鏡映変換に基づく埋め込み空間上の単語属性変換

石橋 陽一　須藤 克仁　吉野 幸一郎　中村 哲
奈良先端科学技術大学院大学
{ishibashi.yoichi.ir3, sudoh, koichiro, s-nakamura}@is.naist.jp

1　はじめに

word2vec[3] や GloVe[5] などの単語埋め込みで得られたベクトルは $\overrightarrow{king} - \overrightarrow{man} + \overrightarrow{woman} \approx \overrightarrow{queen}$ のような加法構成性を持つことが知られている．この演算 (アナロジー) を用いて *king* から *queen* に，また *queen* から *king* に変換することができる．このような単語の属性変換は言語データの拡張や埋め込み空間上の推論などに応用できる．例えばデータ拡張への応用では，*He is a boy.* の各単語の性別属性を変換し *She is a girl.* という新たな文を作ることができる．そこで我々は埋め込み空間上で分散表現の持つ属性を制御して変換する新たな表現学習に取り組んだ．アナロジーによる変換は入力単語が男性か女性かどうかで演算（ベクトルを足すか引くか）が変わるため，入力単語の属性に関する知識が必要となるが，そのような知識は無数にあるため全ての単語に付与することはできない．そこで，そのような知識を使用しない変換を考えると，入力単語が男性であっても女性であっても同一の変換関数で入力単語の性別を反転できるような関数が理想的である．そこで本研究では，そのような理想的な性質を持った写像である鏡映変換を導入することで，埋め込み空間上で単語ベクトルの属性を反転させる汎用的な手法を提案する．

2　埋め込み空間上の単語属性変換

本稿では単語を x, その単語の分散表現を $\mathbf{v}_x$ と表記する．なお，この分散表現は Skip-gram[3] などで事前に学習されているとする．本研究で扱うタスクでは入力として単語 x と，変化させたい属性の one-hot ベクトル $\mathbf{z}$，そして x の属性を反転した単語 t が与えられる．またこれらをまとめた集合を $(x,\ t,\ \mathbf{z}) \in \mathcal{A}$ とする（例 $(man,\ woman,\ \mathbf{z}_{\text{gender}}) \in \mathcal{A}$）．本タスクでは $(x, t, \mathbf{z}) \in \mathcal{A}$ が与えられたとき，属性 $\mathbf{z}$ について反転させる変換関数 $f_{\mathbf{z}}$ に x の分散表現 $\mathbf{v}_x$ を入力し，出力 $\mathbf{v}_y$ が目的語の分散表現 $\mathbf{v}_t$ になるようにする:

$$\mathbf{v}_t \approx \mathbf{v}_y = f_{\mathbf{z}}(\mathbf{v}_x). \tag{1}$$

単語属性変換においては次の性質: (1) 属性 $\mathbf{z}$ を持つ単語はその属性を反転し (2) 属性 $\mathbf{z}$ を持たない単語は変換しない を満たす必要がある．例えば

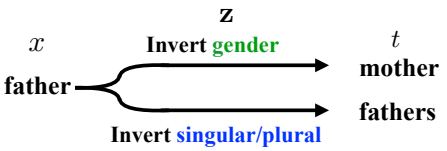

図 1: 単語属性変換タスク

属性語 $(man,\ woman,\ \mathbf{z}_{gender}) \in \mathcal{A}$ が与えられた場合，性別の属性変換 $f_{\mathbf{z}_{gender}}$ によって $\mathbf{v}_{man}$ を $\mathbf{v}_{woman}$ に変換し，属性 $\mathbf{z}_{gender}$ を持たない非属性語 $(apple,\ \mathbf{z}_{gender}) \in \mathcal{N}$ が与えられた場合は変換せず $\mathbf{v}_{apple}$ を出力する．

3　アナロジーに基づく属性変換

自己相互情報量 (PMI) に基づく単語埋め込みは式 2のような加法構成性を獲得することが明らかになっている [1]:

$$\mathbf{v}_{queen} \approx \mathbf{v}_{king} - \mathbf{v}_{man} + \mathbf{v}_{woman}, \tag{2}$$

$$\approx \mathbf{v}_{king} - (\mathbf{v}_{man} - \mathbf{v}_{woman}). \tag{3}$$

式 3より，$\mathbf{v}_{king}$ から差分ベクトル $\mathbf{v}_{man} - \mathbf{v}_{woman}$ を引くことで $\mathbf{v}_{queen}$ に変換できる事がわかる．このように，性別属性を持つ単語ペアの差分ベクトルによって入力単語の性別を変換できる．性別属性以外について議論するために，性別のような二値属性 $\mathbf{z}$ を持つ単語の中で，片方の属性（男性）を持つ単語の集合を $\mathcal{M}$, もう片方の属性（女性）を持つ単語の集合を $\mathcal{F}$，そして $m \in \mathcal{M}$ と $w \in \mathcal{F}$ の差分ベクトルを $\mathbf{d} = \mathbf{v}_m - \mathbf{v}_w$ とすると，アナロジーに基づく単語属性の変換関数は

$$f_{\mathbf{z}}(\mathbf{v}_x) = \begin{cases} \mathbf{v}_x - \mathbf{d} & \text{if} \quad x \in \mathcal{M} \\ \mathbf{v}_x + \mathbf{d} & \text{if} \quad x \in \mathcal{F} \end{cases} \tag{4}$$

となる．式 4より，入力単語 x が $\mathcal{M}$ に属すか $\mathcal{F}$ に属すかによって演算が変わることがわかる．例えば性別属性の場合，x が男性であれば差分ベクトル $\mathbf{d}$ を引き，女性であれば $\mathbf{d}$ を加える．これはつまり，アナロジーに基づく変換には入力単語 x の属性に関する知識が必要であることを示している．

4　鏡映変換に基づく属性変換

そこで属性知識を用いない理想的な変換を考える．属性知識を用いず単語の属性を変換する理想的な変換関

数 $\phi_{\mathbf{z}}$ は

$$\forall m \in \mathcal{M}, \qquad \mathbf{v}_m = \phi_{\mathbf{z}}(\mathbf{v}_w), \tag{5}$$
$$\forall w \in \mathcal{F}, \qquad \mathbf{v}_w = \phi_{\mathbf{z}}(\mathbf{v}_m), \tag{6}$$

のような性質を持つことが望ましい．つまり変換関数 $\phi_{\mathbf{z}}$ は入力単語 m または w が $\mathcal{M}$ に属すか $\mathcal{F}$ に属すか考慮することなく，同じ写像によって変換を行う．式 5, 6をまとめると，

$$\forall m \in \mathcal{M}, \qquad \mathbf{v}_m = \phi_{\mathbf{z}}(\,\phi_{\mathbf{z}}(\mathbf{v}_m)\,), \tag{7}$$
$$\forall w \in \mathcal{F}, \qquad \mathbf{v}_w = \phi_{\mathbf{z}}(\,\phi_{\mathbf{z}}(\mathbf{v}_w)\,), \tag{8}$$

となる．したがって理想的な写像 $f_{\mathbf{z}}$ とは二回適用すると恒等写像となるような変換である．このような写像は対合と呼ばれている (例 $\phi\colon \mathbf{v} \mapsto -\mathbf{v}$) ．

4.1 鏡映変換

鏡映変換は対合の一種であり，鏡と呼ばれる超平面によって 2 つのベクトルの位置を相互に反転させる．したがって同一の鏡による鏡映変換（$\mathrm{Ref}_{\mathbf{a},\mathbf{c}}$）を二回繰り返すと恒等写像となる:

$$\forall \mathbf{v} \in \mathbb{R}^n, \qquad \mathbf{v} = \mathrm{Ref}_{\mathbf{a},\mathbf{c}}(\,\mathrm{Ref}_{\mathbf{a},\mathbf{c}}(\mathbf{v})\,). \tag{9}$$

標準内積が与えられた n 次元実ユークリッド空間 $\mathbb{R}^n$ における鏡映変換は

$$\mathrm{Ref}_{\mathbf{a},\mathbf{c}}(\mathbf{v}) = \mathbf{v} - 2\frac{(\mathbf{v}-\mathbf{c})\cdot\mathbf{a}}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a} \tag{10}$$

と定義される．ここで $\mathbf{a}\cdot\mathbf{a}$ は内積を表す．また $\mathbf{a}$ および $\mathbf{c}$ はそれぞれ鏡（超平面）を決定するパラメタであり，$\mathbf{a}$ は鏡に直交するベクトル，$\mathbf{c}$ は鏡が通る $\mathbb{R}^n$ 上の点である．

4.2 単語属性変換への適用

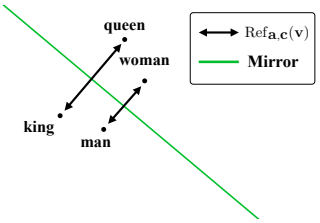

図 2: 鏡映変換に基づく単語属性変換

埋め込み空間上で鏡映変換を行い特定の属性（例 性別）を持つ単語の位置を反転させる．このとき鏡映変換における鏡を属性 $\mathbf{z}$ から推定する．ここで鏡は 2 つのベクトル $\mathbf{a}$, $\mathbf{c}$ によって一意に決まるため，$\mathbf{a}$ と $\mathbf{c}$ を属性 $\mathbf{z}$ から推定する．本研究では全結合の多層パーセプトロン（MLP）によって各属性ごとに $\mathbf{a}$ と $\mathbf{c}$ を推定する（式 11, 12，図 2）:

$$\mathbf{a} = \mathrm{MLP}(\mathbf{z}), \tag{11}$$
$$\mathbf{c} = \mathrm{MLP}(\mathbf{z}). \tag{12}$$

そして入力単語ベクトル $\mathbf{v}_x$ の属性を反転させたベクトルを鏡映変換し $\mathbf{v}_y$ を得る:

$$\mathbf{v}_y = \mathrm{Ref}_{\mathbf{a},\mathbf{c}}(\mathbf{v}_x). \tag{13}$$

4.3 Parameterized Mirrors

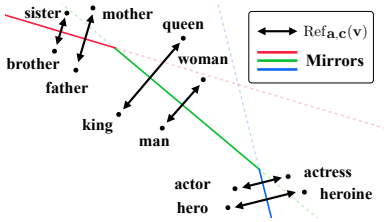

図 3: Parameterized Mirror を用いた鏡映変換

式 11，12によって鏡を推定する場合，図 3の *actor-actress* のような線形分離不可能な単語を正しく変換できない．これは属性 $\mathbf{z}$ から $\mathbf{a}$ および $\mathbf{c}$ を推定しているため鏡が一つに固定されてしまうことが原因である．この問題を解決するため，属性 $\mathbf{z}$ に加えて入力単語ベクトル $\mathbf{v}_x$ も $\mathbf{a}$ および $\mathbf{c}$ の推定に用いる:

$$\mathbf{a} = \mathrm{MLP}([\mathbf{z};\mathbf{v}_x]), \tag{14}$$
$$\mathbf{c} = \mathrm{MLP}([\mathbf{z};\mathbf{v}_x]). \tag{15}$$

ここで $[\cdot;\cdot]$ はベクトルの列方向の連結を表す．このようにして鏡を学習対象化（Parameterized Mirror）することで，鏡の固定化を防ぎ未学習データに対して鏡を逐次推定できる．例えば図 3の性別属性の変換において，まず $\mathbf{v}_{hero}$ から $\mathbf{v}_{heroine}$ への変換が成立するように鏡（青線）を学習する．ここで未学習である $\mathbf{v}_{actor}$ が $\mathbf{v}_{hero}$ と類似している場合，$\mathbf{v}_{hero}$ で学習した鏡と類似した鏡が推定されるため，$\mathbf{v}_{actor}$ から $\mathbf{v}_{actress}$ へ変換することが可能になる（図 3）．

4.4 損失関数

損失関数 $\mathcal{L}$ を以下のように定義する:

$$\mathcal{L}(\Theta) = \frac{1}{|\mathcal{A}|}\sum_{(x,t,\mathbf{Z})\in\mathcal{A}}(\mathbf{v}_y - \mathbf{v}_t)^2 \tag{16}$$
$$+\frac{1}{|\mathcal{N}|}\sum_{(x,\mathbf{Z})\in\mathcal{N}}(\mathbf{v}_y - \mathbf{v}_x)^2. \tag{17}$$

ここで上の項 (式 16) は，属性語 $\mathbf{v}_x$ を鏡映変換して得られたベクトル $\mathbf{v}_y$ が $\mathbf{v}_t$ と近づくように学習を行うことを表す．下の項 (式 17) は非属性語を変換関数によって変化させないための制約である．Θ は学習される MLP の重み集合である．

5 実験

提案手法によって所望の単語属性変換ができるかどうかを検証するため実験を行った．

5.1 実験設定

表 1の 4 つの属性を変換する実験を行った．属性語のデータセットを構築するためインターネットから独自に収集した単語に加え Analogy test set [3, 2] や [4] からデータを取得した．学習済みの単語埋め込みとして word2vec[3] と GloVe[4] を使用した．

表 1: データセット（属性単語のペア数）

Dataset $\mathcal{A}$	Train	Val	Test	Total
男性-女性 (MF)	29	12	12	53
単数形-複数形 (SP)	90	25	25	140
首都-国 (CC)	59	25	25	109
反意語 (AN)	1354	290	290	1934

5.2 評価方法

評価は変換精度（Accuracy）と安定性（Stability）で行った．

$$\delta(\mathbf{v}_y, t) = \begin{cases} 1 & \text{if} \quad \arg\max_{k \in \mathcal{V}}(\cos(\mathbf{v}_y, \mathbf{v}_k)) = t, \\ 0 & \text{otherwise}, \end{cases} \tag{18}$$

$$\text{Accuracy} = \frac{1}{|\mathcal{A}_{\text{test}}|} \sum_{(x,t,\mathbf{Z}) \in \mathcal{A}_{\text{test}}} \delta(\mathbf{v}_y, t), \tag{19}$$

$$\text{Stability} = \frac{1}{|\mathcal{N}_{\text{test}}|} \sum_{(x,\mathbf{Z}) \in \mathcal{N}_{\text{test}}} \delta(\mathbf{v}_y, x). \tag{20}$$

ここで $\cos(\mathbf{v}_y, \mathbf{v}_k)$ は $\mathbf{v}_y$ と $\mathbf{v}_k$ のコサイン類似度を表す．変換精度は属性語が正しく変換されているかを表し（例 *boy* → *girl*），安定性は非属性語が変換されないかを表す（例 *human* → *human*）．なおテスト時の非属性語集合 $\mathcal{N}_{\text{test}}$ は，語彙から $\mathcal{N}_{\text{train}}$ と $\mathcal{A}$ を除外してランダムに 1000 単語サンプリングして使用した．

5.3 実験結果

表 2は提案手法および比較手法の変換精度と安定性の結果である．Knowledge は変換の際に属性知識（例 $actor \in \mathcal{M}$）を用いるかを示している．Ref は単一の鏡を用いる鏡映変換，Ref+PM は Parameterized Mirror を用いた鏡映変換である．また，Diff はアナロジーに基づく手法であり，訓練データ中の単語ペアの差分ベクトル $\mathbf{d}$ を用いて変換する（最高精度のみ記述）．MeanDiff もアナロジーに基づく手法で，訓練データ中の単語ペアの差分ベクトルを平均したものを用いて変換する．Diff$^+$ や Diff$^-$ は，どのような単語ベクトルに対しても $\mathbf{d}$ を足すか引くかのどちらか一方の操作を行うことで知識を用いず変換している．また Ref や Ref+PM に対抗する学習ベースの手法として MLP: $\mathbf{v}_y = \text{MLP}([\mathbf{v}_x; \mathbf{z}])$ とも比較している．

表 2より変換精度と安定性の両方で最も優れている手法は Parameterized Mirror を用いた提案手法（Ref+PM）であることがわかった．知識を用いない手法の中で，次に変換精度が高いアナロジーに基づく手法に 20%以上の差をつけている．例えば GloVe の国-首都の変換では Diff$^+$ が 26%であるのに対し，提案手法の Ref+PM は 76%を獲得している．また興味深いことに安定性における鏡映変換の性能は顕著で，いずれも 99%以上を獲得している．このことから鏡映変換に基づく提案手法は，変換精度を維持しつつ高い安定性を持つことがわかる．一方で，総合的に変換精度および安定性が最も低い手法は MLP であった．

反意語 (AN) に関しては $\mathcal{M}$ と $\mathcal{F}$ に対応する単語が存在しないため鏡映変換と MLP のみの比較を行っているが，MLP は変換精度で鏡映変換を上回っているものの安定性は 1%台と非常に低い．一方で鏡映変換は変換精度をある程度維持しつつ安定性で 100%を獲得している．非属性語 $\mathcal{N}$ の学習量の増加で安定性が向上するか検証した結果 (表 3)，MLP の安定性は $\mathcal{N}$ の学習データ増加に応じてほとんど改善しない (最大 4%) 一方で，鏡映変換は $|\mathcal{N}| = 0$ の時点で 100% の安定性を獲得していた．このような結果となった原因として，鏡映変換を Analogic な空間に適用すると鏡の上に非属性語が分布するように鏡が学習されるため非属性語のベクトルが鏡映により移動しなかったためと推測した．そこで入力単語ベクトルとその鏡の距離 $\frac{|(\mathbf{v}_x - \mathbf{c}) \cdot \mathbf{a}|}{\|\mathbf{a}\|}$ を調べた結果，実際に非属性語は鏡に近く，属性語は鏡から離れて分布していることがわかった (図 4)．この傾向は反意語のように早い段階で安定性が高くなった属性ほど強かった．

また Ref+PM で学習された鏡のベクトル $\mathbf{a}$ を PCA で可視化した結果，各属性ごとにクラスタができており，属性ごとに使われる鏡が異なる事や，Parameterized mirror によって単語ごとにも異なる鏡が使われていることがわかった (図 5)．

表 4は文 $X = \{x_1, x_2, ...\}$ を与え，一単語ごとに変換関数に入力した結果である．MLP は全ての単語を誤って変換してしまっているが，鏡映変換は性別属性を持つ単語のみを変換させている．例えば $\text{Ref}(\mathbf{v}_{father})$ は *mother* に変換されているが，$\text{Ref}(\mathbf{v}_{when})$ は変換されず *when* のままである．アナロジーに基づく手法は男性から女性，もしくは女性から男性のどちらかの変換しかできていない．一方で鏡映変換は入力 x が男性であるか女性であるかといった知識を用いずに x の性別を反転させている (例 *father* → *mother*，*mother* → *father*)．

6 関連研究

本研究と類似した研究として文のスタイル変換が存在する [6]．このタスクでは文が与えられ教師なしでその文の言い回しなどの表現を変換する．これらは文単位の変換であるのに対し本研究では単語単位の変換を

[3] https://code.google.com/archive/p/word2vec/
[4] https://nlp.stanford.edu/projects/glove/

表 2: 変換精度と安定性の比較

Method	Knowledge	word2vec Accuracy (%)				word2vec Stability (%)				GloVe Accuracy (%)				GloVe Stability (%)			
		MF	SP	CC	AN	MF	SP	CC	AN	MF	SP	CC	AN	MF	SP	CC	AN
Ref		20.83	0.00	36.00	0.00	99.80	**100.00**	**99.80**	**100.00**	12.50	2.00	26.00	0.00	**100.00**	**100.00**	**100.00**	**100.00**
Ref + PM		**41.67**	**22.00**	**58.00**	28.79	**99.90**	99.40	99.40	**100.00**	**45.83**	**50.00**	**76.00**	33.54	99.70	99.10	99.20	**100.00**
MLP		8.33	4.00	12.00	**35.86**	2.20	0.00	2.70	1.90	4.17	10.00	18.00	**36.72**	5.10	7.00	5.20	1.20
Diff $^+$		25.00	2.00	32.00	-	72.10	77.90	53.90	-	25.00	2.00	26.00	-	99.30	94.20	99.30	-
Diff $^-$		25.00	2.00	30.00	-	49.60	78.20	56.30	-	25.00	2.00	24.00	-	100.60	99.90	99.50	-
MeanDiff $^+$		4.17	0.00	22.00	-	98.60	99.40	87.60	-	0.00	0.00	22.00	-	**100.00**	**100.00**	**100.00**	-
MeanDiff $^-$		8.33	0.00	14.00	-	97.20	99.30	92.40	-	0.00	0.00	0.00	-	**100.00**	**100.00**	**100.00**	-
Diff	✓	62.50	4.00	64.00	-	-	-	-	-	50.00	4.00	44.00	-	-	-	-	-
MeanDiff	✓	12.50	0.00	36.00	-	-	-	-	-	0.00	0.00	0.00	-	-	-	-	-

表 3: $|\mathcal{N}_{\text{train}}|$ を増加させた際のスコアの変化

		Accuracy (%) $\|\mathcal{N}_{\text{train}}\|$ 0	4	10	50	Stability (%) $\|\mathcal{N}_{\text{train}}\|$ 0	4	10	50
MF	Ref	12.50	12.50	12.50	12.50	**100.00**	**100.00**	**100.00**	**100.00**
	Ref + PM	**45.83**	**41.67**	**37.50**	**41.67**	99.70	99.90	99.90	99.90
	MLP	0.00	4.17	0.00	4.17	0.00	0.40	1.00	5.00
SP	Ref	0.00	0.00	2.00	0.00	**100.00**	**100.00**	**100.00**	**100.00**
	Ref + PM	**48.00**	**40.00**	**50.00**	**46.00**	53.30	99.10	99.10	99.80
	MLP	4.00	6.00	6.00	10.00	0.00	0.50	1.70	7.00
CC	Ref	24.00	26.00	24.00	20.00	**100.00**	**100.00**	**100.00**	**100.00**
	Ref + PM	**76.00**	**72.00**	**74.00**	**74.00**	99.20	**100.00**	99.90	99.90
	MLP	16.00	10.00	14.00	18.00	0.00	0.40	1.00	5.20
AN	Ref	0.00	0.00	0.00	0.00	**100.00**	**100.00**	**100.00**	**100.00**
	Ref + PM	26.90	26.72	33.54	25.69	**100.00**	**100.00**	**100.00**	**100.00**
	MLP	**29.48**	**29.66**	**36.72**	**36.55**	0.10	0.50	1.20	4.60

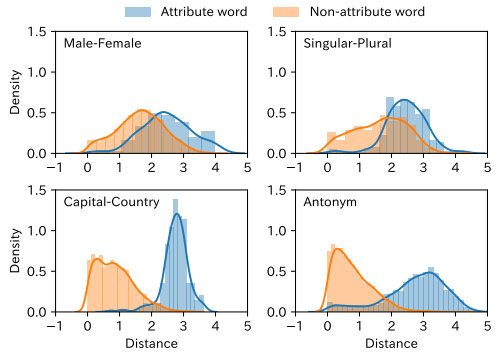

図 4: 単語ベクトルと鏡の距離の分布

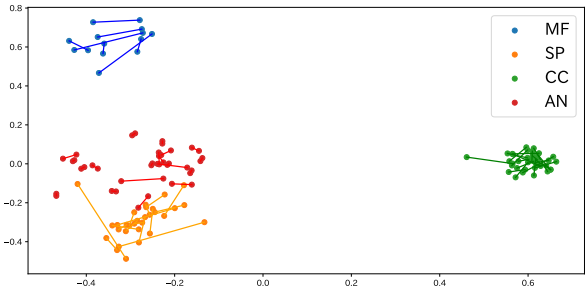

図 5: PCA による **a** の可視化 (線分:ペアの単語)

表 4: 文 X が与えれらたときの変換結果

X	the woman was married when your grandfather was (a) boy (.)
Ref(x)	the **man** was married when your **grandmother** was (a) **girl** (.)
Ref(Ref(x))	the **woman** was married when your **grandfather** was (a) **boy** (.)
MLP	By_Katie_Klingsporn **girlfriend** she fiancee
	Doughty_Evening_Chronicle ma'am **daughter** she (a) **mother** (.)
Diff $^+$	the **man** was married when your **grandfather** was (a) **boy** (.)
Diff $^-$	she **woman** was married she your **grandmother** was (a) **girl** (.)

行っている．また，Soricut[7] らは文字情報に基づいて英語の単数形-複数形のような単語の形態学的変換を行う手法を提案している．本研究では Analogic な埋め込み空間であれば形態学的な変換以外にも適用可能 (例 首都-国, 反意語) な汎用性の高い手法を提案した．

7 結論

本研究では単語埋め込み空間上で単語の属性を変換する新たな表現学習に取り組んだ．アナロジーによる変換では入力単語が持つ属性を知識として与える必要があったが，本研究ではそのような知識を用いることなく変換するために鏡映変換を導入した．実験の結果，提案手法は目的の属性を持つ単語であれば最大 76%の精度で変換し，その属性を持たない単語であれば 99%を変換せず，知識を用いることなく目的属性を持つ単語のみ安定して変化させることに成功した．

謝辞

本研究は JST CREST（課題番号: JPMJCR1513）の支援を受けて行った．

参考文献

[1] Carl Allen and Timothy M. Hospedales. Analogies Explained: Towards Understanding Word Embeddings. In *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, pp. 223–231, 2019.

[2] Anna Gladkova, Aleksandr Drozd, and Satoshi Matsuoka. Analogy-based detection of morphological and semantic relations with word embeddings: what works and what doesn't. In *Proceedings of the Student Research Workshop, SRW@HLT-NAACL 2016, The 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego California, USA, June 12-17, 2016*, pp. 8–15, 2016.

[3] Tomas Mikolov, Ilya Sutskever, Kai Chen, Gregory S. Corrado, and Jeffrey Dean. Distributed Representations of Words and Phrases and their Compositionality. In *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States.*, pp. 3111–3119, 2013.

[4] Kim Anh Nguyen, Sabine Schulte im Walde, and Ngoc Thang Vu. Distinguishing Antonyms and Synonyms in a Pattern-based Neural Network. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017, Valencia, Spain, April 3-7, 2017, Volume 1: Long Papers*, pp. 76–85, 2017.

[5] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. Glove: Global Vectors for Word Representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL*, pp. 1532–1543, 2014.

[6] Shrimai Prabhumoye, Yulia Tsvetkov, Ruslan Salakhutdinov, and Alan W. Black. Style Transfer Through Back-Translation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers*, pp. 866–876, 2018.

[7] Radu Soricut and Franz Josef Och. Unsupervised Morphology Induction Using Word Embeddings. In *NAACL HLT 2015, The 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Denver, Colorado, USA, May 31 - June 5, 2015*, pp. 1627–1637, 2015.