

Subject World における概念体系表示のための一検討 —BSH4 の表示方式—

村上 晴美[†] 上田 洋[†]

[†]大阪市立大学大学院創造都市研究科

我々は概念間の関連を可視化することにより情報検索を支援するシステム Subject World を開発している。本稿では、Subject World における概念の表示方式の設計を検討する。概念間の関係の表記をシソーラスから借用した基本件名標目表第4版(BSH4)をとりあげる。データの中から可視化の対象とする概念を定め、図式表示の採用を検討し、表示方式を設計し、ユーザによる調査結果をふまえて実装を行った。

An Investigation into the Representation of Concepts in Subject World - Representation of BSH4 -

MURAKAMI Harumi[†] UEDA Hiroshi[†]

[†]Graduate School for Creative Cities, Osaka City University

We have been developing a system called Subject World, which assists in information retrieval by visualizing concepts and relations between concepts. In this paper, we investigate the representation of concepts and relations in Subject World. In the first step, we consider the visualization of BSH4, which is one of the subject headings lists in Japanese libraries. We decided on objects to be displayed in BSH4, discussed existing methods of displaying thesaurus, and designed the representation of concepts and relations. In addition, we conducted a user study, and implemented the design of the representation of BSH4.

1 はじめに

我々は概念間の関連を可視化することにより情報検索を支援するシステム Subject World を開発している[1,2]。Subject World では、ユーザは概念を自由に操作することができ、概念を選択してデータベース等の検索を行う。本稿では、Subject World における概念の表示方式を検討する。

情報検索を支援するための概念体系に関しては、「統制された索引言語の語彙であって、あらかじめ概念間の先駆的な関係を明示するように組織化されたもの」と定義されるシソーラス[3]が代表的である。シソーラスが電子化される以前には、図上で適切な索引語を選択するために、系統樹構造、アローグラフ、ターミノグラフの3種類の図式表示が考案されてきた[3, 4]。これらの図式表示を参考にして設計することが考えられる。ただし、概念の操作を前提としない、紙の時代の図式表示が適切かどうかは自明ではない。

また、シソーラスにおいては「を見よ参照あり(UF)」のように、シソーラスに関する知識のない一般のユーザには理解しにくい言葉が使われている。シソーラスに関する知識の有無に関わらず、ユーザにとって理解しやすい表示

方式の設計が必要であると考えられる。

本稿では、概念間の関係の表記をシソーラスから借用した基本件名標目表第4版(以下、BSH4)[5]の表示方式を検討する。BSH4 自体だけではなく、シソーラスをはじめとする概念体系の可視化に関する知見を得られると考える。

2章では、Subject World に求められる機能、3章では BSH4 で可視化の対象となる概念、4章では図式表示の採用に関する検討と仮設計、5章では設計に関して行った調査、6章では調査に基づいた再設計と実装、7章では今後の課題を述べる。

2 Subject World に求められる要件

Subject World は概念間の関連を可視化することにより情報検索を支援するシステムである。ユーザは、概念を自由に操作することにより、主題に関する知識を得ながら情報検索を行える。現在のところ、BSH4 と、日本十進分類法新訂9版(以下、NDC9) [6]などの概念を扱っている。

概念の表示方式に関して、本システムでは以下の機能要件がある。

- (1) 利用者の操作による、概念の入力、選択、編集、移動、削除ができる。

- (2) 上位概念が下位概念の下方に移動されても、概念間の関係を表現できる。
- (3) 概念体系が異なっても、共通的な操作ができる。

BSH4 や NDC9 などの概念体系にはそれぞれ固有の目的があり、含まれているデータの内容や構造が異なる。次章では、BSH4 のデータの中で何を可視化の対象とするか検討する。

3 BSH4 で可視化の対象とする概念

図書館における件名目録編成のための基準として選択された件名標目と参照語を列挙し、標目相互間に必要な連結参照を設定して、一定の方式に配列した表を件名標目表という[5]。BSH4 は、多くの図書館が共通に採用すると予想される件名標目を採録した標準件名標目表の一つである。

表 1: BSH4 とシソーラスの対応と検討対象

BSH4([5]による)	シソーラス([3]による)	略語	対象
限定注記	スコープノート	SN	—
→	を見よ参照	USE	○
を見よ参照あり	を見よ参照あり	UF	○
最上位標目	最上位語	TT	○
上位標目	上位語	BT	○
下位標目	下位語	NT	○
関連参照	関連語	RT	○
参照注記	—	SA	—
—	全体部分関係の上位語	BTP	—
—	種類関係の下位語	NTG	—
—	全体部分関係の下位語	NTP	—

注)本研究では○のみ検討対象とした。

BSH4 のデータには、件名標目、参照語、説明つき参照、細目の 4 種類がある。それぞれ、7,847, 2,873, 93, 169 項目が含まれている。本稿では「件名標目」と「参照語」を対象とする。

一つの件名標目から参照できるものとして「限定注記 (SN)」、「を見よ参照あり (UF)」、「最上位標目 (TT)」、「上位標目 (BT)」、「下位標目 (NT)」、「関連参照 (RT)」、「参照注記 (SA)」の 7 種類がある。この中で、限定注記は件名作業に関する説明であり概念を表すものではないので対象から省く。参照注記も、細目を対象から省いたのと同様に対象から省く。すなわち、「を見よ参照あり (UF)」、「最上位標目 (TT)」、「上位標目 (BT)」、「下位標目 (NT)」、「関連参照 (RT)」の 5 種類を検討対象とする。

参照語の場合は「→ (USE)」1 種類であり、検討対象とする。

上記の中で「を見よ参照あり (UF)」と「→ (USE)」、「上位標目 (BT)」と「下位標目 (NT)」

は相互に関連している。BSH4 の場合は「関連参照 (RT)」は相互に関連参照の関係にある。

さて、BSH4 とシソーラスを比較すると、意味はほとんど同じだが表現が異なるものがある。たとえば、BSH4 における件名標目は、シソーラスでは「索引語」や「ディスクリプタ」であり、BSH4 における「参照語」はシソーラスでは「非ディスクリプタ」である。本章以降では略語またはシソーラスで用いる表現に統一する(表 1 参照)。たとえば、上位語という場合、BSH4 においては上位標目である。

4 図式表示の検討と仮設計

4.1 図式表示の概要

図式表示には系統樹構造、アローグラフ、ターミノグラフの 3 種類がある。本章では[4]より、これらの概要を抜粋する。

系統樹構造では、最上位語は表示の上部に置かれ、下位語はそれに続く位置に印刷され、関係は垂直の連結線により指示される。

アローグラフでは、格子によって全体が区画される。最上位語は、太字で中心に置かれる。下位語は格子内の他の位置に印刷し、上位語から下位語に向けた矢印(アロー)で上位/下位関係を表す。連想関係にある用語は、双方向の矢印または破線で連結して示す。すなわち、上位語が下位語の下方に位置することもあるが、上位語から下位語に矢印を向けることにより表現できる。

ターミノグラフ(またはボックスチャート)では、優先語をその下位語と一緒に枠(ボックス)で囲む。枠中では、最上位語に下線がひかれ、下位語は字下げして階層的に示される。優先語間の連想関係は線で示される。ただし、最上位語間の連想関係は太い線で示される。

4.2 図式表示の検討

2 章の機能要件を実現するためには、概念を自由に操作できないターミノグラフの採用は難しいため、検討対象から除く。

上位語が下位語の下方に移動される可能性があるため、概念の位置のみで概念間の関係を表す一般的な系統樹構造の採用は難しい。系統樹構造の一部とみなすことができる PRECIS のシソーラスネットワークでは、BT/NT 関係、等価関係、連想関係を示す記号を、概念を結ぶ実線の中央においており、この方法は検討の対象となる。しかし、上位語が下位語の下方に配置される場合にどちらが上位語かわからなくなるため、矢印等による明示が必要であると思われる。

アローグラフでは概念間の関係の表示方式

として矢印が採用されているが、概念間の関係の種類が明示されないため、概念間の関係の理解が難しい。

上記の検討より、概念間の関係については、系統樹構造とアローグラフの長所を組み合わせ、概念間の関係を表す記号を線の中央に配置し、線として矢印を用いる方法がよいのではないかと考える。

4.3 仮設計

上記の検討を踏まえて、以下のように仮設計を行った。

- (1) 概念と概念を線で結び、線の中央に意味を表す記号を配置する。これは、概念間の関係を表すためである。
- (2) ユーザが概念を選択して検索を実行（本稿では「近傍検索」と呼ぶ）すると、選択した概念の「を見よ参照 (USE)」「最上位語 (TT)」「上位語 (BT)」「下位語 (NT)」「関連語 (RT)」の5種類を表示する。「を見よ参照あり (UF)」はシソーラス知識のないユーザにとってわかりにくい概念であると考え「を見よ参照 (USE)」と統合した。
- (3) 概念を表すために、概念を表す語の周囲に枠をつける。これは、概念の選択状態や、概念間の線（後述）がどの概念につながっているかをわかりやすくするためである。
- (4) 索引語またはディスクリプタ（以下ディスクリプタ）は一重の四角で表す。

ここで、(a) どのような概念間の線が妥当か (b) 概念にどのような枠をつけるとよいか、という疑問が生じる。次章では、この2点に焦点をあてて検討を行う。

(a) については、実線、片方向矢印（左右）、両方向矢印の4つの選択肢を用意し、概念間の関係を表す一つまたは二つの文に対して適当な表示方式を被調査者に選ばせる。

(b) については、最上位語をどのように表すか4種類の枠の形状から適当なものを被調査者に選ばせる。また、BSH4でもシソーラスでも明示されない概念である最下位語についても検討する。これは、ある概念を展開してもその下方に概念が存在しないことを示すものである。最下位語であることを表示することにより、ユーザの無駄な操作を減らす効果を期待する。

5 調査

5.1 方法

5.1.1 被調査者

大阪市立大学大学院創造都市研究科修士課程院生9名（男性：6名、女性：3名）。被調査者は女性1名（元図書館司書）を除いて、情報科学の基礎的な知識は有するがシソーラスに

はなじみがないことを確認した。

5.1.2 件名標目の選定

例題として「補償」を選び、図1に示す件名標目表[5]の記述をどのように表示すればよいか確認した。「補償」を選択した理由はUF, TT, BT, NT, RTの全てを持つことと、NTが「災害補償」であり、文字列を含むため下位語であることが被調査者にわかりやすいだろうと考えたからである。

ホショウ	補償*	③311.4	③23.96
	UF:	国家補償, 損失補償	
	TT:	行政法 51	
	BT:	行政法	
	NT:	災害補償	
	RT:	土地収用	

図1: 件名標目「補償」

5.1.3 手続き

質問紙を用いて、「補償」に関する一つまたは複数の文に対して、文の内容を示す表示方式を4種類提示して一つだけ選択させた。文と表示方式の対を問とする。問は7つあり、以下のとおりである。すべての問の最後にこの調査に関する自由記述をさせた。

問1: 「補償」を「国家補償」の代わりに検索に使い。

問2: 「補償」の最上位語は「行政法」である。

問3: 「補償」の上位語は「行政法」である。

問4: 「補償」の下位語は「災害補償」である。

問5: 「補償」の関連語は「土地収用」である。

「土地収用」の関連語は「補償」である。

問6: 「行政法」は最上位語である。

問7: 「労災保険」は最下位語である。

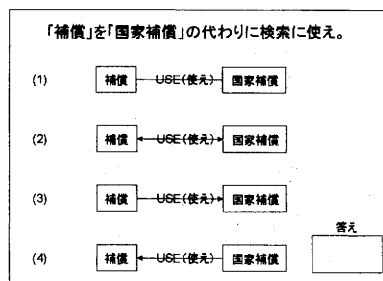


図2: 問1の画面

図2に問1の画面を示す。付録に問2以降の画面を示す。教示は以下のように行った。

「この調査は、情報検索の検索語の表示に関する設計を目的としています。語の種類や、語と語の関係を表すために、どれが最も妥当と思うか、上に示される文について、下の4つの図の中から1つだけ選択して「答え」欄に書いてください。どうしても選択できないときは「×」

を書いてください。できるだけ直観的に選択してください。後で気がかわっても、前に戻って修正しないで下さい。」

5.2 結果と考察

4つの選択肢に対して選んだ人数と割合を示した結果を表2に示す。問毎にカイ2乗検定を行った。問2 ($\chi^2=9.22$, $df=3$, $p<0.05$), 問3 ($\chi^2=9.22$, $df=3$, $p<0.05$), 問4 ($\chi^2=14.56$, $df=3$, $p<0.01$), 問6 ($\chi^2=13.67$, $df=3$, $p<0.01$)で有意差が見られた。

表2：実験結果

問1: USE/UF				
	実線	両方向矢印	矢印左から	矢印右から
USE/UF	0(0%)	1(11%)	3(33%)	5(56%)
問2: TT				
	実線	両方向矢印	矢印上から	矢印下から
TT*	1(11%)	0(0%)	2(22%)	6(67%)
問3: BT				
	実線	両方向矢印	矢印上から	矢印下から
BT*	0(0%)	1(11%)	2(22%)	6(67%)
問4: NT				
	実線	両方向矢印	矢印上から	矢印下から
NT**	2(22%)	0(0%)	7(78%)	0(0%)
問5: RT				
	実線	両方向矢印	矢印左から	矢印右から
RT	3(33%)	5(56%)	1(11%)	0(0%)
問6: 最上位語の枠				
	一重四角	二重四角	上線	下線
最上位語*	0(0%)	7(78%)	1(11%)	1(11%)
問7: 最下位語の枠				
	一重四角	二重四角	上線	下線
最下位語	3(33%)	0(0%)	3(33%)	3(33%)

* $p<0.05$ ** $p<0.01$

有意差がみられた問に関して、TT, BT, NTを現す関係については、矢印利用が利用者の直感にあると考えられる。最上位語の枠については二重四角が他の選択肢と比べてよく選ばれた。

USE/UF, RTに関しては、有意差は見られなかったが、過半数をしめる選択肢があった。最下位語の枠については意見がわかれた。なお、元図書館司書の回答は、最下位語の枠以外は、もっともよく選ばれた回答と一致した。すなわち、シソーラスに関する知識の有無を問わず、最下位語の枠以外は、ある程度ユーザにとって理解できる表現であると考えられる。

自由記述に「矢印に対しての印象なのですが、下向きの矢印(↓)右向きの矢印(→)がしっくりくる。逆に上向き(↑)左向きの(←)の矢印には違和感というほどではないが、しっくりこない気がする。」「最上位語、最下位語で、トランプなどの6・9のような線を入れても上下のイメージをしにくい」「矢印は基本的に下に向いているほうが、良いのではないかと個人的には思いました。」という意見があった。

以上の結果より、実線が矢印については矢印を使う方向が望ましいと考える。アイコンと矢印の形状は、もっとも多くの利用者が選択したものとする。可能な範囲で、左から右、上から下への矢印となるようにする。

6 再設計と実装

上記の結果をふまえて、図3に示すような方式を設計した。

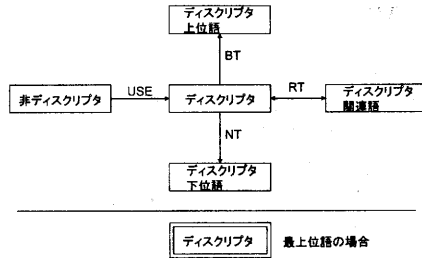


図3: 設計

調査ではシソーラス知識のない被調査者のために略語とその意味を付記したが、コンピュータの画面に概念を複数表示すると煩雑になるため、略語のみ線を中心に配置した。

また、上位語でない最上位語は直接表示しないこととした。これは、たとえば「補償」の場合上位語も最上位語も「行政法」であるが、二つの概念間に複数の線があったり、すべての下位語から最上位語への線があったりすると、画面が煩雑になるからである。若干操作の手間はかかるが、上位語を順番に表示することにより必ず最上位語にはたどりつけるので機能としては補える。詳細を以下に示す。

- (1) 概念の近傍検索を行うと、概念と直接接続されている概念が表示される。(ただし、最上位語は除く。)
- (2) USE/UF 関係は、非ディスクリプタを起点としてディスクリプタを終点とする矢印として、USEを表示する。非ディスクリプタはディスクリプタの左側に配置する。
- (3) RT 関係は(BSH4の場合は常に相互に関連語であるため) 両方向の矢印とし、RTを表示する。選択されたディスクリプタは新たに展開されるディスクリプタの左側に配置する。
- (4) BT/NT 関係は、選択された概念を起点とした矢印で表示し、上位語が起点となる場合にはNT、下位語が起点となる場合にはBTを中心に表示する。上位語は下位語の上側に配置する。
- (5) 最上位語は、BSH4データでは直接接続されているが、概念体系の中では必ずしも直近の上位語ではないことがある。そのため、

概念から最上位語の展開は行わないことにする。最上位語は二重の四角で表す。

- (6) 既表示の概念に対して関係のみかわる場合は、既表示の概念の位置は変えない。
- (7) 「矢印統一」機能を選択すると、BT/NT関係の矢印の向きを統一する。すなわち、BTまたはNTに統一する。

以下、図4から7に実装された画面を示す。

図4は「補償」を選択して近傍検索した例である。非ディスクリプタである「国家補償」と「損失補償」から「補償」に向け矢印が表示される。BT/NT関係に関しては、「補償」を起点として矢印が表示される。「行政法」は上位語であると同時に最上位語であることがわかる。RT関係はBSH4においては常に両方向であるため両方向矢印が表示される。

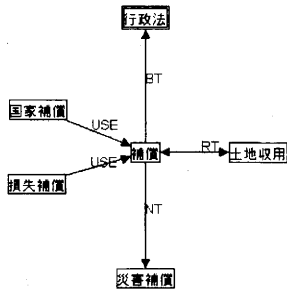


図4: 「補償」から近傍検索

図5は、図4において「災害補償」を選択して近傍検索した例である。「災害補償」を中心として、概念の配置はかえずに「災害補償」から「補償」に向かって矢印をひきなおす。「災害補償」から「補償」を経由して最上位語が「行政法」であることがわかる。「災害賠償」の上位語である「損害賠償」が新たに表示される。

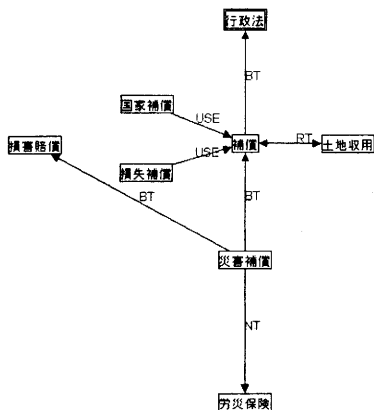


図5: 「災害補償」から近傍検索

図6は、図5から「損害賠償」を選択して近傍検索した例である。「損害賠償」の上位語が最上位語でもある「民法」である。すなわち、「災害補償」から見ると「行政法」だけでなく「民法」も最上位語であることがわかる。

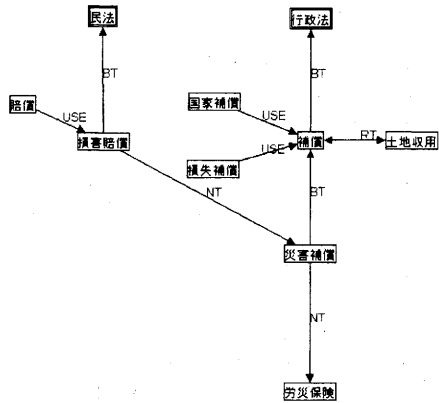


図6: 「損害賠償」から近傍検索

図7は、図6において「矢印統一」機能を実行した例である。画面上のBT/NT関係がすべてNTに統一されていることを示している。

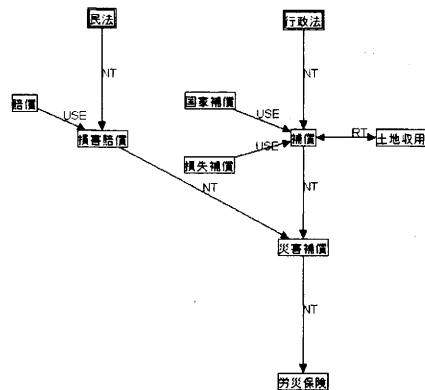


図7: 「矢印統一」

7 今後の課題

今後の検討及びユーザの確認が必要な課題について述べる。

- (1) BT/NT関係の矢印の方向に関して、選択した概念を起点とする表示と、常に上位語を起点とする表示とどちらがよいか。
- (2) 最上位語を簡単な操作で見たいという要望がある場合どうすればよいか。
- (3) 今回は、USEとUFを同じものと考え、USEのみを用いたがこれよりよいか。
- (4) 最下位語を表す枠に適切な形が特定できなかった。「最下位語」はBSH4にもシソー

1) 印刷用に白黒にしているが実際はカラーである。

ラスにも明示されていないため必要ないとも考えられるが、Subject Worldのように概念を探索するシステムにおいては、「この先には展開しても何も存在しない」ことを示すために、表示することが有効であると考へ、適当な表示を引き続き検討したい。

- (5) 現時点では、非ディスクリプタとディスクリプタの区別を明示していない。これはSubject Worldにおいては、非ディスクリプタを用いて情報検索することを許しているからである。本当にこの考え方でよいか。
- (6) 概念から接続される概念の数が増えると、画面上で非常に見にくくなるが、この場合の表示方式をどのようにすればよいか。
- (7) 略語の意味をユーザに提示する必要がある。

最後に、本研究の目的は、Subject Worldにおける情報検索支援のためのBSH4の表示方式の設計であり、件名作業のためのBSH4の表示方式の設計ではない。Subject WorldをBSH4の概念を調べるために使用することは可能であるが、スコープノートをはじめとして取り扱わなかったデータがあること、直接最上位語を表示したいときには操作が多く必要になること、などの制約がある。

8 おわりに

Subject Worldにおける概念の表示方式として、概念間の関係の表記をシンソーラスから借用し

たBSH4について検討した。

本研究で得た知見は、BSH4はもとより、ひろくシンソーラスの可視化に関して、また、一般的な概念体系の可視化に関する知見を与えるものであると考へる。

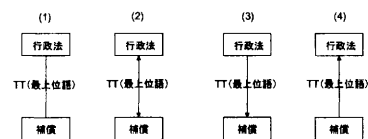
Subject Worldは試作版を公開[7]しており、今後も改良を続けていく予定である。

参考文献

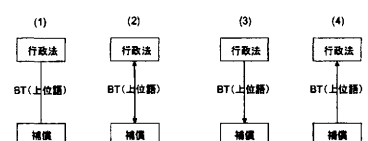
- [1] 村上 晴美, 平田 高志, 北 克一: 主題検索のためのOPAC動的可視化システム, 第49回日本図書館情報学会研究大会, pp.79-92, 2001.
- [2] Murakami, H. Hirata, T., and Kita, K.: Subject World: A System for Visualizing OPAC, 人文科学とコンピュータシンポジウム論文集, pp.237-240, 2002.
- [3] シンソーラスの構成及びその作成方法, JISX0901, 1991.
<http://www.jisc.go.jp/app/pager?id=83460>
- [4] 内藤 衛亨, 中倉 良夫, 影浦 映他訳: シンソーラス構築法, pp.110-116, 丸善, 1989.
- [5] 日本図書館協会件名標目委員会編, 基本件名標目表(BSH)第4版第2刷, 日本図書館協会, 2000.
- [6] もり・きよし原編, 日本図書館協会分類委員会改訂: 日本十進分類法新訂9版本表編, 日本図書館協会, 1995.
- [7] Subject World ホームページ
<http://www.media.osaka-cu.ac.jp/~harumi/research/SubjectWorld/>

付録

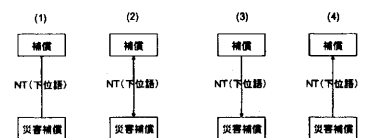
図2:「補償」の最上位語は「行政法」である。



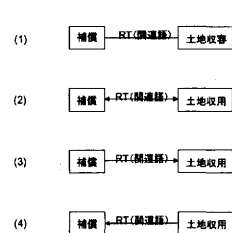
問3:「補償」の上位語は「行政法」である。



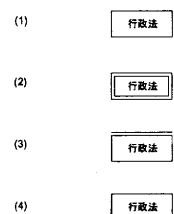
問4:「補償」の下位語は「災害補償」である。



問5:「補償」の関連語は「土地収用」である。「土地収用」の関連語は「補償」である。



問6:「行政法」は最上位語である。



問7:「労災保険」は最下位語である。

