

L T B マスタ辞書の意味記述

田中 裕一
(ICOT)大島 資生
(東京都立大学)

長澤 陽子

1. はじめに

汎用日本語処理系LTBは、日本語の入出力を伴うさまざまな応用プログラムのインタフェースとして、入力文を解析系(形態素解析、構文解析)によりある意味表現に落とし、逆に、同じ形式の意味表現から生成系により日本語文を作り出す機能を提供している。解析系と生成系とは辞書と文法とを共有し、効率性と一貫性を実現している。

マスタ辞書はLTBの日本語文法に基づく文法情報と、語彙と意味構造との対応を与える意味情報から構成されている。ここでは、マスタ辞書にこのような意味情報を記述するための基本的な枠組みとその内容について、特に動詞を中心として述べる。

2. マスタ辞書の意味記述の概要

一般に、文や文脈(談話)の意味を理解し、それに基づいて何か「知的な」処理を行うシステムを構築するためには膨大な世界知識が必要になる。そのような知識をあらゆる分野にわたって収集し、蓄積することは事実上不可能である。そこで個々の専門分野に限って深い知識を記述したり、分野を区切らず、広い領域にわたって一般的な浅い知識を記述したりすることが行われている。

LTBの目的は汎用性にあるので、まず、広い知識を記述することが必要である。しかし、同時に、ICOTで現在開発を進めている談話理解実験システムDUALSのための知識ベースとしての役割も果たさなければならない。現在は、後者に力点を置いて、LTBとしての語彙は約4000語にとどめ、むしろDUALSの評価用テキストに現れる分野に関する深い知識を記述しつつある。

語の意味を記述するにあたって、3つのレベルを想定した。第1は構文的な語の役割か

ら、機械的に語の意味を規定することができるもの。第2は意味を表現するプリミティブを利用して、語の意味を定義するもの。第3は他の概念との関係付けを行うことにより語の意味を記述するものである。第1のレベルには、ごく少数(たかだか数百語)の機能語が属する。第2のレベルは基本的な述語、すなわち和語動詞と形容詞であり、2000語程度の語を含むと考えられる。それ以外の語はすべて第3のレベルに属する。

品詞ごとに、どのように意味を記述するかを次節以下に述べる。

3. 動詞の意味記述

3.1 機能動詞

第1のレベルに属する動詞には機能動詞がある。一般に、機能動詞には「音がする」「期待をかける」「疑いをはさむ」など、名詞との結合が固定化し、実質的な意味は持っていない動詞も含まれるが、ここでは除外する。われわれの体系では、補語として動詞をとり、その動詞の意味に新たな意味を付加したり、その意味を変換したりするものを機能動詞としてとらえていて、そこにはヴォイス的およびアスペクト的動詞が含まれる。この場合の意味の変換の方法は(本動詞として用いられている場合も)ほぼ助動詞と同様の規則により記述することができる。

3.2 基本動詞

ほとんどの和語動詞がこの分類に含まれる。和語動詞を取り上げる理由は、それらが表す意味の抽象度が高く、それだけ基本的な意味を表現していると考えられるからである。基本動詞の意味を記述する枠組みとしては、深層格、シソーラス、形式的定義の3段階を与えた。

深層格はその動詞が必須格として取るアー

Semantic Description of the LTB Master Dictionary
Yuichi Tanaka (Institute for New Generation Computer Technology)
Motoo Oshima (Tokyo Metropolitan University)
Youko Nagasawa

ギメントの意味的な役割を示すものである。ここでは、それに対して表層格との対応関係、必須格要素のカテゴリに関する制約、さらに、意味記述におけるアーギメントとの対応関係を与えている。深層格としては約20の格を設定した。

動詞のシソーラスとしては、約200の項目への分類を行った。ここでは動詞の意味による体系化とともに意志性などのフィーチャを反映させることを考えた。これは名詞的概念のシソーラスにおいて、行為・活動の類の分類を行うために用いられるほか、文生成において、ある動詞概念を表現するための語彙を探索するときなどにも用いられる。

さらに、基本動詞に関しては、意味を表現するプリミティブを設定して、それを用いた形式的表現により意味を定義した。まず、マスタ辞書の語義ファイルには情報子(infon)と呼ばれる形式で意味が記述される。これは動詞がアーギメントとして取るオブジェクトや場所、時間などの間に成立する関係を定義するものである。この情報子の実質的な意味はプリミティブを用いた制約の形で表現され、マスタ辞書とは別の制約辞書に記述される。これが知識ベースに相当するものである。情報子は、それ自身ひとつの意味表現形式として意味構造を構築するための核を提供しているが、さまざまな制約が付けられて初めて限定的意味を表現することができる。

3.3 一般の動詞

一般の動詞に関しても、深層格を与え、シソーラスによる分類を行うことは上と同様である。また、意味の記述については、機能動詞や基本動詞の複合的な意味として表現する(上では機能動詞の意味としてヴォイス・アスペクトだけをあげたが、ここではモダリティ等も含まれることがある)。このような動詞の意味に関する制約には、他の動詞的概念や名詞的概念等への参照が含まれる。これは意味を定義するものではなく、他の概念との関係付けを制約により与えるものである。

4. その他の意味記述

4.1 形容詞の意味記述

形容詞は非常に少数であり、個々の語の持つ意味の抽象度が高く、基本的な意味を表現している。また、動詞とともに述語としてオ

ブジェクト間の関係を与えている。そこで、基本動詞と同様に意味のプリミティブによる定義を行い、さらに、動詞と共通の深層格を用いて格の記述を行った。

4.2 名詞の意味記述

名詞に関しても、それ自身としては実体を表現せず、他の語や句と結び付いて始めて意味を表現するような語を機能名詞として切り出した。機能名詞は、補語としてとる名詞や名詞句の意味を変換して新しい意味を生成するが、その変換の方法は規則化して記述した。

その他の名詞については、まず基本的な意味をシソーラスにより表現した。言うまでもなくこれは非常に粗い近似を与えているにすぎない。実質的な意味は制約辞書に記述される。すなわち、動詞的概念や形容詞的概念を用いて、概念間の関係を制約として記述することにより意味を表現している。

5. おわりに

本稿では、LTBマスター辞書において意味を扱う方式について、動詞を中心に述べた。現在は、DUALS第3版の開発に歩調を合わせて、意味の記述を行っているが、これまでに約300の動詞、数10の形容詞について、制約による意味の記述を終えた。

最後に、議論に加わっていただいたLTB開発グループのメンバーおよびICOT第2研究室の同僚諸氏に感謝する。

— 文献 —

- 村木新次郎：「日本語の機能動詞表現をめぐって」，研究報告集2，国立国語研究所，1980。
- 萩野孝野：「日本語の意味分類試案」，計量国語学会，第31回大会発表資料，1987。
- 奥津敬一郎：「生成日本文法論」，大修館書店，1974。
- 瀬塚孝志，et al：「LTBマスター辞書の構成」，ソフトウェア学会，論理と自然言語研究会，1988。
- 田中裕一，et al：「LTBマスター辞書の意味記述の構想」，ソフトウェア学会，論理と自然言語研究会，1988。