

対話履歴との矛盾を考慮した発話選択

Utterance selection that prevent contradicting dialog history

杉山 享志朗* 水上 雅博 吉野 幸一郎
田中 宏季 鈴木 優 中村 哲

Kyoshiro Sugiyama Masahiro Mizukami Koichiro Yoshino
Hiroki Tanaka Yu Suzuki Satoshi Nakamura

奈良先端科学技術大学院大学情報科学研究科
Graduate school of Information Science,
Nara Institute of Science and Technology

Abstract: When a dialog system outputs an utterance that contradicts the dialog history, users feel that the system does not understand what they say. We propose an dialog system that has a module to prevent contradictions by considering consistency between the utterance and the dialog history. As the first step, we examine the difficulty of contradiction detection in a dialog by creating an inconsistent-utterance dataset from the dialog-breakdown corpus. We constructed a discriminate model based on multi layer perceptron with bag of words features. According to the discrimination result, it is difficult to detect contradictions in high accuracy, even if we only focus on contradictions between the user utterance and one dialog history. We designed a dialog system that uses the contradiction detection result.

1 はじめに

対話システムがこれまで話した内容と整合しない発話を行ってしまうことは、ユーザとの対話を破綻させ、持続可能な対話システムとユーザの関係構築に失敗する。特にこれまでの対話内容と矛盾する発話を行うと、システムが対話内容を理解せず発話を行っていると感じられ、ユーザの満足度を低下させてしまう。こういった問題は、特に非タスク指向対話システムに見られる[1]。

本研究では、発話候補が発話時点までの対話履歴に矛盾するかどうかを識別し、矛盾する内容の発話を抑制する方法を検討する。その足がかりとして、矛盾した発話を対話破綻コーパス¹から抽出し、多層パーセプトロンによる識別器を用いて矛盾しているかどうかの識別を行う。また、その結果から、どのようにして矛盾した発話を行わないようにするかを検討する。

2 矛盾に関する関連研究

矛盾の検出は自然言語処理におけるタスクの一つであり、本研究で利用する文間の矛盾関係は次のように分類することができる。本節では、矛盾の分類を行った関連研究を挙げ、本研究でどの部分に着目するかを述べる。

亀田ら[2]は、矛盾を「論理的矛盾」、「感情的矛盾」、「その他の矛盾」の3つに分類した。東中ら[1]は、対話破綻の分類の一つとして矛盾を挙げている。また、対話破綻を文レベル、応答レベル、文脈レベル、環境レベルの4つに分類しており、矛盾と分類される破綻は主に応答レベルや文脈レベルで起こることが指摘されている。高畠ら[3]は、矛盾を同時矛盾と推移矛盾の2つに分類した。同時矛盾は2つの文が表す事象が同時に起こりえないという関係であり、推移矛盾は時間関係・順序関係のある2つの事象間における矛盾である。

本研究では、まず対話履歴のみを参照して識別可能な論理的矛盾に着目する。すなわち、亀田らの分類では論理的矛盾、東中らの分類では応答・文脈レベル、高畠らの分類では同時矛盾にあたる範囲を扱う。

*連絡先：奈良先端科学技術大学院大学情報科学研究科
奈良県生駒市高山町 8916-5
E-mail: sugiyama.kyoshiro.sc7@is.naist.jp

¹<https://sites.google.com/site/dialoguebreakdown-detection/chat-dialogue-corpus>

3 矛盾を考慮した発話選択

対話システムが対話履歴と矛盾しない発話を生成するため、発話候補を複数用意し、対話履歴との整合性を測り、その整合度合いと発話の適切さによって次の発話を選択するという手法が考えられる。本研究では、発話候補と対話履歴との整合性を測る識別器を作成するため、対話破綻コーパスから矛盾発話のデータセットを作成した。また、多層パーセプトロン (MLP) を用いた識別器を作成し、矛盾の識別に必要な要素や手法について検討した。

3.1 矛盾発話対データセット作成

東中ら [1] によって収集された対話破綻コーパスから、対話履歴と矛盾している発話を抽出し、矛盾発話のデータセットを作成する。対話破綻コーパスは 20 ターンの対話が 1146 対話収録されているが、その内矛盾発話が現れる 117 対話を対象とした。ここで矛盾発話とは、システム発話の内、アノテータのコメントに「矛盾」という単語が現れている発話とした。このアノテータのコメントは破綻アノテーションの理由として任意に付与されたもので、実際にはさらに多くの矛盾が含まれる可能性はある。117 対話に現れた各矛盾発話に対し、対話履歴中の各発話との対を取り、その発話対が矛盾しているかどうかの 2 値を 1 名でアノテーションした。このアノテーションにおいては、複数文を考慮すれば矛盾している場合であっても、各発話対が単独で矛盾していない場合は矛盾しないとした。また、「矛盾」というコメントがないシステム発話に対して、対話履歴中の各発話との対を取り、矛盾していない発話対とした。

3.2 多層パーセプトロンを用いた矛盾識別

3.1 節で作成したデータセットを用いて矛盾識別を行い、矛盾識別の課題や解決に必要な要素などを検討する。本実験では、モデルとして 3 層の MLP を使用した。入力素性は、履歴中の 1 文の Bag of words (BOW) と、発話文の BOW を用いた。データセットに含まれる 1872 発話対の内、8 割の 1497 発話対を学習に用い、残りの 375 発話対をテストに用いた。この識別結果を表 1 に示す。Precision は、識別器が矛盾であると識別した発話対が実際に矛盾している割合であり、Recall は存在する矛盾した発話対の内、矛盾であると正しく識別できた割合である。 F_1 は Precision と Recall の調和平均である。この結果から、人間のアノテータにとっては 1 文対 1 文を比較するだけで判断可能な矛盾であっても、BOW 素性を入力とする MLP では識別は困難であることがわかった。また、発話選択に用いることを

表 1: MLP を用いた識別結果

識別結果 \ 正解	矛盾しない	矛盾する
矛盾しない	333	27
矛盾する	4	11

$$Precision = 11/15 = 0.733$$

$$Recall = 11/38 = 0.289$$

$$F_1 = 0.415$$

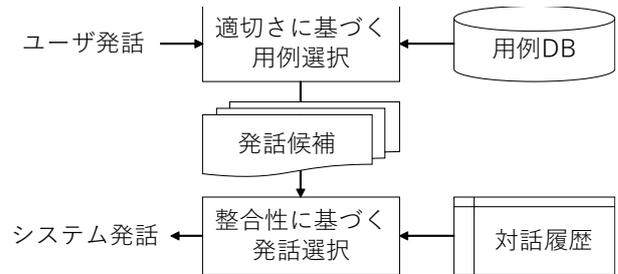


図 1: 対話履歴との矛盾を考慮した対話システム

考えると、矛盾しない発話を矛盾であると誤識別することよりも、矛盾する発話を矛盾しないと誤識別することの方が重大な問題である。言い換えれば、Recall が高い識別器を構築する必要がある。矛盾した発話を識別するためには、表層的な面だけではなく文の意味を捉えることや、外部知識の導入が必要になると考えられるため、今後はそれらについても検討を進める予定である。

4 対話システムにおける矛盾を考慮した発話選択

次に、対話システムにおける矛盾した発話の抑制をどのように実現するかを述べる。提案するシステムの概要を図 1 に示す。このシステムでは、まず通常の用例ベース対話システムの枠組みに沿って、ユーザ発話に対する適切さが高い応答をシステム発話候補とする。次に、各発話候補に対して、対話履歴との矛盾が生じていないか (整合性) を評価し、適切さと整合性の双方を考慮する評価尺度に基づいて最終的なシステム発話を決定する。これにより、ユーザ発話に対するシステム発話の適切さを大きく損なうことなく、対話履歴と矛盾する発話を抑制することが可能になると期待される。

5 まとめ

本研究では、対話システムがこれまでの対話内容と矛盾する発話をしてしまう問題を解決するため、矛盾の検出を行い対話システムで利用する方法について検

討した。具体的には、対話破綻コーパスを元に矛盾発話データセットを作成し、MLP を用いて発話の矛盾を識別した。実験の結果から、1 文対 1 文で判断可能な矛盾であっても、BOW のみを入力とする MLP では識別が難しいことがわかった。加えて、実際の矛盾発話の中には対話履歴中の複数の文を参照しなければ判断できないものや、外部知識が必要なものが含まれており、今後は複数文の意味や外部知識を考慮するための素性や手法が必要であると考えられる。また、今回は対話破綻コーパス中のコメントを利用して矛盾発話を抽出したが、矛盾という単語がコメントに現れていなくても矛盾していると考えられる発話も見られ、より正確にデータセットを作成する必要がある。また、矛盾の識別は含意関係認識のタスクの一部にも含まれており、RITE[4] のようなシェアドタスクで用いられているデータセットなども利用することを検討している。

参考文献

- [1] 東中竜一郎, 船越孝太郎ほか. Project Next NLP 対話タスクにおける雑談対話データの収集と対話破綻アノテーション. *SIG-SLUD*, Vol. 4, No. 02, pp. 45–50, 2014.
- [2] 亀田弘之, 向後英二. 矛盾の処理に関する基礎的考察 (その 1): 分類・分析・根拠・解消法について. *電子情報通信学会ソサイエティ大会講演論文集*, Vol. 1998, p. 160, 1998.
- [3] 高畠悠, 森田一, 河原大輔, 黒橋禎夫, 東中竜一郎, 松尾義博. クラウドソーシングを活用した事態間矛盾の分析と分類, 2015.
- [4] Hideki Shima, Hiroshi Kanayama, Cheng-Wei Lee, Chuan-Jie Lin, Teruko Mitamura, Yusuke Miyao, Shuming Shi, and Koichi Takeda. Overview of ntcir-9 rite: Recognizing inference in text. In *NT-CIR*, 2011.