

対話の構造の可視化のための BERT を用いた先行発話への接続確率の推定モデル

徳永 一輝[†] 中山 隆弘[‡] 白水 始[§] 市川 治[†]

滋賀大学データサイエンス学部[†] 東京大学高大接続研究開発センター[‡] 国立教育政策研究所[§]

1 はじめに

新しい教育手法として生徒が自ら考え学ぶというアクティブラーニングが注目されている。東京大学 CoREF が推進する「知識構成型ジグソー法」では、クラス内の生徒を3～4名の小グループに分け、グループ内外で話し合うことで発見的に学びを進める。この学びの過程を可視化するために、生徒ひとりひとりに接話マイクロフォンを装着してもらい、音声データを音声認識によりテキスト化することが行われる。この記録を教師が授業後に見ることで、生徒が理解を深める過程を追跡することができる。

しかし、教師が生徒の発言を逐一に読み、意図を理解することは大きな労力が必要なことである。そこでいわゆるAI(機械学習)をベースにした分析技術により、教師を補助することが期待されている。

よくデザインされた協調学習では、発言回数が少ない生徒でも、回数が多い生徒と同等の学習成果を収めることが知られており、単なる話量では学習深度を計れないことが知られている。議論の起点となるような発言をしたか、議論をまとめるような発言をしたのであれば、例えば話量が少なくても、学習深度を認めることができるからではないか。本研究では、そのための教師の見取りを助けるべく、生徒のグループ議論を図1のように木構造に可視化することを目標とする。

2 従来手法

対話データの文間の関係や文脈に着目した文章解析の研究は様々なアプローチで行われてきた。工藤は対話文の文脈の乱れに着目し、結束性と一貫性という尺度で文と文の関係を捕捉することを試みた[1]。Nomotoらは1文の中

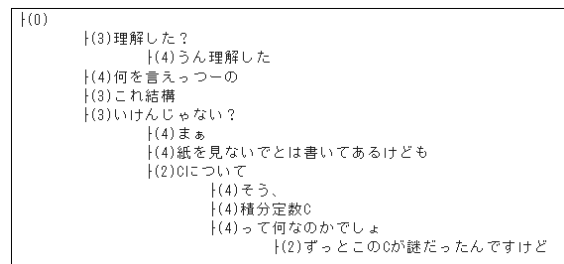


図1 グループ議論の木構造可視化の例. 手動にて作成。()内は話者番号である。

の文脈を考慮して呼応関係を解析した[2]。Cristeaらは教育に関する対話データを使った機械学習で対話文を分析した[3]。

文章を木構造に可視化することに関しても研究が進められている。長尾は発話者をタグ付けし、発言する前に手動でボタンを押すことによって発言と対応させた。この状態で会議を行うことにより生成した議事録と過去の議事録を木構造形式で可視化した[4]。

3 提案手法

BERT モデルをファインチューニングしたモデルを使用し、2文がグループワーク内の議論において係り受けの関係にあるか否かの接続確率を推定する。図2に示すようにn番目の文に対してn-3発話文までの文との接続確率を推定する。推定した接続確率に発話文間の距離を考慮したペナルティを付与する。ペナルティを付与した接続確率が閾値以上かつ一番高い文を接続すべき親と判定する。閾値以下であるならば、どの文も接続していない独立した発話と判定する。

BERT が行う事前学習には Next Sentence Prediction(NSP)が含まれる。本研究が対話の接続を判定するために扱う接続確率は、NSPと似ているが、厳密には意味が異なる。NSPは2文の文脈の連続性を判定するように学習されるため、「トピック」のズレに鋭敏である。本研究では、2文が議論の中で係り受けの関係にあるかを判定するように学習を行う。そのため、前文が話題の提示であったり、主文が同意や否定であったりするケースに鋭敏である。

本報告のモデルは、NSPで事前学習したBERTモデルをベースに、実際のグループワークの対話データの係り受け文ペアを用いてファインチューニングを

Fine-tuned BERT model to estimate the probability of connection to preceding utterances for visualization of dialogue structure.

Kazuki TOKUNAGA[†], Takahiro NAKAYAMA[‡], Hajime SHIROUZU[§], Osamu ICHIKAWA[†]

[†] Faculty of Data Science, Shiga University

[‡] The University of Tokyo - Center for Research and Development on Transition from Secondary to Higher Education

[§] National Institute for Educational Policy Research

行うことにより作成した。本報告ではこの BERT モデルの構築までを主なスコープとし、最終的なディスカッション木構造の推定は試行として紹介する。

3.1 文間距離によるペナルティ

同じ接続確率であれば、主文により近い前文がその親として選ばれるべきである。そのための適切なペナルティを付与するために、以下の式を用いる。

$$\log(p(y | x)) = \log(p(y | x, d)) + \alpha \log(p(d)) \quad (1)$$

y が候補となる d 発話前の文, x が現在文である。 α は任意の係数であって、チューニングにより定める。 $p(d)$ は事前確率であるので、学習データを用いて分布を取得しておく。ただし、 $p(d \leq 3)$ は、小さい係数 ϵ でフロアリングしている。図3にその結果を示す。後に示す実験では α を 0.01, ϵ を 0.0001 に設定した。

3.2 同一話者の連続発話のラベル補正

同一人物が連続で発話した場合に、それが複数の発話に分割されて記録されていることがしばしばある。学習データのラベル付けを行う際には、主文の親として自分自身の発話番号は選択しないルールとしている。同様に接続確率の推論時も、候補となる前文に主文の話者の発話が含まれないようにする。したがって、図2、図3で示した d 発話前の文 ($n-d$ 番目の文) の d のカウントには自分自身が連続で発話した文のカウントは含まれないので注意されたい。

4 学習データの収集と品質評価

中学生のグループワークデータをモデルに学習させるために人力によるラベリングを行った。1つのラベルに対して3~4名でラベリングを行い、多数決により最終的なラベルを決定した。人手によるラベリングの精度を調べるため、収集したデータの一部について、16名でラベリングを行い、その品質を評価した。正解は多数決により定めた。正解率平均は 0.65, 標準偏差は 0.11 であった。単純な推定では、ラベラーを 3名に減らすと標準偏差は約 0.25 となる。

5 先行発話への接続確率の推定モデル

事前学習済み BERT モデルを出発点とし、2つの文が係り受け関係にあるか否かを判別できるように2段階のファインチューニングを行う。まず、対話破綻データベース(DBDC データセット)[5]を使用し、さらに実際の中学生のグループワークの書き起こしテキストデータでファインチューニングを行う。中学生のグループワークデータを使用した評価実験の結果、正解率が 0.74 となった。これは親となる前文が無い、ある場合には前3文のうちどれかという4クラス分類の集計である。なお、独立した発話であるかどうかの閾値は 0.4 に設定した。図4にモデルが推定した木構造の例を示す。

6 おわりに

本報告では、生徒の学びの過程を可視化する一候

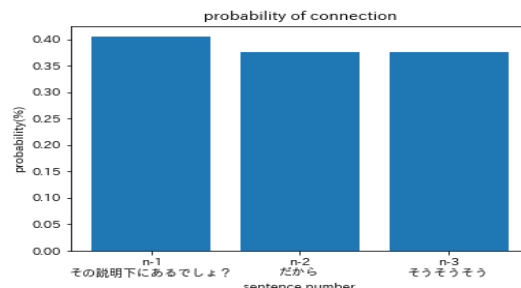


図2 先行文との接続確率の例. n 番目の文「あそうか」

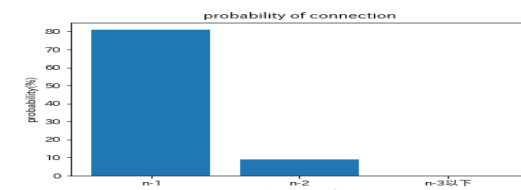


図3 文間距離ごとの接続事前確率

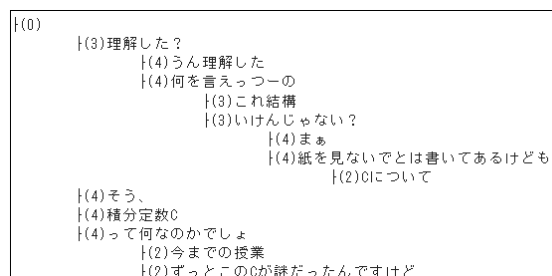


図4 モデルが推定した木構造の例。

補として生徒のグループ議論の木構造に可視化した。提案手法では、DBDC データセット[1]を学習データとしてファインチューニングされたモデルをさらに中学生のグループワークデータでファインチューニングしたモデルを用いた。実際のデータを用いた実験では、正解率 0.74 で対話の木構造を再現することができた。今後、教師にとっての有用性を検討し、木構造による可視化が手段として適切であるかを検証する必要がある。

謝辞

本研究は科研費 20K20816, 17H06107, 19K02999 の助成を受けた。

参考文献

- [1] 工藤 情報処理学会自然言語処理研究会報告 NL-76-7, pp1-8, 1990
- [2] T. Nomoto et al. Proceedings of Workshop on Very Large Corpora, pp. 216-224, 1998.
- [3] D. Cristea, et al. Proceedings of the 18th COLING 2000, vol. 1, 2000.
- [4] 長尾 確 第2回音声ドキュメント処理ワークショップ, 2008
- [5] 東中 竜一郎, 他 人工知能学会 言語・音声理解と対話処理研究会 第72回, pp.45-50, 2014-12-15