

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

梗概

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、顕現性を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。

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[13]。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。

比喩は、ある概念を別の概念でたとえる操作である¹。人間は、比喩を用いることによって、抽象的な概念をより具体的な概念を介して認識できる。また、たとえるという操作によって、実世界の森羅万象の現象を有限の言語で表現することが可能になる。と

ころが、たとえるという操作には推論・連想・記憶などの要素が関わっているため、その計算機による理解は大変困難である。本稿では、視点表現の理解という枠組を用いて比喩理解が部分的に可能になることを示すとともに、その問題点を指摘する。

視点表現は、「概念 A を概念 B という視点から眺める」ことを表現する道具である。ここで、概念 A を target 概念、概念 B を source 概念と呼ぶ。比喩では、target 概念が、たとえられる概念、source 概念が、たとえる概念に相当する。視点表現を理解するとは、「source 概念の性質を、可能なら、target 概念に移す」ことによって、視点表現で表された概念の持つ性質を計算することである。ここで、(1) source 概念のどの性質が target 概念に移されやすいか、(2) 性質が移された結果 target 概念はどのように変化するか、が問題と

¹ 本研究で扱う比喩は直喩・隠喩と呼ばれているものに限る。

なる。本方式では、顕現性という指標でこれらの問題を解決する。即ち、source 概念において顕現性が高い性質ほど target 概念に移されやすく、移された性質は target 概念において顕現性が高くなる。

これまでも比喩を計算機で扱おうとした試みはいくつかある[7,9,8]。これらは、比喩を解析するにあたり、まず、言語表現を、リテラルな文、比喩、非文に分類していた。比喩とリテラルな文を分けるにあたり、選択制限や状況の知識などを用いている。また、比喩と非文の区別はヒューリスティクスにより行なっている。

我々の方法は「リテラルな文・比喩・非文を区別しない」という点でこれまでの研究とは大きく異なっている。実際の言語表現を見てみると、リテラルな文・比喩・非文の境界は曖昧である。また、人間は、その三者を事前に区別して異なる方法で解析しているとは思えない。文が、リテラルな文か比喩か非文かは、理解してみても初めて分かるものであり事前に決定することは不可能である。本稿では、言語表現が与えられたら、まず、視点表現を抽出し、次に、視点表現の理解を行なう。視点表現の理解は、「source 概念の性質を、可能なら、target 概念に移す」という簡単な原則に従う。そして、その移された割合を視点表現の理解の容易性とし、それによりリテラルな文、比喩、非文が判断できると考える。このような方式により、リテラルな文・比喩・非文を区別しない連続的な解析が可能になる。また、従来の解析では、非文と判断されたら何も結果が得られないが、本方式では、限りなく非文に近い文も何らかの結果が得られる。

以下、2 章では、視点表現の理解を扱う概念の変化モデルを示す。まず、準備として、2.1 節で概念を、2.2 節で概念の性質の顕現性を形式的に定義する。本稿では、顕現性を、性質の持つ情報量と差異性によって定義する。2.3 節では、概念の顕現性の変化モデルの概要を説明し、その適用例として形容詞で修飾された名詞句の理解過程を説明する。3 章では、2 章で示した概念の変化モデルを比喩的な表現の理解に応用する過程について説明する。また、視点表現の理解の容易性の定義を行なう。最後に 4 章で、本方式の問題点と今後の課題について述べる。

2 視点表現の理解

本章では、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 $*(T) \setminus *(S)$ で表現する。 $*(T)$ を target 概念、 $*(S)$ を source 概念と呼び、全体としては「source 概念という視点から見た target 概念」という意味を持つ。ここで、視点表現の理解とは、 $*(T) \setminus *(S)$ の性質を求めることに相当する。

例えば、

(1) 望遠鏡で彼女を見た。

では、「道具としての望遠鏡」、 $*(望遠鏡) \setminus *(道具)$ が、

(2) 望遠鏡を持った彼女を見た。

では、「物としての望遠鏡」、 $*(望遠鏡) \setminus *(物)$ が視点表現として抽出される。

従来の我々の視点表現[5]は、target 概念と source 概念の間に上位/下位関係が成立している場合に限られ、視点表現は多重継承による概念の曖昧性を区別するために用いられていた。そのため、

視点表現の理解は性質の継承によって簡単に行なえるが、比喩が扱えないという問題があった。比喩の多くは上位/下位関係がない概念間に「target 概念を source 概念としてみなす」という関係(これを比喩的な関係と呼ぶ)を動的に発見する操作だからである。本稿では、視点表現の理解において

「source 概念の性質を可能なら target 概念に移す」

という規則を用いることで、概念間の比喩的な関係を動的に理解する。

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。視点表現の理解の際、source 概念の性質にも移され易い性質や移され難い性質がある。本稿では、概念の性質の移され易さの度合を顕現性という指標で数値化する。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 の集合からなる。以後、概念を表す場合は概念名を括弧で囲み、その頭に * をつける。

$$*(C) = \{S_1, S_2, \dots, S_n\}$$

性質 S_i は属性名 a_i と属性値集合 V_i の対であり、

$$S_i = a_i : V_i$$

と表現する。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、

$$V_i = \{v_{i,1} \# w_{i,1}, \dots, v_{i,j} \# w_{i,j}, \dots, v_{i,m} \# w_{i,m}\}$$

と表現する。頻度は生起確率と見なすことができる。今、属性値集合 V_i の各要素が排反であるとすると、

$$\sum_{j=1}^m w_{i,j} = 1$$

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

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

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。

概念のこの定義は[20]を参考にしている。各々の属性名に対して、とり得る全ての属性値を羅列し、属性値ごとにその頻度を付加した点が、従来[5]の概念の定義と大きく異なる。頻度は概念の典型性を表現していると言える。例 1 の場合、*(りんご)は典型的に

「色」が「赤」く、「外形」が「球状」であると解釈できる。Smithらは[21]で、頻度の典型性としての性質を利用して、prototype理論[19,17]で提起されたカテゴリ化の典型性を説明している。

また、頻度は、属性名に対する属性値の生起確率ともみなせる。例1の場合、属性名「色」が「赤」、「緑」の*(りんご)はそれぞれ0.8、0.2の頻度で存在し、故に無作為に一個のりんごを抽出した場合、「色」が「赤」、「緑」である確率はそれぞれ0.8、0.2となる。頻度を生起確率とみなすことは、後に顕現性を計算するうえで役立つ。

ここで頻度の決定方法が問題となる。理想としては、概念のインスタンス全てに対して統計をとることが望ましいが、これは実際には不可能である。そこで、[21]では一定の被験者による投票値によって決めている。我々も、投票値から確率を計算し、その値を頻度として使用する。

以上で定義された概念間には様々な関係があるが、以下での説明の都合上、上位/下位関係を定義しておく。

定義2 概念間の上位/下位関係

『*(A)が*(B)の性質を継承できるとき、*(A)と*(B)の間に上位/下位関係があるという。ただし、*(B)が*(A)の上位概念である。上位/下位関係のことをISA関係ともいう。』

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。既に述べたように、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。また、顕現性は概念の典型性の度合ともなる。即ち、ある概念において顕現性が高い性質ほどその概念を特徴付ける性質となる。

本節では、最尤性質の顕現性を定義する。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、以下の二つを扱う。

1. 属性値集合の持つ情報量
2. 他の兄弟概念との差異性

1は属性値集合自体が持つ情報で、概念内で独立に定まる要因である。一方2は、他概念(兄弟概念)との関係から定まる要因である。以下で、それぞれの要因について形式的な定義を与え、これらの要因から顕現性を計算する方法を示す。

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いといえる。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。尤もらしければ典型性も高い。よって、属性値集合の頻度のばらつきの度合は、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3 属性値集合の持つ情報量

『属性値集合 $v_i = \{v_{i,1} \# w_{i,1}, v_{i,2} \# w_{i,2}, \dots, v_{i,m} \# w_{i,m}\}$ 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[2]。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、

$$h(v_i) = \begin{cases} 0 & \text{if } m = 1 \\ \frac{H(v_i)}{\log_2 m} & \text{otherwise} \end{cases}$$

である。 $H(v_i)$ はエントロピーに相当し、

$$H(v_i) = \sum_{j=1}^m w_{i,j} \log_2 \frac{1}{w_{i,j}}$$

である。

エントロピーとは、そもそも無秩序の度合を表す量である。よって、属性値の持つ情報量(冗長度)は属性値集合における属性値の頻度(生起確率)がばらついているほど低くなり、集中しているほど高くなる。

例2 属性値集合の持つ情報量の計算

*(チューリップ), *(桜), *(ばたん), *(カーネーション)の属性名「花の色」に対する属性値集合をそれぞれ、

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

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

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

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

とする。このとき、それぞれの属性値集合の情報量は、

$$\begin{aligned} r(V_{\text{チューリップの花の色}}) &= 1 - \frac{0.35 \times \log_2 \frac{1}{0.35} + 0.3 \times \log_2 \frac{1}{0.3} + \dots + 0.05 \times \log_2 \frac{1}{0.05}}{\log_2 5} \\ &= 0.111143 \end{aligned}$$

$$\begin{aligned} r(V_{\text{桜の花の色}}) &= 1 - \frac{0.95 \times \log_2 \frac{1}{0.95} + 0.05 \times \log_2 \frac{1}{0.05}}{\log_2 2} \\ &= 0.713603 \end{aligned}$$

$$\begin{aligned} r(V_{\text{ばたんの花の色}}) &= \dots \\ &= 1 \end{aligned}$$

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

となる。

(チューリップ)の属性「花の色」に対する属性値の頻度は、(桜)に対してばらついているため、情報量も少ない。よって、属性値集合の情報量のみ考えると、*(桜)の最尤性質「花の色:ピンク」の顕現性の方が*(チューリップ)の最尤性質「花の色:赤」の顕現性より高いことになる。実際、*(チューリップ)の「色」に関して我々は多くの種類を知っているため、「典型的なチューリップの花の色は何色か?」と問われると、「典型的な桜の花の色は何色か?」と問われた時ほどはっきりと答えにくい。

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、「source概念として何故ある概念を使うのか」に対する理由に関係する。即ち、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 4 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。

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

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。

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

』

例 3 差異度の計算

例として、*(植物)の直接の下位概念

{*(ばたん), *(カーネーション), *(チューリップ), *(桜)}

を考える。この集合の要素が互いに兄弟概念となる。「花の色」に関するそれぞれの最尤性質、情報量は例 2 より次のようになる。

*(ばたん)
「花の色:赤」 $r(V_{\text{ばたんの花の色}}) = 1$
*(カーネーション)
「花の色:赤」 $r(V_{\text{カーネーションの花の色}}) = 0.859971$
*(チューリップ)
「花の色:赤」 $r(V_{\text{チューリップの花の色}}) = 0.111143$
*(桜)
「花の色:ピンク」 $r(V_{\text{桜の花の色}}) = 0.713603$

よって、差異度は、

$$\begin{aligned} d(* (ばたん), \text{花の色:赤}) &= \frac{1}{1+0.859971+0.111143} = 0.507327 \\ d(* (カーネーション), \text{花の色:赤}) &= \frac{0.859971}{1+0.859971+0.111143} = 0.436287 \\ d(* (チューリップ), \text{花の色:赤}) &= \frac{0.111143}{1+0.859971+0.111143} = 0.056386 \\ d(* (桜), \text{花の色:ピンク}) &= \frac{0.713603}{0.713603} = 1 \end{aligned}$$

となる。

差異度は、同じ最尤性質を持つ兄弟概念の数が少ないほど、即ち、差異度の計算式において分母の要素数が小さいほど大きくなる。差異度のこのような性質は、「ある概念に特異な性質ほど相対的に際立つ」という直感に合っている。例 3 では、*(桜)の最尤性質「花の色:ピンク」と同じ最尤性質を持つ兄弟概念がないため、他の三つの概念に共通の最尤性質「花の色:赤」より差異度は大きくなり、この場合最大の 1 となる。また一般に、*(植物)から上位/下位関係により継承される性質の差異度は小さくなる。なぜなら、共通の上位概念からの継承により、全ての兄弟概念がその性質を持つからである。

また、差異度の計算式の要素に属性値集合の情報量を用いることで、よりきめ細かい計算結果が得られる。例えば、*(チューリップ)の属性名「花の色」にたいする属性値集合の情報量は例 2 で示したように、他と比べ小さいため、差異度の計算のうえでもあまり影響を与えない。試みに、 $d(* (ばたん), \text{花の色:赤})$ から $r(V_{\text{チューリップの花の色}})$ を除いても 0.537643 にしか変化しないが、 $r(V_{\text{カーネーションの花の色}})$ を除くと 0.899974 と二倍近く変化する。即ち、「強力な対立候補が少ないほど差異度もそれだけ大きい」ことが説明できる。更に、同じ分母でも分子の情報量が大きいほど差異度も大きくなるため、情報量による差異度の差別化が可能になる。例 3 では、同じ最尤性質「花の色:赤」を持つ*(ばたん)、*(カーネーション)、*(チューリップ)でも、属性名「花の色」に対する属性値集合の情報量が最大の*(ばたん)の差異度が最大になる。

差異度は、Smith 等の diagnosticity(診断度)[20]に相当する。また、Winston も[21]で、類推学習での transfer frame(変換フレーム)の構成に兄弟概念との差異性を利用している。Winston は、簡単な規則によって、他の兄弟概念と異なる性質を見つけている。そのため差異性の判断が二値的であり、我々の差異度でいうと 1 に相当する性質しか抽出できない。一方、Smith 等の診断度は二値的ではないが、計算が複雑なため²、動的に変化する頻度情報から動的に診断度を計算するのは困難である。それに比べ、我々の計算方法は、比較的簡単でかつ上で述べたような直感に合っている。

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。

² χ^2 を計算する。

定義 5 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。

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

例 4 顕現性の計算

$$\begin{aligned} \text{salience}(* (\text{ばたん}), \text{花の色 : 赤}) \\ &= 1 \times 0.507327 = 0.507327 \\ \text{salience}(* (\text{カーネーション}), \text{花の色 : 赤}) \\ &= 0.869971 \times 0.436287 = 0.379557 \\ \text{salience}(* (\text{チューリップ}), \text{花の色 : 赤}) \\ &= 0.111143 \times 0.056386 = 0.006267 \\ \text{salience}(* (\text{桜}), \text{花の色 : ピンク}) \\ &= 0.713603 \times 1 = 0.713603 \end{aligned}$$

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。ここで問題となるのは、

1. $*(S)$ のどの最尤性質が $*(T)$ に移される対象となるのか。
2. 最尤性質を移した結果、 $*(T)$ はどのように変化するか。

ということである。1については、前節で定義した顕現性から判断する。即ち、

「顕現性が高い最尤性質は移され易い」

とする。2については、 $*(T)$ の性質の頻度、差異度が変化すると考える。即ち、ある最尤性質を target 概念に移すということは、

「source 概念の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。

あ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の要素(焦点、話題など)によっても変化する。例えば、「緑のりんご」では $*(りんご)$ の性質「色:緑」の顕現性が増す。また、 $*(象)$ 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」の顕現性が増す。視点表現の理解も含め、このような効果を動的な顕現性の変化と呼ぶ。前節で示した顕現性は静的なもので、いわゆる辞書に記述された概念の意味と考えられる。しかし、概念はそれのみで使われることはなく、実際の文・談話内で使われる。そして、その意味は、使われる状況・文脈によって変化する。我々のモデルでは、顕現性という指標を橋渡しとして、概念の意味の動的な変化を統一的に扱

う。よって、視点表現の理解の文脈依存性なども説明できる。たとえば、 $*(象)$ の例で説明すると、「動き」を話題にしている文脈では $*(象)$ の最尤性質「動き:slow」の顕現性が高いため、 $*(男) \setminus *(象)$ の理解においても「動き:slow」が $*(男)$ に移され易くなる。一方、「重さ」を話題にしている文脈では「重さ:大」が移され易くなる。

図1に概念の変化モデルの概要を示す。

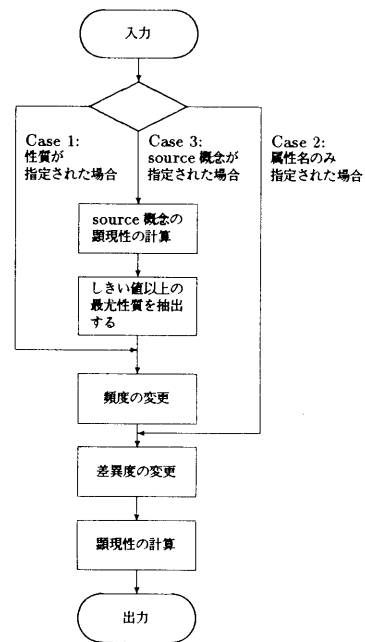


図1: 概念の変化モデル

それぞれの処理について説明する。

[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。なぜなら、たとえば属性名「大きさ」が指定されても、その属性値まではわからないからである。例えば $*(鼠)$ の場合には「大きさ:小」の、 $*(像)$ の場合には「大きさ: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詳しく説明する。

2.3.1 頻度の変更

定義 6 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 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。

例 5 頻度の変更

$*(黄色いチューリップ)$ を計算する。この場合、属性名「花の色」、属性値「黄色」が指定されている。今、例 2 より、指定された属性名「花の色」と同じ名前の属性名が $*(チューリップ)$ に存在する。

花の色 : {赤#0.35, 白#0.3, 黄色#0.2, ピンク#0.1, 紫#0.05}

また、指定された属性名「黄色」も属性値集合の要素として存在し、かつその頻度が 0 ではない。よって、「黄色」の頻度 0.2 は 1 に、他の属性値の頻度は 0 に変化する。最終的な、 $*(黄色いチューリップ)$ の「花の色」に関する性質は

花の色 : {赤#0, 白#0, 黄色#1, ピンク#0, 紫#0}

となる。ここで、最尤性質は「花の色:赤」から「花の色:黄色」に変化し、「花の色」の属性値集合の情報量も 0.111143 から 1 に増える。

一般に、属性値を指定することにより、対象となる属性値集合の情報量は増す(1 になる)、即ち顕現性も増加する。

2.3.2 差異度の変更

定義 7 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。』

属性名が指定されれば兄弟概念も変わるので、これは正しい定義といえる。例えば、 $*(カーネーション)$ の兄弟概念は、属性名「色」が指定されることにより

$\{*(ばたん), *(桜), \dots\}$

から

$\{*(色が赤いカーネーション), *(色が白いカーネーション), \dots\}$

に変わる。ところが、これらの最尤性質はそれぞれ、「花の色:赤」、「花の色:白」、... で共通な要素がない。よって、定義 4 により、それぞれの差異性は最大の 1 となる。一般に、属性名が指定されたときは、その最尤性質の差異度は 1 になることがいえる。

例 6 差異度の変更

例 2 の $*(チューリップ)$ において、話題などで属性名「花の色」が指定されたとする。この場合、最尤性質「花の色:赤」の差異度は 1 になる。その結果、最尤性質「花の色:赤」の顕現性は $0.111143 \times 1 = 0.111143$ になる。何も指定されない時の顕現性は例 4 より 0.006267 であったから、属性名の指定により顕現性が増していることがわかる。

2.3.3 考察

本モデルでは、視点表現の理解において、「候補となる最尤性質を可能なら target 概念に移す」という規則に従うため、その移される割合は、全てが移される場合から全く移されない場合まで連続的である。次章で実際の比喻理解への応用について述べるが、この連続性により、リテラルな文・比喻・非文を区別しない解析が可能になる。

次に、本モデルの問題点について考察する。定義 6 によると、指定された属性値が属性値集合に存在しない場合、頻度の変更はおこなわない。例えば、 $*(黒いりんご)$ の理解では、 $*(りんご)$ の「色」に関する属性値集合には「黒」はないので概念の変化はおこなない。そもそも、属性値集合にない属性とは、その頻度がもともと 0 であることを意味する。即ち、そのような属性値を持つ実体を我々が知らないことに相当する。ところが人間は、推論によって知らない実体でも想像することが出来る。しかし、どんな実体でも想像できるわけではない。例えば、「十本足の人間」などは想像しにくい。これは、性質に関する知識や常識によると考えられる。今後は、性質に関する知識や常識を導入して、モデルを精密化していく予定である。

3 比喻理解への応用

本章では、2 章で示した視点表現の理解の応用例として比喻理解を扱う。ただし、本章で対象とする比喻は、直喩・隠喩の一部である。よって以降、本論文では「比喻」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 比喻理解の概要

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

(3) りんごのような頬

という比喩を解析する。まず, (3)から視点表現*(頬)*(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, *(頬)の性質の頻度・差異度を変更する。ここで, 概念*(頬)の一部を示すと,

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

である。まず, 頻度の変更であるが, 定義 6 により, 指定された属性名「色」, 「表面」, 「外形」が*(頬)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つ 0 でないか調べる。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*(頬)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, *(頬)*(りんご)は

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

となる。ここで, *(りんご)の最尤性質「色: 赤」, 「表面: 滑らか」が*(頬)に移された結果, *(頬)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 7 による差異度の変更により, *(頬)*(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[7,9,8]。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。そのた

め, 分類の基準を設定することは難しく, 様々なヒューリスティクスを導入する必要がある。特に, 比喩と非文を区別することは困難である。本論文では

「リテラルな文, 比喩, 非文は区別できず連続性をもって解析されるべきである」

という立場に立ち, 視点表現の理解を通してこれらを統一的に扱う。

視点表現の理解をおこなうために, まず, 言語表現から視点表現を抽出する必要がある。例えば,

(4) 男は狼だ。

では, 視点表現*(男)*(狼)が抽出され,

(5) アイデアが開花する。

では, 動詞の知識から*(アイデア)*(花)が抽出される。次に, 抽出された視点表現の理解を 2 章の概念の変化モデルで行なう。概念の変化モデルでは, 視点表現*(T)*(S)の理解を行なう際, *(S)の複数の最尤性質を*(T)に移そうとするのだが, 全てが移されるわけではない。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。

ここで, 我々のモデルの評価のため, 視点表現の理解の容易性を定義する。これは, 視点表現がリテラルに近いか非文に近いかを判断する基準になる。

定義 8 視点表現の理解の容易性

『視点表現*(T)*(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。」

即ち, source 概念から target 概念に移される候補となる最尤性質のうち, 実際に移されたものが多いほどリテラルな文に近くなるし, また, 同じ移されるにしろ, 顕現性が高い最尤性質が移されるほうがリテラルな文に近くなる。

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[6]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、

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

(7) アイデアを使い果たす。 $[(\text{アイデア}) \setminus (\text{資源})]$

などは扱えない。なぜなら、(6)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。即ち、target概念にそのまま対応する属性名が存在しなくても適当な変換によって属性値の対応がとれる場合があるということである。このような現象は、共感覚と呼ばれる。共感覚とは、ある様相から別の様相へ属性が転移する現象である。例えば、「なめらかな味」では、「なめらか」はそもそも触覚に関する属性であるが、これを異なる味覚の領域で用いている。ここで、「いかにして異なる様相での対応をとるか」が問題となる。この点については、Marks等の研究[15]が参考になる。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] 安井稔. 言外の意味. 研究社, 1978.
- [2] 宮川洋. 情報理論. 電子通信大学講座第39巻, コロナ社, 1979.
- [3] 佐藤信夫. レトリック感覚. 講談社, 1978.
- [4] 山梨正明. 比喩と理解. 東京大学出版会, 1988.
- [5] 竹内晴彦. 比喩の生成に関する基礎研究. , 日本認知科学会第5回大会 発表論文集, 70-71 ページ, 1988.
- [6] 徳永健伸, 岩山真, 田中穂積. 視点を考慮した概念の同一化とその応用. 情報処理学会 自然言語処理研究会, NL71-8, 1989.
- [7] 楠見孝. 比喩表現の理解過程. 表現研究, 46, 1987.
- [8] 土井晃一, 田中英彦. 隠喩理解—隠喩の検出—. 日本認知科学会 R&I研究会, SIGR&I87(2):3-10, 1988.
- [9] D. Fass. Metonymy and metaphor: what's the difference? In *COLING-88*, 1988.
- [10] D. Fass and Y. Wilks. Preference semantics, ill-formedness, and metaphor. *AJCL*, 9(3-4):178-187, 1983.
- [11] D. Gentner, B. Falkenhainer, and J. Skorstad. Metaphor: the good, the bad and the ugly. In *TINLAP-3*, 1987.
- [12] J. R. Hobbs. Metaphor interpretation as selective inferencing. In *IJCAI-81*, 1981.
- [13] G. Lakoff. *Woman, Fire, and Dangerous Things*. Chicago University Press, 1987.
- [14] G. Lakoff and M. Johnson. *Metaphors we live by*. Chicago University Press, 1980.
- [15] L. E. Marks, R. J. Hammeal, and M. H. Bornstein. Perceiving similarity and comprehending metaphor. *Monographs of the Society for Research in Child Development*, 52(1), 1987.
- [16] J. H. Martin. Understanding new metaphors. In *IJCAI-87*, 1987.
- [17] C. B. Mervis and E. Rosch. Categorization of natural objects. *Annual Review of Psychology*, 32:89-115, 1981.
- [18] A. Ortony. The role of similarity in similes and metaphors. In A. Ortony, editor, *Metaphor and thought*, Cambridge University Press, 1979.
- [19] E. Rosch and C. B. Mervis. Family resemblances: studies in the internal structure of categories. *Cognitive Psychology*, 7:573-605, 1975.
- [20] J. R. Searle. Metaphor. In A. Ortony, editor, *Metaphor and Thought*, Cambridge University Press, 1979.
- [21] E. E. Smith, D. N. Osherson, L. Rips, and M. Keane. Combining prototypes: a selective modification model. *Cognitive Science*, 12(4):485-527, 1988.
- [22] P. H. Winston. Learning by creating and justifying transfer frames. *Artificial Intelligence*, 10(2):147-172, 1978.