

比喩を含む言語理解における視点の役割

A STUDY ON THE ROLE OF VIEWS IN UNDERSTANDING METAPHOR

岩山 真, 徳永 健伸, 田中 穂積

IWAYAMA Makoto, TOKUNAGA Takenobu and TANAKA Hozumi

東京工業大学 工学部

Tokyo Institute of Technology

Abstract

This paper presents a method for understanding natural language including metaphorical expressions. *The view expression* is introduced to regard a concept as another concept. The former is called *the target concept*, and the latter is called *the source concept*. Throughout the paper, the term "understanding" is used in the very restricted meaning, which is to calculate properties of a concept. The understanding procedure consists of two stages. The first stage involves the decision of properties which may be transferred from the source concept to the target concept. In the second stage, the variation of the target concept is calculated. We decide the properties to be transferred by means of *salience*. Salience is calculated by the amount of information which properties have and the difference among the source concept and its sibling concepts. Unlike the methods which have been proposed so far, we do not distinguish metaphors, literals and anomalies. This is also the important feature of our approach.

1 はじめに

日常, 我々が使用する言語には比喩的な表現が必ずといってよいほど含まれている。言語表現のほとんどは比喩的であり, 比喩は人間の認識に深く関わっているという指摘もある[8]。そのため, 計算機によって自然言語理解をおこなうためには, 比喩の問題は避けては通れない問題となる。

本稿では, 比喩を含む言語理解における視点の役割について考察する。まず, 「概念Aを概念Bという視点から眺める」ことを表現するための道具として視点表現を導入する。ここで, 概念Aをtarget概念, 概念Bをsource概念という。本稿でいう理解とは, 視点表現で表された概念の持つ性質を計算するという限られた意味である。理解は, (1) target概念に移されるsource概念の性質の計算, (2) source概念から性質が移されることによるtarget概念の変化の計算の2段階からなる。target概念に移す性質の計算は, 顕現性に基づいておこなうが, 顕現性は, 性質の持つ情報量と, 差異性によって定義する。例として, 形容詞で修飾された名詞句, 比喩的な表現について理解の過程を説明する。また, 本方法では, 従来のように比喩的な表現, リテラルな表現, 非文に区別しない点も重要な特徴である。

2 視点表現の理解

本節では, 視点表現の理解を行なうモデルについて説明する。視点表現は $\ast(T) \setminus \ast(S)$ で表現し, $\ast(T)$ をtarget概念, $\ast(S)$ をsource概念と呼ぶ。

まず2.1節で, 概念を性質の集合として形式的に定義する。次に, 2.2節で, 概念の性質の顕現性を定義する。以上の準備の後, 2.3節で視点表現の理解を扱う概念の変化モデルの枠組を示す。

2.1 概念の表現形式

定義 1 概念

【概念Cは性質 S_i の集合からなり, $\ast(C) = \{S_1, S_2, \dots, S_n\}$ で表現する。性質 S_i は属性名 a_i と属性値集合 V_i の対であり, $S_i = a_i : V_i$ で表現する。ここで, 属性値集合 V_i は, 対応する属性名 a_i の値としてとり得る全ての属性値 $v_{i,j}$ とその頻度 $w_{i,j}$ の対の集合で, $V_i = \{v_{i,1} \# w_{i,1}, \dots, v_{i,m} \# w_{i,m}\}$ で表現する。今, 属性値集合 V_i の各要素が排反であるとする。 $\sum_{j=1}^m w_{i,j} = 1$ である。また, 最大の頻度を持つ属性値を最尤属性値と呼び $v_{i,max}$ で, 属性名と最尤属性値の対を最尤性質と呼び $S_{i,max} = a_i : v_{i,max}$ で表す。】

例 1 概念*(りんご)の記述例

$$\ast(\text{りんご}) = \left\{ \begin{array}{l} \text{色:} \left\{ \begin{array}{l} \text{赤} \# 0.8 \\ \text{緑} \# 0.2 \end{array} \right\} \\ \text{外形:} \left\{ \begin{array}{l} \text{球状} \# 0.9 \\ \text{円柱状} \# 0.1 \end{array} \right\} \\ \vdots \end{array} \right\}$$

最尤属性値はそれぞれ「赤」, 「球状」であり, 最尤性質は「色: 赤」, 「形状: 球状」である。

2.2 顕現性の計算

概念の最尤性質には顕現性が付加される。顕現性は視点表現の理解の際, source概念の性質のtarget概念への移され易さの度合になる。

顕現性には, 我々の文化・経験や記憶の構造によるものもあるが, 本稿では顕現性の要因として, 「属性値集合の持つ情報量」, 「他の兄弟概念との差異性」の二つを扱う。

2.2.1 属性値集合の持つ情報量

まず、顕現性の第一の要因として、属性値集合の持つ情報量を定義する。これは最尤性質の選出のしやすさに関係する要因である。最尤性質は、属性値集合のうち頻度が最大の属性値をその属性値(最尤属性値)として持つ。ここで、最尤属性値を属性値集合から選ぶ際、頻度がばらついている集合から選ぶより、頻度が集中している集合から選ぶ方がその最尤属性値は尤もらしく顕現性が高い。例えば、{0.6, 0.1, 0.1, 0.1, 0.1}の集合から最大値の0.6を選ぶほうが、{0.6, 0.4}の集合から選ぶより尤もらしい。即ち、属性値集合の頻度のばらつき具合は、顕現性の要因となる。本稿では属性値集合の頻度を生起確率とみなすことにより、エントロピーを計算し、属性値集合の情報量を求める。

定義 2 属性値集合の持つ情報量

『属性値集合 $v_i = \{v_{i,1} \# w_{i,1}, \dots, v_{i,m} \# w_{i,m}\}$ の情報量 $r(v_i)$ は次式で定義される。

$$r(v_i) = 1 - h(v_i)$$

$r(v_i)$ は、情報理論での冗長度に相当する[1]。ここで、 $h(v_i)$ は情報源の相対エントロピーに相当するもので、

$$h(v_i) = \begin{cases} 0 & \text{if } m = 1 \\ \frac{\sum_{j=1}^m w_{i,j} \log_2 \frac{1}{w_{i,j}}}{\log_2 m} & \text{otherwise} \end{cases}$$

である。

2.2.2 兄弟概念に対する最尤性質の差異度

顕現性の第二の要因である最尤性質の差異度は、他の概念との関係(差異性)から定まる要因である。例えば、*(男)*(狼)では、なぜ他の動物でなく*(狼)という視点から*(男)を見るのだろうか。この理由として、*(狼)が他の動物と異なっている性質を移したいからだ、ということが一つ考えられる。即ち、*(狼)の性質のうち、他の兄弟概念(動物)との差異性が大きい性質ほど顕現性が高く、移され易いといえる。

定義 3 最尤性質の差異度

『概念*(C)のある性質を $S_i = a_i : V_i$ とすると、その最尤性質 $S_{i,max} = a_i : v_{i,max}$ の差異度 $d(*(C), S_{i,max})$ は次式で計算される。

$$d(*(C), S_{i,max}) = \frac{r(V_i)}{\sum_{*(C_j) \in (*(C) \text{ の兄弟概念 } \cup *(C))} r'(*(C_j), S_{i,max})}$$

ここで、*(C)の兄弟概念とは、概念間の上位/下位関係において*(C)と直接の上位概念を共有する概念のことをいう。 $r'(*(C_j), S_{i,max})$ は次式で計算される。

$$r'(*(C_j), S_{i,max}) = \begin{cases} r(V_k) & \text{if } S_{i,max} = S_{k,max} \text{ なる } S_{k,max} \text{ が } \\ & *(C_j) \text{ に存在する} \\ 0 & \text{otherwise} \end{cases}$$

2.2.3 最尤性質の静的な顕現性

以上の二つの要因から顕現性が計算できる。

定義 4 最尤性質の静的な顕現性

『概念*(C)の最尤性質 $S_{i,max}$ の静的な顕現性は次式で計算される。

$$\text{salience}(*(C), S_{i,max}) = \text{情報量 } r(V_i) \times \text{差異度 } d(*(C), S_{i,max})$$

例 2 顕現性の計算

例として、*(植物)の直接の下位概念

{*(チューリップ), *(カーネーション), *(ばたん), *(桜)}

を考える。この集合の要素が互いに兄弟概念となる。属性名「花の色」に関するそれぞれの属性値集合を

$$V_{\text{チューリップの花の色}} = \{ \text{赤} \# 0.35, \text{白} \# 0.3, \text{黄色} \# 0.2, \text{ピンク} \# 0.1, \text{紫} \# 0.05 \}$$

$$V_{\text{ばたんの花の色}} = \{ \text{赤} \# 1 \}$$

$$V_{\text{カーネーションの花の色}} = \{ \text{赤} \# 0.97, \text{白} \# 0.01, \text{ピンク} \# 0.02 \}$$

$$V_{\text{桜の花の色}} = \{ \text{ピンク} \# 0.95, \text{白} \# 0.05 \}$$

とすると、それぞれの最尤性質、情報量は次のように計算される。

- *(チューリップ) 「花の色: 赤」
- *(ばたん) 「花の色: 赤」
- *(カーネーション) 「花の色: 赤」
- *(桜) 「花の色: ピンク」

$$r(V_{\text{チューリップの花の色}}) = 1 - \frac{0.35 \times \log_2 \frac{1}{0.35} + \dots + 0.05 \times \log_2 \frac{1}{0.05}}{\log_2 5} = 0.111143$$

$$r(V_{\text{ばたんの花の色}}) = 1$$

$$r(V_{\text{カーネーションの花の色}}) = 0.859971$$

$$r(V_{\text{桜の花の色}}) = 0.713603$$

差異度は、

$$d(*(チューリップ), \text{花の色: 赤}) = \frac{0.111143}{1 + 0.359971 + 0.111143} = 0.056386$$

$$d(*(ばたん), \text{花の色: 赤}) = 0.507327$$

$$d(*(カーネーション), \text{花の色: 赤}) = 0.436287$$

$$d(*(桜), \text{花の色: ピンク}) = 1$$

で計算される。よってそれぞれの性質の顕現性は、

$$\text{salience}(*(チューリップ), \text{花の色: 赤}) = 0.111143 \times 0.056386 = 0.006267$$

$$\text{salience}(*(ばたん), \text{花の色: 赤}) = 0.507327$$

$$\text{salience}(*(カーネーション), \text{花の色: 赤}) = 0.379557$$

$$\text{salience}(*(桜), \text{花の色: ピンク}) = 0.713603$$

となる。

2.3 視点表現の理解モデル

本節では、前節までに定義された「概念」、「顕現性」を用いて、概念の変化モデルを示す。このモデルにより、視点表現の理解が可能になる。視点表現 $*(T) \setminus *(S)$ の理解とは

「*(S)の複数の最尤性質を可能なら*(T)に移す」

という操作である。その際、source 概念の最尤性質でも、「顕現性が高い最尤性質は移され易い」。また、ある最尤性質を target 概念に移すということは、「移される最尤性質に相当する target 概念の性質の頻度・差異度、即ち顕現性が変化する」ことに相当する。

概念の性質の顕現性(頻度・差異度)は、視点の他にも形容詞による修飾や文脈的要素(焦点、話題など)によっても変化する。本稿では、視点表現の理解も含め、このような効果を動的な概念の変化と呼び、概念の意味の変化を顕現性という観点から統一的に扱う。以下に概念の変化モデルの概要を示す。

[Case1: 性質(属性名:属性値)の指定による変化]

形容詞などによる修飾に相当する。例えば、「緑のりんご」では、性質「色: 緑」を指定し、頻度・差異度を変更する。

[Case2: 属性名の指定による変化]

話題、焦点などに相当する。この場合、差異度のみ変更する。例えば、*(象)を考えると、「動き」を話題にしている文脈では、最尤性質「動き:slow」の差異度、即ち顕現性が増すし、「重さ」を話題にしている文脈では最尤性質「重さ:大」の差異度が増す。

[Case3: source 概念の指定による変化]

概念の視点表現に相当する。

1. まず、source 概念の最尤性質の顕現性をこのモデルで計算する。
2. 1で計算された source 概念の最尤性質のうち、適当なしきい値以上のものを抽出する。もし、同じ属性名をもつ最尤性質が複数個存在する場合、顕現性が最大なもののみを選ぶ。属性値は互いに排反なため(定義1参照)、同じ属性名で複数の属性値が同時に target 概念に移されることはないからである。
3. 2で抽出されたそれぞれの最尤性質が指定されたものとして target 概念の頻度・差異度を変更する。

次に、概念の変化モデルにおいて基本操作となる、頻度・差異度の変更について定義する。

定義 5 頻度の変更

『属性名 a と属性値 v が指定されたとき、概念 $*(C)$ は以下の手順で変化する。

1. $*(C)$ の性質 $S_i = a_i : V_i$ で $a_i = a$ なる性質が存在すれば 2 へ、存在しなければ $*(C)$ は変化しない。

2. 属性値集合 $V_i = \{v_{i,1} \# w_{i,1}, \dots, v_{i,m} \# w_{i,m}\}$ で $v_{i,j} = v$ なる属性値が存在し、 $w_{i,j} \neq 0$ ならば 3 へ、さもなければ $*(C)$ は変化しない。
3. $v_{i,j}$ の頻度 $w_{i,j}$ の値を 1 に、その他の頻度の値全てを 0 にする。

定義 6 差異度の変更

『属性名 a が指定されたとき、概念 $*(C)$ の性質 $S_i = a_i : V_i$ で $a_i = a$ なるものが存在すれば、性質 S_i の最尤性質 $S_{i,max}$ の差異度は 1 になる。

3 比喩理解への応用: 視点表現の理解の例

本節では、2 節で示した概念の変化モデルの応用例として比喩理解を扱う。ただし、本章で対象とする比喩は、直喩・隠喩の一部である。よって以降、本稿では「比喩」と言うことで「直喩・隠喩」を指す。また、直喩、隠喩の区別は行なわない。

3.1 比喩理解の概要

概念の変化モデルは、概念の意味の変化を顕現性という観点から扱うモデルであった。変化の要因としては、形容詞による修飾、焦点、話題、視点などが考えられるが、比喩もその一つである。比喩は「ある概念を別の概念でたとえる」という操作であるため、我々のモデルでは視点表現の理解に相当する。即ち、たとえる概念は source 概念、たとえられる概念は target 概念に対応し、「比喩を理解する」ことは、「たとえる概念の最尤性質をたとえられる概念に移す」ことに相当する。

例として

(1) りんごのような類

という比喩を解析する。まず、(1)から視点表現 $*(類) \setminus *(りんご)$ を抽出して、概念の変化モデルに入力する。そこではまず、*(りんご)の最尤性質の内、顕現性があるしきい値以上のものを抽出する。今、顕現性がしきい値以上の*(りんご)の最尤性質として、「色: 赤」、「表面: 滑らか」、「外形: 球状」が抽出されたとする。これらの最尤性質が指定された性質(属性名:属性値)になる。次に、指定された性質により、可能なら、*(類)の性質の頻度・差異度を変更する。ここで、概念*(類)の一部を示すと、

$$*(類) = \left\{ \begin{array}{l} \text{色:} \quad \left\{ \begin{array}{l} \text{肌色} \# 0.8 \\ \text{赤} \# 0.1 \\ \text{白} \# 0.1 \end{array} \right\} \\ \text{外形:} \quad \left\{ \begin{array}{l} \text{丸みを帯びた平面状} \# 1 \\ \text{滑らか} \# 0.8 \end{array} \right\} \\ \text{表面:} \quad \left\{ \begin{array}{l} \text{でこぼこ} \# 0.2 \end{array} \right\} \\ \vdots \end{array} \right.$$

である。まず、頻度の変更であるが、定義 5 により、指定された属性名「色」、「表面」、「外形」が*(類)に存在するかどうか調べる。この場合、全て存在する。次に、指定された属性値が存在しかどうか調べる。ここで、属性値「赤」、「滑らか」は存在するが「球状」は存在しないため、*(りんご)の最尤性質「外形:球状」は*(類)に移されない。他の指定された属性値については、それぞれに対応する頻度を 1 に変更する。よって、*(類) \setminus *(りんご)は

$$*(顔)\cdot(りんご) = \left\{ \begin{array}{l} \text{色} \\ \text{外形} \\ \text{表面} \end{array} \left\{ \begin{array}{l} \text{肌色}\#0 \\ \text{赤}\#1 \\ \text{白}\#0 \\ \text{丸みを帯びた平面状}\#1 \\ \text{滑らか}\#1 \\ \text{でこぼこ}\#0 \end{array} \right\} \right\}$$

となる。ここで、*(りんご)の最尤性質「色:赤」、「表面:滑らか」が*(顔)に移された結果、*(顔)の最尤性質「色:肌色」自体も「色:赤」に変化し、属性値集合の情報量もそれぞれ0.418329, 0.278072から1に増加していることに注意されたい。更に、定義6による差異度の変更により、*(顔)\cdot(りんご)の最尤性質「色:赤」、「表面:滑らか」の差異度も1に変化する。よって、最終的にこれらの最尤性質の顕現性は $1 \times 1 = 1$ と最大値に変化する。

3.2 リテラルな文、比喩、非文: その連続的な理解

従来の計算機による比喩解析では、リテラルな文・比喩・非文を明確に区別して扱っていた[4,6,5]。そこでは、まず、言語表現がリテラルな文か比喩か非文かを区別し、それぞれに対する解析を行なう。ところが、実際の言語表現をみてもわかるように、これらの境界は曖昧であり、人間にも区別が困難な場合が多い。我々は、視点表現の理解の結果、最尤性質の移された割合から、文章がリテラルに近いか非文に近いかを連続的に判断できると考える。従来の方式のように、ア・プリオリにリテラルな文・比喩・非文が決まるのではなく、それらは、文を理解してみても初めて分かる副次的なものである。

定義7 視点表現の理解の容易性

「視点表現 $*(T) \setminus *(S)$ の理解の容易性を以下の式で計算する。

$$\frac{\sum_{*(S)から*(T)へ実際に移された最尤性質の顕現性}}{\sum_{*(S)から*(T)へ移される候補となった最尤性質の顕現性}}$$

視点表現の理解の容易性を用いて以下の仮定をおく。

「視点表現の理解の容易性が大きいほど、その視点表現を含む言語表現はリテラルな文に近く、小さいほど非文に近い。」

以上によりリテラルな文と非文を視点表現の理解の結果として計ることができる。それでは比喩はどこに位置するのだろうか。補見は[3]で、たとえる概念とたとえられる概念のカテゴリ的な距離が遠いほど比喩は理解し易く、情緒・感覚的な距離が近いほど比喩は面白く、比喩の良さはその両者によって決まると述べている。我々のモデルで、比喩の面白さ、良さがどこに相当するのかは今後の課題である。

4 おわりに

本稿では、比喩理解を目標として、視点表現の理解を扱うモデルを形式的に定義した。また、このモデルを用いて、比喩の理解が部分的に可能になることを示した。

しかし、本稿で示した概念の変化モデルで扱える比喩の範囲はまだ狭い。例えば、

(2) アイデアが開花する。 $[(\text{アイデア}) \setminus (\text{花})]$

などは扱えない。なぜなら、(2)において、*(花)の最尤性質がそのまま*(アイデア)に移せないからである。例えば、最尤性質「色彩:きれい」を指定して*(アイデア)を変更しようとしても、*(アイデア)には属性名「色彩」を持つ性質がないため「色彩:きれい」は移されない。ところが我々は、実際は、属性名「色彩」を「明瞭度」などの属性名に変換して理解していると考えられる。このような現象は、共感覚と呼ばれる。本稿のモデルに共感覚を導入して、異なる様相間の比喩を扱うのは今後の課題である。

参考文献

- [1] 宮川洋. 情報理論. 電子通信大学講座第39巻, コロナ社, 1979.
- [2] 山梨正明. 比喩と理解. 東京大学出版会, 1988.
- [3] 楠見孝. 比喩表現の理解過程. 表現研究, 46, 1987.
- [4] 土井晃一, 田中英彦. 隠喩理解—隠喩の検出—. 日本認知科学会 RBI研究会, SIGR&I[87(2):3-10, 1988.
- [5] D. Fass. Metonymy and metaphor: what's the difference? In *COLING-88*, 1988.
- [6] D. Fass and Y. Wilks. Preference semantics, ill-formedness, and metaphor. *AJCL*, 9(3-4):178-187, 1983.
- [7] G. Lakoff. *Woman, Fire, and Dangerous Things*. Chicago University Press, 1987.
- [8] G. Lakoff and M. Johnson. *Metaphors we live by*. Chicago University Press, 1980.
- [9] E. Rosch and C. B. Mervis. Family resemblances: studies in the internal structure of categories. *Cognitive Psychology*, 7:573-605, 1975.
- [10] E. E. Smith, D. N. Osherson, L. Rips, and M. Keane. Combining prototypes: a selective modification model. *Cognitive Science*, 12(4):485-527, 1988.
- [11] P. H. Winston. Learning by creating and justifying transfer frames. *Artificial Intelligence*, 10(2):147-172, 1978.