

## 7S-6

## 概念分類項目間の関係記述

横田英司、岸本行生、三池誠司、松川智義、高井貞治

㈩日本電子化辞書研究所

## 1. はじめに

EDRでは、「振る舞いのよく似た概念」を集めるためのカテゴリーとして、概念分類項目を設定した。本発表では、この概念分類項目の「振る舞い」を項目間の概念関係を記述することで具体化したので、記述の目的、方法について報告する。

## 2. 項目間の関係記述の目的

## (1) ボトムアップによる体系

実際の文から、個々の概念間の関係を抽出し、同じ関係を持つ概念をまとめ上げていくことで体系化する方法が考えられる。しかし、例えば、「鳥」と「飛ぶ」とがagent の関係にあるとすると、「鶯」「雀」など、すべての鳥類が実文中に出現してかつ、その文中で「飛ぶ」とagent の係り受け関係にないと正しい体系化はできない。まして、「鳥」は「飛ぶ」とだけに関係しているわけではない。

このように、実際の文例からすべての概念間の関係を抽出して、ボトムアップで体系化するのは現実には不可能である。

## (2) トップダウンによる体系

逆に、鳥類の体系などは常識を使って、簡単に正しく作成可能であるが、抽象的概念や事象を表す概念の常識的な体系は存在しないので、それが実文から抽出した概念記述と一致するとの保証はない。

## (3) 分類項目間の関係記述に基づいた体系化

EDRの目的は自然言語処理用に適した辞書の開発であり、その概念体系辞書は、概念の「振る舞い」の継承関係に基づいて体系化されたものである。

よって、第一ステップとして、分類項目がボトムとなる体系を導くために、その「振る舞い」を具体的に項目間の概念関係として記述する。

## 3. 記述方法

## (1) 「こと概念」をノードとする記述

EDRでは、概念を大きく「もの概念」と「こと概念」の2つに分類している。単語からみると、体言が「もの概念」に用言が「こと概念」にほぼ相当する。

## ①もの-もの関係

例1) 鳥の羽

c#feather -part-of →c#bird

## ②こと-もの関係

例2) 太郎が食べる

c#eat -agent →c#taro

## ③こと-こと関係

例3) 本を読むのをじゃまする。

c#hinder-object →c#read-object →c#book

我々は自然言語処理用辞書に、まず必要なのは用言の振る舞いであると考えて、「こと概念」の分類項目をノードとした概念間の関係を記述した。

## (2) 内省にもとづく記述

人間は「行為」概念なら、「動物」や「人間」など有意志体と行為主体=agent 関係になり、行為対象としてobject関係が成立することは、膨大な実文にあたらなくとも、直感的に知っている。よって、まず直感に基づいた関係記述を行った。

## (3) 2項関係中心の記述

ある分類項目と別の分類項目との関係は、どの関係子で表現されるかという2項関係中心で記述した。

比較表現のように、2項関係で記述できない関係はマーク(rel)を付けるだけで止めた。

## (4) 記述する関係

我々は、「こと概念」の体系と項目間の関係とは次図の様に対応すると仮定した。

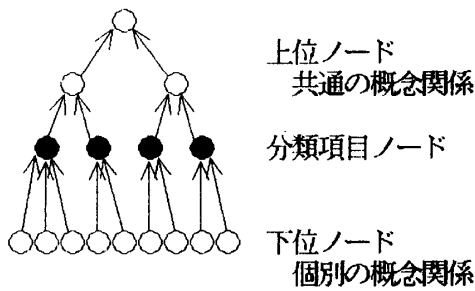


図1. こと概念体系と概念関係

場所格や時間格などの関係は、大多数の概念が共通に持つから、分類項目より上位ノードに記述すべきである。今回は記述しない。

「こと概念」の分類項目は、200~300ぐらいである。ここに、各々の分類項目が区別できる程度の関係を記述する。

分類項目より下位ノードには、個々の概念を特徴づける概念間の関係が記述されると予想している。これらは、今後、実文を個々にあたって、抽出する。

(6) 概念間の関係のゆれ

実文での概念間関係の認定は、辻井が指摘している様にゆれが生じて、困難である。我々は、抽象度の高い概念の分類項目間の関係記述を行ったので、この困難を避けることが出来た。そして、分類項目間に記述した関係と、実文から抽出した概念間の関係の違いは、ランダムなものではなく、分類項目間の関係はプロトタイプであると考えている。

4. 記述データ

(1) 関係子ごとのタプル

関係子	数	関係子	数
agt	1290	cas	845
dgr	494	goa	1911
imp	73	loc	448
man	1	mat	300
obj	30866	plc	507
rel	2645	scp	5944
src	2247		
		合計	47571

(2) データ例

空間関係. 距離s大+dgr	20551/数量の単位
空間関係. 距離s大+dgr	20552/数量の程度
空間関係. 距離s大+obj	111111/民族・種族・人種
空間関係. 距離s大+obj	111112/仲間
空間関係. 距離s大+obj	113131/気体
空間関係. 距離s大+obj	113132/液体
空間関係. 距離s大+obj	113133/粘体
空間関係. 距離s大+obj	113134/固体
空間関係. 距離s大+obj	114111/ひも類
空間関係. 距離s大+obj	114112/針類
空間関係. 距離s大+obj	114113/人形
空間関係. 距離s大+obj	114114/管
空間関係. 距離s大+src	111111/民族・種族・人種
空間関係. 距離s大+src	111112/仲間
空間関係. 距離s大+src	113131/気体
空間関係. 距離s大+src	113132/液体
空間関係. 距離s大+src	113133/粘体
空間関係. 距離s大+src	113134/固体
空間関係. 距離s大+src	114111/ひも類
空間関係. 距離s大+src	114112/針類
空間関係. 距離s大+src	114113/人形
空間関係. 距離s大+src	114114/管

5. まとめ

以上、概念の体系化のための第一ステップとして分類項目間の関係を内省によって記述する方法について述べた。今後の課題は次の2点である。

(1) ボトムアップ・データの組み込み

今回作成したデータを叩き台として、実文から抽出したデータと突き合わせて、修正する。

(2) 多項間関係

比較表現のように多項間関係でないと記述できないものの表現方法を探りだす。

謝辞

本研究の機会を与えていただいた、EDRの横井所長、太細第7研究室長、及び有益なご意見や多大なご協力をいただいた、内田第1研究室長をはじめとするEDRの研究員の皆様に感謝いたします。

参考文献

- 1) 辻井、山梨：NL研52-3, 1985/11  
「格とその認定基準」
- 2) EDR：「概念辞書（第1版）」TR-007, 1988
- 2) EDR：「概念辞書（第2版）」TR-012, 1989
- 3) 高井他：情報処理学会第41回全国大会  
「概念体系の開発思想」
- 4) 岸本他：情報処理学会第41回全国大会  
「概念分類項目の設定」