

B7IM2043

修士論文

単語の極性を埋め込んだ分散表現

中村 拓

2019年2月20日

東北大学 大学院
情報科学研究科 システム情報科学専攻

本論文は東北大学 大学院情報科学研究科 システム情報科学専攻に
修士 (情報科学) 授与の要件として提出した修士論文である。

中村 拓

審査委員：

乾 健太郎 教授 (主指導教員)

周 暁 教授

篠原 歩 教授

鈴木 潤 准教授 (副指導教員)

単語の極性を埋め込んだ分散表現*

中村 拓

内容梗概

単語の分散表現は、類義性や加法構成性において有効性を示しているが、対義や感情極性のポジティブ・ネガティブなど、異なる極性を読み取ることは難しいとされている。本研究は、類義性や加法構成性など単語の文脈的な類似性と、異なる極性の双方を表現できる分散表現の実現を目的として、単語の極性を線形関数で分離できるような分散表現モデル、対義の識別面を学習する手法を提案する。そして、単語の感情極性分類と対義語識別に対する汎化性能、及び極性を埋め込んだ分散表現の類義性・加法構成性を評価した。

キーワード

自然言語処理, 単語分散表現, 極性識別, Transductive SVM, 同時学習

*東北大学 大学院情報科学研究科 システム情報科学専攻 修士論文, B7IM2043, 2019年2月20日.

Embedding Word Polarity into Distributed Representations *

Taku Nakamura

Abstract

Distributed representations of words (word embeddings) show effectiveness in measuring word similarity and in their additive compositionality. However, word embeddings have the challenge to discern difference in word polarity, such as antonymy and sentiment polarity, due to their interchangeability in their context. This research aims to create word embeddings model which can represent both contextual similarity and contrasting polarity of words. This thesis proposes a word embeddings model whose word polarity is separable by a linear function, and a method to learn multiple separating hyperplanes for antonymy. Proposed model and method are evaluated in terms of generalization performance of polarity. Besides, learned embeddings are evaluated in word similarity and additive compositionality.

Keywords:

Natural language processing, Distributed representations, Polarity distinction, Transductive SVM, Joint learning

*Master's Thesis, System Information Sciences, Graduate School of Information Sciences, Tohoku University, B7IM2043, February 20, 2019.

目次

1	はじめに	1
1.1	背景	1
1.2	本論文の構成	2
2	関連研究	4
2.1	単語の分散表現	4
2.2	極性を識別する分散表現	4
3	感情極性の分散表現への埋め込み	6
3.1	初期化モデル	6
3.2	同時学習モデル	6
4	対義の識別面の学習	9
5	対義性の分散表現への埋め込み	13
6	評価実験	15
6.1	感情極性分類	15
6.1.1	実験設定	15
6.1.2	実験結果	16
6.2	対義語・同義語の判定	19
6.2.1	実験設定	19
6.2.2	実験結果	20
6.2.3	同義の訓練データ追加	21
6.3	類義性・加法構成性の評価	26
6.3.1	実験設定	26
6.3.2	実験結果	26
7	おわりに	28
	謝辞	29

付録	35
A 単語間の類似度・関係類推の評価データセット	35

目 次

1	距離を変える (a) と線形関数で分離する (b)	2
2	初期化モデル	7
3	同時学習モデル	7
4	対義の識別面の学習 (ベクトル固定) におけるラベル設定・識別 面更新	10
5	複数識別面を学習する際のラベル設定・識別面更新	12
6	シード文脈ベクトルに対する識別関数値の分布	18
7	対義語・同義語の判定における対義語識別の精度	22
8	対義語・同義語の判定における対義語識別の再現率	23
9	対義語・同義語の判定における対義語識別の F 値	24
10	同義語の訓練データ有無と対義語・同義語の判定性能 (Macro-F1)	25

表 目 次

1	識別関数による単語感情極性の分類正解率	16
2	SGNS のベクトルを特徴量とした SVM による単語感情極性の分類 正解率	16
3	各識別面で識別された対義語ペアの例	21
4	類義性 (類似度)・加法構成性 (関係類推) の評価	27
5	評価データセット別の類義性 (類似度)・加法構成性 (関係類推) の評価	27
6	各評価データセットのサイズ・参考文献	35

アルゴリズムの一覧

1	対義の識別面の学習（ベクトル固定）	10
2	対義性の埋め込み（ベクトル変更）	13

1 はじめに

1.1 背景

単語の分散表現は、単語の意味をベクトル空間で表現する。ベクトル間の類似度や演算によって、単語間の関係を扱えることから、自然言語処理の幅広いタスクに応用されている [1][2][3]。分散表現を学習するモデルは、分布仮説 [4] に基づき、類似した単語は類似した文脈で使われやすい（テキスト中で共起する周辺の単語が似ている）ことを用いて、単語の分散表現を獲得する。そのため、似た文脈で出現する単語は、似た分散表現となる。

しかし、異なる極性の単語（類義語と対義語や、ポジティブな語とネガティブな語）も、似た文脈で出現することが多く、分散表現から極性の違いを読み取ることは難しい。例えば、“large” と “small” は、共に大きさに関する文脈で現れることが多く、ベクトル空間で近い位置になる。

自然言語処理で広く用いられる分散表現から単語の極性を識別することは、自然言語の理解に重要なテキスト間の含意関係認識、対話や生成などの応用タスクで有益だと考えられる。

この問題に取り組む手法として、辞書情報を利用し同義語を近づけ、対義語を遠ざける学習が多く研究されてきたが [5, 6, 7]、距離を直接変える手法は、類義性や加法構成性において有効な単語分散表現の性質を大きく変えてしまう。

そこで、本研究では、距離を大きく変えずに単語の極性を線形関数で分離できるような分散表現モデルを提案する (図 1)。これは、対義語や感情極性の反転した単語が似たような文脈で出現することが多い一方で、対義性や感情極性の反転によって生じる共起語の変化も確かに存在する [8, 9] ことから、辞書情報などの外部知識を教師信号とし、周辺文脈を利用して極性を汎化して学習することを期待している。提案手法では、線形関数で極性を分離しながら意味の類似する単語のベクトルを近づけることができるので、これは既存手法よりも緩い制約である。

また、感情極性については「ポジティブ・ネガティブ」という自然な分離の軸が存在しているが、同じく対照的な意味を捉える対義語については、分離の軸はそれほど明らかではなく、また一つの軸を定めることは難しい。本研究は、まず



図 1: 距離を変える (a) と線形関数で分離する (b)

感情極性の埋め込みに提案手法を検証し，そこで得られた経験を対義語判定に応用する．さらに，感情極性ほど分離の軸が明確でない対義について，分離の軸に対応するような複数の識別面を学習する方法も提案する．

そして，対義語・同義語の判定，及び類義性と加法構成性の評価実験から，提案手法の極性識別に対する汎化性能，分散表現から対義性について複数の分離の軸を学習できる可能性，及び得られた分散表現が文脈的な類似性を保つことを示す．以下に本研究の貢献をまとめる．

- 単語の極性を線形関数で分離できるような分散表現モデルを提案
- 分散表現から、対義の分離の軸にあたる，複数の識別面（分離超平面）を学習する方法を提案
- 感情極性及び対義性を分散表現に埋め込む効果を検証
- 学習した対義の識別面が，どのような対義性を捉えているかを定性的に確認
- 極性を埋め込んだ分散表現の類義性と加法構成性の評価

1.2 本論文の構成

本論文の構成は以下の通りである．まず2章で，本研究に関連する単語の分散表現及び極性識別の研究について述べる．次に3章で，単語の分散表現の学習モデルを，単語の極性を線形関数で分離できるような分散表現モデルに拡張する本研究の提案手法を，感情極性の埋め込みにおいて説明する．4章では，極性の一

つである対義の識別面を学習する手法を提案し，5章で対義性を分散表現に埋め込むためのモデルの拡張について述べる．6章では，提案手法の分散表現及び対義の識別面についての評価実験を感情極性分類，対義語・同義語判定，単語の類似度・関係類推の評価タスクにおいて行い，提案手法の効果を検証・考察する．最後に，7章で本論文を総括する．

2 関連研究

2.1 単語の分散表現

単語は一般に記号として表現され、直接その記号間の類似度や関連性を計算することは難しい。単語の分散表現は、記号である単語に、ある次元数（数百次元など）の対応するベクトルを割り当て、単語をベクトルとして表現する。このように単語をベクトル空間に埋め込み、空間内の一点として捉えることで、距離や演算を定義でき、単語間の類似度や関連性を計算可能にするという利点がある。

分散表現の学習には、「単語の意味は、その単語の周囲の単語（文脈）によって決まる」という分布仮説 [4] に基づく手法がよく用いられ、代表的な分散表現モデルの一つで、自然言語処理において頑健なベースライン手法として知られる SGNS (Skip-gram with Negative Sampling) [10] は、単語の周辺文脈を予測するタスクを通して単語の分散表現を獲得する。

周辺文脈を用いて学習された分散表現において、例えば似た意味をもつ “car” と “automobile” に対応するベクトルは似ており、また “king - woman + man” の演算をした結果のベクトルは “queen” に近くなるなどの性質が見られ、単語の分散表現は、単語間の類似度や関連性を測る類義性や加法構成性の評価タスクで有効性を示すことが報告されている。

一方で、“large” と “small” のように出現する文脈は似ているが、対照的な意味をもつ単語もあり、同義と対義や、感情極性のポジティブとネガティブなど、単語の極性の違いを分散表現で識別することは難しい。そこで、次の 2.2 章で述べるような、この問題に対処するための研究が行われてきた。

2.2 極性を識別する分散表現

分散表現で極性を分離する研究として、単語ベクトル空間を極性の識別に特化させる方法が多く提案されている。この手法には、主に分散表現学習の目的関数に極性の制約を組み込む方法 [5, 6] と、学習済みの分散表現に後処理を加える方法 [7, 11] の 2 つがあり、任意の単語ベクトルに適用でき、学習させた極性の識別性

能が高いという観点から、後処理で極性識別に特化させる研究が多くなっている。一方で、未知の極性に対する汎化性能については、極性の制約を分散表現学習に組み込む手法に比べ、後処理の手法では、まだ改良の余地があるとされている。

これらの手法に共通する課題は、辞書情報を利用して同義語を近づけ、対義語を遠ざける学習を行い単語間の距離を直接変えるため、類義性や加法構成性といった分散表現の元の特性までをも変えてしまう恐れがある。実際、極性を識別するためには、訓練時に周辺文脈を予測する項のウェイトより距離を変える項のウェイトを遥かに大きく設定する必要がある [6] など、学習された単語ベクトルは、もはや周辺文脈の情報を忠実に取り入れたとは言い難い。このように訓練されたベクトルが果たして対義性を汎化できたことになるかについては疑問視する見方もあり、評価時に使われたデータの 90% 以上は「対義語の同義語も対義語」という簡単な推論で訓練データから導き出せるという指摘もある [12]。

感情極性を識別する分散表現にも、同様に単語ベクトル空間を極性に特化させる手法が多く研究されており [13, 14, 15], 「ポジティブ・ネガティブ」という自然な分離の軸が存在する感情極性の識別では成功例が多い。感情極性の埋め込みについては、Voら [16] が本研究と似た提案をしているが、Voらの提案は感情極性だけを学習した 2 次元の短いベクトルであるに対し、本研究は極性情報を分散的に単語ベクトルの全ての次元に保存し、線形関数をもってそれを取り出すものである。これは、単語ベクトルの部分空間に極性や辞書の情報を写像する [17, 18] などの研究とも視点が異なる。また、極性の識別面は線形識別平面であるため、この平面の法線ベクトルが埋め込み空間のどれかの基底と一致するように空間全体を回転させれば、特定の 1 次元の成分の符号を見るだけで極性を判別できる単語埋め込み空間が得られることになる。

3 感情極性の分散表現への埋め込み

いくつかの単語対に対しては，極性辞書によって単語の極性が定義されている．本研究では極性辞書による極性既知の単語をシードと呼び，極性の教師情報として用いる．単語の極性を分散表現学習に取り入れる手法として，(i) シード単語の初期化を極性情報に応じて行う，(ii) シードの極性を分離するための同時学習を行う2つのモデルを提案する．

3.1 初期化モデル

シード単語の極性情報を単語ベクトルの初期化で取り入れるこの提案は，Voらの研究 [16] に倣い各単語のベクトルをネガティブな成分 \mathbf{n} とポジティブな成分 \mathbf{p} に分け，それぞれ前半と後半の次元に対応すると仮定する．モデルの学習を通して，単語の極性がこれらの次元に分散して保存されることを期待する．極性のネガティブなシードに対しては $[\mathbf{n} : \mathbf{p}] = [1, \dots, 1, -1, \dots, -1]$ ，ポジティブなシードに対しては $[\mathbf{n} : \mathbf{p}] = [-1, \dots, -1, 1, \dots, 1]$ となるように単語ベクトル学習時の初期化を行う．各ベクトルは正規化し，シード以外の単語についてはランダムに初期化する．

3.2 同時学習モデル

周辺文脈との共起から単語分散表現を学習する代表的な手法の一つとして Skip-gram with Negative Sampling (SGNS) [10] がある．提案手法では，文脈共起に基づく学習に，SGNS モデルに L2 正則化項を加えた式 1 を最小化すべき目的関数とする分散表現モデルを用いる．本稿では以降 SGNS モデルと表記する．

$$-\sum_{t \in V} \sum_{c \in C_t} \left(\ln \sigma(\vec{v}_t \cdot \vec{v}_c) + \sum_{k=1}^K \ln \sigma(-\vec{v}_t \cdot \vec{v}_c) \right) - \lambda \|\vec{v}_t\|^2 - \lambda \|\vec{v}_c\|^2 \quad (1)$$

ここで， V は単語集合， C_t は単語 t の文脈， \vec{v}_t は単語 t のベクトル， \vec{v}_c は文脈に出現する単語 c のベクトル（文脈ベクトル）， K は負例サンプリング数， \vec{v}_t は学

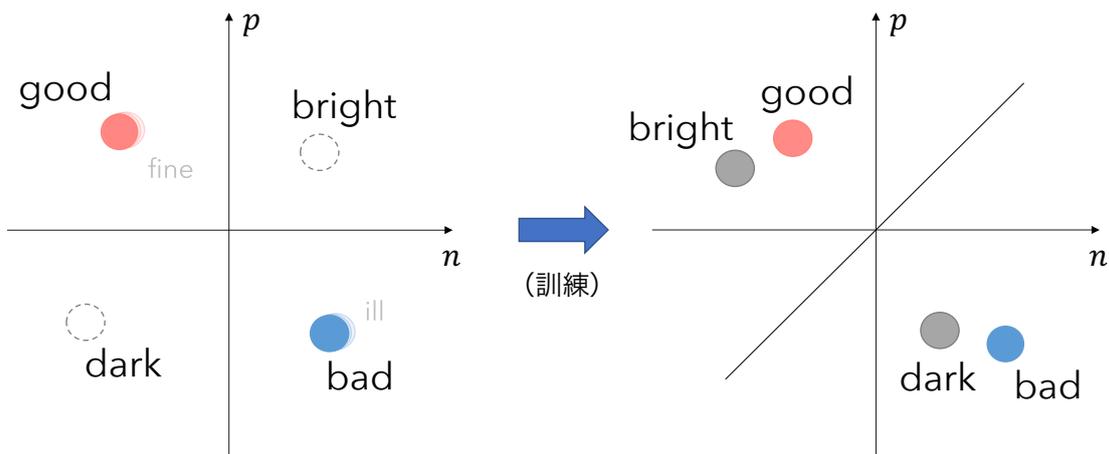


図 2: 初期化モデル

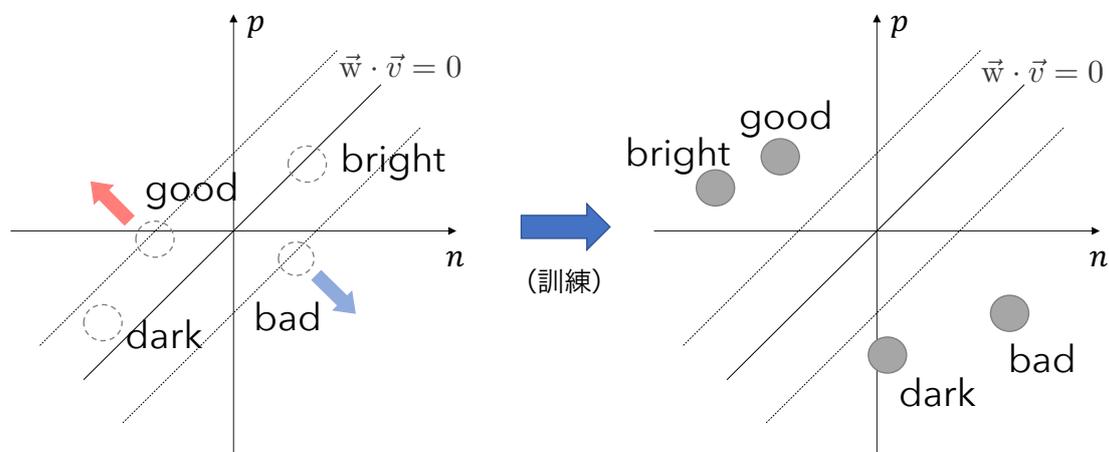


図 3: 同時学習モデル

習データの単語ユニグラム分布からランダムにサンプルした擬似負例単語のベクトル, λ は正則化項のパラメータである.

シード単語の極性を識別するための目的関数として, SVM のヒンジ損失関数 (式 2) を用いる.

$$\frac{1}{n} \sum_{i=1}^n \max(0, 1 - y_i(\vec{w} \cdot \vec{v}_i - b)) \quad (2)$$

ここで, n はシード単語数, y_i はシード単語の極性クラスラベル (ポジティブ 1, ネガティブ -1), \vec{w} は係数ベクトル, \vec{v}_i は単語または文脈ベクトル, b はバイアス項である. シード単語の極性分離と, 周辺文脈共起からの学習を同時に行うため, 係数 \vec{w} を固定して, 単語または文脈ベクトル \vec{v}_i をパラメータとみる. 係数 \vec{w} は前半次元を -1 , 後半次元を 1 として正規化したベクトルを用いる. d 次元の単語ベクトルを学習するとき, 係数ベクトルは

$$\vec{w} = \frac{[\overbrace{-1, \dots, -1}^{d/2}, \overbrace{1, \dots, 1}^{d/2}]}{\|\vec{w}\|^2} \quad (3)$$

となる. バイアス $b = 0$ とし, 全体として式 1 と式 2 の和 (4 式) を目的関数として最小化する.

$$\begin{aligned} & - \sum_{t \in V} \sum_{c \in C_t} \left(\ln \sigma(\vec{v}_t \cdot \vec{v}_c) + \sum_{k=1}^K \ln \sigma(-\vec{v}_t \cdot \vec{v}_c) \right) \\ & - \lambda \|\vec{v}_t\|^2 - \lambda \|\vec{v}_c\|^2 + \frac{1}{n} \sum_{i=1}^n \max(0, 1 - y_i \vec{w} \cdot \vec{v}_i) \end{aligned} \quad (4)$$

SGNS の確率的勾配降下法 (SGD) による単語ベクトル更新と同時に, 式 2 を最小化するための更新を行う. 初期化は全ての単語についてランダムに行う.

4 対義の識別面の学習

対義語のシード（対義の教師情報）については，感情極性におけるポジティブ・ネガティブのような分離の軸が必ずしも定まらないため，対義語辞書の対義語ペアに対して識別のためのラベルを設定する方法を考える必要がある。

本研究では，Transductive SVM (TSVM)[19] のラベル設定及び識別面の学習方法を応用して，対義語シードのペアに対してラベルを設定し，識別面を学習する。TSVM は，半教師あり学習の一つである Transductive 学習を SVM (Support Vector Machine) に適用し，ラベル付きの訓練データが少ない場合でも分類精度を向上させる手法で，学習と推論を交互にを繰り返しながら，学習が収束した時点の予測を最終出力とする。

オリジナルの TSVM では，ラベル付きの訓練データで分類器 (SVM) を訓練し，

1. ラベル無しデータを分類，(正と負の) 仮ラベルを設定
2. 仮ラベルを設定したデータを含めて分類器を再度訓練
3. 仮ラベル (の正と負) を入れ替えた方が分類誤りを減らせるペアを見つけ，ラベルを入れ替える
4. 2, 3 を繰り返す

ことで識別面を学習する。

本研究では，分類器 (SVM) の初期化はランダムに行い，対義語辞書の対義語ペアに対して一方の単語に正のラベル，もう一方の単語に負のラベルを設定する。ラベルの正負はペア毎に (正, 負) と (負, 正) 2通りの仮ラベルで分類誤りの損失を計算し，損失が小さくなるラベルを設定する。よって，対義語ペアの各単語は必ず異なるラベルとなる。識別面の学習は，単語ベクトル空間を固定して行うため，本稿では以降，この手法をベクトル固定と表記する。

ベクトル固定で対義語シードの分類ラベル設定・識別面学習を行う方法をアルゴリズム 1，図 4 に示す。図 4 中の w_1 は識別面を表す。

アルゴリズム 1 対義の識別面の学習 (ベクトル固定)

Input: 単語ベクトル集合 V , 対義シード集合 S

Output: 識別面 w

識別面 w をランダムに初期化

while ラベルの入れ替えがある間 **do**

for each 対義ペア in S **do**

 対義ペア毎に損失計算

▷ 図 4 ①

 損失が小さくなるラベルを設定

▷ 図 4 ②

 識別面 w を更新

▷ 図 4 ③

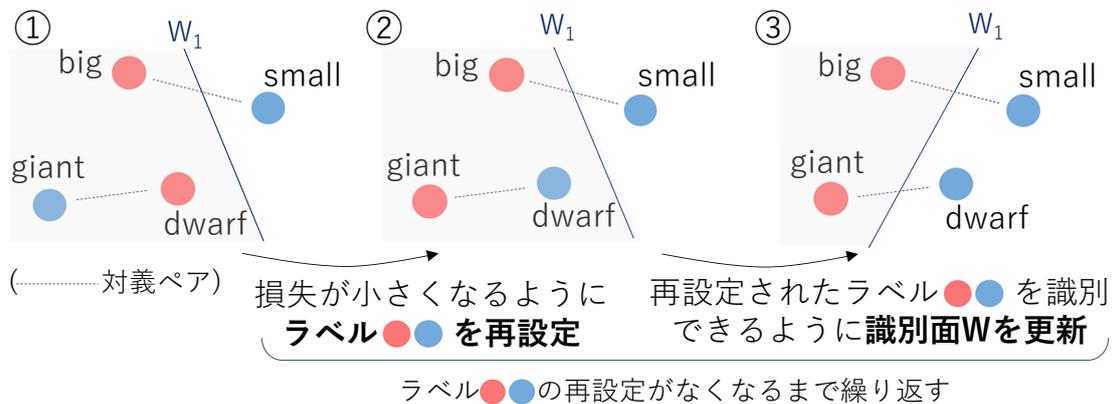


図 4: 対義の識別面の学習 (ベクトル固定) におけるラベル設定・識別面更新

対義語シードを識別するための最小化すべき目的関数として、3.2章における感情極性の識別と同様に、ヒンジ損失関数（式5）を用いる。

$$\frac{1}{N} \sum_{j=1}^N l^k(p_j) \quad (5)$$

ここで、 N はシード単語ペア数、 p_j はシード単語ペア、 $l^k(p_j)$ はシード単語ペア p_j の識別面 w_k に関する損失を表す。

ただし、シード単語ペア p_j の識別面 w_k に関する損失 $l^k(p_j)$ は、

$$l^k(p_j) = \sum_{i \in p_j} \max(0, 1 - y_i(\vec{w}_k \cdot \vec{v}_i - b))$$

となる。 y_i はシード単語のラベル、 \vec{w}_k は識別面 w_k の係数ベクトル、 \vec{v}_i は単語または文脈ベクトル、 b はバイアス項を表す。

対義の識別では、ポジティブ・ネガティブの分離の軸がある感情極性と異なり、分離の軸は1つとは限らない。そこで、複数の識別面を学習する拡張を考える。複数の識別面を学習する場合、ラベルは各識別面ごとに設定し、識別面の更新は、その識別面における損失が最小となるペアのみについて行う（図5）。

識別面 w_1, w_2, \dots, w_m を学習する時、対義語シードを識別するための目的関数は、式6となる。

$$\frac{1}{N} \sum_{j=1}^N \min[l^1(p_j), l^2(p_j), \dots, l^m(p_j)] \quad (6)$$

複数識別面の場合

- 各識別面 W_i でラベル設定
(W_1 ● ●, W_2 ● ●)
- 各ペアで損失が最小となる識別面のみ更新

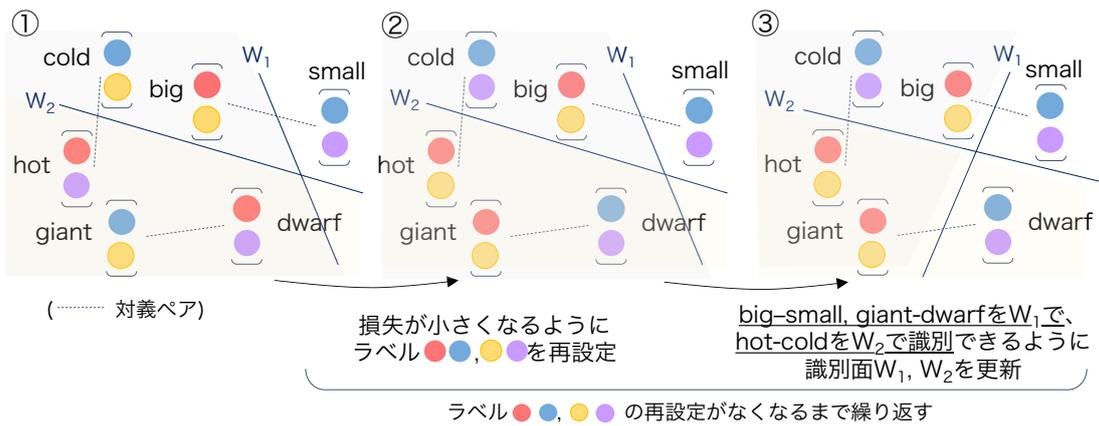


図 5: 複数識別面を学習する際のラベル設定・識別面更新

5 対義性の分散表現への埋め込み

3章の提案モデルによる感情極性を埋め込んだ分散表現の評価実験（6.1章：感情極性分類）において，極性の識別性能が高かった同時学習モデルを用い，分散表現に対義性の埋め込みを行う。

単語ベクトル，及び識別面の初期値は，SGNSモデルで学習した単語ベクトル，及びSGNSの単語ベクトルを用いてアルゴリズム1により学習した識別面とする。初期化後，

1. ベクトル固定で対義シードの単語ペアについてラベル設定
2. 単語ベクトル学習と，1で設定したラベルによる対義シード識別の同時学習で，単語ベクトルを更新
3. 2で得られた単語ベクトルを用いて，ベクトル固定で識別面の更新

1～3を，ハイパーパラメータとして設定したエポック数の間，繰り返す（アルゴリズム2）。対義性の埋め込みは，単語ベクトルをを更新するため，本稿では以降この手法をベクトル変更と表記する。

アルゴリズム 2 対義性の埋め込み（ベクトル変更）

Input: SGNSで学習した単語ベクトル集合 V_{SGNS} ，

V_{SGNS} を用いてベクトル固定で学習した識別面 w_k^{SGNS}

Output: 単語ベクトル集合 V ，識別面 w_k

単語ベクトル集合 $V \leftarrow V_{SGNS}$

▷ 単語ベクトルの初期化

識別面 $w_k \leftarrow w_k^{SGNS}$

▷ 識別面の初期化

for each epoch do

対義シードのペア毎にラベル設定

単語ベクトル学習と同時に識別面 w_k について対義シードを分離

学習した単語ベクトル V 用いて識別面 w_k 更新

分散表現と極性識別の同時学習は，SGNSの目的関数（式1）と対義語識別のためのヒンジ損失関数（式6）の和（式7）を最小化する学習を行う。ただし，式

6における対義語ペア毎の損失 $l^m(p_j)$ (式4) のバイアス $b = 0$ とする.

$$\begin{aligned}
 & - \sum_{t \in V} \sum_{c \in C_t} \left(\ln \sigma(\vec{v}_t \cdot \vec{v}_c) + \sum_{k=1}^K \ln \sigma(-\vec{v}_t \cdot \vec{v}_c) \right) \\
 & - \lambda \|\vec{v}_t\|^2 - \lambda \|\vec{v}_c\|^2 + \frac{1}{N} \sum_{j=1}^N \min[l^1(p_j), l^2(p_j), \dots, l^k(p_j)]
 \end{aligned} \tag{7}$$

SGNS の確率的勾配降下法 (SGD) による単語ベクトル更新と同時に, 式 6 を最小化するための更新を行う.

対義シードのラベル及び識別面の更新は, アルゴリズム 1 : 対義の識別面の学習 (ベクトル固定) に従い, 対義語ペア毎に損失を計算してペア毎の損失が小さくなるようにラベルを設定し, そのラベルを用いて識別面を再学習する.

6 評価実験

6.1 感情極性分類

3章で提案した初期化モデル, 同時学習モデルにおける感情極性の識別性能を確かめるため, 単語に対し, その感情極性がポジティブかネガティブかを判定する二値分類で評価を行った.

6.1.1 実験設定

周辺文脈との共起情報学習データとして, 英語版 Wikipedia を用いた. 出現単語数は約 16 億語, 出現頻度 150 以上の単語を語彙とし語彙数は約 15 万語である.

感情極性の教師データ (シード) として, MPQA[20] からタイプが strongly subjective で, ポジティブまたはネガティブいずれかのラベルが付与された単語のうち, Wikipedia から作成した語彙内の 3499 語 (ポジティブ 1255 語, ネガティブ 2244 語) を用いた.

評価データとして, Opinion Lexicon[21] から, ポジティブまたはネガティブのどちらかで, シードに用いた単語と重複しない 2281 語 (ポジティブ 763 語, ネガティブ 1518 語)¹を用いた.

各モデルについて, ベクトルの次元数は 200 次元, 共起をカウントする文脈の窓幅は前後 3 語, 負例サンプル²数は 3, 正則化項のパラメータ $\lambda = 1/16384$ とした. SGD のエポック数は 1 とした.

単語または文脈ベクトル \vec{v}_i に対する極性の識別関数 $f(\vec{v}_i)$ は, 式 2, 4 における \vec{w} を用いて

$$f(\vec{v}_i) = \vec{w} \cdot \vec{v}_i \quad (8)$$

となり, 感情極性の識別については, \vec{w} を式 3 としたことから, 学習した各ベクトル $\vec{v}_i = (v_{i1}, \dots, v_{id})$ のポジティブな成分の和と, ネガティブな成分の和の, 差 $\sum_{j=d/2}^d v_{ij} - \sum_{j=1}^{d/2} v_{ij}$ に対応する.

¹多数クラス (ネガティブ) が 66.54% となった.

²サンプリングには単語のユニグラム確率の 0.75 乗を用いた.

識別関数の値が正の時ポジティブ、負の時ネガティブを表す。単語 (target) ベクトル、文脈 (context) ベクトルをそれぞれ用い、各々の識別関数によって評価データで単語の感情極性分類を行う。

制約の効果の範囲を確かめるため、シード単語の近傍 (いずれかのシード単語とのコサイン類似度が 0.5 より大きい) 語のみについても分類を行なった。

ベースラインとして、SGNS モデルで学習したベクトルの各次元値を特徴量に、シード単語の極性ラベルを教師として分類器 (SVM) の訓練を行い、評価データで分類 (open test) を行う。SVM は libsvm³ は RBF カーネルを用い、単語、文脈ベクトルともに正規化する。分類器が感情極性の訓練データの情報を十分に捉えていることを確認するため、訓練データにおいても評価 (closed test) を行う。

6.1.2 実験結果

表 1: 識別関数による単語感情極性の分類正解率

		全体	シード近傍
初期化モデル	target	75.31	89.46 (2012 語)
	context	87.59	-
同時学習モデル	target	87.29	90.19 (1947 語)
	context	88.20	-

表 2: SGNS のベクトルを特徴量とした SVM による単語感情極性の分類正解率

		open test	closed test
SVM-SGNS	target	91.10	99.54
	context	91.36	99.51

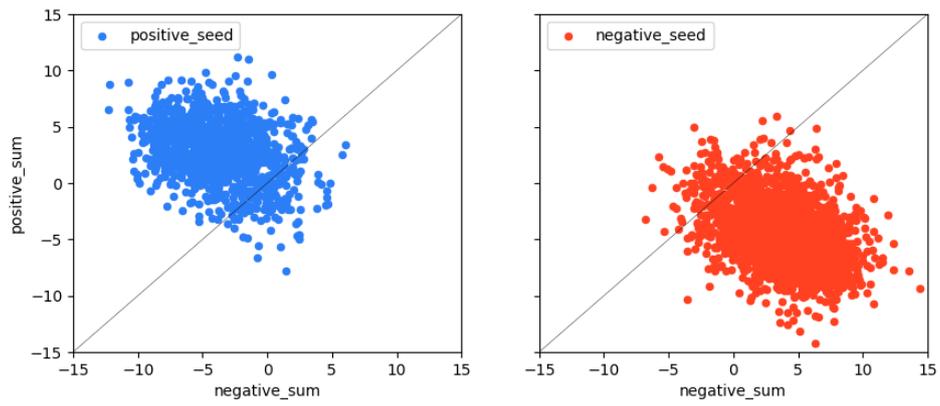
感情極性分類の結果を表 1, 表 2 に示す。表 1 のシード近傍については、評価

³libsvm (<https://www.csie.ntu.edu.tw/~cjlin/libsvm/>) の実装を用い、パラメータは $\gamma = 1$, その他はデフォルトの設定とした。

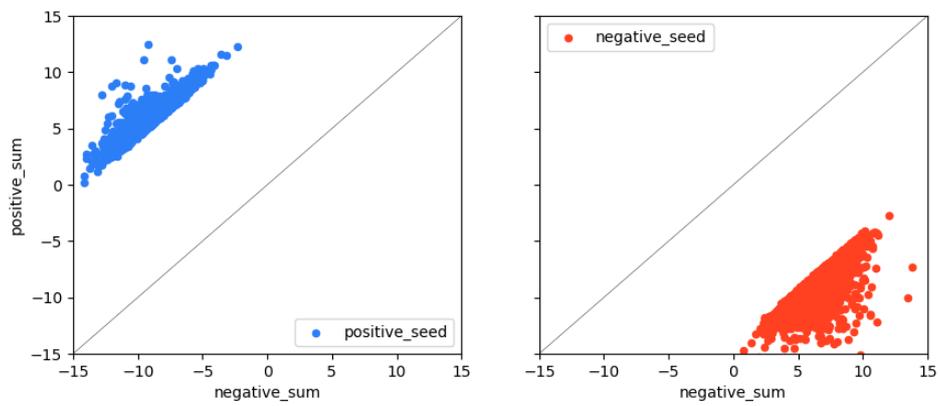
データ全体で評価した場合と比較して変化があった，単語（target）ベクトルの識別関数を用いた結果のみ示す．

SGNSモデルの分散表現を特徴量に，シード単語を教師データとして訓練したSVMが最も高い識別性能を示した．初期化モデルはシード近傍において，評価データ全体で分類を行なった場合と比べ正解率が高く，シード単語に近い単語については感情極性の特徴を捉えていると考えられる．同時学習モデルは単語ベクトル，文脈ベクトルの識別関数を用いた場合ともに評価データ全体においてもSVMに匹敵する分類正解率となり，初期化モデルに比べ極性の汎化能力が高いと推測される．また，両提案モデルの結果に共通して，文脈ベクトルの識別関数の方が単語ベクトルの識別関数に比べて高い識別性能を示し，文脈ベクトルに極性の違いが強く現れる傾向がみられた．

単語の極性を線形分離するように保持すると仮定した成分が各提案モデルでどう学習されているか確かめるため，極性の違いをより保持できると考えられる文脈ベクトルについて，シードのポジティブ成分の和，ネガティブ成分の和をそれぞれ縦軸，横軸として図6のようにプロットした．各図の左側（青）がポジティブ，右側（赤）がネガティブなシードを表し，図中の斜線は識別関数が0となる，すなわち式2, 4における $\vec{w} \cdot \vec{v}_i$ が $\vec{w} \cdot \vec{v}_i = 0$ となる境界を表す．極性のクラスラベルはポジティブを1，ネガティブを-1としていることから，境界の上側がポジティブ，下側がネガティブな領域とみることができる．シードの極性に応じて初期化を行うだけでも，ある程度極性を分離でき，同時学習モデルでは，さらに識別面に関してポジティブな成分とネガティブな成分がよく分離されるように学習できていると言える．



(a) 初期化モデル



(b) 同時学習モデル

図 6: シード文脈ベクトルに対する識別関数値の分布

6.2 対義語・同義語の判定

対義の識別面の学習（ベクトル固定）、及び同時学習による対義性の埋め込み（ベクトル変更）の、対義語の識別における効果を測るため、対義語ペアと同義語ペアからなる評価データを用い、性能評価実験を行った。

ただし、提案手法は単語のラベルを分類するように学習するため、評価データの単語ペアに対し、異なるラベルを予測すれば対義語ペア、同じラベルを予測すれば同義語ペアと判定できたものとした。

なお、ベクトル変更で得られた分散表現と識別面を用いる際の識別関数は、式 8 である。

6.2.1 実験設定

対義性の教師データ（対義シード）には、WordNet[22] 及び Roget's Thesaurus[23] から作成されたデータ⁴より、抽出したペア間の単語重複を除き、語彙内の 4996 ペア（9992 語）を用いた。

周辺文脈との共起情報学習データ、及び学習時のパラメータは、感情極性分類（6.1.1 章）と同じデータ及び設定を用い、SGD のエポック数は 2 とした。

また、感情極性分類の評価実験（6.1.2 章）において、極性の識別性能が高かった文脈ベクトルに対義性の埋め込みを行った。

ベクトル固定（4 章）における識別面の学習には SVM を用い、SGNS の単語ベクトルを特徴量とした。

提案手法のラベル設定に対するベースラインとして、対義語ペアの単語同士は異なるラベルにする条件は変えず、ペアのどちらの単語が正または負のラベルとなるかはランダムに設定し、SGNS の単語ベクトルを特徴量として訓練した SVM を用いた。ベースラインのランダムなラベルは、4 つの異なるラベル設定を行い、結果はその平均を示す。SVM のパラメータ等の設定は感情極性分類の評価実験（6.1.1 章）と同様である。

⁴<https://github.com/antonyms/AntonymPipeline/tree/master/data>

評価データには, [24][25] で使われた対義語・同義語ペアのデータセット⁵から, 形容詞・動詞・名詞の全ペアを合わせた 2100 ペアのうち, 語彙内で対義シードのペアを除く 1954 ペア (対義 971 ペア, 同義 983 ペア) を用いた.

6.2.2 実験結果

図 7, 図 8, 図 9 に対義クラスについての精度 (precision), 再現率 (recall), F 値を示す. それぞれ横軸は識別面の数, 縦軸は各指標で, 識別面の数を増やしていった時の各指標の変化を表し, 破線はチャンスレートを示している. ベースラインとしたランダムなラベル設定の SVM では, precision 0.51, recall 0.46, F 値 0.48 (それぞれ図 7, 図 8, 図 9 中, 識別面の数 1 における Δ , \square , \circ で表示) であった.

ここで, precision は分離されたペアのうち実際に対義であった割合を, recall は実際に対義であるペアのうち分離できたペアの割合を表す.

図 7 から, ベクトル固定では識別面 1 つのときランダムなラベル設定のベースラインに対して, より正確に対義を見分けられており, 学習と推論を交互にを繰り返しながら, ある程度意味味の側面を捉えた識別面を学習できる可能性がある. 一方で, 識別面の数を増やすと, どの識別面でも単語を分離しないことは難しくなり, ほぼチャンスレートとなる.

図 8 から, 識別面 1 つのときランダムなラベル設定のベースラインよりも対義を分離できており, 提案手法のラベル設定に効果があったと考えられる. 識別面の数を増やすと, チャンスレートに沿って対義を分離できる割合は高くなり, 複数の識別面のうちどれか 1 つで分離されたら対義としているためだと考えられる. ベクトル変更における分散表現と対義識別の同時学習による, ベクトル固定の学習で得られた対義の軸の情報を, ある程度分散表現に埋め込むベクトル変更の結果がベクトル固定と比べて低下することについては, 訓練データに過学習している可能性が考えられる.

⁵http://www.ims.uni-stuttgart.de/forschung/ressourcen/experiment-daten/lexical_contrast_dataset.html

提案手法で学習された対義の識別面によって、どのような対義性が分離されるか、対義語・同義語の判定（6.2.2章）でF値が最も高かった、ベクトル固定で識別面の数が4のとき、各識別面で識別された対義語ペアの例を表3に示す。表3は、ベクトル固定で識別面の数が4のとき、各識別面において損失最小で分離された対義語ペアの例である。各識別面で分離されたペアから、対義を分離する軸として、例えば識別面1は「ネガティブ・ポジティブ」、識別面2は「数や安定性」、識別面3は「構成」、識別面4は「大きさや自然」のような基準で分けられていると解釈しうる。このように各識別面で識別されたペアから、どのような対義の軸が存在するか、解釈の可能性があり、またベクトル固定の学習では、SGNSで学習した通常の単語ベクトルを特徴量に用いていることから、通常の単語ベクトルでも対義を識別する軸が存在しているのではないかと考えられる。

識別面 1	識別面 2	識別面 3	識別面 4
unfavorable, favorable	terminal, unfinished	organized, unstructured	shallow, oceanic
unreasonable, philosophical	tribal, singular	manual, visual	arboreal, urban
mortal, immortal	tribal, lone	universal, tribal	external, civic
unlawful, lawful	exoteric, mystical	pneumatic, stuffy	trivial, climatic

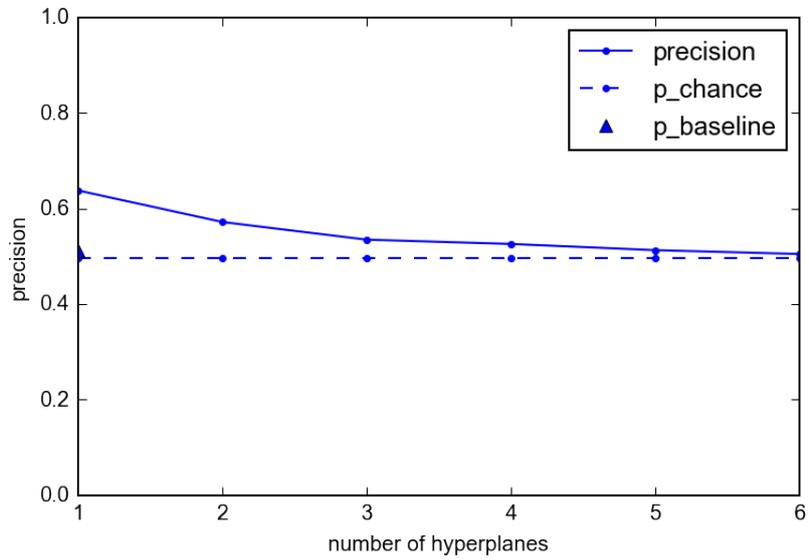
表 3: 各識別面で識別された対義語ペアの例

6.2.3 同義の訓練データ追加

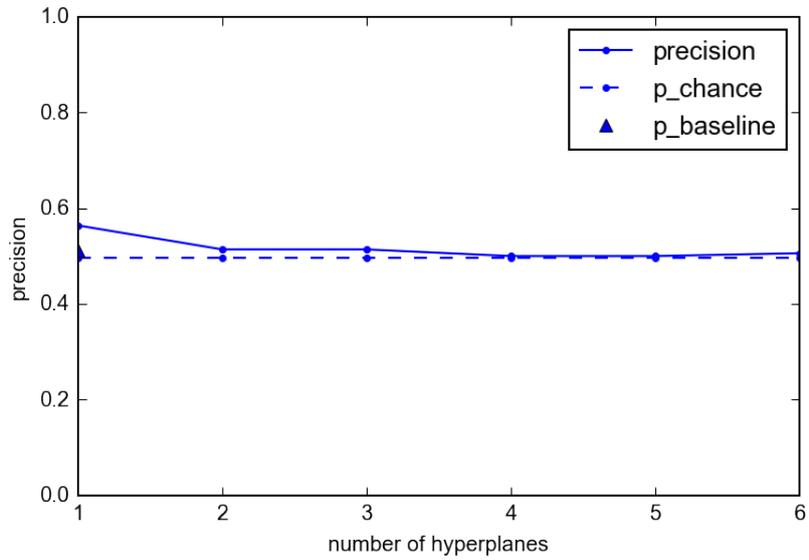
識別面の数を増やすと、対義識別の precision（分離されたペアのうち実際に対義であった割合）が下がる問題に対し、極性の訓練データとして、対義のペア情報に加え、同義のペア情報も利用して学習する設定を試した。同義の訓練データとして、WordNet[26] から、対義の訓練データと単語が重複しない 3546 ペアを用いた。

ベクトル固定で、対義と同義の訓練データを両方用いて識別面を学習したときの、対義語・同義語判定の結果（Macro-F1）を図 10 に示す。

訓練データとして対義情報のみを用いる場合と比べ、同義情報も利用することで、識別面の数を増やしたとき、recall は上がりにくくなるが、precision の低下

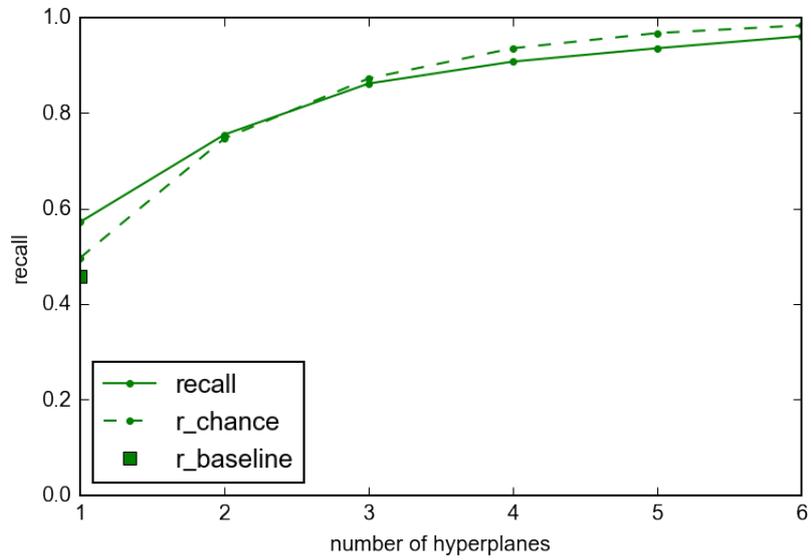


(a) ベクトル固定

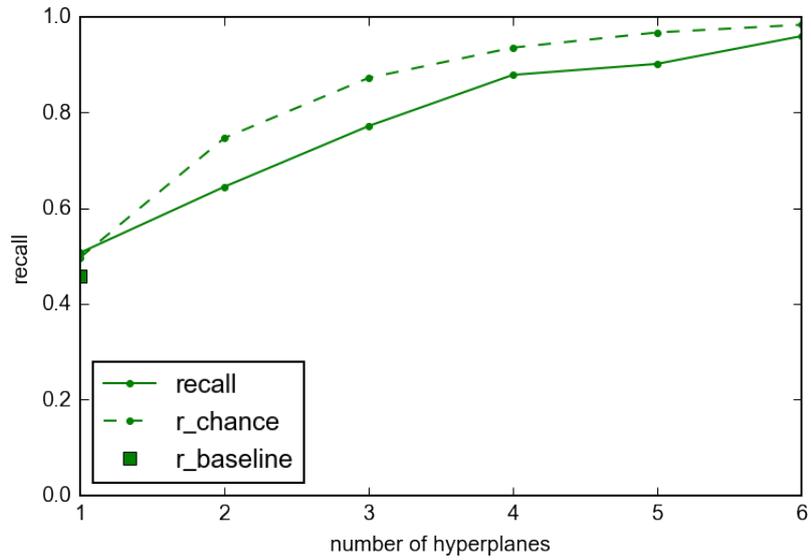


(b) ベクトル変更

図 7: 対義語・同義語の判定における対義語識別の精度

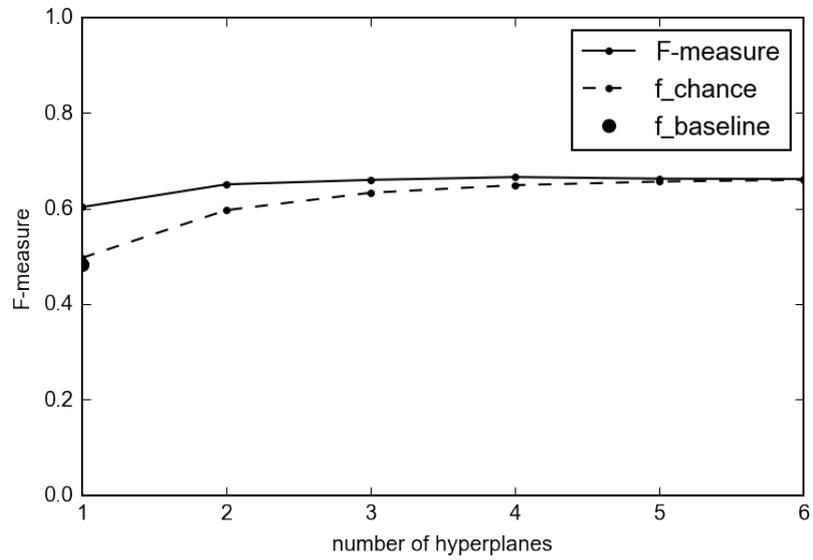


(a) ベクトル固定

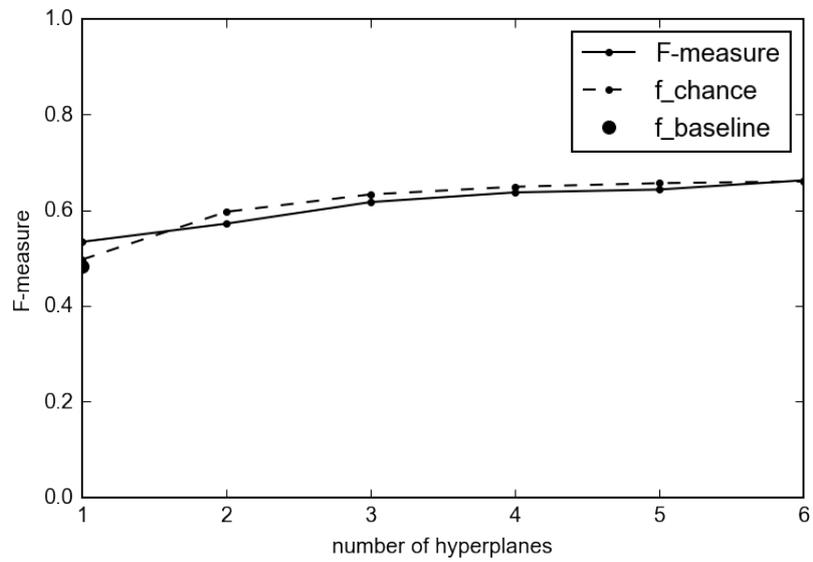


(b) ベクトル変更

図 8: 対義語・同義語の判定における対義語識別の再現率



(a) ベクトル固定



(b) ベクトル変更

図 9: 対義語・同義語の判定における対義語識別の F 値

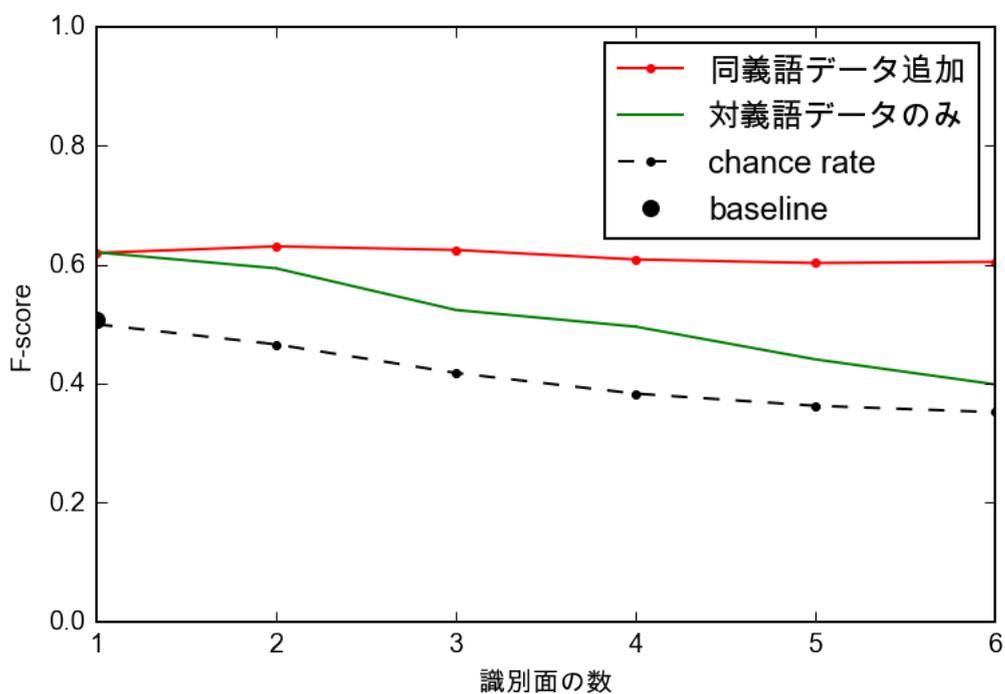


図 10: 同義語の訓練データ有無と対義語・同義語の判定性能 (Macro-F1)

が緩和され、全体として F 値は向上する。同義語を誤って分離してしまうことが少なくなると言え、より正確に対義性を学習するために、同義情報も用いることは有効だと考えられる。

6.3 類義性・加法構成性の評価

ベクトル変更の同時学習で得られた分散表現は，単語分散表現に期待される類義性や加法構成性が保たれるか検証するため，単語間の類似度，及び関係類推の評価データセットを用いて類義性，及び加法構成性の評価を行った。

単語間類似度の評価データセットは，単語ペアについて，人が各単語の類似度に応じて判断したスコア付与したもので，分散表現での類似度との相関で評価する。

関係類推の評価データセットは，例えば「東京」に対する「日本」の関係が「パリ」に対する「フランス」の関係に相当するとき，「東京：日本」と「パリ：？」を与えられ，？に対応する答えは「フランス」であると当てるようなタスクになっている。

6.3.1 実験設定

類似度の評価データは9種類，関係類推の評価データは2種類のデータセットを用いた。

評価指標は，類似度については Spearman の順位相関係数 (ρ)，関係類推については正解率 (Acc) を用いる。

ベクトル変更による単語分散表現は，対義語・同義語判定の評価実験 (6.2章)と同じ単語ベクトルを用いた。ベースラインとして，SGNSで2エポック学習した単語ベクトルを用い，その他の学習設定は，6.1.1章と同様である。

比較手法として，関連研究の一つ Counter-fitting[11]を用いた。Counter-fittingは，学習済みの単語ベクトルに，ベクトル空間で同義語を近づけ対義語を遠ざけるような後処理を加える手法で，著者が公開している実装⁶及び同義語・対義語データを用い，SGNSの単語ベクトルに後処理を行った。

6.3.2 実験結果

類似度評価データセット・関係類推評価データセットにおける，スピアマンの順位相関係数 (ρ)・正解率 (Acc) を表4，表5に示す。表4では，類似度の結果

⁶<https://github.com/nmrksic/counter-fitting>

は全9種類の評価データでのマクロ平均、関係類推の結果は全2種類の評価データを合わせたデータでの正解率を、表5では、各評価データセットでのスピアマンの順位相関係数、正解率を示す。表4から、極性を識別する分散表現学習の既存手法（Counter-fitting）と比較して、ベクトル変更の同時学習により、単語分散表現に期待される類義性や加法構成性における有効性を大きく変えずに、対義性のある程度埋め込める（6.2.2章：対義語・同義語の判定の実験結果）ことがわかる。

	類似度 (ρ)	関係類推 (Acc)
SGNS	63.42	61.22
ベクトル変更	63.25	61.49
Counter-fitting	59.38	47.40

表 4: 類義性（類似度）・加法構成性（関係類推）の評価

特に、同義語に限らず関連性（“relatedness”）の高い単語を類義とする WSR や、関係類推のデータセットにおいて、同時学習による対義性埋め込みは、単語ベクトルの特性を保ち、向上させる場合もある。

	類似度 (ρ)									関係類推 (Acc)
	MEN	MC	MTurK	RARE	R&G	SCWS	Simlex	WSR	WSS	
SGNS	70.2	67.0	78.1	44.7	76.6	65.9	35.0	57.8	75.5	62.75
ベクトル変更 (6 識別面)	70.6	66.4	73.2	45.2	75.6	66.4	35.5	60.7	75.7	62.80
Counter-fitting	65.9	61.3	71.1	41.5	76.4	62.1	35.6	49.9	70.6	49.72

表 5: 評価データセット別の類義性（類似度）・加法構成性（関係類推）の評価

7 おわりに

本論文では、単語の文脈的な類似性と異なる極性の両方を表せる単語分散表現の実現を目的として、単語の分散表現学習と、辞書情報を利用して極性の識別面に関する分離を同時に行う手法を提案した。さらに、異なる極性として感情極性以外の対義関係も識別するため、対義の識別面を複数学習する手法を提案した。感情極性分類の評価実験において、提案手法で単語の極性をある程度汎化して分散表現に埋め込めることを示した、さらに、対義語・同義語判定の評価実験において、異なる対義の性質を捉える複数の対義の識別面を学習しうることを示した。また、単語の類義性・加法構成性の実験において、提案手法で学習した単語分散表現は、元の分散表現の重要な性質である文脈的な類似性も保つことを確認した。

謝辞

本研究を進めるにあたり，ご指導・ご助言くださいました乾健太郎教授，鈴木潤准教授に感謝いたします。また，ご多用の中，本論文の審査をお受けくださいました周暁教授，篠原歩教授に感謝いたします。

研究の理論面，実装面ともに多くのご助言をいただきました田然前研究特任助教に感謝いたします。また，研究に関する日々の議論や学生生活でお世話になりました研究室の皆様感謝いたします。

末筆ながら，多大に支えていただいた家族と友人に感謝いたします。

参考文献

- [1] Ronan Collobert, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. Natural language processing (almost) from scratch. *J. Mach. Learn. Res.*, Vol. 12, pp. 2493–2537, November 2011.
- [2] Danqi Chen and Christopher Manning. A fast and accurate dependency parser using neural networks. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 740–750, 2014.
- [3] Oren Melamud, David McClosky, Siddharth Patwardhan, and Mohit Bansal. The Role of Context Types and Dimensionality in Learning Word Embeddings. In *Proceedings of NAACL-HLT 2016*, pp. 1030–1040, 2016.
- [4] Zellig S Harris. Distributional structure. *Word*, Vol. 10, No. 2-3, pp. 146–162, 1954.
- [5] Manaal Faruqui, Jesse Dodge, Sujay Kumar Jauhar, Chris Dyer, Eduard Hovy, and Noah A. Smith. Retrofitting Word Vectors to Semantic Lexicons. In *NAACL*, 2015.
- [6] Masataka Ono, Makoto Miwa, and Yutaka Sasaki. Word Embedding-based Antonym Detection using Thesauri and Distributional Information. *NAACL*, 2015.
- [7] Nghia The Pham, Angeliki Lazaridou, and Marco Baroni. A Multitask Objective to Inject Lexical Contrast into Distributional Semantics. In *ACL*, 2015.
- [8] Song Feng, Jun Seok Kang, Polina Kuznetsova, and Yejin Choi. Connotation Lexicon: A Dash of Sentiment Beneath the Surface Meaning. In *ACL*, 2013.
- [9] Kim Anh Nguyen, Sabine Schulte im Walde, and Ngoc Thang Vu. Integrating Distributional Lexical Contrast into Word Embeddings for Antonym-Synonym Distinction. In *ACL*, 2016.

- [10] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. In *NIPS*, Vol. 2, 2013.
- [11] Nikola Mrkšić, Diarmuid Ó Séaghdha, Blaise Thomson, Milica Gašić, Lina Rojas-Barahona, Pei-Hao Su, David Vandyke, Tsung-Hsien Wen, and Steve Young. Counter-fitting Word Vectors to Linguistic Constraints. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 142–148, mar 2016.
- [12] Zhigang Chen, Wei Lin, Qian Chen, Xiaoping Chen, Si Wei, Hui Jiang, and Xiaodan Zhu. Revisiting Word Embedding for Contrasting Meaning. In *ACL*, 2015.
- [13] Richard Socher, Brody Huval, Christopher D. Manning, and Andrew Y. Ng. Semantic Compositionality through Recursive Matrix-Vector Spaces. In *EMNLP*, 2012.
- [14] Mohit Iyyer, Varun Manjunatha, Jordan Boyd-Graber, and Hal Daumé III. Deep Unordered Composition Rivals Syntactic Methods for Text Classification. In *ACL*, 2015.
- [15] Duyu Tang, Furu Wei, Bing Qin, Nan Yang, Ting Liu, and Ming Zhou. Sentiment Embeddings with Applications to Sentiment Analysis. *IEEE Transactions on Knowledge and Data Engineering*, 2016.
- [16] Duy Tin Vo and Yue Zhang. Don’t Count, Predict! An Automatic Approach to Learning Sentiment Lexicons for Short Text. In *ACL*, 2016.
- [17] Abhishek Kumar and Jun Araki. Incorporating Relational Knowledge into Word Representations using Subspace Regularization. In *ACL*, 2016.

- [18] Sascha Rothe and Hinrich Schütze. Word Embedding Calculus in Meaningful Ultradense Subspaces. In *ACL*, 2016.
- [19] Thorsten Joachims. Transductive Inference for Text Classification using Support Vector Machines. In *16th International Conference on Machine Learning (ICML-99)*, pp. 200–209, 1999.
- [20] Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. In *HLT-EMNLP*, 2005.
- [21] Minqing Hu and Bing Liu. Mining and Summarizing Customer Reviews. In *KDD '04*, 2004.
- [22] Princeton University. About WordNet. <https://wordnet.princeton.edu>, 2010.
- [23] B.A. Kipper. *Roget's 21st Century Thesaurus*. Philip Lief Group, 2009.
- [24] Michael Roth and Sabine Schulte im Walde. Combining word patterns and discourse markers for paradigmatic relation classification. In *ACL*, 2014.
- [25] Kim Anh Nguyen, Sabine Schulte Im Walde, and Ngoc Thang Vu. Integrating Distributional Lexical Contrast into Word Embeddings for Antonym-Synonym Distinction. In *ACL*, pp. 454–459, 2016.
- [26] Christiane Fellbaum. *WordNet: An Electronic Lexical Database*, Vol. 71. 1998.
- [27] Elia Bruni, Nam Khanh Tran, and Marco Baroni. Multimodal distributional semantics. *J. Artif. Int. Res.*, Vol. 49, No. 1, pp. 1–47, January 2014.
- [28] Kira Radinsky, Eugene Agichtein, Evgeniy Gabrilovich, and Shaul Markovitch. A word at a time: Computing word relatedness using temporal semantic analysis. In *Proceedings of the 20th International Conference*

- on *World Wide Web*, WWW '11, pp. 337–346, New York, NY, USA, 2011. ACM.
- [29] George A. Miller and Walter G. Charles. Contextual correlates of semantic similarity. *Language and Cognitive Processes*, Vol. 6, pp. 1–28, 01 1991.
- [30] Thang Luong, Richard Socher, and Christopher Manning. Better word representations with recursive neural networks for morphology. In *Proceedings of the Seventeenth Conference on Computational Natural Language Learning*, pp. 104–113. Association for Computational Linguistics, 2013.
- [31] Herbert Rubenstein and John B. Goodenough. Contextual correlates of synonymy. *Commun. ACM*, Vol. 8, No. 10, pp. 627–633, October 1965.
- [32] Eric H. Huang, Richard Socher, Christopher D. Manning, and Andrew Y. Ng. Improving word representations via global context and multiple word prototypes. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers - Volume 1*, ACL '12, pp. 873–882, Stroudsburg, PA, USA, 2012. Association for Computational Linguistics.
- [33] Felix Hill, Roi Reichart, and Anna Korhonen. Simlex-999: Evaluating semantic models with (genuine) similarity estimation. *Computational Linguistics*, Vol. 41, No. 4, pp. 665–695, 2015.
- [34] Eneko Agirre, Enrique Alfonseca, Keith Hall, Jana Kravalova, Marius Paşca, and Aitor Soroa. A study on similarity and relatedness using distributional and wordnet-based approaches. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, NAACL '09, pp. 19–27, Stroudsburg, PA, USA, 2009. Association for Computational Linguistics.
- [35] Tomas Mikolov, Kai Chen, Gregory S. Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *CoRR*, Vol. abs/1301.3781, , 2013.

- [36] Tomas Mikolov, Wen-tau Yih, and Geoffrey Zweig. Linguistic regularities in continuous space word representations. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 746–751. Association for Computational Linguistics, 2013.

付録

A 単語間の類似度・関係類推の評価データセット

abbreviation	seen in vocabulary / size (words)	reference
類似度		
MEN	3000/3000	[27]
MTurk	285/287	[28]
M&C	30/30	[29]
RARE	1221/2034	[30]
R&G	65/65	[31]
SCWS	1968/2003	[32]
SLex	998/999	[33]
WSR	238/252	[34]
WSS	196/203	[34]
関係類推		
GL	19364/19544	[35]
MSYN	5926/8000	[36]

表 6: 各評価データセットのサイズ・参照文献

発表文献一覧

国際会議論文

- Taku Nakamura, Rafal Rzepka, Kenji Araki, and Kentaro Inui. Be More Eloquent, Professor ELIZA Comparison of Utterance Generation Methods for Artificial Second Language Tutor. In Proceedings of Linguistic and Cognitive Approaches to Dialog Agents (LaCATODA 2017), pp.34-41, August 2017.

国内会議・研究会論文

- 中村拓, 田然, 乾健太郎. 単語の極性を埋め込んだ分散表現. 言語処理学会第24回年次大会, pp.348-351, March 2018.

ワークショップ

- 中村拓, 田然, 乾健太郎. 単語の極性を埋め込んだ分散表現. 第21回情報論的学習理論ワークショップ (IBIS), November 2018