

会話履歴を考慮する応答手法における BERT の活用 Improvement by BERT of response method considering conversation history

石川大貴†
Daiki Ishikawa

土屋 誠司‡
Seiji Tsuchiya

渡部 広一‡
Hirokazu Watabe

1. はじめに

会話システムでは、システムとユーザーが会話を繰り返していく中で過去の会話を、会話履歴として扱う。この会話履歴というデータを用いたユーザーへの応答を会話履歴考慮応答という。会話履歴考慮応答では、ユーザーが過去に会話で出していた情報をもとに、それらを考慮した発展的な応答を行う。本研究では、より人間に近い自然な会話応答ができるようにすることを研究目的とし、会話システムにおける会話履歴考慮応答の生成手法の一つとして、BERT を活用した応答生成手法を追加した。

2. 関連技術

2.1 意味理解システム

意味理解システムとは、入力文を単語が保持する情報ごとに、7W1H と述語フレームに分割するシステムである。7W1H とは、5W1H(いつ, どこで, 誰が, 何を, なぜ, どのように)に、Who+(誰と), Whom(誰を)を追加したものである。意味理解システムにおいて、「私は昨日、友達と映画館で映画を観た」という例文を7W1H 格に格納したものを表 1 に示す。

表 1 意味理解システム格納例

Who	Who+	What	When	Where
私	友達	映画	昨日	映画館
Whom	How	Why	述語	
			観た	

2.2 会話履歴フレーム

会話履歴フレームとは、話者の会話履歴を意味理解システムと直前のシステムの発話を用いて 7W1H と述語に分け格納したものであり、各フレーム格においての分類も意味理解システムに準拠したものである。

2.3 会話履歴フレームデータベース

ユーザーごとに会話履歴フレームを格納し、保存されるデータベースである。この時、フレーム内のそれぞれの格がすべて一致する場合、会話履歴は格納されない。また、データベースに格納する際に対話が行われた日付を表す Day 格がフレームに格納される。

2.4 BERT

BERT(Bidirectional Encoder Representations from Transformers)^[1]とは、すべての層において左右両方向の文脈を強調させることで、ラベルの存在しないテキストから深い双方向表現を事前学習するように設計された大規模言語モデルである。また、文章を分散表現(ベクトル化)することが可能である。

2.5 コサイン類似度

コサイン類似度とは、異なるベクトル同士が「どれくら

い似ているか」を表す指標である。コサイン類似度は-1 から 1 の間の実数値で表され、類似度が高いほど値は 1 に近づく。

3. 既存システム

3.1 システム概要

既存の会話システム^[2]は、話者の発言に対し適した応答を選択し、会話を展開していくことで、テキスト形式の対話を目指したシステムである。既存システムの概要を図 1 に示す。

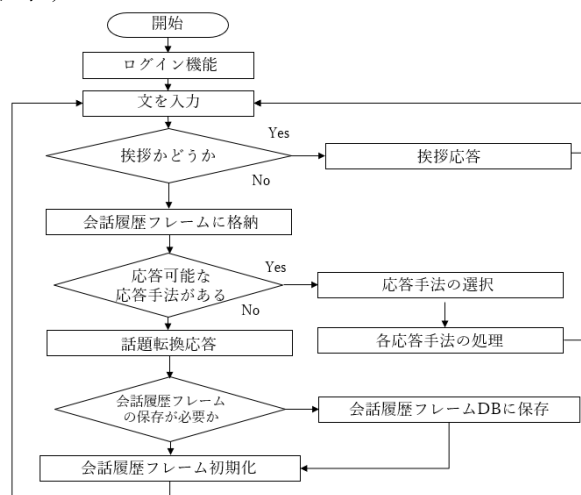


図 1 既存システムの概要

システムを起動するとログイン待機状態となり、任意の名前を入力した後メッセージ入力待機状態へと移行する。その後、図 1 に示すように進み、入力文を会話履歴フレームに格納し応答可能な応答手法の確認を行う。この際、応答可能な手段が存在する場合は状況に基づいた応答を行い、存在しない場合は話題転換応答が行われる。そして、会話履歴フレームの保存が必要であれば会話履歴フレームデータベースに保存し、会話履歴フレームを初期化した後、入力待機状態に戻る。このシステムにおいて、会話履歴考慮応答は、応答手法の一つとして存在しており、入力文と会話履歴フレームデータベースに格納されている文の What 格または Where 格の語の少なくともどちらかが完全一致している場合に生成するというアルゴリズムを用いている。

3.2 既存システムの問題点

既存システムでは、What 格、または Where 格の単語の完全一致のみを判断して会話履歴考慮応答を生成するため、それ以外の時に、意味的に似ている文章が入力されても、文章の意味を考慮した応答を生成することができない。

† 同志社大学大学院理工学研究科

‡ 同志社大学理工学部インテリジェント情報工学科

4. 提案手法

本研究では、BERT を用いた文章の類似度計算による会話履歴考慮応答の生成を追加し、その精度評価を行った。

この会話システムでは、ユーザーの過去の会話文と新規の会話文を BERT を用いて 1 文ずつベクトル化し、それらをコサイン類似度計算することで、システムがユーザーの会話内容・意味表現から理解した文章をもとに適切な返答をするアルゴリズムになっている。これにより、会話は過去の文脈に基づいた内容となり、より自然な会話が可能になると考えられる。また、応答生成には、条件を満たした会話履歴フレームの What 格または Where 格の単語を用いる。提案するシステムの会話履歴考慮応答の生成アルゴリズムの概要を図 2 に示す。

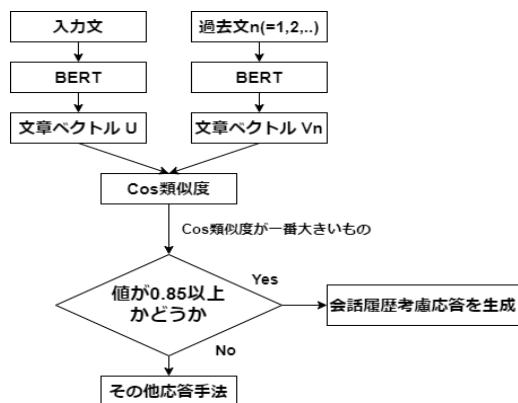


図 2 提案する会話履歴考慮応答の生成フロー

5. 評価

被験者は 5 名で、まず、図 1 で示したシステムで会話履歴フレームが 20 個保存されるまで会話を繰り返した。図 3 に、事前に登録された会話履歴リストを示す。

- 友達と京都の映画館に行く
- 友達と動物園に行く
- 友達と近所の公園でサッカーをする
- 彼女と遊園地に行く
- バイト仲間とバイト先でバイトをする
- 近所のスーパーに行く
- 京田辺の自宅でお菓子を作る
- 自宅のインターネットで服を買う
- 京田辺の学校に行く
- 友達とカラオケに行く
- 友達と海に行く
- 姉と神戸で釣りをする
- 友達と大阪でボウリングをする
- 自宅で音楽を聴く
- 友達と東京でライブに行く
- 遊覧でゴルフをする
- 大阪の美容院に行く
- 友達と三重でドライブをする
- 友達と家でゲームをする
- 大学の研究室で研究をする

図 3 事前に登録された会話履歴リスト

その後、ChatGPT³⁾に「あなたが昨日したことを短い文章で複数生成して下さい」と入力し、短い文章を複数生成した。その中から無作為に文章を選び、システムへの入力(ユーザーの入力)とした。その後、会話履歴考慮応答が 20 個生成されるまでこれを繰り返した。そして、得られた入出力結果について、それぞれ評価をしてもらった。評価の基準は、事前に登録しておいた会話履歴リストを被験者に見せた後、20 個の入出力結果を確認してもらい、「あなたがこの人だった場合、この応答は自然であるか」という質問に対する、「○：過去の内容が考慮されており、自然である」、「△：過去の内容が考慮されている

が、意味的に不自然である」、「×：日本語として成り立っていない」の 3 段階評価である。表 2 に評価結果を、図 4 にそれぞれ評価された出力例を示す。

表 2 評価結果

○	△	×
53%	32%	15%

○と評価された例	△と評価された例
U: 今日映画を観ました S: 京都の映画館で観るのですか?	U: 私は部屋を掃除しました S: 家ですのですか?
U: 私は日記を書きました S: 大学の研究室で書くのですか?	U: 私は新しいレストランに行きました S: 場所は遊園地ですか?
U: 私は昼食にサンドイッチを作りました S: 京田辺の自宅で作るのですか?	U: 私は新しいパズルを解きました S: 遊園地で解くのですか?
×と評価された例	
U: 私は洗濯物を片付けました S: 場所は自宅のインターネットですか?	
U: 私は古い服を整理しました S: 場所は自宅のインターネットですか?	
U: 私はコーヒーを淹れて新聞を読みました S: 場所は自宅のインターネットですか?	

図 4 評価例 (U:ユーザー, S:システム)

6. 考察

本研究では、文章の部分一致における会話履歴考慮応答に加えて、文章の意味一致による応答の生成を目指した。表 2 と図 4 から、本研究の目的としていた、文章の意味一致による会話履歴考慮応答を生成することができたと考えられる。今回、会話履歴フレーム数が 20 という少ない数であったが、○と評価された出力は 53%と半数を超える結果となった。また、△と評価された出力は 32%であり、これは会話履歴フレーム数が少なかったため、一番類似度が大きいものが、会話履歴考慮応答の生成に不適切なものになってしまったと考えられる。そのため、今後さらに会話履歴フレーム数を増やしたり、類似度の閾値を調整することで、さらなる精度向上ができると考えられる。一方で、×と評価された出力は 15%であり、自然な会話システムを目指すにあたってはさらなる改善案が必要な結果となった。このような評価が得られた一つの原因として、図 4 で×と評価された例に挙げているような「インターネット」のように手段として扱うべき情報を、場所の情報として扱っている点が考えられる。そのため、手段の情報を正しく格納できるようなシステムの拡張が必要である。その他にも、本システムでは、会話履歴考慮応答の生成に What 格と Where 格の単語しか用いてないため、他の格を用いた応答も生成できるように拡張することで、より自然な会話履歴考慮応答が生成できると考えられる。

7. おわりに

本研究では、会話履歴考慮応答の生成手法を追加することで、より自然な会話ができるシステムの構築を目指した。既存システムでは生成できない文章の類似度という視点からも応答が生成できるようになったが、不自然だと捉えられてしまう応答も生成された。これらを改善することで、システムの性能をより向上できると考えられる。

参考文献

[1]Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. "BERT: Pre-training of deep bidirectional transformers for language understanding."(2018).
 [2]小山翔平・土屋誠司・渡部広一, “雑談型対話システムにおける自己開示発話の割合増加手法の提案”, 情報科学技術フォーラム(FIT), セッション E-009, (2020/9/2).
 [3]OpenAI. ChatGPT. Version GPT-4, OpenAI, 2024, www.openai.com/chatgpt. Accessed 12 June 2024.