

## 質問への応答に依存しない傾聴対話システムのリアクション生成

青木 裕哉†

伊島 翔大†

中野 大樹†

藤江 真也†

†千葉工業大学 未来ロボティクス学科

## 1 はじめに

傾聴対話ロボットの振る舞いについて検討する。傾聴対話とは、高齢者施設などで高齢者の話を引き出しながら進める対話である。本研究では特に引き出した発話に対するロボットのリアクションについて検討する。

従来の研究では、言語処理結果を用いた応答生成手法が提案されている [1]。高齢者の発話については、音声認識の精度が必ずしも十分でない。そのため、正しい認識結果が得られないことを想定した検討が必要である。

人と人が対話しているとき、人は相手の言葉を聞き取れなくても大きく誤った応答を行うことはない。これは、直前の質問に対する相手の応答がある程度予想されるため、文脈的に違和感のない応答を行えるからである。本研究では質問ごとに適切なリアクションをしばりこみ、文脈に沿ったリアクションを生成する手法を提案する。提案手法と音声認識結果を利用する手法とを比較して、印象にどのような変化があるかを調査する。

## 2 事前調査

## 2.1 傾聴ボランティアのリアクション分析

ロボットが行うリアクションを選定するため、高齢者施設における傾聴ボランティアと高齢者の対話の様子を収録し、従来研究 [1] などを参考にしながら、ボランティアの発話からリアクションを抽出し分類した。傾聴ボランティアは同一人物で、高齢者がそれぞれ異なる 10 対話からデータを分析した。従来研究 [1] で相槌 2 回（「うんうん」など）を応答の対象としていたことから、それより多く行われていたものを抽出した。

その中から対話中に現れた頻度が高く、ロボットでの表現が可能なものである、「理解」「共感」「驚き」を用いる本研究で用いるリアクションとした。また、対話中に複数回行われていた質問を実験に用いる質問として採用した。ただし、文脈に関係なく回答できるものに限った。具体的には、好きな食べ物や、天気についての質問などである。

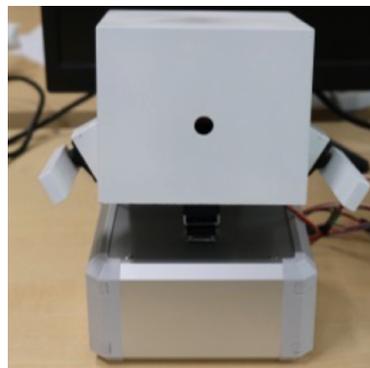


図 1: 卓上対話ロボット TOFU

## 2.2 対話ロボット

本研究では卓上対話ロボット TOFU を用いる。TOFU は図 1 のような外見である。首にロール、ピッチ、ヨー角の 3 自由度と、左右腕部にそれぞれ 1 自由度を持ち、それぞれを動作させることで発話に伴うジェスチャなどを生成できる。

## 2.3 WOZ 法による高齢者の振る舞いの調査

ロボットと対話したときの高齢者の振る舞いの変化と、選定したリアクションで十分に傾聴対話が行えるかを調査するため、高齢者施設で WOZ 法によるロボット対高齢者の対話実験を行った。高齢者にはロボットが自動で動くことと伝えた。実際には、実験者がボタンを押すことにより、予め用意された質問を発話したり、リアクションを生成したりする。また、実験者がタイピングすることで任意の内容の発話を生成することもできる。ロボットからの質問、高齢者の回答、ロボットのリアクションという流れを、用意した全ての質問について終えた時点に対話の終了とした。

実験の結果、被験者がロボットに向かい大声で話しかけたり、人でも内容が聞き取れない発話もあり、音声認識結果が正しく得られる例はほとんどなかった。そのため、言語情報を用いることは難しいということが確かめられた。また、実験者の感想として、ロボットのリアクションに「驚き」の強弱と「すごい」などの「賞賛」があると良いという意見があった。

Generation of reaction independent of an answer to a system question  
in active listening dialog system  
†Yuya AOKI †Shota IJIMA †Daiki NAKANO †Shinya FUJIE  
†Chiba Institute of Technology

R: 行ってみたい場所はある?  
 (音声为正しく認識された場合)  
 U: ロンドンです  
 R: そうなんですね. どうしてですか?  
 (音声を誤認識した場合)  
 U: ロンドンです(「ボンドです」と誤認識)  
 R: 誰のボンドですか?

図 2: 音声認識を用いたシステムのリアクション例

### 3 質問の内容に依存しないリアクション生成

#### 3.1 質問とリアクションの対応づけ

システムのリアクションは, 基本的にはシステムの質問に対する回答をもとに行われる. しかし, 音声認識により回答を理解するシステムでは, 図 2 の例のように, 誤認識や誤理解によりリアクションを大きく誤ることがある. 一方, この例においては, 「そうなんですね」などと理解を示すリアクションであれば, 回答の内容にかかわらず大きな誤りとはならない. 各質問に対して, このようなリアクションを用意できれば, 音声認識に頼らず適切なリアクションを生成できる. しかし, どの質問にどのリアクションが適切かは自明ではないため, 人手で適切かどうかを判断する.

具体的には, 10 通りの質問に対して, 想定される回答を 3 つずつ用意し, 各回答に対して 5 つのリアクション(理解, 弱い驚き, 強い驚き, 賞賛, 共感)がふさわしいかについて, アンケートをとった. ふさわしさについては(そのリアクションをしてくれると)嬉しい, 構わない, 攻撃的に感じる, 文脈的に違和感を感じる, の中から選ばせた. 5 名の作業員から得られた結果について重み付き kappa 値を求めた結果, 考えられる全ての組み合わせで中程度の一致が見られた.

#### 3.2 システムへの実装

前節のアンケートの結果において, 想定した回答がどの場合であっても攻撃的に感じず, 文脈的に違和感のないリアクションを採用し, その質問とリアクションを生成するシステムを図 1 のロボットに実装した. システムの質問にユーザが回答したかどうかの検出のために音声認識を用いるが, 認識された内容は利用せず, リアクションの内容は質問にのみ依存して決定する.

## 4 評価実験

#### 4.1 実験条件

提案手法を実装したロボットと, 音声認識を利用する従来手法 [2] を実装したロボットとを, 実際の対話を通した主観評価により比較する. 対話では, ロボットが被験者に対し質問を行い, 被験者が回答をし, ロ

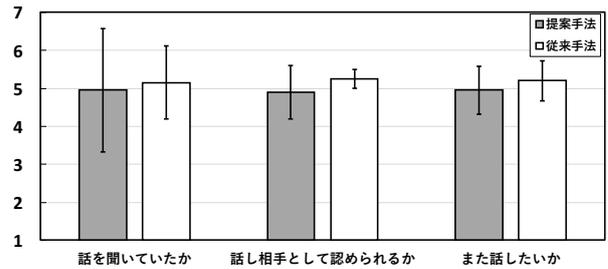


図 3: 対話実験の評価結果

ボットがリアクションを行う, ということを繰り返す. これを用意した質問全てを行うまで続ける. 1 被験者に対して従来手法と提案手法の 2 通りの対話を行う. 各対話終了後, 被験者は「ロボットは話を聞いていたか」「ロボットは話し相手として認められるか」「また話したいか」の項目を 7 段階で評価する. 被験者は 20 代の男女 20 名である.

#### 4.2 結果

実験結果を図 3 に示す. t 検定の結果, どの項目に関しても提案手法と従来手法に有意な差は見られなかった.

概ね想定通りの動作をしていたが, 各手法において数例ずつ極端に悪印象を持たれた対話があった. 提案手法においては, 将来の夢に関する質問への「特にないです」という回答に対して, 「すごい」と発話した例があった. 従来手法においては, 音声認識に失敗し応答を生成できない例や, 認識結果が正しいにもかかわらず文脈的に整合していない応答をする例があった.

各手法における悪印象の原因はそれぞれ異なり, 相補的である. 従来手法において音声認識の信頼度が低い場合に提案手法によるリアクション生成を行うといったように, 両手法を組み合わせることでより印象の良いシステムを構成できる可能性がある.

## 5 まとめ

質問に対する回答の内容に依存しないリアクション生成方法を提案し, 対話ロボットに実装し, 実験を行った.

今後の課題としては, 今回人手で行ったリアクションの絞り込みを, 言語理解による矛盾検出などを応用して自動化することが挙げられる.

## 参考文献

- [1] 山口貴史, et al.: 傾聴対話システムのための言語情報と韻律情報に基づく多様な形態の相槌の生成, 人工知能学会論文誌 31(4), C-G31-1 (2016).
- [2] 石田真也, et al.: 傾聴対話システムのための発話を促す聞き手応答の生成, 人工知能学会研究会資料 SIG-SLUD-B504, pp. 1-6 (2016).