

## 共感に着目した音声対話に基づく問診システムの構築

## Development of Medical Interview System Based on Spoken Dialogue Focusing on Empathy

山本雄樹  
Yuki Yamamoto西田昌史  
Masafumi Nishida

## 1. はじめに

体調不良を自覚しながらも、面倒に感じて受診を怠る人は多い。中には重篤な状態に陥るケースが存在する。このケースを回避するには、症状が重篤になる前に受診の必要性を促すことが有効だと考えられる。

受診を促す手法の一つとしてテキスト対話型の問診システムが挙げられる。Ubic 株式会社提供「症状検索エンジンユビー」[1]や株式会社ギミック提供の「Doctors File」[2]などがそれに該当する。これらのシステムは患者の体調から、疑いのある疾病や受診の必要性を提示することができる。

ただし、使用者が電子機器を操作する必要があるサービスは高齢者にとって使いにくい。高齢者が問診システムを気軽に使用するためには、対話を音声で行うことが有効だと考える。また音声による対話は認知症予防にもつながる可能性がある。

近年、音声対話にかかわる研究は多く行われている。東中ら[3]は人前で実際に対話システムを動作させ評価を行う対話システムライブコンペティションを毎年行っている。また武田ら[4]はコンペティション参加時の経験を踏まえ、対話システムの設計方針についてまとめている。井上ら[5]は人間同士の会話で行われる相槌や共感、ターンテイキングなどの特徴をアンドロイドで再現することでより人間に近い音声対話を行う手法を提案している。また、高津ら[6]はニュース記事のようなまとまりのある情報を会話によって伝える対話システムを提案している。高齢者の認知症予防を目的とする対話システムとして、下岡ら[7]は音声対話ロボットのための傾聴システムの開発を行った。ユーザ発話の不足格などに対しての繰り返し、問い返し応答や、ユーザ発話の感情推定により共感応答や相槌応答を生成するアルゴリズムをシステムに組み込み、評価実験を行った。

しかしながら、シチュエーションが問診である音声対話システムの研究はあまり行われていない。問診は他の会話と比較して、ユーザとの信頼関係を築くことがより重要である。そのため、従来の対話システムとは異なるアプローチが必要だと考えられる。このような背景のもと、本研究ではユーザとの信頼関係を築くことを目的として、共感に着目した音声対話に基づく問診システムの提案を行う。

## 2. 従来システム

まず、提案システムの前身である、従来システムについて説明する。従来システムは株式会社シルバコンパスとの共同研究プロジェクトの一環として開発された。

従来システムではユーザから聞いた情報をもとに受診の必要性

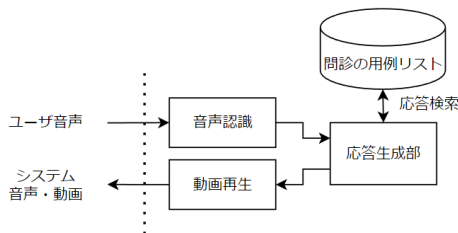


図 1 従来システムの構成

や、どの科を受診すべきか提案する。なお、病気の特定を行うことは医療行為に該当するため、行わない。

従来システムの構成図を図 1 に示す。従来システムでは音声認識に Google Cloud Platform の Speech-To-Text API を使用しユーザの発話を文字として認識する。患者と医師のリアルな対話を再現するため、システムの出力はあらかじめ撮影した映像を再生する手法をとっている。応答生成には用例ベースを採用している。Elasticsearch に問診の対話データを格納し、これを参照してユーザへの応答を選択する。対話データとユーザの発話とをコサイン類似度を用いて比較し、最も類似度の高かったデータの応答部分をシステムの発話として出力する。

## 3. 提案システム

提案システムではユーザとシステムの信頼関係を築くことを目標として、共感に着目した対話システムの構築を行った。

提案システムの構成図を図 2 に示す。状況に応じた多様な発話を行うために、システムの発話には合成音声を使用した。ユーザの発話内容に対して共感を示す振る舞いを行うことでユーザとの信頼関係の構築を図る。音声認識には従来システムと同じ手法を、音声合成には日本語の音声合成を行うソフトウェアである Open JTalk を使用した。

対話制御には用例ベースに加えて、フレームというデータ構造を用いた状態遷移の 2 つを併用して対話制御を行う。対話データは従来のものに共感を示す発話を追加している。

フレームを用いた状態遷移では症状に関する質問を行う対話の生成を行う。具体的には「お腹が痛いです」に対して「お腹のどのあたり？」のように尋ねる対話である。用例ベースではどの情報についてユーザが話したかを管理し対応することは難しい。しかしフレームを用いることで 図 3 に示すように「(お腹の) 下の方」「下腹部」などのキーワードを認識し、柔軟に対話を行うことができる。

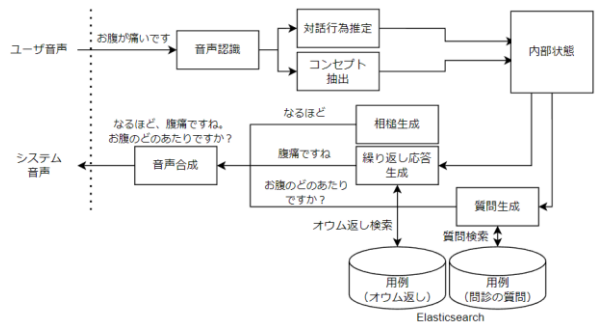


図 2 提案システム構成図

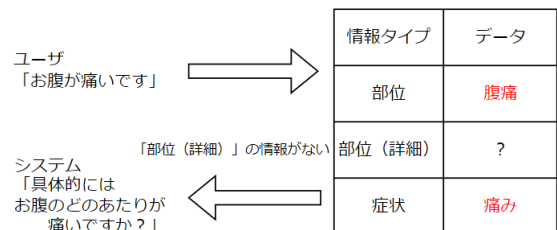


図 3 フレームを用いた状態遷移

本システムはキーワードの認識に系列ラベリングの手法である CRF を利用する。

次に問診システムが生成する発話内容について説明する。本研究ではユーザとの信頼関係を構築するのに有用な発話を優先して採用している。まず相槌の生成についてである。ユーザの発話に対しての応答で「そうですね」や「なるほど」というように相槌をシステム発話の初めに出力する。相槌の形態はその都度ランダムに選出する。次に繰り返しについてである。繰り返しはユーザが語った内容を、「そのまま」「要約して」「言い換えて」繰り返しすることを指す。ユーザの発話が「お腹が痛いです」ならば「お腹が痛いのですね」や「腹痛ですね」などと応答する。

#### 4. 評価実験

提案システムについて評価実験を行った。実験の詳細を説明する。従来システムと提案システムでは、出力が映像から音声のみになることに加え、応答生成の手法を変更した。これらのシステムで比較実験を行う場合に、評価値の違いがどの要因によるものか明らかでない。そのため、両システムの間位置するシステムが必要である。そこで従来システムの出力を映像から音声のみに変更した比較用システムを用意した。

被験者は 9 名である。被験者は各システムと 3 回ずつ、与えられた異なるシナリオに沿って対話を行い、システムごとにアンケートに回答する。実験で使用するシナリオの例を図 4 に示す。

アンケートは以下 9 つの項目について 7 段階評価を行った。

- ① 対話の内容が自然だと感じましたか？
- ② 対話の内容が適切だと感じましたか？
- ③ 対話に満足していますか？
- ④ 症状を伝えられたと感じましたか？
- ⑤ 問診がスムーズだったと感じましたか？
- ⑥ 信頼できる問診システムだと感じましたか？
- ⑦ システムが共感していると感じましたか？
- ⑧ 問診に満足していますか？
- ⑨ もう一度対話したいと感じましたか？

実験結果について説明する。対話についてシステムごとに平均ターン数、タスク達成率、音声認識率について集計したものを表 2 に、アンケートの結果をグラフにしたものを図 5 に示す。

まず対話実験の結果について考察する。平均ターン数と音声認識率についてはあまり差異が見られなかった。タスク達成率は提案システムが他と比べて、7%以上良い結果を得た。これは提案システムに対話行為推定とコンセプト抽出を取り入れたことにより、対応できる発話の幅が広がったのだと考えられる。

次にアンケートについて考察する。システム間で有意な差が見られたものは②、⑤、⑦である。②について、提案システムが他より

部位	頭
部位 (詳細)	-
症状	痛み
程度 (5段階評価, 1: min, 5: max)	5 (max)
発熱	なし
備考	・ 頭をぶつけるなどの外傷はなし

図 4 実験で使用したシナリオ例

表 1 対話実験の結果

	従来システム	比較用システム	提案システム
平均ターン数	4.4回	4.4回	4.6回
タスク達成率	67%	63%	74%
音声認識率	86%	86%	90%

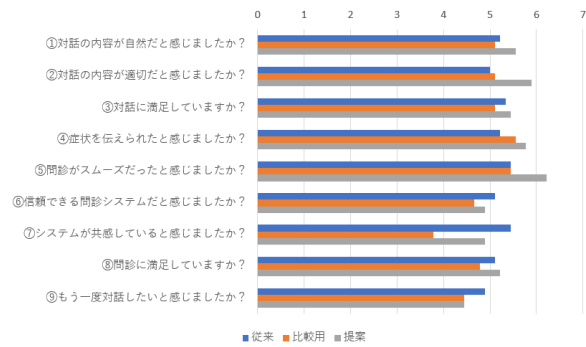


図 5 アンケート結果

約 0.7 高い結果となった。これは提案システムがフレームを用いて適切に応答できたことに加え、オウム返しにより、システムが症状を理解したと分かっているためにシステムの応答が尤もらしく聞こえたのではと考えられる。⑤については提案システムが他システムより約 0.8 高くなった。これは提案システムが再質問の回避により、効率良く問診を行うことができたことを示している。⑦では従来システムが 5.4、提案システムが 4.9 と、比較用システムの 3.7 を大きく上回る結果となった。従来システムは唯一出力に動画を用いており、ユーザに返される情報量が多いため、共感されていると感じたのだと考えられる。提案システムは他システムと比較して、発話に「そうですね」や「頭が痛いのですね」等、相槌やオウム返しを多く行う。これらの振る舞いにより、ユーザが共感されたと感じたものと思われる。

これらの結果から、システム応答が音声だけでも十分に共感した対話の実現できると考えられる。

#### 5. おわりに

システムの出力を動画から音声のみに変更するとともに、システム発話に共感を含む振る舞いを行う音声対話型問診システムを提案した。結論として提案システムは従来システムと比較しても遜色ない結果となった。つまり提案システムにより、問診の品質を維持しながら動画撮影のコストを削減することができた。

一方で人がいたほうが話しやすい、フィードバックが音声のみだと不安になる等のコメントを被験者の方からいただいた。今回は実装したシステムは出力が音声のみであったが、CG エージェントを用いる等、音声以外でもユーザにフィードバックを行えるよう、システムの拡張を進める予定である。

#### 参考文献

- [1] 症状検索エンジンユビエ, <https://ubie.app/>
- [2] Doctors File, <https://doctorsfile.jp/>
- [3] 東中 竜一郎, 高橋 哲朗, 堀内 颯太, 稲葉 通将, 佐藤 志貴, 船越 孝太郎, 小室 允人, 西川 寛之, 宇佐美 まゆみ, 港 隆史, 境 くりま, 船山 智, 対話システムライブコンペティション 5, 人工知能学会研究会資料 言語・音声理解と対話処理研究会, 96 巻, 96 回 (2022/12), p. 93-100, 2022.
- [4] 武田 龍, 駒谷 和範, 中島 圭祐, 中野 幹生, 複数の対話システムコンペティションにおけるシステム開発の設計指針, 人工知能学会論文誌, 37 巻, 3 号, p. IDS-B\_1-9, 2022.
- [5] 井上 昂治, ラーラー ディベッシュ, 山本 賢太, 中村 静, 高梨 克也, 河原 達也, アンドロイド ERICA の傾聴対話システム-人間による傾聴との比較評価-, 人工知能学会論文誌, 36 巻, 5 号, p. H-L51\_1-12, 2021.
- [6] 高津 弘明, 福岡 維新, 藤江 真也, 林 良彦, 小林 哲則, 意図性の異なる多様な情報行動を可能とする音声対話システム, 人工知能学会論文誌, 33 巻, 1 号, p. DSH-C\_1-24, 2018.
- [7] 下岡 和也, 徳久 良子, 吉村 貴克, 星野 博之, 渡部 生聖, 音声対話ロボットのための傾聴システムの開発, 自然言語処理, 24 巻, 1 号, p. 3-47, 2017.