

傾聴の応答で繰り返される語句の検出性能の向上

川本 稔己 長谷川 駿 上垣外 英剛 船越 孝太郎 奥村 学
東京工業大学
kawamoto@lr.pi.titech.ac.jp

1 はじめに

対話において相手の話を聞いているという意思表示を行うことは円滑なコミュニケーションを行う上で重要である。その手段として、ジェスチャや発話をする事が挙げられる。ジェスチャの具体例は頷くことや首をかしげることであり、発話だと相槌を打つことが代表的である。相槌以外にも相手の発話に賛同を示すために話し手が使用した語句を反復し発話をする時がある。以下に例を示す。

話し手: 僕はやっぱり野球は無理だなと思ってきましたね。

聞き手: 野球は無理でしたか。

この例では、話し手の発話に含まれる「野球は無理」が応答で反復され、賛同を示している。このような応答を繰り返し応答という。人が傾聴するときと同様に、対話システムに傾聴をさせる際にも相槌だけではなく繰り返し応答も行うことで会話をより円滑に進める事が出来る。その生成を行うにはどの語句を反復するかを決める必要がある。よって、本稿では話し手の発話の中から繰り返し応答に用いられる文節を検出する研究に取り組む。

伊藤ら [1] は個々の文節に関する素性を主に用いて SVM で検出する手法を示しているが、発話における文節間の関係を十分に考慮していない。この問題を解決するために、本稿では文節間の関係を捉えることが可能な BERT [2] を使用する。さらに、発話をまたぐ文脈を考慮する目的で直前の話し手の発話も入力として用いる。

一般に繰り返し応答に対するデータを大量に準備することは難しい。その理由としては、賛同を示す発話の種類である繰り返し応答が対話時に多用されるものではないという点が挙げられる。そのため、本稿では傾聴対話コーパス [3] の話し手の発話を用いて、その発話に対して取りうる繰り返し応答のバリエーションを人手でクラウドソーシング

表 1 先行研究で使用された素性一覧

1.	文節の主辞の品詞
2.	文節の主辞の品詞細分類
3.	文節の主辞の活用形
4.	文節の機能語の品詞
5.	文節の機能語の品詞細分類
6.	文節の機能語の活用形
7.	文節の形態素数が以下の 4 分類のいずれであるか (1, 2, 3-6, 7 以上)
8.	文節が属する節境界単位の種類
9.	節境界単位内で先頭から何番目の文節か
10.	文節が固有名詞を含むか否か
11.	文節が数詞を含むか否か
12.	文節が固有表現を構成する形態素を含むか否か
13.	IREX の分類について、文節がその固有表現を構成する形態素を含むか否か
14.	文節内に内容語が存在し、文節の重要度 I が全形態素の中央値 13.56 より大きいかな否か
15.	直前の文節の最後の形態素がフィラーかな否か

サービスを用いて作成することによりデータを用意する。

作成したデータを対象とした評価実験の結果、提案手法はベースラインである SVM を用いた手法より 24 ポイント高い F 値を達成した。それにより文脈を考慮した提案手法の有効性を確認できた。

2 先行研究

伊藤らは、文節ごとに与えられる素性を元に各文節が繰り返し語句を含む文節であるかを、分類器を用いて分類している。その際に用いられている素性を表 1 に示す。

先行研究から、特に固有表現・フィラーの後の語句、重要度が高い語句は繰り返しの対象になりやすいことが分かっている。固有表現の抽出および IREX タグの付与には CaboCha[4] が用いられている。フィラーの直後の語句は話し手が言いよどみした後に発言したものであり、比較的重要な意味を持つと考えられる。語句の重要度は以下の指標で計られている。

$$I(w_i) = \log_2 \frac{\sum_j F(w_j)}{F(w_i)+1}$$

ここで、 w_i は形態素、 $F(w_i)$ は『日本語話し言葉コーパス』短単位語彙表 ver 1.0¹⁾における w_i の出現頻度である。語の異なり総数は 42,545、語の総頻度 $\sum_j F(w_j)$ は 7,479,773 である。素性 14 の文節の重要度 I は、文節に属する内容語の重要度 I のうち最大のものとした。

3 提案手法

本節では文節間の関係を捉え、発話をまたぐ文脈を考慮し、繰り返し文節を検出する手法を提案する。図 1 に提案モデルの概要を示す。

提案手法では、事前学習された、日本語を対象とした Bidirectional Encoder Representations from Transformers (BERT)²⁾ をエンコーダとして使用する。BERT を使うことで文節間の関係を捉えることが可能となる。入力には、発話をまたぐ文脈を考慮するために、話し手の対象発話 $W = (W_{1,1}, \dots)$ と対象発話の直前の話し手の発話 $W^* = (W_{1,1}^*, \dots)$ を用いる。以後話し手の対象発話を SP、対象発話の直前の、同じ話し手の発話を PreSP と呼ぶ (例は 4.1 節を参照)。入力順序は時系列に沿い、PreSP の後 SP とする。発話の前後には発話の開始記号として <CLS>、発話の終端記号として <SEP> を挿入する。なお、BERT は特徴量抽出のみに利用し、fine-tuning は行っていない。

次に、予め分割しておいた文節ごとに、文節が表現しているベクトルを得る。そのために BERT でエンコードした検出対象発話の単語単位のベクトルを、文節ごとに self-attentive span extractor^[5] を利用して文節単位のベクトルに変換した。そうして得たベクトル h_i^s に表 1 に示す素性を結合し、MLP (multi-layer perceptron) の入力とした。MLP は最終層に sigmoid 関数を使用しているため、出力は [0,1] の連続値になる。その値に対する閾値を使用して二値に分類する。閾値は開発データを用いて最適化した。

4 データ

本節では人手で繰り返し応答を作成し、繰り返し対象語句を含む文節のタグ付けを行う手順について説明する。

繰り返し応答は傾聴時に多く出現することから、話し手の発話についてはデータとして高齢者と介護

士による傾聴対話 [3] を用いた。この対話は、話し手となる高齢者と聞き手となる介護士が情報案内および傾聴を行うコーパスである。本コーパスの発話の総数は 9716 である。

4.1 繰り返し応答データの作成

繰り返し応答のデータがある程度の量用意するためにクラウドソーシングを利用して繰り返し応答を手で作成した。³⁾ クラウドソーシングには Lancers⁴⁾ を使用した。具体的な手段としては、データの話し手の発話で、内容語 (名詞、動詞、形容詞、副詞のいずれか) を含む文に対して以下のような設問を提示した。内容語を特定するために必要な形態素解析器には MeCab⁵⁾ を用いた。作成してもらう際には、50 個の設問を 1 つのパートとし、合計 36 個のパートをそれぞれ 3 名に依頼した。したがって作成者はのべ 108 名である。以下が作成作業時に提示された設問と発話の例である。以後、聞き手の発話を LI と呼ぶ。

Q: 下線が引いてある SP の語句を使い応答文を作成してください。それ以前の対話は文脈理解に使用してください。

PreSP: はい私の名前が。

LI: 何が出るんです？釣った大きさ？

SP: 大きさとかね、そういう何匹釣ったとか、そんなん。

以上の設問に対する 3 名からの回答が以下である。

A-1: 何匹釣ったとかということが出ているのですね。

A-2: 釣った魚の大きさとか数の情報がでるのですね。

A-3: 大きさと何匹釣ったかとかですか。

これらの例では下線の話し手の文で使われている内容語「大き」「何」「匹」「釣っ」を一つ以上含む繰り返し応答が作成できていることがわかる。

そして、提示した話し手の文と、人手で作成された、上のような 3 文のうち 1 文をランダムに選択し、それらを文対とする。以降、本稿において LI とは作成した応答とする。

1) https://pj.ninjal.ac.jp/corpus_center/bccwj/freq-list.html

2) <https://huggingface.co/cl-tohoku/bert-base-japanese>

3) 傾聴対話コーパスからルールベースで繰り返し応答となる発話対を抽出しようとしたが、十分な量得られなかった。

4) <https://www.lancers.jp/>

5) <https://taku910.github.io/mecab/>

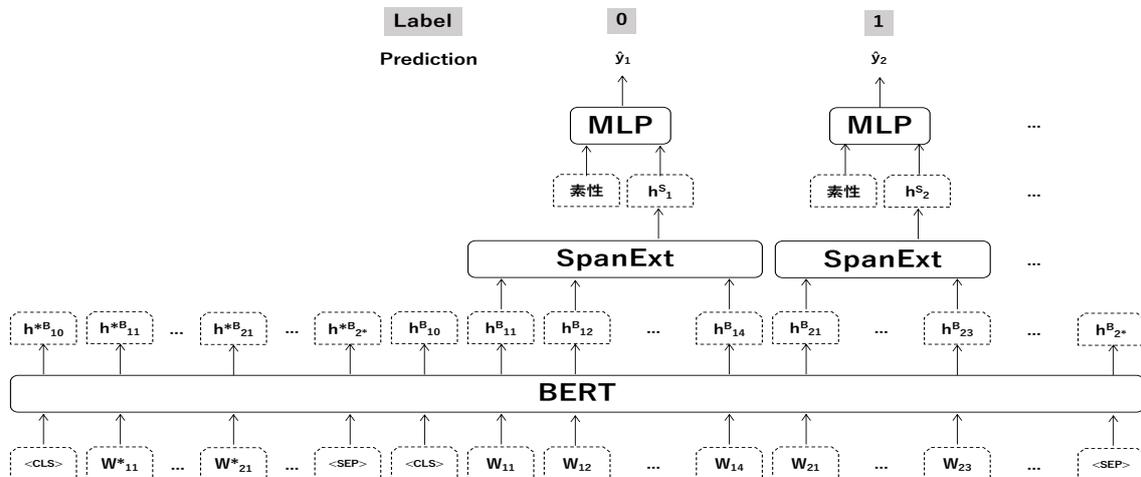


図1 提案手法で使用するニューラルネットワークの構造. BERTには時系列順に話し手の対象発話 (SP) とその直前の話し手の発話 (PreSP) を入力する. MLPの入力には, 文節ごとに SpanExtractor を使用したベクトルに表1の素性を追加している.

その結果 1760 の文対が得られた.

4.2 繰り返し対象語句を含む文節のタグ付け

4.1 節で得られた文対に対して, 繰り返し対象となっている語句を含む SP の文節を以下の手続きで同定しアノテーションした. 文節の分割には CaboCha を用いた.

- LI 中と一致している内容語を含む SP の文節を繰り返し対象となる文節とする.
- 同じ内容語に対して複数の文節がある場合は, 他の繰り返し対象となる文節に最も近いものを選択する. 他の繰り返し対象文節が存在しない場合は, 最初に出現したものを選択する.

以下にタグ付けの例を示す. 下線の箇所が繰り返し対象となる語句を含む文節である. なお, スラッシュは文節間の区切りを表す.

SP: それまで/はね, /ちょっと/今/しばらく/冬眠中です./冬眠さしても/うてます.

LI: 冬眠明けが/楽しみです.

5 実験

先行研究では, SP を節単位に分割し, 繰り返し対象文節を1つも持っていない節は実験対象から除外

している. しかし, 対話システムやチャットボットで実際に繰り返し応答を生成する際には, 節よりも文が理解の単位として重要であり, また, 繰り返し文節が存在しない文も対象として実験を行うことが望ましい. そこで本稿では, 先行研究の再現という目的で, 参考として先行研究と同じ節単位での実験も行うが, 文単位で実験を行う.

5.1 節単位での実験

先行研究の実験設定に沿うように, SP を節境界で分割し, 繰り返し対象文節を1つも持っていない節は実験の対象から除外した. このため実験対象の節には必ず1つ以上の繰り返し対象文節が含まれる. 節境界の検出には節境界解析ツール CBAP[6] を用いた.

上述したように, 節には必ず少なくとも1つ繰り返し文節が含まれるという設定であるので, 繰り返し対象となる文節を検出する際, 繰り返し対象文節と判定される文節が1つもない節については, 二値分類する際のスコアが最大の文節を繰り返し対象文節とする.

表2 実験対象データのサイズ

発話文数	1760
発話文中の文節数	13207
繰り返し文節を含む節数	2531
繰り返し文節を含む節内の文節数	10970

5.2 文単位での実験

本設定では文を一つの単位とみなす。また、傾聴において繰り返し応答をしない場合も存在する方が自然である。よって、一つも繰り返し対象文節と判定されるものがない場合も、節単位での実験のように、後処理は行わないこととする。これにより、先行研究より自然な実験設定になる。

5.3 実験概要

実験には4節で作成したデータを用いた。データのサイズを表2に示す。実験では訓練データ:開発データ:テストデータ=3:1:1として5分割交差検証をモデルの初期値を変更して3回ずつ実験し、計15回の平均を実験結果とした。

繰り返し対象文節を正解とし、正解に対する再現率・適合率・F値により評価した。再現率・適合率・F値はそれぞれ、

$$\begin{aligned} \text{再現率} &= \frac{\text{正しく検出できた繰り返し対象文節数}}{\text{正解の繰り返し対象文節数}} \\ \text{適合率} &= \frac{\text{正しく検出できた繰り返し対象文節数}}{\text{繰り返し対象として検出された文節数}} \\ F \text{ 値} &= \frac{2 * \text{再現率} * \text{適合率}}{\text{再現率} + \text{適合率}} \end{aligned}$$

のように計算される。

ベースラインには Support Vector Machine (SVM) による先行研究を用いた。表1の素性に基づいた先行研究の再現を可能な限り行った。SVMはscikit-learnライブラリ⁶⁾のSVMを、初期設定されたハイパーパラメータに基づいて使用した。節単位での実験は、先行研究からの本手法による改善を確認するために行うので、ベースラインであるSVMと提案手法の2種類のモデルで比較を行う。

文単位の実験においてはその2種類に加えて、ablation studyとして提案手法の入力にPreSPを組み込まないw/o PreSP、提案手法のMLPの入力に表1の素性を追加しないw/o 素性、そのどちらも組み込まないw/o PreSP and 素性の3種類の設定下でも実験を行った。

表3 実験結果 (節単位) (%)

	再現率	適合率	F 値
SVM	54.15	61.54	57.59
提案手法	82.27	58.55	68.39

表4 実験結果 (文単位) (%)

	再現率	適合率	F 値
SVM	22.39	56.95	32.01
提案手法	68.24	48.13	56.21
提案手法 w/o PreSP	76.00	42.90	54.82
提案手法 w/o 素性	67.70	42.86	52.46
提案手法 w/o PreSP and 素性	68.06	42.56	52.33

5.4 実験結果

節単位での実験結果を表3、文単位での実験結果を表4にそれぞれ示す。

節単位では、ベースラインのSVMのF値が57.59%であったのに対し提案手法のF値が68.39%となり、約11ポイントの向上を確認した。これは先行研究と同様の設定において提案手法が有効であることを示している。

文単位では、ベースラインのSVMのF値が32.01%であったのに対し提案手法のF値が56.21%となり、約24ポイントと大幅に向上したことが確認された。その上で、提案手法と提案手法w/o PreSPを比較すると、PreSPを入力に加えることで発話をまたぐ文脈を考慮することが出来ていることがわかる。また、提案手法と提案手法w/o 素性を比較すると、素性を入れた手法がより高いF値を得ていることから、先行研究の素性はSVMだけでなく提案手法においても効果的であることが分かる。

6 おわりに

本稿では、文節間と発話間の文脈を考慮することで傾聴の応答で繰り返される語句の検出性能を向上させる手法を提案した。提案手法は従来の手法とは異なり、BERTを用いることで文節間の関係を考慮している。その上で直前の話し手の文(PreSP)を加えることで発話をまたぐ文脈を考慮することが可能であり、また、MLPの入力に先行研究の素性を結合することによりさらなるスコアの向上も実現している。実験の結果、F値の大幅な向上が確認された。

今後は本手法で検出した繰り返しとなる語句を用いて、傾聴における応答の生成を目指したい。

6) <https://scikit-learn.org/stable/>

参考文献

- [1]伊藤, 村田, 大野, 松原. 傾聴を示す応答で繰り返される語りの語句の検出. 言語処理学会, 2019.
- [2]Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805, 2018.
- [3]Yoshino, Tanaka, Sugiyama, Kondo, and Nakamura. Japanese dialogue corpus of information navigation and attentive listening annotated with extended iso-24617-2 dialogue act tags. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), 2018.
- [4]工藤拓, 松本裕治. チャンキングの段階適用による日本語係り受け解析. 情報処理学会論文誌, Vol. 43, No. 6, pp. 1834–1842, 2002.
- [5]Kenton Lee, Luheng He, Mike Lewis, and Luke Zettlemoyer. End-to-end neural coreference resolution. arXiv preprint arXiv:1707.07045, 2017.
- [6]丸岡, 柏岡, 熊野, 田中. 節境界自動検出ルールの作成と評価. 言語処理学会, 2004.