

## 回答文をアノテートするエージェントの提案 ～チャットボットを含むQAサイトの分析～

森 愛理花 遠山 紗矢香 竹内 勇剛

静岡大学 総合科学技術研究科

gs18053@s.inf.shizuoka.ac.jp , {tohyama , takeuchi }@inf.shizuoka.ac.jp

### 1 はじめに

近年, Siri などの音声エージェントシステムが普及している. また, 高齢世帯や独居世帯の増加に伴い, ユーザと雑談が出来るシステムのニーズも増加している[1]. そして, ロボットに対して道具としての機能とともに共存者としての役割も求めることが示唆されている. しかし現状インタラクションが円滑に行われているとはいえない. そのため人-エージェント間にインタラクションの不自然さを解消する必要がある.

チャットボットとのコミュニケーションを行う上でチャットボットが抱える問題に, 表面的にはインタラクションが成立しているようにみえても, 意図や文脈の理解あるいは共有がされておらず, 適切なやりとりが行えないというものがある. これらの課題を解決するための研究は数多く行われている. 本節では, 多面的な評価手法の提案を行っている研究のみについて述べる.

対話の破綻に焦点を絞った研究では, 雑談対話の収集と対話破綻アノテーションとその破綻原因の類型化について述べ, 単語情報のみで破綻を検出する手法の限界を示唆している[2, 3]. また, 人-ロボット間相互作用にかかわる心理学的評価[4]や, QA サイトにおける質問の多面的評価[5]がある. しかし, これらの研究では評価者がどのような印象をもったのかについての検討は行われていない.

本研究では, 人が対話において相手にどのような印象を抱いているかを明らかにするために, 人/チャットボットの判別テストを実施した. その結果を踏まえ, 新たなインタラクションモデルの提

案を行う.

なお, 本研究では, 対話データとして, 教えて! goo に投稿されたテキスト対話ログを用いている.

### 2 背景

#### 2.1 QA サイト

QA サイトとは, あるユーザが匿名で投稿した質問に対して他のユーザが回答を行うことで成立しているコミュニティサービスである. 様々な知識を多くの人々で共有することやその場でコミュニケーションをとることが目的となっている.

質問と回答をデータとして内容的あるいは数量的に分析し, 質問に対して適切な回答を得るために利用することを目的としている研究では, QA サイトにおけるベストアンサーを計算機が推定可能か検証したものや[6]質問に対して適切な回答を得るために考慮する要素として質問のタイプの提案をおこなった研究がある[7].

また2016年9月よりコミュニティサービス教えて! goo にて恋愛相談 AI「オシエル」として, チャットボットが回答を行うサービスが開始された.

#### 2.2 チャットボット

チャットボットとは, 会話を提供する人工知能で, 自然な会話を目指すものとしてりんな[8]やオシエル[9]などがある.

「オシエル」はユーザの投稿する質問に回答をしており, 2018年現在約3万件の回答を行っている. 回答の特徴としては, 質問への回答の投稿時間が1分以内であること, 自己紹介, 偉人の言葉の引用などがある. また, ユーザは「オシエル」からの投稿以外に他のユーザからも回答を受け取

ることができる。

本研究では、実際にチャットボットと人がやりとりを行っている「教えて！goo」を対象に分析を実施する。

### 3 方法

教えて goo 内で運用されているチャットボット（オシエル）と人のやり取りをデータベースとして、人/チャットボットの判別テストとして以下の2件を実施した。

- ① 質問と回答のやりとりを提示
- ② 回答文のみを提示

そして、これらの結果を用いて人/チャットボットの特徴と違いを探る。

#### 3.1 データの収集

**【対象】** 教えて goo! の恋愛相談カテゴリから人工知能オシエルが回答した質問と質問に寄せられた回答のすべて

**【時期】** 2016年10月23日～2016年10月29日

**【手続き】** 教えて goo! でチャットボットが恋愛相談カテゴリ内で回答した質問文とそれに寄せられたユーザの回答を2016年10月23日から2016年10月29日の7日間のデータを記録した。

#### 3.2 質問紙の構成

**【データの抽出方法】** 収集したデータは質問151件、回答586件であった。収集したデータからチャットボットの回答のみ寄せられた質問とチャットボットの回答を特定する要因として投稿時間・名前・偉人の言葉の引用・自己紹介文などを削除した。これに合わせて人（ユーザー）の回答からも同項目を削除した。そして、チャットボットを特定する要因を削除したデータからランダムに抽出したデータで作成した質問紙で予備調査を実施した。予備調査により回答文の文字数が多い、一つの質問に対して寄せられた回答数が多いと回答者の負担になるという意見が出たため回答の文字数と回答数も考慮した。文字数は回答586件の文字数の平均を算出し414字以上の回答文はデータから削除

した。回答文は、回答者の負担を減らし読みやすくするため4件以上のものを削除した。本調査では、チャットボットを特定する要因と回答者の負担となる要因を削除したデータから質問紙を作成した。

#### 3.3 質問紙調査

##### ① 質問と回答のやりとりを提示

**【器具】** 質問紙を作成し用いた。質問紙30部作成し、ナンバリングを行った。

**【調査回答者】** 大学生の男女30名。

**【時期】** 2019年1月7日から2019年1月11日

##### ② 回答文のみを提示

**【器具】** 質問紙を作成し用いた。質問紙30部作成し、ナンバリングを行った。

**【調査回答者】** 大学生の男女30名。

**【時期】** 2018年11月26日～2018年12月26日

#### 3.4 質問紙調査の結果と考察

回答文の正答率と人、チャットボット、全体との比較を表1に示す。いずれの条件においても、正答率は85%を超えており、人はチャットボットと人の文章の違いをある程度判断できたといえる。

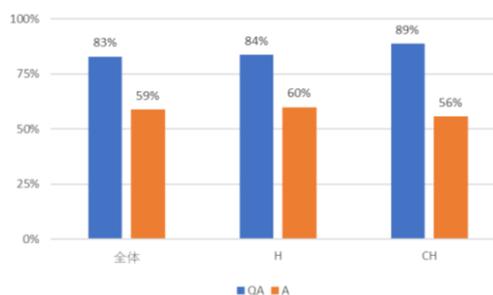


図1 正答率(正解数/回答者数)

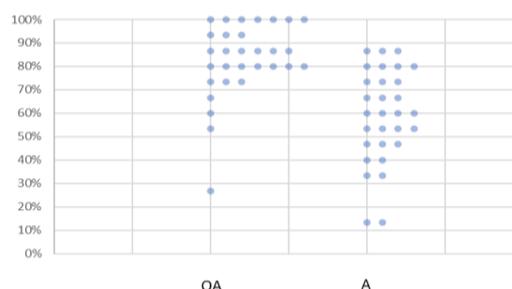


図2 回答者ごとの正答率(正解数/問題数)

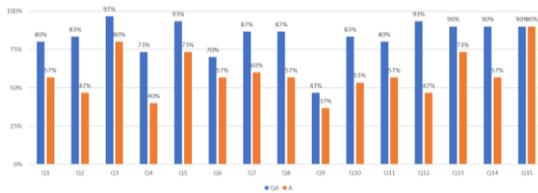


図3 回答ごとの正答率(正解数/問題数)

#### 4 提案手法

エージェントが人同士のコミュニケーションに参加するために従来の研究では、1体のエージェントが人とどのようにしてかかわるのかという点に焦点が当てられているものが多い。これに対し本研究では、1体のエージェントが人と与えられた役割の機能を分担することで、多人数会話における参与の枠に介入することが可能であるかどうかを検討する。

対象とする多人数会話の場を、会議や議論など問題解決をおこなう対話場面に限定する。また、集団の中で共有されている仕事のことを「役割」とし、役割を満たすために必要な条件を「機能」と定義した。1体のエージェントと人を組み合わせて、1つの役割に統合することを本研究では、「ニコイチ」と呼ぶ。ニコイチとは、本来、自動車業界などで複数の個体から1つの個体を作り出す修理・製造方法であり、日常では、親しい友人のことをニコイチと呼称する場面で用いられる。本研究のアプローチの比較を図4に示す。

多人数会話において、その場にいる人間の役割は、参与者と傍参与者に分かれる。参与者は話し手と聞き手という役割があり、傍参与者は会話に参加する自由度が存在し、会話に参加する以外にも会話場面の制御といったファシリテーションの役割を担う場合もある。しかし、このファシリテーションという役割を人間が行う場合、本来の機能を十分に発揮できなくなってしまうことがある。

物も言いようで角が立つという、ことわざが示すように、人に何かを伝える場合、言い方によっては、相手の気持ちを不快にさせたり、怒らせたりす

ることがある。これを日常場面では、意見をもつていても場の空気によって発言できない場、隣の席に上司がいるため遠慮して発言を控えようと思ってしまうものや長時間話し合っても結論のないまま終了したりすることがある。このような人間だと社会的身分が関係に影響して十分に機能を発揮できないという場面が存在する。

場の空気によって人が発言できないという場面において、エージェントによって、人が発しにくいリアクションを行うことで人同士の対話に介入をする。エージェントに人の発揮し辛い機能(F)を補助させることで役割の機能の一部を担い、ニコイチでファシリテーションの役割を満たすことを想定している。ただし、エージェントの行う人の発揮し辛い機能の補助は、人の機能(F1)をそのまま置き換えたものではなくエージェントがもつ能力の中で置き換えたものになる。そのため、人が道具的にエージェントを活用するのではなく自立的に人の補助を行えると考えている。

本研究のアプローチを行うことによって、人同士の対話場面にエージェントが介入することが可能となり、対話を補助することで人同士の対話を円滑に維持することが望まれる。

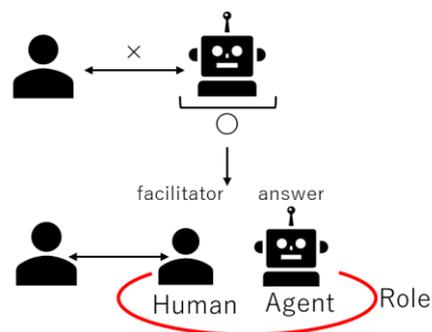


図4 提案手法

#### 参考文献

- [1] 柴田崇徳：癒し系ロボットとソフトマテリアル，日本ゴム協会誌，第78巻，第8号，pp.313-320，2005
- [2] 東中竜一郎，船越孝太郎，荒木雅弘，塚原裕

- 史, 小林優佳, & 水上雅博. (2015). Project Next NLP 対話タスク: 雑談対話データの収集と対話破綻アノテーションおよびその類型化. 言語処理学会大 21 回年次大会ワークショップ論文集.
- [3] 芝原優真, 中野幹生, & 山本一公. (2015). 単語情報に基づく雑談対話における破綻の検出の検討. 研究報告自然言語処理 (NL), 2015(8), 1-7.
- [4] 神田崇行, 石黒浩, 石田亨: 人間 - ロボット間相互作用にかかわる心理学的評価; 日本ロボット学会誌, vol. 19, no. 3, pp362-371 (2001)
- [5] 渡邊直人, et al. コミュニティ QA における質問の多面的評価法の検討. 情報知識学会誌, 2011, 21. 2: 163-168.
- [6] 石川, et al. QA サイトにおけるベストアンサー推定の分析とその機械学習への応用. 情報知識学会誌, 2010, 20. 2: 73-85.
- [7] 栗山和子, et al. QA サイトにおける質問と回答の分析. 研究報告データベースシステム (DBS), 2009, 2009. 19: 1-8.
- [8] Wu, X., Ito, K., Iida, K., Tsuboi, K., & Klyen, M. (2016). りんな: 女子高生人工知能. 言語処理学会第 22 回年次大会発表論文集, 306-309.
- [9] 中辻真, et al. ディープラーニング活用事例と使いこなしの勘所:[言語処理分野] 4. AI による恋愛相談への回答生成-答えのない回答生成への試み. 情報処理, 2018, 59. 11: 978-982.