

概念分類項目の設定

7S-5

岸本行生、松川智義、三池誠司、横田英司、高井貞治
(株)日本電子化辞書研究所

1. はじめに

概念体系は他の概念との関係の観点からみて振舞いが同じ概念をグループ化し、そのグループ毎に上位概念を設定することによって概念を体系化したものである。その概念体系の開発にあたって、まず上位概念を導き出すための項目(概念分類項目)を設定する。

ここではその概念分類項目の設定方法について報告する。また、大量の概念を扱うことの開発効率を考慮して、概念分類項目に他の概念との関係の振舞いが類似していると考えられる概念を分類する作業を行なったので、その概念分類作業についても報告する。

2. 概念分類項目の設定

概念分類項目を上位項目を導き出すための項目とするために、以下の方針に基づいて設定した。

他の概念との関係が類似する要因となり、概念に共通して存在すると考えられる要素ないしは概念を抽出し、それらを項目設定の枠組として設けて、その枠組に基づいて概念分類項目を設定する。

項目設定の枠組を設けるにあたっては、概念間の関係の類似性という観点から項目設定の見通しを良くするために、「具体物概念」、「抽象概念」、「こと概念」の3つのグループに分けて行なった。

以下にそれぞれのグループの項目設定の枠組と概念分類項目の設定について示す。

(1) 「具体物概念」

「具体物概念」については(図1)に示す基本的な概念を枠組とした。概念分類項目はその枠組の項目を細分化することによって設定した。その際に以下の2つの属性についても考慮した。

- ・個別属性 : 具体物そのもの及びその性質によって示される
- ・関係属性 : 他のものや「こと」との関係によって示される

この属性によって、1つの具体物概念に対しても複数の観点で分類できるようにしている。

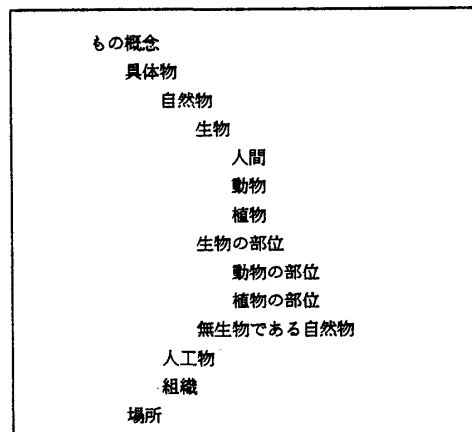


図1. 「具体物概念」の枠組

(2) 「抽象概念」

「抽象概念」については他の概念との関係の振舞いが他の概念によるところが大きいので、まず類似した概念を整理できるように以下の3つの観点を大きな枠組として設け、さらにその観点を詳細化して概念分類項目を設定した。

- ・具体物に結び付く概念
- ・思考内容に関する概念
- ・その他の抽象概念

(3) 「こと概念」

「こと概念」の枠組を設けるに当たって以下の2つの特性を考慮した。

- ・「ことの性質」(図2)

その“こと”の概念が何に関する概念かを示す特性

- ・「様態」(図3)

どのような変化を表す概念かを示す特性

この2つの特性を「軸」と呼び、それぞれの「軸」を組み合わせたものを「チャンク」と呼ぶ。「こと概念」についてはこの「チャンク」を項目設定の枠組とした。

個々の「軸」を詳細化し、その「軸」の組合せによる「チャンク」を基準として、必要に応じて他の概念との関係を記述する上で必要と考えられるレベルへ細分化して概念分類項目として設定した。チャンクと概念分類項目の例を以下に示す。

概念(概念分類項目名)	ことの性質	様態	(下位分類)
#散らす	空間関係	散	(距離)
(具体物が空間的に距離を伴って分散すること)			
#集める	空間関係	集	(距離)
(具体物が空間的に距離を伴って集中すること)			
#遠ざかる	空間関係	距離s大	
(空間的に、始点からある距離離れること)			
#近づく	空間関係	距離g小	
(空間的に、終点にある距離接近すること)			

ON CONCEPT CATEGORIALIZATION

Yukio KISHIMOTO, Tomoyoshi MATSUKAWA, Seiji MIIKE,
Eiji YOKOTA, Sadaharu TAKAI

Japan Electronic Dictionary Research Institute, Ltd.

さらに、「チャンク」が概念分類項目として広範囲の概念を示しており、分類作業上の項目としても粗いと考えられるような場合には、より直観的に概念を捉えられるよう「チャンク」の下位分類も設けた。

例	チャンク	下位分類名	概念例
	空間関係 & 格値(場所s)	<距離>	#離れる
		<着脱>	#脱ぐ
		<付着>	#切る

このようにして1単語で表せないものは簡単な句で表現した概念分類項目を設定し、他の項目との違いが明確になるようにそれぞれの項目に定義を与えた。

3. 概念分類作業

概念分類項目に概念を分類する際には、項目数が多いために作業が困難になることを回避できるようにする必要がある。そこで概念分類項目へ分類するためのガイドラインとしての上位項目(大項目・中項目)を設け、3段階の階層を持った分類項目のセット(概念分類体系)とし、分類作業を階層的に行なうことにした。さらに、項目を階層化したことによって、分類作業上たどってきた上位項目の配置によって分類が中断してしまうという問題を回避するため、1つの項目を複数の上位項目下に配置(多重配置)した。

日本語・英語合わせて約30万概念を抽出し、上記の概念分類項目のセットに分類する作業を行なった。複数の概念分類項目に分類する(多重分類)を許すようにして、できる限り概念分類項目レベルに分類するように作業を進めた。その結果約8割の概念が概念分類項目レベルまで分類され、約1割の概念が多重分類された。

4. まとめ

以上、概念分類項目の設定方法とその項目に基づく概念分類作業について述べた。設定の枠組を設けたことによって、他の概念との関係の振舞いという観点から一貫した概念分類項目を設定することができた。また、分類作業もその項目を多重配置したことによって、ほとんどの概念をいずれかの項目に分類することができた。

今後は今回の分類作業結果をさまざまな観点から検討することが課題となるが、特に概念の体系化という観点から項目の再検討及び分類された概念の妥当性を検討していきたい。

【謝辞】

本研究の機会を与えていただいた、EDRの横井所長、吉田第4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) 日本電子化辞書研究所:「概念辞書(第2版)」TR-012, 1989
- 2) 高井、奥村、中原:「要素概念の組合せによる概念分類体系」EDR TM-009, 1989
- 3) 高井、他:「概念体系の開発思想」(41回情報処理学会全国大会), 1990

空間: ものと空間の関係
・ 空間関係: もの間の空間的な関係
・ 空間属性: ものとその空間的な属性との関係
時間: 時間(回数、期間)
経過: 事象の経過
所有: 所有物と所有者の関係
情報: 情報と情報所有者の関係
交際: 有意志体とその社会的、心理的関係
機能: ものとその機能との関係
種類: ものどもの種類、区別などによる関係
数量: ものと数量の関係(属性や関係の程度を含む)
属性: ものとそのものが持つ属性との関係 (空間的、時間的な属性を除く)
評価: ものとそのものに対する評価との関係
音 : 音と音の発生体との関係
光 :
可能性:
存在:

図2. 「ことの性質」の種類

順序: 順序の変化
距離: 距離的な変化(方向性をもつ変化)
集散: 変化が集散として起こる変化
状態: 状態、継続

図3. 「様態」の種類