

直喩生成データセットを用いた比喩表現の自動選択

蜜石 湧斗[†] 嶋田 和孝[†]

[†]九州工業大学大学院 情報工学府 先端情報工学専攻, 〒 820-8052 福岡県飯塚市川津 680-4

E-mail: †{y_mitsuishi,shimada}@pluto.ai.kyutech.ac.jp

あらまし 比喩は心象を利用して説明や記述を分かりやすくし、協調や誇張の効果がある。平易な文を比喩表現を含む文に書き換えることで情報をより円滑に伝達することが可能になる。本研究ではさまざまな文に対応した比喩書き換えを行うために、文脈に適した比喩表現を自動で選択する手法について提案する。まず、比喩選択モデルの学習に用いる直喩生成データセットの作成する。比喩の適切さ(比喩度)を用いて、文脈に適している比喩表現を含むデータを収集する。自動比喩選択を行うために、Neural Networkによる比喩選択モデルを提案し、作成したデータセットで学習し、自動比喩選択を行う。実験では比喩選択システムを用いた比喩選択の適切さについて検証し、既存の連想辞書を用いる手法では行えない新規性のある比喩書き換えの有効性を確認した。

キーワード 直喩生成データセット, 自動選択, 深層学習

Selection of Figurative Expression Using Automatically Generated Data Set

Yuto MITSUISHI[†] and Kazutaka SHIMADA[†]

[†] Graduate School of Computer Science and System Engineering, Kyushu Institute of Technology, 680-4, Kawazu, Iizuka, Fukuoka, 820-8502 Japan

E-mail: †{y_mitsuishi,shimada}@pluto.ai.kyutech.ac.jp

Abstract Figurative expressions are one of effective approaches to lead readers to understand. In this paper, we propose a selection method of a suitable figurative expression about an adjective in a sentence by using neural networks (LSTM with attention). NNs usually need a large amount of training data for generating a model. We first construct the training data set automatically by computing a simile-likelihood score of each figurative candidate. Next, we select sentence-figurative expression pairs with high simile-likelihood scores as the training data. Then, we generate a LSTM-based selection model by using the training data. We obtained sentences with some interesting figurative expressions that cannot be selected by dictionary-based methods.

Key words Automatically Generated Data Set, Automatic Selection, Deep Learning

1. ま え が き

文学的な表現において比喩は心象を利用して説明や記述を分かりやすくし、協調や誇張の効果がある。平易な文を比喩表現を含む文へと書き換えることにより、新しい概念を理解させることができ、情報をより円滑に伝達することが可能になる[1]。比喩表現を含む文への書き換えを行う研究では、主に確率生成モデルや意味ネットワークを作成し、書き換えられる語から連想される名詞を用いて書き換える比喩表現を選択する手法が提案されている[2][3]。しかし、これらの研究では入力文の形容詞、名詞の数が設定されていたり、単純な構文しか書き換えができない。

本研究では、自由な構文に対応し、文脈を考慮した比喩表現を含む文の自動書き換えを行うことを目的とする。近年、高機

能な言語モデルとして知られている Neural Network (NN) を活用した文書分類モデル[4]の研究が行われている。NNの一種である Recurrent Neural Network (RNN)により、文を単語の時系列データとして用いることで文脈情報を考慮した学習ができる。この RNN を用いた文書分類モデルを応用し、比喩の生成をある文に対する適切な比喩表現を選択する問題とおきかえれば自動書き換えが可能になる。すなわち、RNN 分類モデルを用いた比喩選択を行うには、入力文と入力文中の単語に置き換え可能な比喩表現のペアデータが必要になる。しかし、学習に用いることができる十分な量のデータが存在しない。

そこで、本論文では NN 分類モデルを用いた比喩選択を行うため、大きく分けて2つの処理を行う。まず、比喩選択モデルの学習に用いる直喩生成データセットを作成する。直喩生成データセットを作成する際、比喩文の適切さ(比喩度)を用い

て作成する。比喩度とは認知心理学における比喩の処理過程を工学的に取り入れた計算手法である。そして、RNNによる比喩選択モデルを提案する。直喩生成データセットにより学習した比喩選択モデルを用いて比喩選択を行う。選択された比喩表現に対して人手評価し、提案した比喩選択モデルの有効性を検証する。

2. 関連研究

平易な文を比喩表現を含む文に書き換える研究が様々行われている。Abeら[2]は統計的分析を用いて比喩文に書き換えている。大規模な言語データから Probabilistic LSI (PLSI)[5]という確率生成モデルを用いて単語と概念関係の確率モデルを構築し、平易な文に含まれる3つ形容詞と、1つの名詞から、名詞を説明する直喩表現を生成している。北田ら[3]は電子辞書の知識ベースを用い、置き換え元から喩辞を生成し書き換えを行っている。比喩文に書き換えるアルゴリズムには認知心理学における実験から得られた比喩理解、生成に関する知見[6]を工学的に導入している。Rezepkaら[7]はOrtonyの比較理論[8]を数値で表す計算手法を提案し、比喩文の生成を行っている。Ortonyの比較理論とは喩辞と被喩辞の共有する属性概念を検索し、最も距離の近いものを解とする理論である。新聞コーパスから取得した形容詞-名詞のbi-gramから喩辞と被喩辞に関連する目的語の距離を求める。設定した閾値以上の距離の喩辞を正しい比喩文として計算している。

また、比喩理解のために比喩表現を修飾語に書き換える研究も行われている。坂口[9]は連想概念辞書[10]を用いて比喩表現の書き換えを行っている。連想概念辞書とは人間を被験者として大規模な連想実験を行い、刺激概念と連想概念の対、および両者間の距離が定義されている辞書である。この辞書をIntegrated-and-Fireモデル[11]のニューラルネットワーク上に実装し、喩辞と被喩辞の属性概念を計算する。共通する属性概念の距離を計算し、最も距離の近い属性概念を置き換える修飾語として、文の書き換えを行っている。富山ら[12]はWikipediaから取得した大規模コーパスを用いて単語の分散表現を学習し、喩辞の書き換えを行っている。富山らはword2vec^(注1)で計算した分散表現を用いて、喩辞と被喩辞に類似した形容詞を得る。Ortonyの比較理論を元に、最も類似度が高い形容詞で喩辞を書き換えている。Maoら[13]は単語分散表現とWordnet^(注2)の知識を使用し、書き換えを行っている。Wordnetから喩辞の上位概念と類義語を取得したものを置き換え候補語とし、喩辞以外の文脈語の単語ベクトルと書き換え候補語の単語ベクトルの類似度が最も高いものを置き換えに用いる。

3. 直喩生成データセットの自動作成

本節では比喩選択モデルの学習に用いる直喩生成データセットの作成手法を説明する。図1に直喩生成データセットの作成手法を示す。本研究では、直喩表現から連想される修飾語の

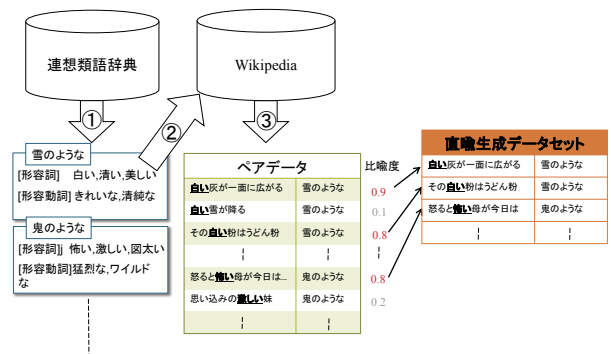


図1 直喩生成データセットの作成手法

アと、大規模コーパスの文を用いて、文と比喩表現のペアデータを作成する。我々は比喩文がどのくらい適切であるかを測る指標の比喩度[14]を提案している。本研究では、この既存の比喩度の計算手法を改良した新たな比喩度を提案する。新たに提案する比喩度を用いて、ペアデータより作成した比喩文の比喩度を求める。そして、設定した閾値以上の比喩度であるペアデータを直喩生成データセットに用いる。

3.1節では、文と比喩表現のペアデータ作成方法について説明し、3.2節では、比喩度計算手法について説明する。また、3.3節では、比喩度を考慮した直喩生成データセットの作成について説明する。

3.1 文と比喩表現のペアデータ作成

文と比喩表現のペアデータ作成の取得方法について説明する。比喩文作成時に置き換え元から連想される連想語が必要になる。本研究では、連想類語辞典^(注3)を用いて連想語を取得する。連想類語辞典は言葉を指定すると、その言葉の類語、同義語、連想語が表示されるWebサービスである。まず、連想類語辞典内の「○○のような」という比喩表現を全て獲得し、比喩表現から連想する形容詞と形容動詞を取得する(図1-①)。次に、大規模コーパスから連想される修飾語を含む文を検索する(図1-②)。そして、大規模コーパスから検索した文と比喩表現からペアデータを作成する(図1-③)。本研究では、2018年4月20日付でウィキメディア財団より提供されているデータベース・データ^(注4)から入手したものをを用いる。

3.2 比喩度計算手法

3.1節で作成したペアデータは単純なマッチングで生成したものであり、全て正しい比喩文であるとはいえない。例えば、図1にあるように「白い雪」を「雪のような雪」と置き換えるのは不適切であり、「思い込みの激しい妹」を「思い込みの鬼のような妹」とするのも適切とはいえない。すなわち、ある形容詞や形容動詞を連想類語辞典中の比喩表現に置き換える際には、その文の文脈を考慮しなければならない。文脈を考慮した正しい比喩であるペアデータのみを後述するRNNの学習データとして用いるために、ペアデータ内の比喩表現がどのくらい適切かを計算する必要がある。本研究では、認知心理学にお

(注1) : <https://radimrehurek.com/gensim/models/word2vec.html>

(注2) : <https://wordnet.princeton.edu/>

(注3) : <http://renso-ruigo.com/>

(注4) : <https://dumps.wikimedia.org/jawiki/>

る比喩の処理過程を工学的に取り入れ、比喩表現の適切さを測る比喩度を導入する。楠見ら [6] は心理実験により、比喩の処理過程について以下のようにまとめている。

(I) 直喩、隠喩理解の出発点は、喩辞と被喩辞の字義どおりのカテゴリ的意味における特徴対応によって、非類似性を検出することにある。カテゴリ的非類似性は比喩としての斬新さに影響を与える。

(II) 喩辞と被喩辞の情緒・感覚的意味における特徴対応によって、類似性を検出する。この情緒・感覚的類似性が高いほど比喩が理解しやすくなる。

(III) 喩辞と被喩辞の共有特徴の修飾語で修飾すると、情緒・感覚的類似性が高まり、その結果として、比喩の理解容易性が高まる。

このように喩辞、被喩辞、置き換え元^(注5)の3つの単語の関係により、比喩の斬新さや理解しやすさが変化する。

そこで、本研究では (I) よりカテゴリ的類似性 $simCat$, (II) より情緒・感覚的類似性 $simAss$, (III) より喩辞と被喩辞の共起度 $metPMI$ を求め、比喩度 M を算出する。これに新たな指標として置き換え元と喩辞の類似度 Sim , 喩辞の置き換えやすさ Rel , 置き換え元と喩辞の極性相異判定 $pDif$ を導入し、これら3つの指標を結合した比喩度 $ProposedM$ を提案する。以降、これらを順次説明する。

3.2.1 カテゴリ的類似性

カテゴリ的類似性 $simCat$ は喩辞 ($source$) と被喩辞 ($target$) の大まかなカテゴリ特徴が同じだとスコアを高くする指標である。大まかなカテゴリ特徴の類似性を測るため、日本語 WordNet^(注6) により得た喩辞と被喩辞の上位概念の一致率を計算する。 $comCat$ は喩辞、被喩辞の上位概念の一致数を表す。 $sourceCat$ は喩辞の上位概念数、 $targetCat$ は被喩辞の上位概念数を表す。全データ中の最大値 $tempCat^{max}$ で正規化した式 (2) の値をスコアとして用いる。

$$tempCat = comCat \left(\frac{comCat}{sourceCat} + \frac{comCat}{targetCat} \right) \quad (1)$$

$$simCat = \frac{tempCat}{tempCat^{max}} \quad (2)$$

3.2.2 情緒・感覚的類似性

情緒・感覚的類似性 $simAss$ は喩辞と被喩辞の特徴対応が類似しているほどスコアを高くする指標である。情緒・感覚的類似性を測るため、連想類語辞典より得た喩辞と被喩辞の連想語の一致率を計算する。 $comAss$ は喩辞、被喩辞の連想語の一致数を表す。 $sourceAss$ は喩辞の連想語数、 $targetAss$ は被喩辞の連想語数を表す。全データ中の最大値 $tempAss^{max}$ で正規化した式 (4) の値をスコアとして用いる。

$$tempAss = simAss \left(\frac{comAss}{sourceAss} + \frac{comAss}{targetAss} \right) \quad (3)$$

(注5): 「恐ろしい母」を「鬼のような母」と書き換え可能な場合
喩辞: 鬼のような, 被喩辞: 母, 置き換え元: 恐ろしい

(注6): <http://compling.hss.ntu.edu.sg/wanja/>

$$simAss = \frac{tempAss}{tempAss^{max}} \quad (4)$$

3.2.3 共起度

共起度 $metPMI$ は、日常的に頻繁に使われる喩辞と被喩辞の組み合わせほどスコアが高くなるように導入する。これは、日常的に頻繁に使用される言葉を比喩候補として用いた方が理解しやすい比喩文を作成できるためである。本研究では日本語 Google Ngram [15] を使用し自己相互情報量を計算する。 $C(word)$ は日本語 Google Ngram 内での $word$ の出現数である。 N は日本語 Google Ngram 内での出現した直喩表現数を表す。

$$tempPMI = \log_2 \frac{C(target, source) \cdot N}{C(source)C(target)} \quad (5)$$

$$metPMI = \frac{tempPMI'}{tempPMI'^{max}} \quad (6)$$

3.2.4 置き換え元と喩辞の類似度

置き換え元 ($original$) に類似する喩辞ほど、置き換えを行った場合に意味の相違が生まれにくいと仮定し、置き換え元と喩辞の類似度が高いものほどスコアを高くする。 Sim は $word2vec$ によって得た単語の分散表現の \cos 類似度を求める。 $w2v(word)$ は $word2vec$ により求める単語 $word$ の単語ベクトルである。

$$Sim = \cos(w2v(original), w2v(target)) \quad (7)$$

3.2.5 喩辞の置き換えやすさ

喩辞から連想される置き換え元が多いほど、様々な解釈を持つ喩辞である。喩辞の持つ解釈が多いほど、それぞれの喩辞と置き換え元の関連性が低くなる仮定する。下記の2つの文は「軽いスーツケース」の「軽い」を置き換えた文である。

- 羽根のようなスーツケース
- ミジンコのようなスーツケース

「羽根のような」と「ミジンコのような」は互いに「軽い」という解釈を持つ喩辞である。連想類語辞典から「羽根のような」に連想される形容詞を検索した場合「軽い」のみが検索されるが、「ミジンコのような」は「軽い」のほかに「小さい」や「弱い」なども検索される。このことから、「ミジンコのような」のほうがより多くの解釈を持ち、「ミジンコのようなスーツケース」から「軽いスーツケース」を連想しにくいといえる。そこで、喩辞から連想される置き換え元の数より、喩辞と置き換え元の関連度を求める。喩辞と置き換え元の関連度を式 (8) に示す。 mod は直喩表現から連想される修飾語の数を表す。

$$Rel = \max \left(\frac{1}{mod}, 0.01 \right) \quad (8)$$

3.2.6 極性相異判定

置き換え元と喩辞との極性が異なる場合、文脈に適していない置き換えを行う可能性があるためペナルティを設ける。本研究では高村ら [16] の感情極性対応表を用いて単語の極性を取得する。置き換え元の極性 $repPol$ を式 (9) に表し、喩辞の極性 $metPol$ を式 (10) に表す。 $senti(word)$ は感情対応表から取得

した単語 *word* の極性値を表す。 *metPol* は喩辞の極性を表し、喩辞から連想される語 *Ass* の極性の平均を用いる。置き換え元の極性 *orgPol* と喩辞の極性 *metPol* との極性が異なるものはペナルティとしてスコアを下げる。

$$orgPol = senti(original) \quad (9)$$

$$metPol = \log \frac{1}{n} \sum_{i=0}^n senti(Ass_i) \quad (10)$$

$$pDif = \begin{cases} 1 & (\text{Same polarity}) \\ 0 & (\text{Different polarity}) \end{cases} \quad (11)$$

3.2.7 比 喩 度

楠見らの心理実験に基づき提案する指標である、カテゴリ的類似性 *simCat*、情緒・感覚的類似性 *simAss*、喩辞と被喩辞の共起度 *metPMI* から、比喩度 *M* を求める。比喩度 *M* を式 (12) に示す。

$$M = simCat + simAss + metPMI \quad (12)$$

また、新たに提案する指標である置き換え元と喩辞の類似度 *Sim*、喩辞の置き換えやすさ *Rel*、置き換え元と喩辞の極性相異判定 *pDif* を追加した比喩度 *ProposedM* を式 (13) に示す。

$$ProposedM = M \times Rel \times pDif \quad (13)$$

そして、スコアが高い順に文脈を考慮した喩辞が含まれる文とする。

3.3 比喩度を考慮した直喩生成データセット作成

3.1 節で作成したペアデータは、文脈情報を考慮せず作成しているため、全て正しい比喩表現であるとはいえない。全てのペアデータを NN 比喩選択モデルに用いると、ノイズとなるデータを多く学習してしまうため、正確な比喩選択が行えない。そのため、提案した比喩度により、文脈を考慮したペアデータをデータセットに使用し、ノイズデータを削除することが重要になる。まず、新たに追加した指標の有効性を検証する。そして、直喩生成データセットに用いるペアデータ選択手法を説明する。

3.3.1 比喩度の有効性検証

楠見らの心理実験に基づき提案する比喩度 *M* をベースラインとし、新たに提案する指標の有効性を検証を行った。連想類語辞典から獲得した 508 件の喩辞を用いて、文と比喩表現のペアデータ 5,333,001 件を獲得した。獲得したペアデータを用いて比喩書き換えを行い、比喩文を生成した。生成した比喩文の比喩度を 5 つの計算手法^(注7)で求め、それぞれの比喩度上位 30 文をに対し人手で評価した。評価者 3 人に置き換えた比喩文の適切さを 5 段階^(注8)で評価した。

評価者の平均評価値を表 1 に示す。最も評価が高い手法は

表 1 比喩度評価結果

提案手法	平均評価値
<i>M</i> (ベースライン)	2.76
<i>M</i> × <i>Sim</i>	3.04
<i>M</i> × <i>Rel</i>	2.91
<i>M</i> × <i>pDif</i>	2.76
<i>ProposedM</i>	3.16

表 2 比喩度分布別評価結果

分布	平均評価値
上位 30 件	3.16
中位 30 件	1.21
下位 30 件	1.10

ProposedM であった。評価結果から、ベースラインの手法より指標 *Sim*、*Rel* を用いた手法の方が評価が高いため、これらの指標が有効であると考えられる。

また、比喩度の大小関係が正確であることも重要である。比喩度の大小関係が正確であれば、閾値による足切りをすることで、ノイズデータのみをデータセットから削除することが可能である。そこで、最も評価が高い *ProposedM* の比喩度上位 30 文、比喩度中位 30 文、比喩度下位 30 文それぞれを評価者 3 人が同様に 5 段階で評価した。評価結果を表 2 に示す。上位、中位、下位を順番に比較すると、比喩度が低くなるほど平均評価値も低くなることからランキングは正確であったと考えられる。特に下位 30 件のデータは *pDif* のペナルティによってスコアが下がったデータであるため^(注9)、*pDif* も有効である。

3.3.2 直喩生成データセットに用いるペアデータ選択手法

機械学習は学習するデータセットに偏りがある場合、正しい学習を行うことができない。直喩生成データセット内の比喩表現に偏りがあった場合、入力文の文脈や置き換える語に関係なく特定の比喩表現を選択してしまう。そこで、選択する比喩表現に偏りが無い直喩生成データセットを作成するため、比喩表現ごとに直喩生成データセット内に上限値を設定し、比喩度が高いものから順にペアデータを選択する。本論文では、比喩表現ごとに上位 500 件、比喩度が 0 以上のペアデータを学習データとして使用する。

4. Neural Network を用いた比喩選択モデル

本節では自動比喩選択を行うための Attention 付き Bi-directional Encode-Decoder モデル [17] について述べる。Attention とは予測モデルに入力データのどの部分に注目するか知らせる機構である。Bi-directional Encode-Decoder モデルは通常のエンコーダ (RNN) の中間層に入力文を逆向きに入力するエンコーダ (RNN) の中間層を連結し、将来の単語に関する情報を追加したものから予測を行うモデルである。本研究では RNN の拡張モデルである LSTM をエンコーダに用いる。直喩生成データセットでこのモデルを学習することで、入力文から喩辞を選択する際の重要になる単語を学習する。未知の文か

(注7) : *M*, *M* × *Sim*, *M* × *Rel*, *M* × *pDif*, *ProposedM*

(注8) : 5. 最高評価 (理解可能) ~ 1. 最低評価 (理解不可能)

(注9) : *pDif* が 0 になった場合、*ProposedM* (式 13) は 0 になる。

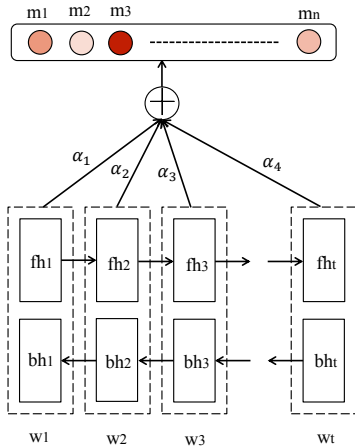


図2 Attention 付き Bi-directional Encoder-Decoder モデル

ら喩辞の出現確率を求めて、選択を行うことが可能になる。比喩選択モデル図を図2に示す。まず入力文を表す系列ベクトル $S = \{w_1, w_2 \dots w_t\}$ をそれぞれの入力層に入力する。通常のエンコーダーの中間層 fh に入力文を逆向きに入力するエンコーダーの中間層 bh を連結した双方向エンコーダー LSTM で文書を変換する。

$$fh_t = LSTM_{fw}(w_t, h_{t-1}) \quad (14)$$

$$bh_t = LSTM_{bw}(w_t, h_{t+1}) \quad (15)$$

$$h_t = [fh_t; bh_t] \quad (16)$$

$$h = \{h_1, \dots, h_t\} \quad (17)$$

各単語に対応する隠れ層 h_t を入力とし予測の際その単語に注目すべき確率を計算する。各隠れ層の状態 h_t は発話全体の情報を含み、各単語の重要度を考慮した文の表現ベクトル α_t を出力する。 W_a, B_a, u_a はモデルパラメーターである。

$$u_t = \tanh(W_a h_t + b_a) \quad (18)$$

$$\alpha_t = \frac{\exp(u_t^T u_a)}{\sum_t \exp(u_t^T u_a)} \quad (19)$$

$$\alpha = \sum_t \alpha_t h_t \quad (20)$$

LSTM の構造により、入力の全時系列を学習する。そして表現ベクトル α を softmax 分類器によって分類を行い、確率値がもっとも高い喩辞 $\{m_1, m_2 \dots m_n\}$ を選択する。

5. 比喩選択モデル評価実験

比喩書き換えを行うにあたって書き換えた比喩文が日本語として受理できることが最も重要である。喩辞は一般的に置き換え元となる語よりも抽象的な表現であり、様々な解釈が含まれる。そのため、選択した喩辞が被喩辞を修飾する語として適切でない場合、日本語として受理できない文に書き換えてしまうことになる。そこで、書き換えた文が理解可能であるか人手評価実験を行う。

また、比喩書き換えを行うにあたって置き換え元の情報を保

表3 比喩文理解容易性評価結果

比喩選択手法	平均評価値
比喩度 (<i>ProposedM</i>) のみ	1.66
NN による比喩選択モデル	1.75

持した比喩を選択することも重要である。書き換え前の文の情報が欠落することで本来の意味で伝達できなくなる。そのため、置き換え元の情報を保持している喩辞を選択することが必要である。そこで、自動選択された喩辞が置き換え元の内容を保持しているかを評価する置き換え内容保持評価実験を行った。

本実験では、4節で提案した比喩選択モデルの有効性を検証するため、3節で提案した比喩度 *ProposedM* のみを用いて選択する手法と比較を行う。以降、2種類の人手評価実験について説明する。

5.1 比喩文理解可能性評価

自動選択された喩辞を用いて書き換えた比喩文が、理解可能であるかを評価する実験である。評価データを例を以下に示す。書き換えた比喩文を見て文の内容が理解可能であるかを3段階(3: 理解可能, 2: 文脈には適していないが理解可能, 1: 理解不可)で評価した^(注10)。

例) 比喩文理解容易性評価データの例

- 「般若のような外面に反して温厚な性格」 評価値 [3]
- 「通夜のような外面に反して温厚な性格」 評価値 [1]

提案手法の学習に使用していない比喩度上位の文からランダムで選択した100文に対し、3人の評価者が評価した。

評価結果を表3に示す。書き換えを行った100文の平均評価値を示している。比喩選択モデルのほうが比喩度 (*ProposedM*) のみの場合よりわずかに高い評価を示した。しかし、どちらの手法とも低い評価であった。学習に用いる前の比喩度 *ProposedM* のみの場合でも平均評価値が低いため、比喩度自体の改良が必要である。

5.2 置き換え内容保持評価

自動選択された喩辞が置き換え元の内容を保持しているかを評価する実験を行った。評価データの例を下記に示す。置き換え前の文を見て、選択した比喩が元の文の内容を保持しているかどうかを2値(1: 内容を保持している, 0: 内容を保持できていない)で評価した。

例) 置き換え内容保持評価データ

いかつい外面に反して温厚な性格

- 「般若のような」 評価値 [1]
- 「通夜のような」 評価値 [0]

比喩選択を行った文は比喩文理解容易性評価と同じ100文を使用した。

評価結果を表4に示す。100件中、内容を保持している比喩表現を選択した精度を表している。比喩度 (*ProposedM*) のみの選択手法のほうが評価が高かった。喩辞と被喩辞のカテゴリ的類似度や、Google Ngram による喩辞と被喩辞の組み合わせ

(注10) : 元の表現が何であったかは被験者は知らないことに注意

表 4 内容保持評価結果

選択手法	recall
比喩度 (<i>ProposedM</i>) のみ	0.31
NN による比喩選択モデル	0.18

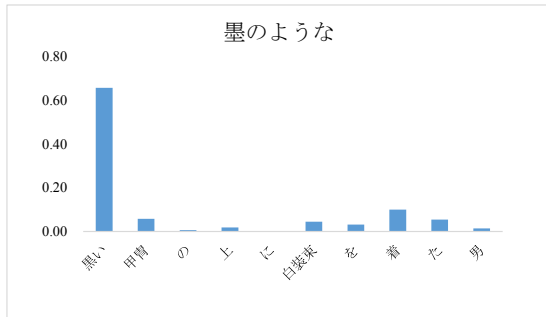


図 3 attention 例

の使用頻度など、人間の感覚に近いスコア算出方法を用いる比喩度に対し、時系列データより大まかに文脈情報を学習し比喩選択を行う比喩選択モデルのほうが置き換える語の内容を保持が困難である。よって、評価結果に差が生まれたと考える。

5.3 考察

今回の比喩選択モデルでは、内容を保持した比喩表現が選択することができなかった原因として、Attention が正しく機能していないことが挙げられる。図 3 に Attention の例を示す。「黒い甲冑の上に白装束を着た男」が「黒曜石のような」に置き換えることができる文を入力した場合、Attention は、ほぼ「黒い」という単語のみにしかされていなかった。このように文長が長い文では置き換える場所の部分しか Attention がされていないため、置き換え元や被喩辞の前後の単語を考慮せず、「黒い」に置き換えることが多い「墨のような」を選択してしまう。そのため、「黒い」が含まれるものには必ず「墨のような」という喩辞が選択されていた。このことより、被喩辞や喩辞の前後単語は特に Attention を強くするなどの改良が必要である^(注11)。

5.2 節の評価では、比喩度のみを用いる手法の方がよい結果だったが、NN を用いる手法の方が評価の高い比喩書き換えを行う例も存在した。人手評価が高かった比喩文に「きれいな盆栽を買って帰ってきた」を「見本のような盆栽を買って帰ってきた」と置き換えた文がある。連想類語辞典では「見本のような」は「代表的な」や「標準的な」など適正なものを連想する比喩表現とされており、「きれいな」という形容詞は取得できない。このように、比喩度のみでの書き換えでは選択できない比喩を選択でき、新規性のある比喩書き換えを行うことが可能であるといえる。したがって、より精度の高い比喩度の計算を行い、適切なペアデータを学習のための直喩生成データセットとして獲得することが重要である。

(注11)：ただし、実験では Attention を使用しない比喩選択モデルでもそもそも適切なモデルを構築できなかった。そのため Attention は比喩選択モデルに必要な機構である。

6. むすび

本研究では文脈を考慮した比喩文への書き換えを行うため、比喩選択モデルの学習に用いる直喩生成データセットを作成し比喩選択を行った。人手評価実験から提案した比喩選択モデルより、比喩度のみの方がより文脈の内容を保持した書き換えを行えた。しかし、比喩選択モデルを用いる場合連想類語辞典では連想されなかった新規性のある比喩書き換えを行うことが可能である。今後の課題として、比喩度計算手法の改良や調整、比喩選択モデルの改良を行う。

文 献

- [1] George Lakoff and Mark Johnson. *Metaphors we live by*. University of Chicago press, 2008.
- [2] Keiga Abe, Kayo Sakamoto, and Masanori Nakagawa. A computational model of the metaphor generation process. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 28, pp. 937-942, 2006.
- [3] 北田純弥, 萩原将文. 電子辞書を用いた比喩による文章作成支援システム. 情報処理学会論文誌, Vol. 42, No. 5, pp. 1232-1241, may 2001.
- [4] Siwei Lai, Liheng Xu, Kang Liu, and Jun Zhao. Recurrent convolutional neural networks for text classification. In *AAAI*, Vol. 333, pp. 2267-2273, 2015.
- [5] Thomas Hofmann. Probabilistic latent semantic analysis. In *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*, pp. 289-296. Morgan Kaufmann Publishers Inc., 1999.
- [6] 楠見孝. 比喩の処理過程と意味構造. 風間書房, 1995.
- [7] Rafal Rzepka, Pawel Dybala, Kohichi Sayama, and Kenji Araki. Semantic clues for novel metaphor generator. In *Proceedings of 2nd International IJCAI Workshop of Computational Creativity*, pp. 937-942, 2013.
- [8] Andrew Ortony. Beyond literal similarity. *Psychological review*, Vol. 86, No. 3, p. 161, 1979.
- [9] 坂口琢哉, 石崎俊. 連想概念辞書のニューラルネットワークへの符号化と比喩理解システムへの応用. 安田女子大学紀要, Vol. 38, pp. 169-179, 2010.
- [10] 岡本潤, 石崎俊. 概念間距離の定式化と既存電子化辞書との比較. 自然言語処理, Vol. 8, No. 4, pp. 37-54, 2001.
- [11] Wolfgang Maass and Christopher M Bishop. *Pulsed neural networks*. MIT press, 2001.
- [12] 富山翔司, 松尾豊. 大規模コーパスを用いた word2vec による比喩の書き換え. 人工知能学会全国大会論文集, 2016.
- [13] Rui Mao, Chenghua Lin, and Frank Guerin. Word embedding and wordnet based metaphor identification and interpretation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, pp. 1222-1231. Association for Computational Linguistics, 2018.
- [14] 蜜石湧斗, 嶋田和孝. 連想類語辞典を用いた直喩生成コーパスの作成. 電子情報通信学会信学技報, Vol. 118, No. 122, pp. 57-60, 2018. NLC2018-6.
- [15] 工藤拓, 賀沢秀人. Web 日本語 N グラム第一版. 言語資源協会, 2007.
- [16] 高村大也, 乾孝司, 奥村学. スピンモデルによる単語の感情極性抽出. 情報処理学会論文誌, Vol. 47, No. 2, pp. 627-637, 2006.
- [17] Al-Sabahi Kamal, Zuping Zhang, and Kang Yang. Bidirectional attentional encoder-decoder model and bidirectional beam search for abstractive summarization. *CoRR*, 2018.