であることなどが明示できる。

4 フレームを用いた深層意味構造

以上、深層意味構造の基本となる述語フレームと名詞句フレームを説明したが、述語フレームは別の述語フレーム及び名詞句フレームを参照することができるし、名詞句フレームは名詞句修飾情報で述語フレームを参照することができる。したがって、フレームから別のフレームを参照することにより、図6の様な、いわばネットワーク状の複雑な深層意味構造を作ることができる。

量化情報の「定」は \forall 、「不定」は \exists 、「全」は \forall に解釈されるものとし、これらのオペレータで、すでにみた述語フレームの命題としての解釈、および名詞句フレームの述語としての解釈を結び付けること

によって、名詞句フレームを参照する述語フレームの重文・複文情報を除いた解釈ができる。オペレータのスコープ関係は適切に定義される必要があるが、ここでは扱わないこととする。重文・複文情報で結ばれている場合は、単純に連言として解釈すると図6の解釈は以下の様になる。

$$\exists x_2[\text{人}(x_2) \wedge \text{男}(x_2) \wedge \exists x_3[\text{器}(x_3) \wedge \text{飯}(x_3) \wedge \text{深}_{c_1}(x_3) \wedge \text{大}_{c_2}(x_3) \wedge \text{陶器}(x_3) \wedge \text{小}_{c_3}(x_3) \wedge \text{貸借行為}(\text{太郎}, x_2, x_3) \wedge \text{昨日}]] \wedge \exists y_1[\text{人}(y_1) \wedge \text{男}(y_1) \wedge \exists y_2[\text{器}(y_2) \wedge \text{飯}(y_2) \wedge \text{深}_{c_1}(y_2) \wedge \text{大}_{c_2}(y_2) \wedge \text{陶器}(y_2) \wedge \text{貸借行為}(\text{太郎}, y_1, y_2) \wedge \text{消失行為}(y_1, y_2)]]]$$

5 あとがき

従来のIPALプロジェクトにくらべて、今回は人員、期間、予算すべての点で大幅に規模を縮小されており、制約が大きい。

このため、副詞的表現など、枠組み上でもっと工夫改良が望まれるものも現在のところ残っている。しかし、重文(並列文など)や複文(埋め込み文)が可能な枠組みに立っており、十分に複雑な表現を取り扱うことができるようになってきていると考える。

参考文献

- Devlin, Keith. 1991. *Logic and Information*. Cambridge University Press.
- Karttunen, Lauri and Stanley Peters. 1979. "Conventional Implicature." Oh and Dinneen (eds.) *Syntax and Semantics vol. 11: Presupposition*. Academic Press, NY.