

談話理解のための 辞書／言語知識ベース

I C O T

第 2 研 究 室

田 中 裕 一

談話を理解するのに必要な 知識を辞書の形で与える

- 対象：談話理解実験システム (DUALS)
汎用日本語処理系 (LTB)
- 内容：言語内知識 (文法情報・機能的意味)
言語外知識 (世界知識)
- 特徴：解析・生成で共通の辞書を用いる

語彙の大きさ

DUALS-I	~100 語
DUALS-III	~2,000 語
.....目標.....	
机上版辞書	~60,000 語
EDR辞書	~200,000 語
CYCプロジェクト	~1,000,000 (概念)

構文的情報の例

読み：	...
表層表現：	...
品詞：	動詞
活用型：	強変化, 弱変化, ...
語義：	...
態：	[直接受動, 間接受動, 使役, 授受]
深層格：	AGT, OBJ, ...
表層格(能動)：	(AGT:が, OBJ:を)
表層格(受動)：	(OBJ:が, AGT:に)
相：	[状態, 準状態, 瞬間, 継続]
自他：	[絶対他動, 絶対自動, 相対他動, 相対自動]
可能動詞化：	[可, 不可]
意志性：	[あり, なし]
派生名詞：	...

語のクラス

- (1) 言語運用上の機能を担う語 (数百語)
 - 構文的処理が可能
- (2) 基本的かつ抽象的な意味をもち、
比喩も含め広く利用される語 (数千～1万語)
 - STASS, タイプ理論による解析
- (3) 具体的かつ複合的な意味をもつ語 (何十万語)
 - STASS, 制約の枠組みで記述

語の意味の記述方法の特徴

- (1) 機能語
 - 機能名詞, 機能動詞
 - 語のfunctionalな意味に注目
 - 論理式で表現可能
- (2) 基本語
 - 形容詞, 基本的な和語動詞,
一部の名詞, など
 - シソーラスで概念を近似的に表現し,
意味構造に対するindexを与える
 - 形式的表現により意味を定義する
- (3) それ以外の語
 - 漢語動詞, 大部分の名詞
 - 語の意味を定義するのは困難
 - 概念間の関係および制約で表現

グループ	類	機能名詞の例
時間	前後 ∴	前, 以前, 先 (さき), ... 後, 以後, 以降, ...
場所	上下 ∴	上, 上側, 上方, 上部, ... 下, 下側, 下方, 下部, ...
順序	前後 ∴	前, 先頭, ... 後ろ, 後尾, ...
程度	上下 ∴	上, 以上, ... 下, 以下, ...

「視点を」導入した名詞のシソーラスの例

- 1.2 [物]
- 1.2.1 [無生物]
- 1.2.1.1 [人工物]
- 1.2.1.1.1 [部品, 材料]
- 1.2.1.1.2 [道具]
- 1.2.1.1.2.1 [(用途) 工事, 工作に使用する道具]
- 1.2.1.1.2.2 [(用途) 農作業, 造園に使用する道具]
- ∴
- 1.2.1.1.2.22 [(特徴) 電気を使う道具]
- 1.2.1.1.2.23 [(特徴) 刃物]
- ∴
- 1.2.1.1.3 [製品, 生産物]
- 1.2.1.1.3.1 [(特徴) 身に付ける物]
- 1.2.1.1.3.1.1 [(着用場所) 頭]
- 1.2.1.1.3.1.2 [(着用場所) 体]
- 1.2.1.1.3.1.3 [服の部分]
- 1.2.1.1.3.1.4 [(用途) 防寒]
- ∴

必須格の概要

行為者(AGT), 経験者(EXP), 対象(OBJ),
生産物(PRD), 記述対象(ODS), 道具(IMP),
材料(MAT), 対象空間(LOC), 通過点(VIA),
始点(SOU), 終点(GOA), 存在場所(PEX),
所有者(POS), 方向(DIR), 用途(PUR),
原因(CAU), 随伴者(PAR), 対抗者(OPP),
内容(COT), 比較基準(COM),

今後の課題

- 量と質のバランス
ICOTの立場としては「質」
EDRの大規模辞書との関連
- 作成・運用・利用の形態／方法
知識表現, 学習の研究
知識ベースとの関連
- 将来の談話理解のために
照応, 談話構造,
比喩, 発話行為, …