

比喩の特徴発見における概念辞書の利用*

7H-2

今井 豊† 石崎 俊‡
慶應義塾大学 政策・メディア研究科§

1 はじめに

従来の概念辞書には、上位一下位概念のような概念の包含関係を表す情報や、部分-全体関係のような概念の隣接関係を表す情報は記述されているが、異なる概念間の類似性に関する情報が不足している。比喩理解のためには、任意に与えられた二つの概念の類似性を表現する共通の顕著な特徴を的確に発見する方法を考える必要がある。共通の顕著な特徴は一つとは限らず、複数個存在する場合も多い。また、概念間の類似性が大きい比喩ほど多くの情報を伝達するため、理解が容易な比喩といえる[1]。

そこで本論文では、二つの概念 T (Topic, 被喩辞) と V (Vehicle, 喩辞) に共通の顕著な特徴を、COFFS (Common Features Finder System) というコネクショニストモデルを用いたシステムによって複数個抽出し、特徴についての収束値の重みつき和を類似性の大きさとして、比喩文「 T は V だ」の理解容易性、比喩の良さと関連づけることを検討する。

2 比喩の特徴発見の方法

比喩における意味は、カテゴリー的意味、情緒・感覚的意味、スクリプト的意味に分類することができる[1]。このうち、カテゴリー的意味に基づいて定量的な概念空間を構成する手法としてMDS法があげられ、情緒・感覚的意味に基づく定量的な意味空間を導入する手法としてはSD法が有用であると考えられる[2]。比喩文「 T は V だ」における二つの概念 T と V の類似性は、主に情緒・感覚的意味を表す特徴 (feature) によって説明されると仮定し、SD法の実験を行った。具体的には、弱い→強い、機械的→人間的などの両極状態概念対を35個 (ペア) 用意し、すべての概念に共通の特徴の束とした。分析対象の概念として、コンピュータ、ハッカー、海など66個を用意し、上記特徴を-3から+3までの7段階で被験者20人に評定させた。第1回目の実

験では、概念を一つずつ独立で提示して35個の特徴について評定し、第2回目の実験では比喩文中で評定した。概念を独立に提示された第1回目の実験結果をもとに、比喩特徴発見システムCOFFSによって、任意に与えられた二つの概念 T と V に共通の顕著な特徴を抽出する[3]。図1に、COFFSの構造を示す。featureの個数を m とするとき、 $m+2$ 個の節点を持ち、 f_1 から f_m は feature の節点を表し、 T と V はそれぞれ被喩辞と喩辞の節点を表す。

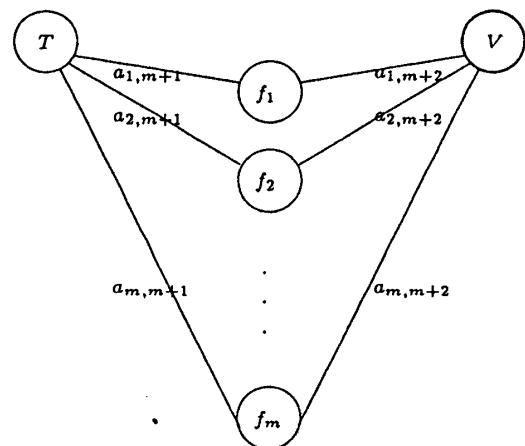


図1 COFFSの構造

このネットワークを用いて、次のような定式化を行う[4]。

$$x(t+1) = A * x(t) + c \quad (t=0, 1, 2, \dots) \quad (1)$$

$$x(0) = c \quad (2)$$

ここで $x(t)$ は節点 ($f_1, f_2, \dots, f_m, T, V$) の時刻 t における活性値を表す。活性値が大きいほど顕著な特徴を示すと考えられる。

要素数 $m+2$ の定常入力 c の i 番目 ($i=1, 2, \dots, m$) の要素には、 f_i に関する T と V の分布の形状の一致度が反映される。また、 T に相当する $m+1$ 番目の要素は1、 V に相当する $m+2$ 番目の要素は2とした。それは、比喩文「 T は V だ」において T よりも V の顕現性の方が重要であるという理由による。そのために図1では、 V の節点をほんの少し太い線で表した。

$(m+2) \times (m+2)$ の正方行列である遷移行列 A には評定平均値の絶対値が反映され、その要素は次のようにし

*Usage of Concept Dictionary in Finding Common Features of Metaphor

†Yutaka IMAI

‡Shun ISHIZAKI

§Graduate School of Media and Governance, Keio University

で決める。「TはVだ」が与えられたとき、まず $f_i (i = 1, 2, \dots, m)$ に対する T の評定平均値 $M_{T,i}$ と V の評定平均値 $M_{V,i}$ を求める。両者が同符号である場合には、 $a_{i,m+1} = a_{m+1,i} = |M_{T,i}|$ 、 $a_{i,m+2} = a_{m+2,i} = |M_{V,i}|$ とし、異符号である場合には、 $a_{i,m+1} = a_{m+1,i} = 0$ 、 $a_{i,m+2} = a_{m+2,i} = 0$ とする。 f_i と $f_j (i \neq j)$ は結合していないから、 $a_{i,j} = 0 (i \neq j; i, j = 1, 2, \dots, m)$ とし、T と V も同様の理由で $a_{m+1,m+2} = a_{m+2,m+1} = 0$ とする。また自分自身との結合があると考え、対角成分をすべて 1 にする。さらに収束のために正規化も行う。

(1)(2) に従って反復計算し、 x が収束したとき、 f_1 から f_m の収束値を調べ、活性値の大きい上位 5 個の feature を比喩文「TはVだ」における T と V の類似性を表現するものと判断する。さらに、上位 5 個の feature の正規化された収束値に対し、 $[1 - (T \text{ の評定値 } 0 \text{ の相対頻度})] \times [1 - (V \text{ の評定値 } 0 \text{ の相対頻度})]$ を各々掛け、5 個をあらためて足し合わせた重みつき合計値 TOTAL を計算する。この重みづけによって、重要度が低く真中で尖った分布パターンを除去することが可能である。TOTAL は、比喩の良さに関するパラメータとして用いる。

3 実行結果と考察

$m = 35$ のときの実行例を上位 5 個の TOTAL の大きい順に表 1 にまとめた。また、図 2 は、SD 法の平均化

表 1 比喩の解析例 (12 例中、典型的なもの 6 例)

例文 TはVだ	活性値の大きい feature 上位 5 個	上位 5 個 TOTAL	MDS 距離
ハッカーは 手品師だ	上手な、応用的、成熟した、人工の、難しい	1.25	1.09
コンピュータは 奴隷だ	重い、暗い、人工の、多い、冷たい	1.11	1.55
交通渋滞は 海だ	大きい、深い、多い、継続的、連続的	1.07	1.50
テクノロジーは 胎児だ	複雑な、新しい、尊い、良い、本当の	0.97	1.79
発明者は 開拓者だ	難しい、良い、広い、人間的、強い	0.88	0.41
専門家は エイリアンだ	強い、難しい、賢い、冷たい、暗い	0.82	0.92

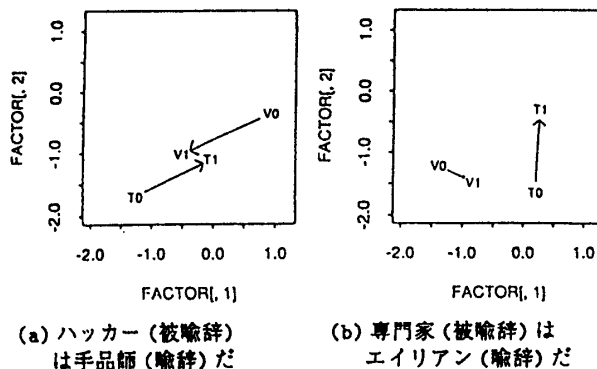


図 2 喩辞と被喩辞の因子分析結果におけるダイナミクス

されたデータに因子分析 (主因子法) を行い、第 1 因子と第 2 因子の空間で表示したものである。(a) には表 1 の中で TOTAL が最大の「ハッカーは手品師だ」について、(b) には表 1 の中で TOTAL が最小の「専門家はエイリアンだ」について概念の動きが示されている。独立提示で T と V を評定したときは T0 と V0 で、比喩文「TはVだ」中で T と V を評定したときは T1 と V1 で表示されている。(a) では「ハッカー」と「手品師」の T1 と V1 が十分に近づいているが、(b) では「専門家」と「エイリアン」の T1 と V1 は方向も異なり近づいていないのがわかる。TOTAL の大きさと比喩の良さとはある程度結びつくと考えられる。

表 1 には、SD 法の実験に用いた 66 個の概念を対象にして行った MDS 法の実験から得られた概念間の定量的な距離も記してある。これらのパラメータ間の相関は低く (相関係数は 0.28)、比喩の良さを MDS 法の距離データだけから明示することは難しい。しかし、TOTAL がほぼ同じ値であるならば、MDS 法の距離が大きい方がカテゴリー的に離れているということになり、比喩としての面白さ・斬新さが効いて良い比喩になると考えられる [1]。また、TOTAL の値が小さくて、MDS 法の距離が大きいならば、アノマラスな文であると考えられる。以上のような規則を今後は検証する予定である。

4 まとめと今後の課題

両極状態概念の束を用いて比喩における類似性を説明することの有効性が示された。しかし 35 個の特徴だけでは類似性の記述力としては必ずしも十分とはいえない。また、大規模な概念辞書を想定したとき、必要な特徴の数のオーダーを予測することが重要である。概念の特徴や動作環境などの新たな概念情報を効率的に収集できる連想実験システムが構築されているので [5]、これを比喩理解に役立てることを考えている。

本研究は文部省科研費重点領域「音声対話」の助成を受けて行われた。

参考文献

- [1] 楠見孝: "比喩の処理過程と意味構造", 風間書房 (1995).
- [2] 石崎俊, 田中茂範, 今井むつみ: "柔軟な意味解析のための概念空間の定量化", 情報処理学会自然言語処理研究会資料 100-3 (1994).
- [3] 今井豊, 石崎俊: "比喩理解のための概念辞書における意味構造記述の検討", 情報処理学会第 50 回全国大会論文集 (3), pp127-128 (1995).
- [4] Hasida, K. et al.: "A Connectionist Approach to the Generation of Abstracts", Natural Language Generation, Martinus Nijhoff Publishers, pp.149-156 (1987).
- [5] 大熊智子, 石崎俊: "連想実験に基づく概念辞書の構築と検索", 情報処理学会第 51 回全国大会論文集 (1995).