

文章内の時間関係を捉える脳内メカニズム解明への取り組み

須藤 百香 (指導教員：小林 一郎)

1 はじめに

本研究では、ヒト脳内における日本語時間に関する処理メカニズムを解明するために日本語時間関係を識別する自然言語処理モデルを開発、援用することでヒト脳内の状態を推定することを目的とする。時間関係識別には、日本語書き言葉コーパス (BCCWJ) の時間関係を示す TimeBank コーパスの文章中に表れる6種類の時間 (過去, 現在, 未来, 等) を文または文間の時間関係を対象とする。とくに、関係識別においては、文中の単語間の依存関係を捉えた Shortest Dependency Path によって規定された単語から時間を識別する手法 [1] を word2vec [2] や BERT [3] を単語の潜在トークンとして使用し時間関係の識別を行う。実験を通じて、開発したモデルは人の時間識別とほぼ同等の精度をもった結果であることを確認した。また、ヒト脳内における時間関係の認識状態を調べる予備実験として、DVD鑑賞をした際の脳活動データと動画内の自然言語のBERTによる意味表現との間の相関性を調査した。

2 時間関係識別

2.1 時間関係

Allen [4] は、ヒトが認識する時間関係を13種類に分類しているが、区別が明確になるように、本研究ではその時間関係を Before, Overlap, After, Vague の4種類とそれに Before-or-Overlap と After-or-Overlap の2つを加えた6種類の2つにまとめて時間関係の識別を行う。また、1文中または文間に表れる時間関係の識別を行う課題として、文中の時間関係を表すイベントとその文書が作成された時間関係を表す **DCT** (Document Creation Time), 同じ文中の二つのイベント間の時間関係を表す **E2E** (Event-Event), 同じ文中のイベントと時間に関する表現との時間関係を表す **T2E** (Time-Event), 隣接文の二つのイベント間の時間関係を表す **MAT** (Matrix) の4つの設定を対象とする。

2.2 時間識別深層学習

一般に関係認識においては、統語構造の依存関係における対象となる語彙間の最短依存経路 (SDP: Shortest Dependency Path) を用いる研究が高い精度を挙げている [5]。このことを踏まえて、Cheng ら [6] は、他の研究に先駆けて SDP による関係認識を時間関係識別に導入し、高い精度で時間識別に成功した。耿ら [7] は、Cheng らの手法を元に BCCWJ-TimeBank コーパス [8] を用いて、日本語の時間関係識別深層学習モデルを構築した (図1参照)。本研究では、耿らによって開発された日本語時間識別モデルを用いて実験を行う。

2.3 実験

実験概要 時間識別において、耿ら [7] は BERT の最終層を使った時間識別結果を得ている。一方で、彼らが見つかった BERT は 12 層のものであり、BERT 各層

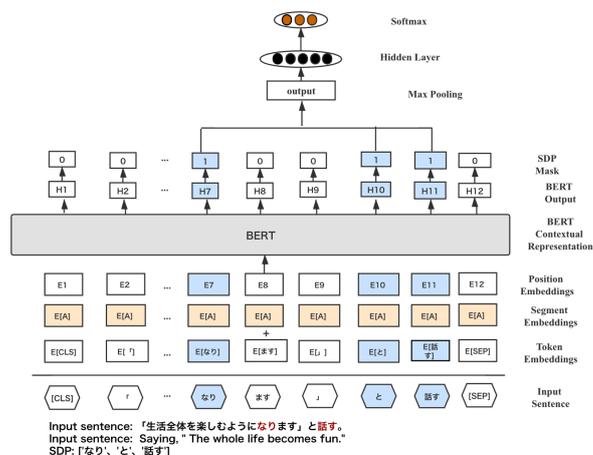


図 1: BERT + SDP モデル

にはそれぞれ自然言語の意味を処理する個別の機能が表れることが指摘されている。一般にファインチューニングをする最終層にそのタスクに有効な機能が表れていると言われているが、本研究では、最終層の一つ前の層を使った時間識別を調査する。

実験設定 モデルとして、事前学習のみの BERT を使用したモデルおよびそれにファインチューニングをしたモデルの2つをベースにして、SDP をそれぞれに追加する場合としない場合の4つを設定した。前者の事前学習のみの BERT を用いたモデルでは学習率を 10^{-2} , 最適化アルゴリズムに Adam を採用し、後者のファインチューニングを施したモデルは学習率を $2e^{-5}$, 最適化アルゴリズムに AdamW を採用している。全てのモデルにおいて、使用した事前学習文脈埋め込み言語モデルとして、国立国語学研究所において作成された NWJC-BERT [9] を用い、隠れ層として 768 次元を設定し、学習時のエポック数は 20, バッチ数は 16 とした。

実験結果と考察 表1に BERT をベースにして、ファインチューニング (FT) を追加した場合、SDP を取り入れた場合、FT と SDP の両方を追加したモデルを用いて時間関係識別の4つのタスクに対して推定した精度を示す。

表1より、最終層または最終層一つ前の層を用いて時間関係識別を行った精度は、ほぼ同等であることがわかった。手法の比較においては、DCT の識別結果を除くと、概ね BERT+SDP+FT の組み合わせの手法による精度が一番高く、BERT のみのものと BERT+FT, BERT+SDP とを比べると、BERT < BERT+SDP < BERT+FT の順で精度の向上が見取れる。これにより、SDP および FT の有効性を確認できた。

表 1: それぞれのモデルにおける 4 つの時間関係識別タスクの識別精度

タスク	モデル	BERT		BERT+FT		BERT+SDP		BERT+SDP+FT	
		層\ラベル数	4	6	4	6	4	6	4
DCT	最終層一つ前	76.0%	74.8%	84.5%	82.2%	75.8%	74.1%	82.0%	82.0%
	最終層	75.2%	74.0%	83.4%	81.7%	74.7%	75.8%	82.0%	81.9%
T2E	最終層一つ前	47.3%	45.8%	54.9%	52.0%	52.4%	49.9%	55.4%	53.0%
	最終層	47.7%	45.0%	54.9%	51.7%	52.7%	48.4%	55.3%	53.3%
E2E	最終層一つ前	51.7%	54.4%	60.4%	61.0%	55.3%	52.8%	62.2%	62.6%
	最終層	53.1%	53.0%	62.5%	61.6%	54.5%	53.7%	62.7%	62.6%
MAT	最終層一つ前	49.6%	49.6%	50.7%	50.8%	51.0%	50.7%	51.0%	52.3%
	最終層	47.2%	50.3%	51.4%	50.8%	51.0%	50.4%	52.5%	52.4%

表 2: BERT による発話特徴量から脳活動データを推定した際の相関係数

	正則化項 α	0.5	1.0	5.0	10.0	10 ²	10 ³	10 ⁴	2.5×10 ⁴	5.0×10 ⁴	10 ⁵	10 ⁶	10 ⁷
設定	4 秒先	0.345	0.345	0.333	0.324	0.269	0.176	0.090	0.067	0.053	0.045	0.035	0.034
	4,5 秒先	0.308	0.320	0.340	0.343	0.319	0.236	0.131	0.097	0.076	0.061	0.041	0.039
	4,5,6 秒先	0.285	0.302	0.333	0.340	0.318	0.216	0.112	0.085	0.068	0.054	0.034	0.031

3 言語から脳活動への回帰

3.1 DVD 視聴脳活動データ

実験で使用するデータは、情報通信研究機構脳情報通信融合センターにおいて取得された約 9 時間の動画視聴時の脳活動を計測したデータとそれに対する発話情報を紐付けたデータである (図 2)

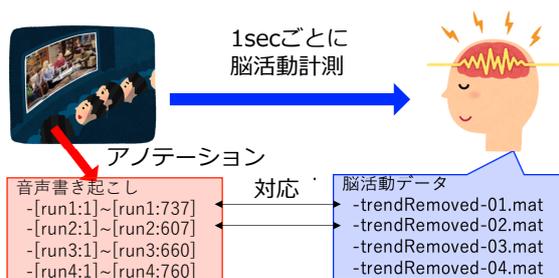


図 2: DVD 視聴脳活動データ

3.2 実験

上記, DVD 視聴脳活動データを用いて, 動画中の発話を書き起こした文章と動画を視聴した被験者の脳活動データ間の相関性を検証する実験を行った.

実験設定 動画内の発話情報として, 数秒間に発話された内容すべてをその発話の最初から最後まで 1 秒ごとに記録したものをを用い, 1 秒ごとの発話内容を一つの文として BERT による埋め込みベクトルによって表現した. この埋め込みベクトルと脳活動データのペアデータを作成し, Ridge 回帰にかける. この際, 脳活動データの観測遅れに対応するように 4 秒, 5 秒, 6 秒先の時点の埋め込みベクトルを含む場合を設定しペアデータを作成した.

実験結果と考察 実験結果を示す表 2 から 4,5 秒先かつ正則化項 $\alpha = 10.0$ の時, BERT による発話特徴量と脳活動データとの相関係数が一番高くなり, 0.343 となった. 正則化項 α が低い値のときに相関係数が高くなった. また, 言語の特徴量ベクトルは 4, 5 秒先の時点のものを用いると推定精度がよくなることがわかつ

た. 相関係数 0.3 は, 脳神経科学の分野において有意な数値であると判断される. BERT は言語に基づく脳活動を推定するのに有効であることが確認できた.

4 おわりに

本研究では, 時間識別深層学習モデルを使用して, 最終層の一つ前の層を使った時間識別を調査し, ファインチューニングおよび SDP の有効性を確認した. また, DVD 視聴脳活動データと動画中の発話を書き起こした文の言語刺激との相関性を検証する予備実験において, 言語による脳状態の推定に BERT は有効であることが確認できた. 今後はこの時間関係識別深層学習モデルから脳内状態を推定する回帰モデルを構築し, ヒト脳内における時間関係の認識状態を調べていきたい.

参考文献

- [1] Razvan Bunescu and Raymond Mooney. A shortest path dependency kernel for relation extraction. In *Proc. of HLTC and EMNLP*, pp. 724–731, October 2005.
- [2] 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, Z. Ghahramani, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, Vol. 26, pp. 3111–3119, 2013.
- [3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proc. of NAACL2019*, pp. 4171–4186, June 2019.
- [4] J.F. Allen. Maintaining Knowledge About Temporal Intervals. *Communications of the ACM*, Vol. 26, No. 11, pp. 832–843, 1983.
- [5] Yan Xu, Lili Mou, Ge Li, Yunchuan Chen, Hao Peng, and Zhi Jin. Classifying relations via long short term memory networks along shortest dependency paths. In *Proc. of EMNLP2015*, pp. 1785–1794, September 2015.
- [6] Fei Cheng and Yusuke Miyao. Classifying temporal relations by bidirectional LSTM over dependency paths. In *Proc. of ACL2017*, pp. 1–6, July 2017.
- [7] 耿晨, 程飛, KANASHIRO Pereira Lis, 浅原正幸, 小林一郎. 依存関係と文脈表現を用いた日本語時間関係識別. 人工知能学会全国大会論文集, Vol. JSAI2020, pp. 4Rin115–4Rin115, 2020.
- [8] Masayuki Asahara, Sachi Yasuda, Hikari Konishi, Mizuho Imada, and Kikuo Maekawa. BCCWJ-TimeBank: Temporal and event information annotation on Japanese text. In *Proc. of PACLIC 27*, pp. 206–214, November 2013.
- [9] 浅原正幸, 西内沙恵, 加藤祥. Nwjc-bert: 多義語に対するヒトと文脈化単語埋め込みの類似性判断の対照分析. 言語処理学会第 26 回年次大会発表論文集, pp. 961–964, 3 2020.