

P1-43: The 38th annual meeting of Japanese cognitive science society

## 単語の意味空間を心的操作する = 射影

浅川伸一 (あさかわ しんいち) 近藤 公久 (こんどう ただひさ)

礼和 3 年 9 月 1 日

# 要約

文脈効果を心的操作として捉えるモデルを提案した。我々は、意味空間のうち部分空間を文脈とし、部分空間への射影変換によって文脈効果を扱うことができると考えた。このような射影変換を用いることの利点は、階層構造で表現された意味概念では、詳細な記述、操作、検索過程などと直接関連させることが困難な心的操作を表現可能なことである。射影変換が文脈効果を表す心的操作の一部と見なすことで推論や検索過程が記述可能と考えられる。そこで本発表においては、日本語 wikipedia を用いて訓練した単語のベクトル表現集合を意味空間とし、これらに対して射影変換によって文脈効果が得られるかを検討した結果を報告する。本手法は、深層学習との親和性も高く、今後発展が期待できる。

# 文脈効果

文脈効果を心的操作として捉えるモデルを提案する。文脈効果とは、言語理解における事前情報の影響によって解釈が異なることに代表されるが、単語命名課題におけるプライミング効果や、文章の多義性・曖昧性の解消、さらには、視覚における恒常性、錯視、だまし絵、視覚探索における制約など、認知的処理における時間的前後に限らない、状況などを文脈にとらえることが可能である。



# THE CAT

図 1: 視覚的文脈の例。左: Gregory のダルマシアン犬

<https://michaelbach.de/ot/cog-Dalmatian/>. 右: THE CAT

# 単語埋め込みモデル

自然言語処理においては、単語埋め込みモデル、または、ベクトル空間モデルと総称されるモデル (word2vec:[5]; GloVe:[6]; ELMo:[7]; BERT:[1]) が存在する。すでに、翻訳、分類、要約などの種々の言語課題 (GLUE:[8]) で性能向上が認められている。これは、単語や文を多次元ベクトルとして意味空間を表現していると捉えることができる。我々は、意味空間のうち部分空間を文脈とし、部分空間への射影変換によって文脈効果を扱うことができると考えた。

# 射影

語彙を構成するベクトルからなる行列を  $X$  で表す。 $X$  が張る空間  $P(X)$  に各単語を射影し、射影された部分空間内で各語間の距離を測れば所与の目的が近似的に達成される。このような射影変換を用いることの利点は、階層構造で表現された意味概念では、詳細な記述、操作、検索過程などと直接関連させることが困難な心的操作を表現可能なことである。

# 単語埋め込みモデル

Mikolov らのベクトル空間モデルは, 'king' - 'men' + 'women' = 'queen' のようにベクトルの加減算として類推, 推論課題を定義した [5]。単語埋め込みモデルでは, 意味のみならず, 統語についての推論も可能である。図 2 左では, 意味が, 同図右では統語情報, この場合複単語の単数-複数の関係を表している [5]。加えて, 図 3 では, 各国の国名と対応する首都名との関係を主成分分析して布置したものである。同図では, 横軸左に国名, 右に首都名が対応する首都名が布置されている。興味深いことに縦軸は, 上から下にほぼユーラシア大陸を東から西に並んでいる。主成分分析は, 各項目に対応するベクトルを集めた行列を  $X$  とすれば, 主成分分析の目的関数は以下で与えられる:

$\max_w \{w^T X^T X w - \lambda (w^T w - 1)\}$ . ここで  $\lambda$  は Lagrange 乗数である。主成分分析は  $w$  によって張られる空間への射影行列を求めることに相当する。従って, 適切な射影行列を定義できれば, 国名と首都との関係と, 各国の地理的關係をも表象されていると考えることが可能である。

# 単語埋め込みモデル 例 1 意味と文法の図示

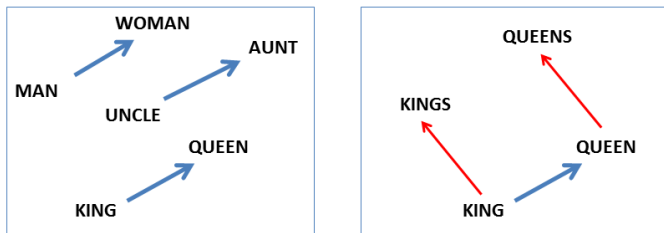


図 2: Word2vec による推論の例: 左は意味。右は文法情報とみなすことができる。[5] Fig. 2 より

# 単語埋め込みモデル 例 2 対応関係のある意味の図示

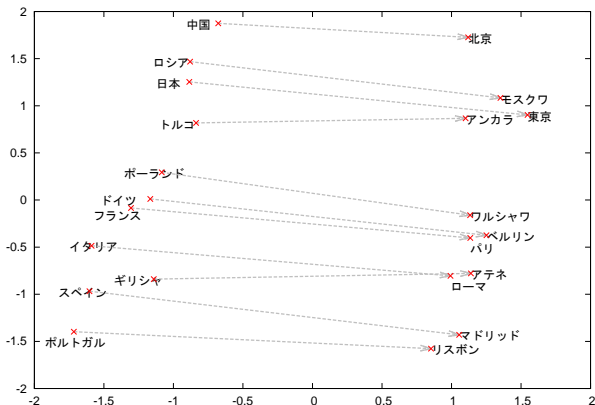


図 3: [4] Fig. 2 を改変。横軸は国とその首都。縦軸はユーラシア大陸を西から東へ大まかな地理情報に対応する



# 単語埋め込みモデル 例 3 構文解析木

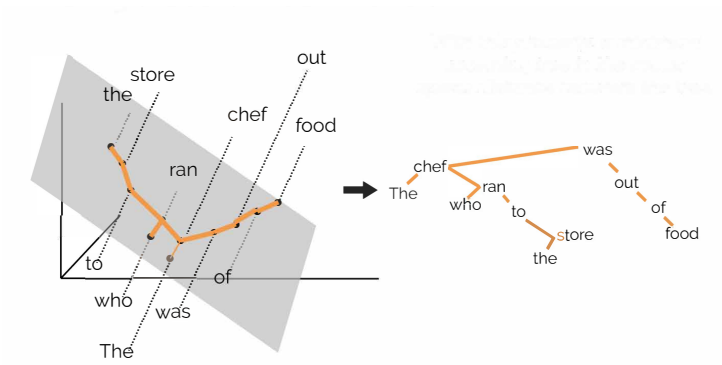


図 4: 構文解析木のルートからのノード数に対応した距離への変換例 [2]。

# 線形射影空間

実数ベクトル  $X \in \mathbb{R}^{n,m}$  が張る空間への射影行列は次式で定義される:

$$P(X) = X(X^\top X)^{-1}X^\top \quad (1)$$

本稿では、単語埋め込みモデルで得られる単語ベクトルを用いて意味空間を用いて心的操作を説明することを試みた。

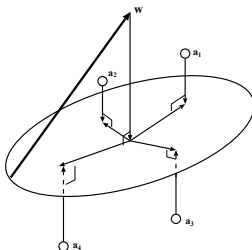


図 5: 射影の概念図

# 実験

日本語ウィキペディア全文<sup>1</sup>を MeCab<sup>2</sup>を用いて分かち書きにした。得られた全単語について、word2vec[4]で訓練した結果を用いた。訓練に用いたモデルはCBOW[3]でウィンドウサイズを20とし、埋め込み次元数を200とした。従って得られた単語の意味表現は各語とも200次元のベクトルとなる。

---

<sup>1</sup>[http:](http://dumps.wikimedia.org/enwiki/latest/enwiki-latest-pages-articles.xml.bz2)

[//dumps.wikimedia.org/enwiki/latest/enwiki-latest-pages-articles.xml.bz2](http://dumps.wikimedia.org/enwiki/latest/enwiki-latest-pages-articles.xml.bz2)

<sup>2</sup><https://taku910.github.io/mecab/>

# 多肢選択課題

手がかりを与えられ、4つの選択肢の中から最も近いものを選ぶようなピラミッドパームツリー検査 (PPTT) を考える。このとき、与えられた4つの選択肢からなる意味空間へ射影行列を定義し、この部分空間への射影したベクトル間の布置と、もとのベクトル間との比較を示した。

# 多肢選択課題 結果 1

課題：「兜」「麦わら帽子」「帽子」「ヘルメット」の中から、ターゲット語：「オートバイ」にもっとも近いと思われるものを選び。図右のように、通常の word2vec 空間では傾向が認められない。だが、4 つの帽子類で構成される空間へ「オートバイ」を射影することにより、「ヘルメット」との関係が明確となっている。

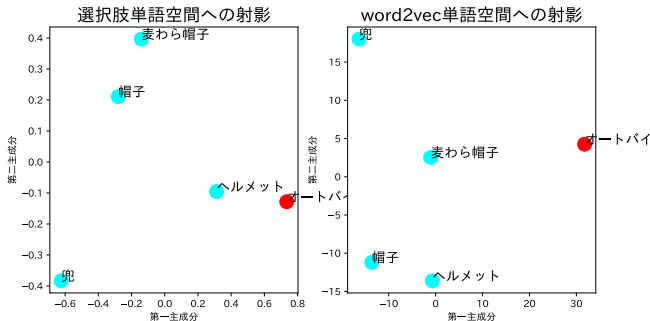


図 6: 左: 選択肢で構成された部分空間への射影。右: 元の埋め込みベクトルの布置

## 多肢選択課題 結果 2

課題：水田，池，滝，海，の中から，ターゲット語：かもめ，にもっとも近いと思われるものを選べ。「かもめ」に対する，水に関する景観「水田」「池」「滝」「海」で構成される空間へと射影することにより，「海」と「かもめ」との関係が word2vec による空間附置よりも，射影空間の方が明確となる。

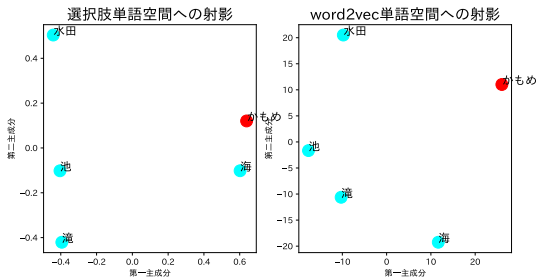


図 7: 左: 選択肢で構成された部分空間への射影。右: 元の埋め込みベクトルの布置

## 多肢選択課題 結果 3

課題：コップ，コーヒーカップ，湯呑み，杯，の中から，ターゲット語：急須，にもっとも近いと思われるものを選べ。ターゲット語「急須」と「湯呑」との関係は word2vec 空間でも同様に近いが，射影空間ではコーヒーカップの附置が明確に分離されていることが見て取れる。

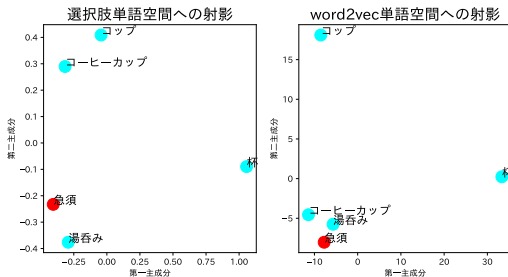


図 8: 左: 選択肢で構成された部分空間への射影。右: 元の埋め込みベクトルの布置

## 多肢選択課題 結果 4

課題：「学校」「工場」「デパート」「郵便局」の中から、ターゲット語「黒板」にもっとも近いと思われるものを選び。4つの建物選択肢空間へ射影した場合，「黒板」と「学校」とが明確に近づいている。

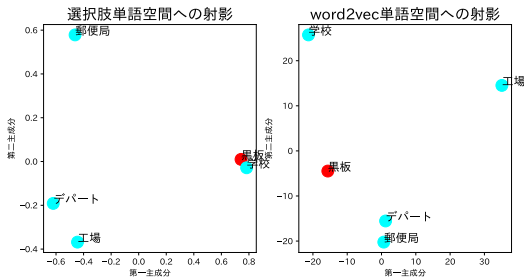


図 9: 左: 選択肢で構成された部分空間への射影。右: 元の埋め込みベクトルの布置



# 単語連想課題

神経心理学検査の一つに、動物や野菜を思いつく限り表出する課題がある [10]。この単語連想課題を、射影行列を用いた提案手法により模倣することを試みた。単語  $k$  と  $v$  との間の類似度  $S(k, v)$  を、単語埋め込みモデルによって表現された単語ベクトルのコサイン距離とみなし、手がかり単語  $k$  が与えられたとき、単語  $v$  を想起する確率は、次式ソフトマックス関数に従う密度関数に従うものと仮定した:

$$P(v|k) = \frac{e^{S(v,k)/\beta}}{\sum_{z \in \mathcal{Z}} e^{S(z,k)/\beta}}, \quad (2)$$

ここで  $\beta$  は温度パラメータであり、 $\mathcal{Z}$  は全関連語である。

## 単語連想課題 2

ソフトマックス関数は、ボルツマン分布でもあるので、温度パラメータ  $\beta$  によって、その挙動が変化する。 $\lim \beta \rightarrow 0$  の極限では、単語の産出確率は決定論的になり、最近隣語のみが出力される。一方で  $\beta$  が高くなるに従って、すべての単語は等確率で算出されるようになる。 $\beta$  が大きければ、(従って温度が高ければ) 想起語はデタラメに近くなる。すなわち  $\beta$  によって、ある種の被験者や患者の応答の変動を記述可能であると考えることが可能である。

$S(x, x) = 0$  と仮定することは、同じ単語を想起しないために必要な措置である。しかし、 $S(x, x) \neq 0$  であれば、同じ単語を何度も算出する、ある種の痴呆症などに見られる保続反応 (preservation) を想定する場合には必要となると考えられる。

# 単語想起課題 産出例

上式に従って、自由連想を行った例を以下に示す:

手がかり語: イヌ → ['カワウソ', 'ワニ', 'ザリガニ', 'ザリガニ', 'ウミヘビ', 'カワウソ', 'ニホンアナグマ']

オリジナル word2vec による イヌ の最近隣語 → ['イヌ', 'ネコ', 'オオカミ', 'ブタ', 'ウサギ', 'ウシ', '犬', 'アライグマ', 'ニワトリ', 'ハツカネズミ']

手がかり語: サンショウウオ → ['ザリガニ', 'アリクイ', 'ワニ', 'ヘビ', '猛禽類', 'アカギツネ', 'ザリガニ']

「ザリガニ」と連続して表出していることが見て取れる。健常者であれば、このような想起語の連続は起こり得ないと予想される。種々の記憶障害、痴呆症、あるいは保続と言った神経心理学的症状を想定した場合には、既述の産出例のような想起系列が考えられると思われる。

## 単語連想課題 2

水野らは健常者に対して日本語 300 語を用いて対応する連想語を調べた [9]。彼女らのデータを説明することを試みた。水野らの用いた刺激語 300 語のうち 1 節で用いた word2vec に存在する 296 語の単語を用いた。‘あたま’，‘しにせ’，‘だんご’，‘はんこ’，の 4 語は word2vec に登録が無いので除外した。これらの 4 単語は，水野らが，ひらがな表記語刺激として用いた単語である。水野らのデータには，ターゲット語に対して想起された単語の総頻度で除した連想強度が記されている。水野らの連想強度と word2vec の最近接語との関係を次表に示す：

## 単語連想課題 2 行動データとの比較

Table 1:

条件	単語数	割合
top1 が同じ	10	0.034
top10 以内に水野 top1 が存在	17	0.057
top100 以内に水野 top1 が存在	50	0.169
100 以外に水野水野 top1 が存在	142	0.480
上記以外の順位	77	0.260

行動データと今回用いた単語埋め込みモデルとの間には、なお乖離があることが推察される。人間の単語想起過程と、単語埋め込みモデルとの間での齟齬は、今後の課題である。

# オノマトペ

市販のオノマトペ辞典 [11] に記載されている索引項目 2761 語のうち、単語埋め込みモデル、日本語ウィキペディアによる word2vec に登録項目のある 1761 語を用いて、オノマトペ空間を作成した。

このオノマトペ空間へ任意の単語を射影することで、任意の単語に類似するオノマトペを得ることができると考えた。ランダムに単語を抽出し、その単語を上記オノマトペ空間へ射影することで、近傍のオノマトペを求めることができると考えた。結果の一部を以下に示した:

基本単語からオノマトペ空間での隣接語の検索例を以下に示す (カッコ内はコサイン類似度):

- ① 球茎 :(びっしり:0.332) (ぬめり:0.317) (でこぼこ:0.281)
- ② 吐露 :(うんざり:0.551) (やきもき:0.480) (がっかり:0.457)
- ③ サイエントロジー :(ザック:0.249) (トロン:0.235) (ペロン:0.212)

# 考察

適切な射影空間を定義することで、4 肢選択課題、単語想起課題、オノマトペ課題で示したような結果を得た。今後は、心理空間への (射影) 変換を考えることで、心理モデルとしての妥当性を検証する必要がある。その他にも、(1) ELMo[7] や BERT[1] などより性能の高い単語埋め込みモデルを用いる。(2) 訓練データセットの拡大を図る。今回用いた日本語 wikipedia 全文のデータと英語版 wikipedia のデータサイズを比較すれば、およそ 10 倍の違いがある。(3) 射影空間を適切に設定する。オノマトペの意味情報から、カタカナ語の部分情報を部分射影空間として削除し、音韻情報との共有空間へと射影するなど、の改善策が考えられる。これらの方向性は、自然な拡張のように思われる。

# 文献

- [1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint*, 2018.
- [2] John Hewitt and Christopher D. Manning. A structural probe for finding syntax in word representations. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4129–4138, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.
- [3] Tomas Mikolov, Quoc V. Le, and Ilya Sutskever. Exploiting similarities among languages for machine translation. *arXiv:1369.4168*, September 2013.
- [4] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S. Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In C. J. C. Burges, L. Bottou, M. Welling, Zoubin Ghahramani, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pages 3111–3119. Curran Associates, Inc., 2013.
- [5] 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 NAACL*, Atlanta, WA, USA, June 2013.
- [6] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. GloVe: Global vectors for word representation. In *Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, Doha, Qatar, Oct. 2014.
- [7] Matthew E. Peters, Mark Neumann, Mohit Iyyery, Matt Gardnery, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. Deep contextualized word representations. *arXiv preprint*, 2018.
- [8] Alex Wang, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, and Samuel R. Bowman. GLUE: A multi-task benchmark and analysis platform for natural language understanding. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 2019.
- [9] 水野 りか, 清河 幸子, 川上 正浩, and 柳谷 啓子. 連想語頻度表 3 モーラの漢字・ひらがな・カタカナ表記語. ナカニシヤ出版, 東京, 2011.
- [10] 杉下 守弘. WAB 失語症検査 日本語版. 医学書院, 1986.
- [11] 小野 正弘, editor. 日本語オノマトベ辞典. 小学館, 東京, 第一版 edition, 2007.