

# Sentence-BERT と語義定義文を利用した語義間の類義判定手法

石井佑樹<sup>1</sup> 佐々木稔<sup>2</sup>

<sup>1</sup>茨城大学大学院理工学研究科情報工学専攻 <sup>2</sup>茨城大学工学部情報工学科

{22nm702y, minoru.sasaki.01}@vc.ibaraki.ac.jp

## 概要

自然言語処理技術を開発する上でシソーラスから得られる意味知識は有用である。しかし、既存研究では主に単語間の関係を検出することに注目し、語義を対象とした類義関係の検出は行われていない。本研究では日本語辞書に記述された語義定義文と Sentence-BERT を用いた類義判定手法を提案する。岩波国語辞典の見出し語、語義定義文及び分類語彙表の分類番号を用いて作成した評価データセットを対象として、学習データを用いて fine-tuning した類義判定モデルによる類義判定実験を行った。実験の結果、提案手法における Sentence-BERT や変更した語義定義文を用いることによって、ベースライン手法よりも効果的に類義判定ができることを示した。

## 1 はじめに

自然言語処理において単語同士の上位下位関係や同義関係、類義関係等の意味知識を獲得することは機械翻訳や質問応答等の技術を開発する上で重要な課題である[1]。意味知識を獲得するのにあたり、膨大な語が存在する中から手動で知識を獲得するのは困難であり知識獲得の自動化が行われてきた[2][3][4][5][6]。また機械学習が発展してきた近年では、特に文章中の多義語に対して適切な意味を識別するタスクである語義曖昧性解消 (WSD) において、シソーラスから得られる意味知識を学習データに利用することで教師あり学習に必要なアノテーションのコスト削減に成功し[7][8]、シソーラスを活用した知識ベースの学習手法の有用性が示されてきた。

しかし、類義語などの語彙知識獲得を行う研究では単語同士の関係に着目するのみであり、語義同士の関係は考慮されていない。単語同士の類義関係であると多義語のどの意味で類義関係なのかが分からないが、どの語義で類義なのかが分かると語義識別や言い換えに役立つ。語義レベルの関係の例として以下の「うまい」と「じょうず」という単語を例として挙げる。今までの研究では単語間の類似性や分

布仮説[9][10]を用いて「うまい」と「じょうず」という単語ペアを同義語であるとして獲得することを行ってきた。語義レベルとは単語の語義に注目し、どの語義が類義であるかという語義間の関係を獲得することである。本研究では岩波国語辞典第五版[11]において意味区分された語義を対象に以下の例において「うまい」の語義 2 と「じょうず」の語義 1 が類義であると判定することを目指す。

- うまい【甘い・旨い】

語義 1: ”味がよい。「一汁を吸う」...

語義 2: ”よい。すぐれている。...

語義 3: ”自分にとって都合がよい。...

- じょうず【上手】

語義 1: ”ある物事をする技術がすぐれている...

語義 2: ”《多く「おー」の形で》世辞。世辞を...

(岩波国語辞典第五版より引用)

また日本語においては英語 WordNet を翻訳した日本語 WordNet[12]が存在するが、英語 WordNet に存在しない類義関係をどう追加するのかといった課題があり、実用可能な完成度に至っていない。

本研究では入力した 2 単語に対して語義レベルで類義判定することを目的とする。目的の実現のため語義定義文の変更と Sentence-BERT による深層距離学習を利用した類義判定手法を提案する。語義定義文に対して語義を表す記述以外の削除や不足する内容の追加等の変更を行うことで、変更前の語義定義文に比べて適切な語義の特徴を捉えた埋め込みベクトルが得られることに期待できる。また深層距離学習とは、埋め込み空間上において同じラベル付けされたデータ同士の距離を小さくし、異なるラベル付けされたデータ同士の距離を大きくする手法であり、これを利用することで本研究の目的とする語義レベルの類義判定用に埋め込み空間の調節がなされ、類義判定精度の向上に期待できる。

関連研究として文脈情報を用いた単語間の同義判定手法[13]、文章エンコーダを fine-tuning することで単語間関係のラベル推測をする手法[14]が挙げられる。後者の研究では単語の関係分類に文章エンコ

ーダの fine-tuning が一定の効果があると示しており、単語に注目したこの手法を語義定義文に対応させるよう変更することで語義の類義判定にも効果が得られると考えられる。

## 2 語義間の類義判定手法

### 2.1 類義判定システム

語義間の類義判定システムを図1に示す。システムの流れとしては、まず fine-tuning を施した類義判定モデルに対象の語義定義文を入力し文章ベクトルを得る。得られた語義定義文の文章ベクトルは mean pooling 操作し BERT 出力を平均ベクトルに変換する。次に得られた2つの文章ベクトルに対してコサイン類似度を求め閾値を越していた場合2つの語義は類義であると判定するシステムである。

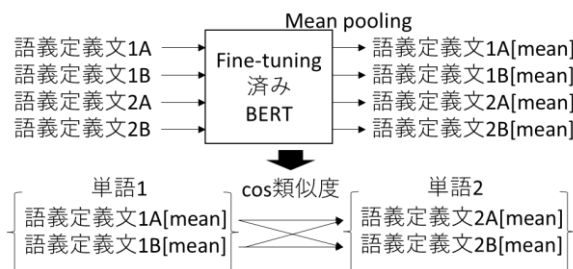


図1: 語義の類義判定システム

BERT 出力の平均を採用した理由としては、Reimers ら[15]が BERT 出力の CLS 及び平均を用いて Semantic Textual Similarity (STS) ベンチマーク [16]を行った結果、平均を用いた方が優れていたためである。また STS ベンチマークは2文間の類似度を測定するベンチマークであり、Reimers ら[15]は本研究との提案手法と同様に2つの文章ベクトルに対してコサイン類似度を求め2文間の類似度を求めているためである。

事前学習済み BERT モデルとして東北大学乾・鈴木研究室が公開している日本語 BERT モデル (bert-basejapanese-whole-word-masking)を使用した。

### 2.2 事前学習済みモデルの fine-tuning

#### 2.2.1 Sentence-BERT (SBERT) 手法

fine-tuning 手法に Reimers ら[15]が提案した Sentence-BERT (SBERT) 手法を採用する。意味ベクトル空間上で類似した文章のベクトルを近くに分布させる SBERT 手法により類義判定に適した文章ベクトルを得るためである。SBERT によるモデル学

習の流れを述べる。まず2つの語義定義文を日本語 BERT に入力する。得られた出力に対して pooling 操作を行い語義定義文の文章ベクトルを得る。Reimers ら[15]は pooling 手法において、出力された全ての単語埋め込みの平均を用いる mean pooling を推奨しているためこれを適用して文章ベクトル  $u, v$  を得た。また、BERT を利用して文章ベクトル得る方法として CLS トークンのベクトルを用いるという方法があり、mean pooling に加え CLS ベクトルを用いた場合の実験も行う。これらのベクトルを  $(u, v, |u-v|)$  として結合し全結合層に入力し学習を行った。学習する語義定義文のペアは類義であるか類義でないかを 0,1 でラベル付けしており、損失関数には対象損失を採用した。SBERT によるモデル学習の流れを図2に示す。また、SBERT 手法のベースライン手法として単語を対象とした類義語判定実験を行う。

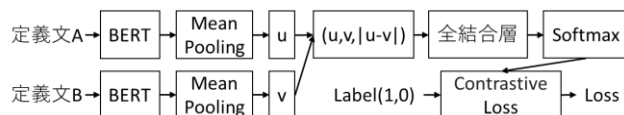


図2: SBERT 手法によるモデル学習

#### 2.2.2 SBERT+Infersent 手法

1章で示した Zhang ら[14]の研究は類義判定においても一定の成果が得られると考えられる。そこで Zhang ら[14]の手法の基礎部分を SBERT 手法と組み合わせた SBERT+Infersent 手法を同時に提案する。Zhang ら[14]の提案する学習手法は、Conneau ら[17]が提案した Infersent を文章エンコーダの学習に採用しており、学習された文章エンコーダを更に意味関係の推測が行えるよう学習させるものである。本研究ではコサイン類似度を用いて類義の判定を行うため、Infersent の文章エンコーダを学習させる部分を参考に学習に適用する。SBERT+Infersent 構造は図3の流れで類義判定モデルを fine-tuning する。

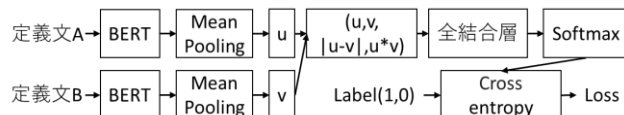


図3: SBERT+Infersent 手法によるモデル学習

### 2.3 語義間の類義ラベル付け

語義間の類義判定モデルを構築するために、語義同士が類義か否かをラベル付けした学習データを手作業で作成した。対象は岩波国語辞典に記載されて

いる単語とその語義定義文とし、対象単語の類義語ペアの選出には分類語彙表のレコード<sup>i</sup>に含まれる分類番号と段落番号を利用する。

具体的には、まず分類番号と段落番号が同じ単語A,Bを類義語ペアとして選出する。このとき一方の単語は必ず語義が2つ以上存在する。単語Aは3つの語義を持ち、単語Bは2つの語義を持つ場合は、AとBの語義定義文の全ての組み合わせ(6組)に対して人手によって類義であると判断したペアは1、類義でないと判断したペアに0を付与し2値ラベル付けを行った。

ペアの中には類似した意味を複数持つ定義文が存在する場合があります、この例を図4に示す。「くどい」の定義文は「しつこい」の2つの語義の意味を含むため、21578-0-0-1-0と21578-0-0-2-0の定義文は13612-0-0-0-0と類義であるとラベル付けしている。

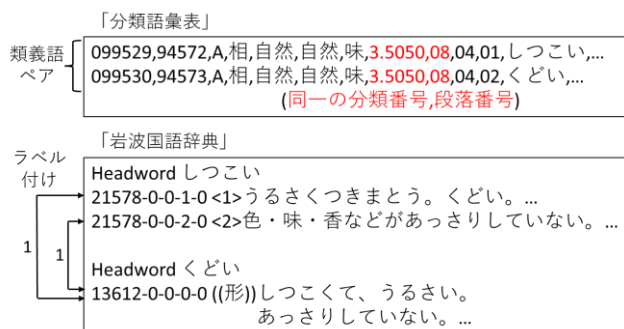


図4：語義間の語義ラベル付けの例

## 2.4 定義文の変更

定義文の変更については、石井ら[18]が提案した変更方法を適用する。石井ら[18]は、語義の類義判定に有効な語義定義文の文章ベクトルを得る変更方法として以下に示す設定A①～④を岩波国語辞典の定義文に適用しており、この変更方法を本研究に採用する。

- ・語義説明の無い定義文の削除(文変更無し) A
- ・辞書内表現・記号の削除・置換, 全角英字を半角英字に置換, 用例文内のハイフン「—」の置換(ひらがな置換)①
- ・用例文内のハイフン「—」の置換(一対一置換)②
- ・追記情報の削除③
- ・全角括弧「( )」表現を削除④

<sup>i</sup> <https://clrd.ninjal.ac.jp/goihyo.html>

## 3 実験

### 3.1 実験データ

2.3, 2.4節をもとに、岩波国語辞典に記載された単語から名詞、動詞、形容詞に分類される単語ペアそれぞれ50組のデータを選出した。この50組のデータを訓練用に30組、検証用に10組、テスト用に10組に分けた。また、各品詞の訓練データ、検証データ、テストデータをすべて組み合わせたALLを用意した。

更に語義定義文を用いた類義判定の有効性を示すため単語データセットを用意し、単語を対象とした類義判定を提案手法(SBERT)で行い結果を比較する。単語は語義データセットに含まれる名詞、動詞、形容詞を用いる。単語数は各品詞において訓練データに60個、検証・テストデータにそれぞれ20個存在する。学習及びテストを行う類義単語ペアの数は各品詞において、訓練データは1770組、検証・テストデータは45組となる。また語義データセット同様全ての品詞データを合わせたALLを用意した。

### 3.2 評価指標

本研究の評価は類義の判定をする二値分類タスクであるためmacro-F1スコアを評価指標として求めた。類義の判定においては、類似度を測るコサイン類似度の閾値を0.3から0.95まで0.01刻みで変更し検証データを用いてmacro-F1が最高値となる閾値を求め、テストデータに適用しmacro-F1を求めた。

### 3.3 実験設定

SBERT手法及びSBERT+InferSent手法による類義判定モデルfine-tuning時のハイパーパラメータを表1に示す。

表1：各モデルのハイパーパラメータ (左表:SBERT手法, 右表 SBERT+Infersent手法)

| Hyperparameter             | 値         | Hyperparameter      | 値        |
|----------------------------|-----------|---------------------|----------|
| Adam-学習率                   | 2e-5      | 全結合層-中間層次元数         | 512      |
| warmupステップ                 | ※         | 全結合層-中間層数           | 1        |
| 評価ステップ数                    | {10,100}  | 全結合層-Adam, SGD学習率   | 0.01     |
| エポック数                      | {1, 2, 3} | 全結合層-SGD荷重減衰        | 0.99     |
| ミニバッチ数                     | 16        | 全結合層-Adam荷重減衰       | 0        |
| ※訓練データ数÷ミニバッチ数の余り×0.1の整数部分 |           | BERT最終層-Adam,SGD学習率 | 2e-5     |
|                            |           | エポック数               | {10, 20} |
|                            |           | ミニバッチ数              | 64       |

表 2 : 実験結果 (左表 : SBERT 手法, 右表 : SBERT+Infersent 手法)

| 設定                          | epoch | step | train pooling | test pooling | 名詞            | 動詞            | 形容詞           | ALL           | 特徴ベクトル   |           |       |           |
|-----------------------------|-------|------|---------------|--------------|---------------|---------------|---------------|---------------|--|-----------|-------|-----------|
|                             |       |      |               |              |               |               |               |               | (u,v)  | (u-v u*v) | epoch | optimizer |
| fine-tuning無し<br>(ベースライン)   |       |      |               | CLS          | 0.6124        | 0.6558        | 0.4809        | 0.5585        | (u,v)  | 10        | adam  | 0.6138    |
|                             |       |      |               | mean         | 0.6789        | 0.5787        | 0.5550        | 0.6734        |  |           |       |           |
| fine-tuning済み<br>(ベースライン)   |       | ※    | mean          | mean         | 0.7922        | 0.6783        | 0.5890        | 0.7284        | (u,v, u-v )  | 10        | SGD   | 0.7489    |
|                             |       |      |               | mean         | 0.7897        | 0.6371        | 0.6033        | 0.7185        |  |           |       |           |
|                             |       |      |               | mean         | <b>0.8327</b> | <b>0.7664</b> | <b>0.6526</b> | 0.7231        |  |           |       |           |
|                             |       |      |               | mean         | 0.8257        | 0.7172        | 0.6473        | <b>0.7383</b> |  |           |       |           |
| fine-tuning無し               |       |      |               | CLS          | 0.5676        | 0.6479        | 0.7535        | 0.6720        | (左表)   |           |       |           |
|                             |       |      |               | mean         | 0.5506        | 0.5547        | 0.7056        | 0.6810        |  |           |       |           |
| fine-tuning済み<br>(CLSを用いる)  | ※     | ※    | CLS           | CLS          | 0.7028        | 0.8002        | 0.7253        | 0.7332        | train poolingはモデル学習時, test poolingは類義判定時に用いられる poolingを示す. stepは評価ステップを示す. 「※」は各品詞でF値が最高値となったepoch, stepの組み合わせを同じ品詞に適用していることを示す. |           |       |           |
|                             |       |      |               | mean         | 0.7268        | 0.7680        | <b>0.8315</b> | 0.7747        |  |           |       |           |
| fine-tuning済み<br>(定義文の変更無し) | ※     | ※    | mean          | CLS          | 0.7112        | 0.7860        | 0.8008        | 0.7382        | 「(ベースライン)」は4.1.2節で示した単語データセットを用いた実験結果を示し, 「(定義文の変更無し)」は3.4 節で示した定義文の変更を行わず, 岩波国語辞典の定義文をそのまま用いた実験結果を示す.                           |           |       |           |
|                             |       |      |               | mean         | 0.7231        | 0.8422        | 0.7710        | 0.6959        |  |           |       |           |
| fine-tuning済み               |       |      | mean          | mean         | 0.8218        | 0.7940        | 0.8152        | 0.7657        |  |           |       |           |
|                             |       |      |               | mean         | <b>0.8493</b> | 0.7935        | 0.8262        | <b>0.8172</b> |  |           |       |           |
|                             |       |      |               | mean         | 0.7356        | 0.8579        | 0.8237        | 0.7762        |  |           |       |           |
|                             |       |      |               | mean         | 0.8206        | 0.8081        | 0.8233        | 0.7855        |  |           |       |           |
|                             |       |      |               | mean         | 0.8263        | <b>0.8780</b> | 0.7997        | 0.7613        |  |           |       |           |
|                             |       |      |               | mean         | 0.8143        | 0.8624        | 0.8055        | 0.7737        |  |           |       |           |

## 5 実験結果・考察

SBERT 手法, SBERT+Infersent 手法の実験結果を表 2 に示す.

SBERT 手法においては, 提案手法が ALL において最高値の 0.8172 を出した. 石井ら[18]の実験では, 類義判定が難しい例として以下に示す「すがお」「じがお」の単語ペアを挙げている.

### ● すがお【素顔】

語義 1 : ”ふだんの顔。”

語義 2 : ”化粧していない顔。地顔。転じて、ありのままの状態。「東京の素顔」”

語義 3 : ”酒に酔っていない時の顔。しらふ。”

### ● じがお【地顔】

語義 1 : ”素顔。”

この例での類義判定では, 「すがお」の語義 2 と「じがお」の語義 1 のペアが類義であると判定したい.

表 3 にこの例の定義文の組み合わせから得られた cos 類似度を示す. fine-tuning 無しの日本語 BERT モデルを用いて求めた場合は正解ペアの cos 類似度が語義ペアの組み合わせ中で最低値となるが, fine-tuning 済み類義判定モデルを用いた場合は最高値となった. 閾値を設けた類義判定において, 理想的なコサイン類似度の数値を得られたことから判定精度が向上しているとわかる.

表 3 : 「すがお」「じがお」語義ペアの cos 類似度

| 語義ペア                     | cos類似度        |
|--------------------------|---------------|
| fine tuning無し            |               |
| 「すがお」語義1-「じがお」語義1        | 0.8257        |
| <b>「すがお」語義2-「じがお」語義1</b> | <b>0.7763</b> |
| 「すがお」語義3-「じがお」語義1        | 0.7869        |
| fine tuning済み            |               |
| 「すがお」語義1-「じがお」語義1        | 0.5760        |
| <b>「すがお」語義2-「じがお」語義1</b> | <b>0.7642</b> |
| 「すがお」語義3-「じがお」語義1        | 0.5207        |

SBERT+Infersent 手法においては, 特徴ベクトルを (u,v,|u-v|) にした設定で最高値 0.7532 となった. 結果から最適化手法に交差エントロピー誤差を用いるよりも対象誤差を用いた方が, 特徴ベクトルにおいては (u,v,|u-v|,u\*v) を用いるよりも (u,v,|u-v|) を用いた方が有効であるといえる. 更に特徴ベクトルにベクトルの距離である |u-v| を用いなかった場合の精度が一番低い距離学習が有効であるといえる.

## 6 おわりに

本研究では, 単語の語義を対象に類義であるか否かを判定する類義判定手法を提案し, 語義の類義判定を効果的にできることを示した. 今後は, 専門家の協力の元でさらに大規模な類義データセットを構築し, 新たなシステムの考案を行いたい.

## 参考文献

- [1] 乾健太郎 (2007). 自然言語処理と言い換え. 日本語学, Vol.26, No.11, pp. 50-19.
- [2] 呉鍾勲, 山田一郎, 内元清貴, 鳥澤健太郎, 橋本力 (2011). 言語横断共訓練による単語間の上位下位関係の獲得. 情報処理学会, 情報処理学会論文誌 52 (12), pp. 3558-3573.
- [3] 山田一郎, 鳥澤健太郎, 風間淳一, 黒田航, 村田真樹, ステイン デ・サーガ, フランシス ボンド, 隅田飛鳥, 橋本力 (2011). 分布類似度と Wikipedia から獲得した構造情報を利用した上位下位関係獲得. 情報処理学会, 情報処理学会論文誌 52 (12), pp. 3435-3447.
- [4] 新里圭司, 鳥澤健太郎 (2005). HTML 文書からの単語間の上位下位関係の自動獲得. 言語処理学会, 自然言語処理 12 (1), pp. 125-150.
- [5] 寺田昭, 吉田稔, 中川裕志, 同義語辞書作成支援システム. 言語処理学会, 自然言語処理 15 (2), pp. 39-58.
- [6] 伴凌太, 高橋宏季, 位野木万里 (2019). word2vec を用いた同義語辞書自動作成手法の提案と適用評価, 情報処理学会, 第 81 回全国大会講演論文集 2019 (1), pp. 265-266.
- [7] Sawan, K. Sharmistha. J. Karan, S., and Partha T. (2019). "Zero-shot word sense disambiguation using sense definition embeddings." In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp. 5670-5681, 2019.
- [8] Bevilacqua, M., and Navigli, R. (2020). "Breaking Through the 80% Glass Ceiling: Raising the State of the Art in Word Sense Disambiguation by Incorporating Knowledge Graph Information" In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp. 2854-2864.
- [9] Harris, Z. (1985). "Distributional Structure" The Philosophy of Linguistics. Oxford University Press, pp. 26-47.
- [10] 王玉馨, 清水伸幸, 吉田稔, 中川裕志 (2008). 単語類似度ネットワークを通じた自動同義語獲得. 情報処理学会自然言語処理研究会, NL-185, pp. 7-14.
- [11] 西尾実, 岩淵悦太郎, 水谷静夫 (1994). 岩波国語辞典 第五版, 岩波書店.
- [12] Hitoshi Isahara, Francis Bond, Kiyotaka Uchimoto, Masao Utiyama, and Kyoko Kanzaki. (2008). "Development of the Japanese WordNet" In Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08), pp. 2420-2423.
- [13] Hagiwara, M. (2018). "A Supervised Learning Approach to Automatic Synonym Identification based on Distributional Features" In Proceedings of the ACL-08: HLT Student Research Workshop, pp. 1-6.
- [14] Zhang, L., and Moldovan, D. (2019). "Classification of Semantic Relations between Pairs of Nominals Using Transfer Learning" In Proceedings of FLAIRS 2019, the Thirty-Second International Florida Artificial Intelligence Research Society Conference (FLAIRS-32), pp. 92-97.
- [15] Reimers, N., and Gurevych, I. (2019). "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks." In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pp. 3982-3992.
- [16] Cer, D., Diab, M., Agirre, E., Lopez-Gazpio, I., Specia, L. (2017). "SemEval-2017 Task 1: Semantic Textual Similarity Multilingual and Cross-lingual Focused Evaluation" In Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017), pp. 1-14.
- [17] Conneau, A., Kiela, D., Schwenk, H., Barrault, L., and Bordes, A. (2017). "Supervised Learning of Universal Sentence Representations from Natural Language Inference Data" In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pp. 670-680.
- [18] 石井佑樹, 佐々木稔 (2023). BERT の学習済みモデルを用いた語義定義文の類義判定に有効な日本語辞書の変更方法. 言語処理学会第 29 回年次大会発表論文集, pp. 2254-2258.