

反義の対立軸を用いたシソーラスの構築

高橋 孝典 横山 晶一 西原 典孝

山形大学 工学部

1. はじめに

意味処理においては、知識概念の上下関係や、全体部分関係を考慮して構築されたシソーラスなどからそれらの関係を抽出して処理を行うことが多い。しかしながら、意味処理にはその他にも種々の観点が必要とされる。反義概念はその中でも重要であるが処理の難しい概念の一つである。

反義対には、「男/女」、「表/裏」といったものから、「出場/入場」、「出場/退場」、「出場/欠場」のような一義的でないものまで様々なものが存在する。反義対にペアになる各単語の意味とペアとしての意味が混在することにより、反義対自体の定義が不明瞭なものになっているからである。

現在のシソーラスでは、反義概念は適切に表現できていない。例えば、日本語語彙大系[1]で「点火/消火」は

++点火++

[1名詞/1000抽象/1235事/2304自然現象/2305非生命現象/2306物象/2307非刺激/2308化学現象/2310火/2311発火]

++消火++

[1名詞/1000抽象/1235事/1236人間活動/1560行為/1920労働/1936仕事/2001操作/2023充電・録音等/2024消防]

となり、反義性を明示的に見ることはできない。このような問題を解決し反義性を適切に表現できれば、これまでの上下、全体部分関係に反義関係が加わり、意味の視野を広げることができる。

本稿では、現在のシソーラス上で適切に表現できていない反義概念を反義の対立軸を用いることで実現しようと試みた。その結果、反義概念に対立軸を付与することにより、シソーラス上で反義概念の利用を可能にした。また、反義の対立軸から反義がグループ化されるが、これによる意味上視野が広がることを確認する。反義概念の対立軸の付与による検索視野の拡大を利用して、反義の対立軸を用いた語彙検索システムへの応用を試みた。その結果、シソーラス上では遠い位置関係にある反義を近くにあるものとして捉えることや、全体部分関係のような上下概念以外の反義概念という性質を用いてシソーラス上を移動できることが可能となった。

2. 先行研究

反義概念に関してこれまで行われてきた研究は言語学的な観点からのものがほとんどで、機械処理を視野に入れたものとして荻野、横山他[2]が挙げられる。ここでは、EDR概

念体系[3]において、多重リンクの不整合性や、記述の不統一のために反義性が十分に表現できていないことを指摘している。また、既存の反対語辞典を参考に、反義対に対して共通成分と差異成分の抽出を行っている。本稿は、この研究を基本的に継承する。高橋他[4]では反義の対立軸を用いることで概念体系を拡張できる可能性のあることを指摘した。

3. 反義の対立軸

3.1 対立軸の定義

反義について「意味特徴の共通成分と相違成分から成り立ち、何らかの意味特徴が相互に向かい合っている単語ペア」[5]とする。その上で、対立軸について「共通成分を支点とした、対極を意識した相違成分を表わしたもの」とする。相違成分に反義概念の性質が現れると考え、この相違成分を対立軸として取り扱う。実際には、「遠洋漁業/沿岸漁業」に対しては、対立軸を「遠洋:沿岸」のように A:B の形式で表す。

3.2 対立軸の付与

反対語辞典[6]のうち、EDR日本語単語辞書に登録してある語彙 7393 対を取り出し、反対語辞典に記載されている語義を参照し、対立軸の付与を恣意的に行った。

EDR日本語単語辞書に存在するもののみ取り扱ったのは、概念辞書中に語義が登録されているため、今後の機械処理への応用に使えると判断したからである。本来は電子化されたデータから統計的に反義概念を抽出できればよい。しかし、上で述べたように、反義の情報としてどんな情報を抽出すればよいか、またどんな情報を持っていれば反義として機械的に扱えるのかがまだ分かっていない。従って、本稿では対立軸を恣意的に付与した。

以下に、対立軸の付与に関する方針について述べる。

まず、対立軸の数は限定しない。反義対は、反義を構成する語彙それぞれの意味と、反義としての意味があり、また、反義ペアとしては同じでも、使われる状況により意味合いの異なる場合があるので、様々な状況に対処する必要があるためである。

また、類似した対立軸について、まとめることはあえて行わない。「本:支」や「本:節」といった類似した対立軸が付与される場合もあるが、これらを一つの対立軸で代表させるといった作業を行わず、そのまま残す。語彙からのリンクが増えることで意味上の視野が広がることが予想されるためであ

は以前よりも複雑にはなるが、シソーラス上で反義概念が適切に表現されることを目指した。本システムでは、語彙に反義と対立軸を付与することにより、これまで上位・下位概念間の移動が主であったシソーラス上で、それとは異なる反義に関する移動が可能となる。その結果、語彙からのリンクが増え、柔軟にシソーラス上を移動できるため、語彙検索が容易になった。本システムで参照する辞書は以下のとおりである。なお、EDR電子化辞書から必要に応じて、情報を取得した。

1. 反義辞書:

語彙に対して反義と反義の対立軸を付与したもの

形式：語彙、反義、対立軸(複数)

例：叔父、叔母、男、女、父、母

2. 対立軸をキーにした反義対リスト:

反義辞書と整合性の保たれた辞書であるが、対立軸をキーとするため、別に用意した。

形式：対立軸、反義対リスト

例：拙/巧、稚拙/老巧、拙劣/巧妙、拙技/巧技、拙速/巧遅、拙さ/巧みさ、拙/巧

3. EDR概念見出し辞書:

EDR概念見出し辞書から不要な情報を削除して使用する。

形式：概念識別子、概念見出し語、概念語義

例：1e9c8e, 造り酒屋, 酒を醸造して卸すのを業とする職業

4. EDR概念体系辞書:

概念間の上位/下位関係を記述したもの

形式：上位概念識別子, 下位概念識別子

例：3aa877, 3a2993

5. EDR日本語単語辞書

形式：単語見出し語, 概念識別子, 概念見出し語, 概念語義

例：どしん, 3d1ce0, 地響きする, 重い物の移動で地響きする

システムの処理の流れは図2ようになる。

処理の流れを順に説明する。まず、ユーザが初期語彙を入力する。EDR日本語単語辞書に存在しない場合、その旨を表示して繰り返す。入力語がEDR日本語単語辞書に存在する場合、概念識別子を取得して、これをもとに上位概念、下位概念、反義、反義間に存在する反義の対立軸を取得、表示する。ユーザが次に移動する方向を番号で選択する。上位概念、下位概念を選択した場合には、繰り返す。反義を選択した場合、EDR日本語単語辞書を検索する。EDR日本語単語辞書に存在する場合は概念情報を取得して繰り返す。対立軸を選択した場合、対立軸をキーにした、反義対リスト辞書から同じ対立軸を持つ反義対を取得して、選択する。その後、反義対から、語彙を選択して繰り返す。

以上の処理を繰り返すことで、概念体系上を移動する。これにより、上位、下位とは異なる柔軟な移動が可能となる。

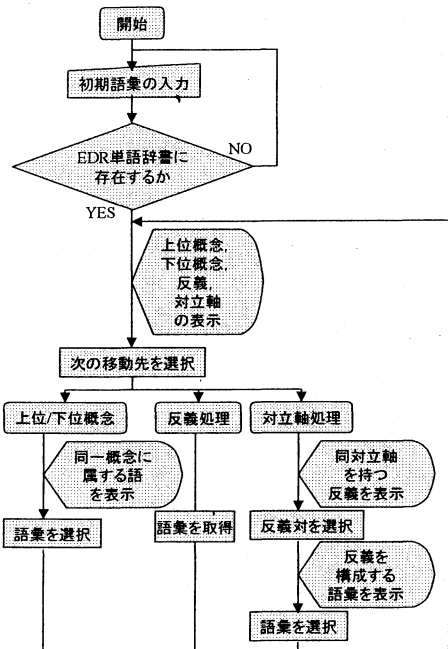


図 2 システムの処理の流れ

4.2 結果

実際に語彙を入力した結果を示す。

まず、初期語彙として「近い」を入力する。すると、以下のように表示される。

語彙を入力してください：近い

0:<上位>:444d25:“”:いろいろな人間の属性

1:<反義>:遠い

2:<対立軸>:遠近

次の移動先を番号で選択してください。:_

ここで、“2”と入力すると、対立軸「遠:近」を持つ反義対が表示される。

次の移動先を番号で選択してください。:2

0:間遠い/間近い

1:耳遠い/耳近い

2:手遠い/手近い

3:程遠い/程近い

4:間遠/間近

5:浅近/深遠

6:近海魚/遠海魚

...

このようにして、EDR概念体系上で、上位/下位概念間以外

の選択肢である反義に関しての移動が可能となった。また、対立軸を設定し、システム上で対立軸を選択できるようにしたこと、ユーザが対立概念を意識できれば、他の語彙を入力しても同対立軸を持つ語彙を提示することで、シソーラス上では遠い位置にある語彙への移動が可能となった。

問題点としては以下が挙げられる。

1. 対立軸を持つ概念がまだ少ない。

「上:下」「遠:近」「大:小」など、同一の対立軸を持つ反義対が多いものは選択肢が増え、他の概念に移動することが簡単にできるが、対立軸から、または反義からの移動に選択性が少ないものは移動先が限られ、結果的に他の概念への移動に上位概念を選択するしかなくなる。

2. 中間ノードに移動した場合、上位・下位概念が多数出力され、次の移動先を選択するのが難しくなる。

解決策としては、反義を持つ概念のみをシステム上で利用することが考えられる。また、多数検出された場合、全てを一度に表示しないなどユーザに配慮した表示方法が必要となる。

5. おわりに

本稿では、反義の対立軸を付与することで、シソーラス上で適切に反義を表現する方法を提案し、反義の対立軸を付与した語彙検索システムについて述べた。

反義の対立軸を語彙に付与することで、全体部分関係、または上位・下位関係とは異なる反義関係をシソーラス上に実現することができた。また、反義の対立軸を用いて、シソーラス上を柔軟に移動できる語彙検索システムの作成を試みた。その結果、上位/下位概念間の移動に加え、反義、さらに同対立軸を持つ語彙へのリンクが可能となった。

現段階での問題点として次のことが挙げられる。

紙媒体辞書での反義対と、EDR概念辞書中の見出し語とのズレがあり、EDR概念辞書上で反義として判断することが難しい場合がある。また、概念語義をみる限り、反義のペアを構成するにふさわしい概念ペアを見つけない反義もある。

反義を意識させる反義対と意識させない反義対との対立軸のすり合わせに関する問題もある。「現実的/理想的」や「西岸/東岸」など、表層的な情報から反義を意識できるものと、「大納会/大発会」のように「最後:最初」という対立軸が表層的には判断できず、語義から判断する場合があります。手作業で付与した対立軸が他の対立軸とのグループを形成できないものがあつた。今回は、対立軸の付与に際し、類似するものはまとめない方針を採ったが、グループ化ができないと、他の語彙へのリンクが途切れてしまうことになるので、グループを構成できる大きさになるまでは、対立軸をまとめていく必要がある。

今後の方針として以下のことが挙げられる。今回、反義の対立軸を手作業で付与したが、その際、反義対の見出しのみを観察することが多く、深い情報を付与することが難しかった。語義情報は紙媒体の反対語辞典やEDR概念辞書の概念語義を参照したが、EDR概念辞書の語義情報は反義性について十分考慮されて付与されているわけではないことや、反対語辞典の参照が不便だったことがあり、対立軸の付与の作業がスムーズに行われなかった。今後、紙媒体の辞書から語義を電子化することも考慮すべきである。

応用として、以下のようなことがある。接辞と反義との関係を検証する。今回は「大きい:小さい」「乾く:湿る」などを「大:小」「乾:湿」などと、反義の対立軸を比較的簡単な形で表した。反義の対立軸を接辞として語彙に付与することで、新たな反義が形成できると思われる。また、反義概念は類義概念と類似することが多いため、反義の対立軸を類義概念の差異成分として扱うことができれば、類義概念の取り扱いが容易になると考えられ、対立軸などの反義表現が、発想の転換や連想記憶に繋がる可能性がある。

謝辞

本研究の一部は、(財)山形県企業振興公社との共同研究として行われました。元顧問の鎌田仁先生はじめ関係者に深く感謝致します。

参考文献

- [1] 池原悟, 宮崎正弘, 白井論, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦: 日本語語彙大系, 岩波書店(1999).
- [2] 荻野孝野, 大久保佳子, 横山晶一, 大島玲子: 概念体系における反義概念の検討, 言語処理学会第6 回年次大会論文集, 419-422 pp. (2000).
- [3] (株) 日本電子化辞書研究所: EDR 電子化辞書第2.0 版.
- [4] 高橋孝典, 横山晶一, 西原典孝: 反義の対立軸を用いた語彙検索システム, 平成14年度電気関係学会東北支部連合大会(2002)
- [5] 村木新次郎: 対義語の輪郭と条件, 日本語学 VOL.6, 6月号, 明治書院(1987).
- [6] 反対語対照語編纂委員会: 反対語対照語辞典, 柏書房(1998).