

非対話データを用いた対話形式コンテンツの自動生成

Conversational Content Generation from Non-Conversational Data

岩橋 千穂^{1*} 稲葉 通将¹
Chiho Iwahashi¹ Michimasa Inaba¹

¹ 電気通信大学

¹ The University of Electro-Communications

Abstract: Conversational content has an advantage over non-conversational content in that it is easier to understand and provides a more immersive experience for the reader. In this paper, we propose a method for automatic generation of conversational content from non-conversational data. First, we extract key sentences from non-conversational data. Next, we generate conversational content based on the extracted key sentences. To generate conversational content, we use question or response generation models learned from conversational data. The proposed method enables automatic generation of conversational content that is more interesting and promotes understanding of the content compared to the original non-conversational data.

1 はじめに

複数の話者が対話する形式を取ることで読者に情報を伝える対話形式コンテンツは、ニュース記事などの非対話形式コンテンツと比較して、簡潔な発話の応酬で構成されるため内容を理解し易いという利点がある。また、対話という身近に感じられる形式を取ることで、読み手に没入感を与えることも可能である。

そこで本研究では、非対話データを用いた対話形式コンテンツの自動生成手法を提案する。ここでは、非対話データ内の情報を提示する人物と、質問などを通してその情報を聞き出すような発言をする人物による2者の対話を想定している。その具体例を図1に示す！これは、ニュース記事のような非対話データおよびその関連文書から対話形式コンテンツを生成する例である。Aが「首相が来年夏に延期された東京五輪・パラリンピック開催への決意を表明した」等の非対話データ内の情報を提示し、Bはそれを聞き出すような役割を担っている。

非対話データ

第203臨時国会が26日召集され、菅義偉首相は衆参両院本会議で内閣発足後初の所信表明演説に臨んだ。首相は、来年夏に延期された東京五輪・パラリンピック開催への決意を表明。新型コロナウイルス対策と経済活動を両立させる方針を明言し……

関連文書

首相が臨時国会、特別国会で当面の内閣の基本的姿勢や重要課題について見解を示すために行う。首相の演説に対し……

対話形式コンテンツ

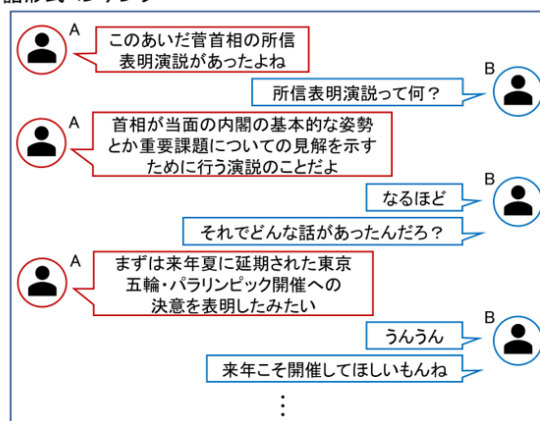


図 1: 対話形式コンテンツの例

2 対話形式コンテンツ生成

本手法ではまず非対話データから重要文抽出を行う。次に、対話データを用いて学習した質問・応答生成モデル

を使用し、抽出した重要文が答えとなる質問や、情報提示の発話に対する応答を生成する。重要文から成る発話とそれに対する質問・応答を交互に提示することで、対話形式のコンテンツを形成する。その際、一般的に意味が知られていない単語や人名の説明など、自然な対話の流れで必要となる情報が対象の非対話データ

*連絡先： 電気通信大学
〒 182-8585 東京都調布市調布ヶ丘 1-5-1
E-mail: i1710081@uec.ac.jp

¹<https://www.jiji.com/jc/article?k=2020102601011&g=pol>, <https://www.jiji.com/jc/article?k=2020102600634&g=tha> より引用。

タ内に含まれない場合には、図1の「所信表明演説」に対する説明のように、関連文書を参照し、その情報を利用する。さらに、より実際の対話に近づけるため、本題に入る前の導入と話題を終えた際のまとめ部分の対話を生成し、ここまでで生成された対話の前後に追加する。また、非対話形式データは書き言葉で書かれているが、実際の対話は話し言葉で行われる。そのため、非対話形式データから生成した発話の口調は話し言葉へと変形する必要がある。加えて、情報提示の発話に対して生成した応答の口調に一貫性がない場合にも、口調統一のために変形が必要である。そこで最後に、このような口調変形等の調整を行い、コンテンツとしての体裁を整える。

2.1 応答生成

重要文抽出には、抽出型要約タスクで高い性能を達成したBERTSUMEXT[1]などのTransformerベースの手法を使用する。

質問・応答生成については、Wangらの提案した会話型質問生成手法[2]やTianらの提案した応答生成手法[3]を参考に、本研究に適した手法を検討中である。

関連文書の参照に関しては、TF-IDF値などを用いて非対話データ内で注目する単語を決定し、それをキーワードとしてWikipediaを参照することを考えている。

導入部分の生成手法としては、現時点では非対話データの冒頭部分に基づくルールベースの手法を検討している。

まとめ部分に関しては、非対話データに生成型要約を適用し、得られた要約文を発話として使用した1ターン程度の対話を生成することを考えている。

2.2 口調の統一

口調を統一するためには、期待する口調でない発話を適切な口調へ変形すればよい。口調の変形はテキストのスタイル変換と捉えることが可能である。

本研究では、Heらの提案したニューラルネットワークベースのスタイル変換モデル[4]を使用する。これは、非並列コーパスを部分的な並列コーパスと見なし、スタイル変換器を学習する手法である。

通常のスタイル変換ではポジティブな文からネガティブな文への変換など、文全体を大きく変形する機会が多い。予備的な検討を行ったところ、Heらのモデルの損失関数を修正することで、口調のみの変形が可能であることを確認している。

3 関連研究

3.1 ユーザシミュレータ

ユーザシミュレータは擬似的なユーザとして使用される対話システムであり、対話システムの学習や性能評価のために広く利用されてきた。ユーザシミュレータを用いた研究は、最終的に対話形式のデータが生成されるという点で本研究と類似している。

近年の研究では、Kreyssigらが、コーパスから対話行為を学習し応答を生成するニューラルネットワークベースのユーザシミュレータを提案している[5]。また、Shiらは汎用的なユーザシミュレータ構築のための指針を示した[6]。

このような手法は、あくまで対話システムの学習や評価のために使用されるものである。一方、本研究はコンテンツとしての対話形成を目的としており、通常のユーザシミュレータを用いた研究とは異なる。

3.2 質問・応答生成

本研究では、情報提示の発話の後に、聞き役となる話者の発話として質問や応答に相当する発話を生成する。このような点で、質問・応答生成は本研究と強く関連している。

Wangらは意味的一貫性を保ち興味を引くような質問を生成する会話型質問生成手法を提案した[2]。この手法では、モデル学習時の入力として生成する質問への回答を与えることで意味的一貫性を保持する。さらに、より明示的に意味的一貫性を保つため、強化学習やGANも利用されている。また、モデル学習時に質問ラベルを入力することで、入力された発話に対して適切かつ興味喚起を促す質問の生成を実現している。

Tianらは情報量に富み多様性のある応答生成のための手法を提案した[3]。この手法では、まず学習コーパスの入力・応答ペアをクラスタに分割し、各クラスタを代表する入力ベクトルおよび応答ベクトルを生成する。そして応答生成時には、これらのベクトルを用いて入力発話に対する応答ベクトルを新たに生成し、応答生成の入力の一部として使用する。このように各クラスタから抽出した特徴を応答生成に利用することで、入力発話のトピックに即した応答生成が実現される。また、学習コーパスをクラスタに分割することにより、生成する応答全体の多様性を増加させることが可能である。

これらの手法はどちらも直前の発話に対する質問または応答を生成するものである。一方、本研究では情報提示を行う複数の発話が予め決まっており、聞き役の発話としてこれらの複数発話の間に入るような質問や応答を生成する。よって、情報提示の発話を自然に

導くような質問または応答の生成手法を検討する必要がある。

4 おわりに

本研究では重要文抽出と質問・応答生成を用いて非対話データから対話形式コンテンツを自動生成する手法を提案した。本手法により、元の非対話データと比較して、より興味を引き内容理解を促すコンテンツの自動生成が可能になると期待される。今後は、情報提示を含む自然な対話を形成するための手法を検討し、提案手法に基づくシステムの構築を目指す。

参考文献

- [1] Yang Liu and Mirella Lapata. Text summarization with pretrained encoders. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 3730–3740, Hong Kong, China, November 2019. Association for Computational Linguistics.
- [2] Weichao Wang, Shi Feng, Daling Wang, and Yifei Zhang. Answer-guided and semantic coherent question generation in open-domain conversation. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 5066–5076, Hong Kong, China, November 2019. Association for Computational Linguistics.
- [3] Zhiliang Tian, Wei Bi, Xiaopeng Li, and Nevin L. Zhang. Learning to abstract for memory-augmented conversational response generation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 3816–3825, Florence, Italy, July 2019. Association for Computational Linguistics.
- [4] Junxian He, Xinyi Wang, Graham Neubig, and Taylor Berg-Kirkpatrick. A probabilistic formulation of unsupervised text style transfer. In *International Conference on Learning Representations*, 2020.
- [5] Florian Kreyssig, Iñigo Casanueva, Paweł Budzianowski, and Milica Gašić. Neural user simulation for corpus-based policy optimisation of spoken dialogue systems. In *Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue*, pp. 60–69, Melbourne, Australia, July 2018. Association for Computational Linguistics.
- [6] Weiyang Shi, Kun Qian, Xuewei Wang, and Zhou Yu. How to build user simulators to train RL-based dialog systems. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 1990–2000, Hong Kong, China, November 2019. Association for Computational Linguistics.