

雑話システムにおけるユーザの趣味を 考慮した発話手法に関する検討

宅和 晃志[†] 吉川 大弘[†] ジメネス フェリックス[†] 古橋 武[†] (吉川 大弘)
[†] 名古屋大学大学院 工学研究科

1. はじめに

近年、人間との自然な対話を目指した非タスク指向型対話システム(雑話システム)が注目されている。雑話システムが実際に使用される場面として、独居高齢者の話し相手や教育支援のための学習者パートナーロボットなどが想定される。このような一人の個人と長期的、かつ親密に対話をするような状況においては、自然な対話を行うことは当然ながら、ユーザに飽きを感じさせない多様な発話と、ユーザに親密感を与えるような魅力溢れる発話を両立させる必要がある。そこで本稿では、ユーザの趣味を考慮した発話を行うことでユーザに親密感を与えられるとの仮定のもと、ユーザの趣味に関連した多様な発話文を生成する手法について検討する。

2. Twitter より発話文を取得する手法

多様な発話文の生成手法として、Twitter を用いる手法が提案されている[1][2]。これらの手法では、大量のツイートを取得しておき、フィルタリングと機械学習的アプローチにより、実際に使用可能な発話文のみを取得することができる。本稿でも、これらの手法をベースに Twitter より発話文を取得する。しかし、Twitter 全体を対象に発話文を取得し、趣味に関連する発話文を選択すると、下記のような期待しない発話文までもが大量に取得されてしまう。ここで、本稿で取得したい発話文は、既知であることを前提とした文脈や事実を含まず、特定の話題を含む発話文である。例えば、「読書量と年齢が釣り合わない人がいますよね」という発話文は、「読書」や「読書量」という話題に対して使用可能であるが、「年齢」という話題には適切ではない。

3. 趣味に関する発話文を取得する手法

Twitterを始めたSNSは、投稿者の近況報告や他ユーザとの交流のために主に使用される。そのため、Twitter 全体を対象に発話文を取得すると、上記の期待しない発話文までもが大量に取得される。そこで、「ボット」と「あるある検索」を使用する方法を検討する。

3.1 「ボット」を使用する方法 この方法では、ユーザ名が「趣味名+bot」であるユーザを検索し、それらのユーザのツイートを発話文として取得する。例えば「バイクbot」で検索すると、5 ユーザほどヒットする。これらのボットはそれぞれ趣味であるバイクに関するツイートのみを行っている。そのため、その趣味を

もつ人にとって興味を抱く発話を効率的に取得できると考えられる。

3.2 「あるある検索」を使用する方法 この方法では、ツイート全体において「#+趣味名+あるある」で検索を行い、そのツイートを発話文として取得する。例として、「#読書あるある」で検索を行った結果を図1に示す。読書する人にとって「確かにこういうことあるよね」と共感できるツイートがなされていることがわかる。そのため、その趣味をもつ人にとって共感できる発話を効率的に取得できると考えられる。

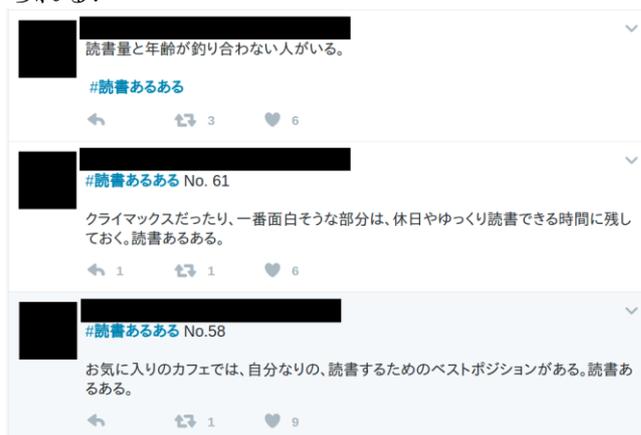


図1. あるある検索結果の例

4. 今後の課題

「ボット」と「あるある検索」によって発話文を取得するだけでは、まだ期待しない発話文が紛れる可能性があるため、それらの除去方法を検討していく予定である。また、取得した発話文の使用方法についても検討する必要がある。最終的には、本手法を対話システムに実装し、被験者実験等によりその有効性の検証を行っていく予定である。

参考文献

- [1] 稲葉通将, 神園彩香, and 高橋健一. "Twitter を用いた非タスク指向型対話システムのための発話候補文獲得." 人工知能学会論文誌 29.1 (2014): 21-31.
- [2] Higashinaka, Ryuichiro, et al. "Syntactic filtering and content-based retrieval of Twitter sentences for the generation of system utterances in dialogue systems." Situated Dialog in Speech-Based Human-Computer Interaction. Springer International Publishing, 2016. 15-26.