

# WordNet Lexicographer カテゴリ推定による語義サイズ縮約を用いた 語義曖昧性解消

橋口卓弥<sup>1</sup> 佐々木稔<sup>1</sup>

<sup>1</sup>茨城大学大学院理工学研究科情報工学専攻  
{22nm740a, minoru.sasaki.01}@vc.ibaraki.ac.jp

## 概要

語義曖昧性解消は、文中に出現する多義語が辞書中のどの語義として使われているかを推定するタスクである。このタスクに対して様々な手法が研究されている中、語義間の上位下位関係などを使って少ないサイズの語義で語義曖昧性解消を行うという手法が存在する。しかし、この手法を用いた場合、正解語義とは lexicographer カテゴリが異なる語義ラベルを選択して誤った語義ラベルに判定しやすいという問題がある。そこで、本稿では上位下位関係を使った語義サイズの縮約手法に lexicographer カテゴリの推定モデルを追加し、推定した lexicographer カテゴリを推定された語義ラベルの一部を書き換える手法で精度の向上を試みた。実験の結果、学習に古典的な埋め込みを用いた場合での名詞単語の語義識別精度が上がった。

## 1 はじめに

我々が普段文書や発話などで用いている言語には、文脈に応じて複数の意味を持つ単語が多く存在する。例えば、“mouse”という英単語は、ねずみや(電子デバイスの)マウスなど複数の意味を持っている。このような複数の意味を持つ単語が、特定の文脈ではどの意味で使われるのかを識別し、文の意味の曖昧性を無くすことを目的とした自然言語処理のタスクを語義曖昧性解消という。

語義曖昧性解消は、今日に至るまで様々な研究が行われており、多くのアプローチ方法が提案されてきたが、その中でも教師あり手法が最良のスコアを出しており、現在の手法では教師あり手法が優勢である。この教師あり手法の問題点として、人の手によって作成された意味注釈付きコーパスは量が限られていることがあげられる。最大のコーパスである SemCor では、単語は 33760 種類の意味キーで注釈

されているが、これは語義曖昧性解消タスクで広く使われている語義データベースの意味リストの約 16%に過ぎない。そこで、Loïc Vial ら[1]は、WordNet に存在する語義間の意味関係を利用して、語義の縮約を行うという手法を提案した。語義縮約による手法を用いて最新の語義曖昧性解消ニューラルネットワークで実験を行った結果、最新技術を大幅に上回るスコアが得られた。

しかし、Loïc Vial らの手法を用いて語義曖昧性解消タスクを行い、推定結果を調べると、正解語義とは lexicographer カテゴリが異なる語義を選択して誤った語義が多く見られた。そこで、本稿では、Loïc Vial らの手法に lexicographer カテゴリの推定システムを追加し、推定した lexicographer カテゴリで語義ラベルの一部を書き換える手法を提案する。また、この手法が語義推定の精度向上に有効かどうかについて分析をする。なお、lexicographer カテゴリとは、語義の種類を表す番号であり、詳細は 3.4 節で説明する。

## 2 関連研究

言語モデルに基づく語義曖昧性解消は、Yuan ら[2]によって最初に作られ、Le ら[3]によって改良された。膨大な量の注釈つきでないデータで学習したりカレントニューラルネットワークによって、注目単語の周りの単語を考慮した予測ができるニューラル言語モデルが中心的な構成である。

ソフトマックス分類器に基づく語義曖昧性解消の主なニューラルネットワークは、ソフトマックス関数によって計算された確率分布によって、入力された各単語を直接分類し、意味を付与する。このニューラルネットワークは 2 つの手法に分けられる。1 つは、Iacobacci[4]らが提案した、辞書に含まれる様々な単語に対して、トークンごとに異なる複数のニューラルネットワークを持ち、それぞれが特定の単語

と意味を管理できる手法である。もう 1 つは、Raganato ら[5]が提案した。全ての単語を管理して辞書に存在する全ての意味セットで意味を割り当てることができる大規模で一般的なニューラルネットワークを有する手法である。

Ciaramita と Altun [6]、Izquierdo ら[7]は複数の Wordnet の意味タグをグループ化することで、より粗い意味リストを作成し、自然言語処理タスクでより有効になる可能性があるということを利用した研究であるクラスタリング手法を提案している。

### 3 提案手法

本稿で提案する lexicographer カテゴリの推定による語義曖昧性解消の手法は、Loïc Vial らの語義サイズ縮約手法に我々が考案した lexicographer カテゴリの推定システムを追加して語義曖昧性解消を行う。

まず、訓練データと WordNet3.0 の語義データを用いて語義の縮約を行う。次に、縮約した語義で作成された訓練データを用いて語義ラベル推定モデルの学習を行い、学習したモデルでテストデータ内の各単語の語義ラベルの推定を行う。そして、WordNet3.0 の語義定義文を用いて lexicographer カテゴリ推定モデルの学習を行い、lexicographer カテゴリ推定システムでテストデータ内の各単語の lexicographer カテゴリを推定する。最後に、各単語の語義ラベルの推定結果(に含まれる lexicographer カテゴリ)を推定した lexicographer カテゴリに書き換える。なお、3.1 節と 3.2 節は Loïc Vial らの語義サイズ縮約手法の説明、3.3 節と 3.4 節は我々が提案する lexicographer カテゴリ推定システム追加手法の説明である。

提案手法の大まかな処理の流れを図 1 に示す。

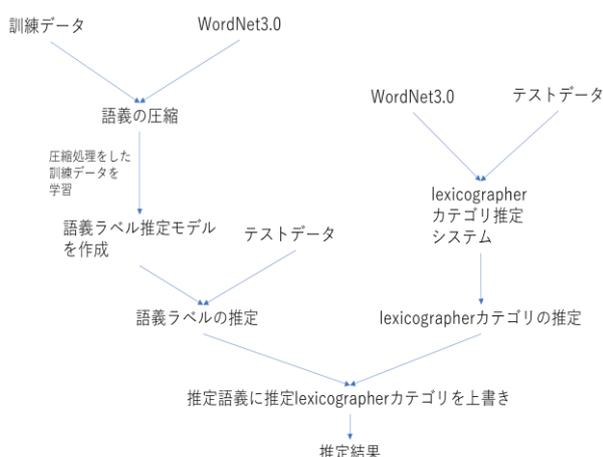


図 1：提案手法の処理順序

### 3.1 同義語に基づく語彙サイズの縮約

語義の辞書である WordNet にて、語義は synset と呼ばれる同義語の集合で構成されている。synset とは、同じ語義を持つグループで、例えば”person”, “individual”, “someone”, “somebody”の 1 番目の語義は全て”a human being”という共通の語義で定義されている。WordNet には各単語のそれぞれの語義に語義ラベルがついているが、語義ラベルを共通の語義、つまり synset へマッピングすると、例で示した”person”, “individual”, “someone”, “somebody”の 1 番目の語義ラベルを 1 つの synset キーで表すことができ、語彙を縮約することができる。

### 3.2 上位下位概念による語義サイズの縮約

WordNet の語義には、上位概念と下位概念という意味の包含関係が存在し、上位概念は一般化、下位概念は特定化を表している。

本手法では、同じ上位概念関係の語義をグループ化して、語義サイズを縮約する手法を紹介する。実際に、”mouse”の 1 番目の語義(小さな齧歯類)と 4 番目の語義(電子機器)、またこれらの上位概念からなる小さな木構造で考える。これを図 2 に示す。

”mouse#1”の上位概念である”living\_thing#1”と”mouse#4”の上位概念である”artifact#1”によって”mouse”の語義を識別できることから、”artifact”と”living\_thing”よりも下位の概念は全て取り除くことができる。また、”mouse#1”を”living\_thing#1”にマッピングしても”mouse”の語義を曖昧さなく識別することができ、また他の生き物や動物の単語も同様

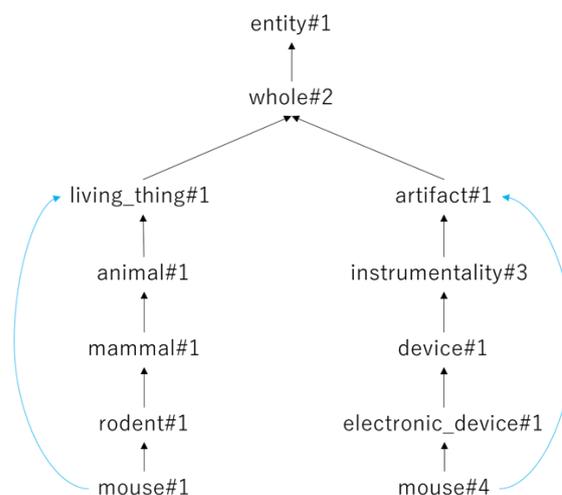


図 2: mouse#1 と mouse#4 の上位概念からなる木構造

にマッピングできる。このことから、語義サイズの縮約が可能となる。

本手法は、WordNet の全ての語義を上位概念の階層で最も高い祖先にマッピングすることであるが、制約がある。1 つは、祖先は対象単語の異なる語義を全て識別できなければならず、もう 1 つは、WordNet に含まれる全ての語義を判別するために不可欠な上位概念を無くしてはいけないということである。これらを踏まえて、本手法は次の 2 つのステップで処理を行う。1. WordNet の全ての語義のペアの最初の共通親の子を” necessary” としてマークする。2. 全ての語義を、語義から上位概念を辿っていった際の最初の” necessary” とマークされた親にマッピングする。

### 3.3 lexicographer カテゴリ推定システム

WordNet の語義ラベルには、語義の種類ごとに lexicographer カテゴリが付与されている。語義ラベルの 1 つ目のコロンの次に該当する番号が lexicographer カテゴリであり、2 桁の番号で表される。例をあげると語義ラベル” time%1:11:00:” の lexicographer カテゴリは 11 である。なお、名詞の lexicographer カテゴリの一覧は付録に示す。

本手法では、テストデータ内の各単語の lexicographer カテゴリを推定することが目標である。本手法の処理の流れを図 3 に示す。

WordNet の全語義(全 synset)を入力とし、それに対応する lexicographer カテゴリを出力するように BERT をファインチューニングしてモデルを作成する。しかし、作成したモデルは語義定義文から lexicographer カテゴリを推定するモデルであり、このモデルでテストデータの用例文から lexicographer カテゴリを推定することはできない。そこで、このモデルを活かすために、各語義定義文と(用例文中の)各単語のベクトルを利用する。

作成したモデルを用いて、用例文の各単語の埋め込みベクトルを作成する。モデルに用例文を入力することで、モデルの隠れ層から用例文中の各単語のベクトルを取得することができる。このベクトルを用いて、lexicographer カテゴリを推定する。

用例文中の注目単語 1 つの lexicographer カテゴリ推定を行うとする。WordNet 上での注目単語の各語義定義文を作成したモデルに入力し、モデルの隠れ層から各語義定義文の[CLS]ベクトルを取得する。用

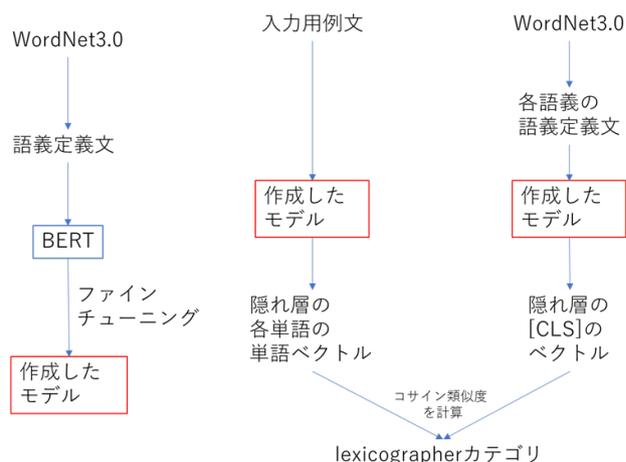


図 3: lexicographer カテゴリ推定システム

例文中の注目単語の単語ベクトルと各語義定義文の[CLS]ベクトルとのコサイン類似度を計算し、最も類似度が高い語義定義文の lexicographer カテゴリを推定結果とする。

コサイン類似度とは、2 つのベクトルがなす角のコサイン値のことで、すなわち 2 つのベクトルの類似性を表す尺度である。最大値の 1 に近づくほど、ベクトル同士が似ていると言える。コサイン類似度は、以下の式で求めることができる。

$$\cos(\vec{p}, \vec{q}) = \frac{\vec{p} \cdot \vec{q}}{|\vec{p}| \cdot |\vec{q}|}$$

### 3.4 lexicographer カテゴリの上書き

語義ラベル推定モデルで推定した語義ラベルを、lexicographer カテゴリ推定システムで推定した lexicographer カテゴリによって上書きする処理について説明する。

推定した語義ラベル内の lexicographer カテゴリがシステムで推定した lexicographer カテゴリと異なる場合に上書き処理を行う。その後、上書き処理をした語義ラベルが Word-Net 上に存在するか調べる。存在する場合、上書きした語義ラベルを最終的な推定結果とする。存在しない場合、上書き処理したラベルではなく元の推定語義ラベルを最終的な推定結果とする。

## 4 実験

本実験は、Loic Vial らの語義サイズ縮約手法のみで語義曖昧性解消タスクを行った場合(①)と、lexicographer カテゴリの推定手法を追加した語義曖

味性解消タスクを行った場合(②)のそれぞれの推定結果を比較し、lexicographer カテゴリ推定手法の有効性を検証することを目的としている。

## 4.1 実験の詳細

3節で説明した方法に基づいて実験を行う。

モデルの学習は WordNet3.0 とコーパス SemCor を用い、評価は コーパス SemEval2015 を用いた。

語義ラベル推定モデルの学習は Vial らの方法に従った。また、語義ラベル推定モデルは、古典的な埋め込みを用いたものと BERT の一種である bert-large-cased model を用いたものの 2 種類を作成した。それぞれのモデルについて個別に実験を行った。

①と②それぞれについて F1 スコアを求め比較することにより、lexicographer カテゴリ推定手法の評価を行った。F1 スコアについては、品詞ごとの F1 スコアと全体の F1 スコアを求めた。

## 4.2 実験結果

表 1 と表 2 に実験結果を示す。

表 1: 語義ラベル推定モデルの学習に古典的な埋め込みを用いた場合の実験の結果

	名詞	動詞	形容詞	副詞	全体
①	64.69	56.34	74.17	83.54	65.54
②	<b>65.84</b>	55.13	68.21	83.54	64.93

表 2: 語義ラベル推定モデルの学習に BERT を用いた場合の実験の結果

	名詞	動詞	形容詞	副詞	全体
①	81.68	66.00	86.71	86.79	79.01
②	74.05	62.78	80.38	86.79	73.27

語義ラベル推定モデルの学習に古典的な埋め込みを用いた場合の実験では、名詞のスコアが高くなった。動詞と形容詞のスコア、および全体のスコアは低くなった。

語義ラベル推定モデルの学習に BERT を用いた場合の実験では、副詞以外の品詞のスコアが下がり、全体のスコアも低くなった。

## 5 考察

古典的な埋め込みを用いて語義ラベル推定モデル

を学習させた場合、lexicographer カテゴリの推定システムを追加して行った語義曖昧性解消は、lexicographer カテゴリの推定システムを使わずに行った語義曖昧性解消に比べて名詞単語のスコアが高くなり、名詞単語の識別精度が上がったことが分かる。したがって、lexicographer カテゴリ推定システムを追加する手法はこの点において有効であることを示すことができた。形容詞の識別精度が他の品詞と比べて大きく下がっている点については、推定結果を調べると書き換えた 12 の単語の内 7 単語が同じ単語で正解語義とは誤った上書きをしており、推定語義が偏ると精度に大きな影響が出る事が分かった。推定結果を詳細に分析すると、レキシコグラファーカテゴリ推定システム単体での正答率は 76.85 だったが、lexicographer カテゴリが正解と一致していても、語義ラベルの末尾の番号が異なっていたり、そもそも書き換え時に存在しない語義ラベルとなったりするため、実際の語義ラベル正答率は大きく落ちると考えられる。しかし、各単語の推定語義と正解を比較すると、lexicographer カテゴリのみ誤っている語義がまだ多く存在しているため、lexicographer カテゴリ推定システムでの推定の精度を上げることで、語義推定の精度を更に高めることができるのではないかと考える。

BERT で語義ラベル推定モデルを学習させた場合で全体的に精度が大きく下がったのは、レキシコグラファーカテゴリ推定システムの推定精度がシステムを追加しない場合の語義推定精度より低いためであると考えられる。

総じて、lexicographer カテゴリ推定の精度を上げられるようにアプローチを改善、または新たに考案することが必要であると考ええる。

## 6 おわりに

本論文では、語義曖昧性解消タスクにおいて、Loïc Vial らの語義サイズ縮約手法に、lexicographer カテゴリ推定システムを加えることで、古典的な埋め込みを用いた場合に名詞の精度が上がることを確認した。

今後は、lexicographer カテゴリ推定精度を向上させるため、lexicographer カテゴリ推定システムの改良や、別のアプローチでの lexicographer カテゴリの推定などを考えていきたい。

## 参考文献

- [1] Loïc Vial Benjamin Lecouteux Didier Schwab. 2019. Sense Vocabulary Compression through the Semantic Knowledge of WordNet for NeuralWord Sense Disambiguation. In proceedings of the 10th Global WordNet Conference - GWC 2019. arXiv:1905.05677v3 [cs.CL].
- [2] Dayu Yuan, Julian Richardson, Ryan Doherty, Colin Evans, and Eric Altendorf. 2016. Semi-supervised word sense disambiguation with neural models. In COLING 2016.
- [3] Minh Le, Marten Postma, Jacopo Urbani, and Piek Vossen. 2018. A deep dive into word sense disambiguation with lstm. In Proceedings of the 27th International Conference on Computational Linguistics, pages 354–365. Association for Computational Linguistics.
- [4] Ignacio Iacobacci, Mohammad Taher Pilehvar, and Roberto Navigli. 2016. Embeddings for word sense disambiguation: An evaluation study. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 897–907, Berlin, Germany, August. Association for Computational Linguistics.
- [5] Alessandro Raganato, Claudio Delli Bovi, and Roberto Navigli. 2017b. Neural sequence learning models for word sense disambiguation. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pages 1167–1178. Association for Computational Linguistics.
- [6] Massimiliano Ciaramita and Yasemin Altun. 2006. Broad-coverage sense disambiguation and information extraction with a supersense sequence tagger. In Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, EMNLP '06, pages 594–602, Stroudsburg, PA, USA. Association for Computational Linguistics.
- [7] Rubén Izquierdo, Armando Suárez, and German Rigau. 2007. Exploring the automatic selection of basic level concepts. In Proceedings of RANLP, volume 7. Citeseer.