

Transformer を用いたフィルタリングアーキテクチャに基づく非タスク指向対話システムの設計

2019SE007 林祐杜

指導教員：野呂昌満

1 はじめに

近年、機械学習を用いた非タスク指向対話システムの正確な文章生成、会話の流れに沿った自然な応答に関する研究が盛んに行われている [3]. 非タスク指向対話システムとは、雑談などの自由な話題でユーザと会話を行う対話システムである。非タスク指向対話システムの目的は、自由な話題でユーザと会話を行うことでユーザを退屈させないことである。

大規模対話モデルでは応答に矛盾、話題の飛び、キーワード関係誤りが発生することが多い [1]. 対話システムでは、会話の流れに沿った自然な応答と正確な文章生成が重要視されているので、上述の問題を解決することは重要である。

本研究の目的は、ユーザに対してより適切な応答生成を行う対話システムの設計である。適切な応答とは、上で述べた問題が発生しない応答のことを指す。

研究目的を詳細化すると研究課題は以下の通りとなる。

1. フィルタリングアーキテクチャに基づく対話システムの設計
2. 対話システム設計の妥当性検証

ユーザに対してより適切な応答生成を行うために、生成文の中から不適切文を除去する必要がある。さらに、ユーザに適した応答をすることが重要である。杉山ら [1] は、不自然な応答を除去するために、生成文のフィルタリングを行った。小林ら [2] は、感情推定を行うことで、ユーザの嗜好や人間関係を推定した。杉山ら、小林らの対話システムを融合することで、それぞれのユーザに適した応答ができる対話システムを設計できるので、研究目的を達成できると考える。研究目的を達成できているか確認するために、対話システムの設計について妥当性検証をする必要がある。本研究では、既存の対話システムの融合について考察をすることで設計する対話システムの妥当性検証を行う。

2 既存研究

杉山ら [1] は、Transformer を用いて応答文の候補を生成し、候補文からフィルタを用いて不自然な文を除去する対話システムを構築した。杉山らが実装した非タスク指向対話システムは、応答生成モデルに対話データセットを学習させることで様々な対話状況でも精度の高い応答生成を実現している [3].

小林ら [2] は、感情推定モデルを構築してユーザの嗜好を推定することで、ユーザに適した応答を行う対話システムを実装した。小林らは、深層格推定、極性推定、素性抽

出、人称推定、極性判定規則からユーザの嗜好や人間関係を考慮する非タスク指向対話システムの実装をしている。

3 アプローチ

本研究では、杉山ら小林らの応答生成モデル、フィルタリングモデル、感情推定モデルを融合することで、より適切な応答生成を行う対話システムを設計する。応答生成モデルでは、対話文を学習することで自然な雑談を行えるようにする。フィルタリングモデルでは、データセットとユーザ情報格納クラスを使用することで、応答生成モデルが出力した不適切文を除去する。本研究のフィルタリングモデルは、Transformer によって対話中に入力された人物や場所とユーザの関係を記憶することで、ユーザに対して適切な応答を行えているか推定する。感情推定モデルは入力文とユーザ情報格納クラスから、ユーザの感情を推定する。杉山らの対話システムでは、応答生成モデル、フィルタリングモデルを用いて会話状況に関係なく状況に合わせた応答を行う。小林らの対話システムでは、感情推定モデルを用いて、様々なユーザに適した応答を行う。本研究では、杉山ら小林らの対話システムを融合することで、様々なユーザに適した自然な雑談を行う対話システムの設計を試みる。本研究では、設計した対話システムの構造に関する考察をすることで対話システムの妥当性検証を行う。

4 フィルタリングアーキテクチャに基づく非タスク指向対話システムの設計

Transformer, 応答生成モデル, フィルタリングモデル, 読み取りモデル, 感情推定モデルを用いて、ユーザの嗜好を推定してユーザに対してより適切な応答生成を行う。設計する対話システムを静的に表した図を図 1 に示す。

Transformer と各モデルを用いて、対話システムを設計する。本研究では、複数の既存の対話システムを融合することで、ユーザに対して適切な応答生成の精度向上を目指す。複数の既存の対話システムを融合するために、Transformer を用いて各モデルを設計する。各モデルは、結果を出力するためにモデルに対応したデータセットを使用する。図 1 の読み取りモデルでは、入力文を SentencePiece を用いてサブワードに分割する [4]. ユーザ情報格納クラスは、SentencePiece によってサブワードに分割されたユーザ情報を格納する。感情推定モデルでは、ユーザ情報格納クラスからユーザ情報を受け取り感情を推定する。応答生成モデルは、話者性、共感といった雑談対話に必要なスキルを学習する。フィルタリングモデルは、応答生成モデルが出力した応答を、感情推定モデルと過去

に出力した応答からユーザに適した応答かを推定して不適切文を除去する。応答生成モデルの出力がフィルタリングモデルに不適切と判断された場合、もう一度応答生成モデルからフィルタリングモデルに回答候補が渡される。これらの各モデルを用いて、ユーザに対してより適切な応答生成の実現を目指す。設計する対話システムを動的に表した図を図 2 に示す。

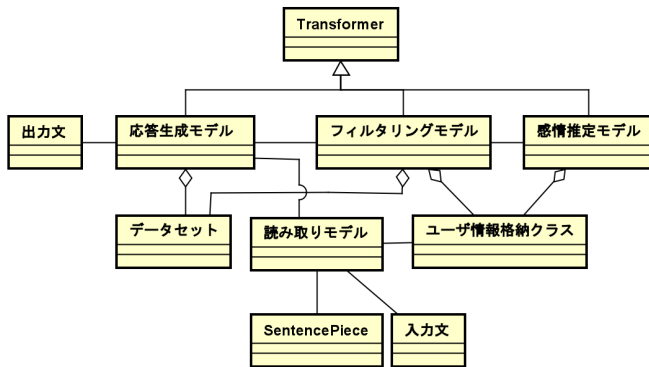


図 1 非タスク指向対話システムのアーキテクチャ A

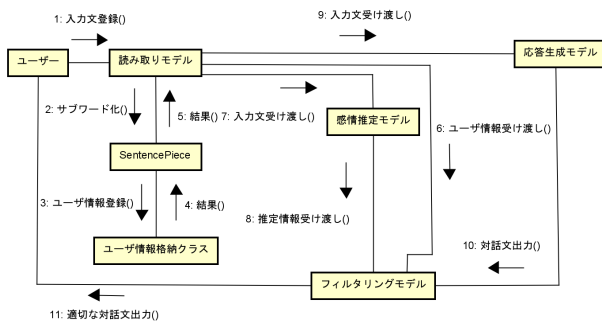


図 2 非タスク指向対話システムのアーキテクチャ B

5 考察

本研究で設計する対話システムは、追加するモデルと追加するモデルに対応したデータセットとの間に has-a 関係を持たせることで新しいモデルを利用できる。本研究の対話システムは、新しいモデルに対応した既存のデータセットを再利用することが出来る柔軟な構造をしている。これより、データセットが十分にない状態でも新しいモデルの学習を行えると考える。本研究で設計する対話システムに新しいモデルを追加した場合、他のモデルと同じように新しいモデルからフィルタリングモデルへ結果が出力される。このように設計する対話システムに新しいモデルを追加した場合でも、現在設計する対話システムとほぼ同じ動的処理で新しいモデルを追加することができる。と考える。

設計する対話システムを試行運用する場合、対話データセットは、Twitter 対話データセット、スキル学習用データセットを使用する。ユーザに対して適切な応答を

推定する上で、ユーザ対話文中にある単語同士の関連性を Transformer に学習させる必要がある。単語同士の関連性を学習させるには大量の対話データが必要になるので、対話データ数の多い Twitter 対話データセットを使用する。他方、自然な応答文を作成するためには、Transformer に自然な対話文を学習させる必要がある。自然な対話文を多く含むスキル学習用データセットを使用する。応答生成モデルでは、雑談対話に必要なスキルの学習と様々な対話状況に適した対話文を生成するために、Twitter 対話データセット、スキル学習用データセットを使用する。フィルタリングモデルでは、それぞれのユーザに対して適切な応答を推定するために Twitter 対話データセットを使用する。

6 おわりに

本研究では、ユーザに対してより適切な応答生成を行う対話システムの設計を目的として、既存の非タスク指向対話システムの融合を行った。既存の対話システムを融合することで、応答生成で発生する矛盾、話題の飛び、キーワード関係誤りを軽減する対話システムの設計をした。本研究では、複数の既存の対話システムを融合するために、Transformer を用いて各モデルの設計を行った。Transformer を用いて全てのモデルの設計を行うことで、新規モデルの追加を可能にした。

今後、設計した対話システムを定性的に評価するために実装を行う必要がある。既存の非タスク指向対話システムの応答と比較検証を行うことで、設計した対話システムの有効性を検証する。本研究の対話システムを実装した後に本研究の対話システムと、既存の非タスク指向対話システムの応答を定性的に評価をする。評価スコアから、設計した対話システムが既存の非タスク指向対話システムと比べて矛盾、話題の飛び、キーワード関係誤りを軽減するか確認を行う。さらに、文章評価モデル、人格評価モデルのような新しいモデルを導入することで、より上記の問題を軽減したユーザに対して適切な応答生成を行う対話システムの実装を目指す。

7 参考文献

- [1] 杉山弘晃, et al, “Transformer encoder-decoder モデルによる趣味雑談システムの構築,” 人工知能学会研究会資料 言語・音声理解と対話処理研究会, 2020.
- [2] 小林峻也, 萩原将文, “ユーザーの嗜好や人間関係を考慮する非タスク指向型対話システム,” 人工知能学会論文誌, 2016.
- [3] 東中竜一郎, et al, “対話システムライブコンペティション 3,” 人工知能学会研究会資料 言語・音声理解と対話処理研究会, 2020.
- [4] T.Kudo, et al, “SentencePiece: A simple and language independent subword tokenizer and detokenizer for Neural Text Processing,” *Arxiv*, 2018.