

会話エージェントを利用した認知症患者のためのコミュニケーション支援

比企野 純太[†] 中野 有紀子[†] 安田 清[‡]

成蹊大学理工学部情報科学科[†] 千葉労災病院リハビリテーション科[‡]

1 はじめに

近年、我が国における認知症患者は年々増加傾向にあり、患者数に対して介護者が全く足りておらず、認知症患者に対する支援はまだ十分とは言えない。また、けがや疾病の治療のために入院する認知症患者も多く、病院内でのケアが必要となり、看護師の負担が大きくなっている。

本研究では、アニメーションキャラクターを用いた会話エージェントが非課題遂行対話を通して患者に対応することにより、認知症患者の精神安定、介護者の負担軽減を目指している。本稿では、認知症患者のための話し相手となるエージェントの設計指針を得るために、まず、Wizard-of-Oz 法による予備実験を行った。次に、そこでの観察にもとづき、相槌や頷きを返しながらユーザの発話を促す機能を持つ会話エージェントを実装した。

2 予備実験

2.1 WoZ 実験によるデータ収集

認知症患者と対話する会話エージェントを作成するにあたり、認知症患者に対して非課題遂行対話を行うエージェントが有効であるのか、エージェントのアニメーションデザインが妥当であるか、患者はエージェントとのインタラクションを快く受け入れてくれるか等を確認しておくことが重要となる。そこで、実験者がエージェントの応答を制御する WoZ 実験を行うことにより、エージェントと認知症患者との理想的なインタラクションのあり方を検討する基礎データを収集した。

実験材料: 被験者に尋ねる質問は、千葉労災病院で認知症患者の介護にあたっているボランティアの方と認知症患者との対話の書き起こしから抽出した。表 1 に質問の例を示す。システムからの全質問が終わったところで、実験は終了となる。別室より患者の様子や発話をモニターし、実験者が WoZ システムを操作して、エージェントの応答を選択した。聴力に不安のある被験者にはイヤホンを使用していただいた。

WoZ システム: エージェントアニメーションは ADOBE

表 1: エージェント発話例

今、体の具合はどうですか。
ここに来てどれくらい経ちますか。
ご飯は食べてますか。
よかったです。
頑張りましょうね。
わかりました。
じゃあ今度歩く練習をしましょうね。

FLASH CS4 PROFESSIONAL バージョン 10.0 を用いて作成し、ブラウザ上で実行した(図 1)。動作選択を行う GUI は Java で実装した。瞬き等の待機モーションと発話中のリップシンクのアニメーションは自動で実行し、相手の発話中に行う頷き動作と発話音声の再生は GUI により実験者が操作した。エージェントの音声は合成音声器 Hitachi HitVoice により作成した。



図 1: 会話エージェント

収集データ: 被験者は、千葉労災病院の整形外科に入院中の認知症患者 2 人と物忘れ外来に通院中の認知症患者 3 人の計 5 人である。被験者は全員女性であり、年齢は 60~92 歳、認知症のレベルを表す MMSE は 5~20、平均 12.2 であった。被験者の正面と横からビデオ撮影を行った。また、実験終了後に被験者に対して簡単なインタビュー形式のアンケートを行った。対象者とその家族には実験の目的と内容を説明し、インフォームド・コンセントを得た。

2.2 インタラクションの観察から得た知見

全員の患者において、全ての質問が実行され、会話は問題なく遂行された。実験後のインタビューやビデオ観察から以下のような知見が得られた。

- 全員の患者において、エージェントとの会話が成り立つことが確認された。
- 実験後のインタビューより、認知症患者は会話エージェントを好感を持って受け入れていることが確認された。
- 会話エージェントに対して認知症患者から質問してくるという場面は 1 度も見られなかった。従って、興味を引くような質問を適切なタイミングで行い、相手の返答を聞く機能が有効であると考えられる。
- エージェントの発話タイミングが悪く被験者の発話とオーバーラップした場合に被験者が発話を止めるという場面が複数回見られた。

以上の観察結果より、認知症患者に対してアニメーションエージェントが語りかけ対話をするということは十分可能であり、有効であるという印象を持った。

3 語りかけエージェントの実装

3.1 システム概要

前述の WoZ 実験を行った結果、操作者の適切な相槌や適切なタイミングでの質問によって会話がより円滑になる印象を受けた。また、会話において、相槌や頷きは聞き手からの話し手へのフィードバックと言われ、話し手の発話継続を促す機能を持つ。さらに、Gratch[1]らは頷き等の適切な非言語行動をエージェントが表出することにより、ユーザのエージェントに対する信頼度が向上することを示している。

そこで、本研究では以下の 3 つの特徴を持つ、語りかけエージェントを実装した。

- ユーザが話し終わってから次の質問に進む
- 適切なタイミングで相槌を打つ
- 適切なタイミングで頷く

相槌と頷きの自動生成については、Gratch[1]とTakahara[2]を参考にして、表 2 に示す相槌/頷き生成ルールを作成した。

Communication support for persons with dementia using a conversational agent

[†]Junta HIKINO, [†]Yukiko NAKANO, [‡]Kiyoshi YASUDA;

[†]Seikei University, Faculty of Science and Technology;

[‡]Chiba Rosai Hospital, Department of Rehabilitation;

表 2: フィードバック生成ルール

現在のピッチ n が $n-1$ から $n-16$ までの平均ピッチより 5%低下した場合	うなずく
相手の発話後、2 秒以内の無音があり再度発話した場合	あいづち(ええ)
相手の発話後、2 秒以上の無音があった場合	あいづち(はい)
相手の発話の最後のピッチが同 1 発話中の平均ピッチと比べ下り幅が 2%以内だった場合	あいづち(はい)
相手の発話の最後のピッチが同 1 発話中の平均ピッチと比べ 10%以上上がった場合	あいづち(はい)

システム構成を図 2 に示す。認知症患者の音声認識は現段階では難しいため、今回は音声言語理解は使用せず、マイクに対して音声入力があったか否かをユーザの入力判定に用いた。マイク感度は適宜被験者に合わせて変更した。閾値以上の大きさの発話音声で認識されると、Java の動作決定部に信号が送られ、被験者が発話状態であることをシステムが認識する。ピッチ取得にはフリーソフトの Pitch&Rhythm Monitor[3]を使用した。これによって検出されたピッチ [Hz] は Java の動作部に送られ、表 2 のルールが適用され、エージェントの動作が決定される。wav ファイルの再生と同時にファイル再生終了イベントが確認されるまで、アニメーション部ではリップシンクの動作を行う。wav ファイル再生が終了すると、待機モーションに戻る。エージェントによる質問は前回の質問に加え、千葉労災病院で使用されている患者へのアンケート「自分史マニュアル」から抽出した。

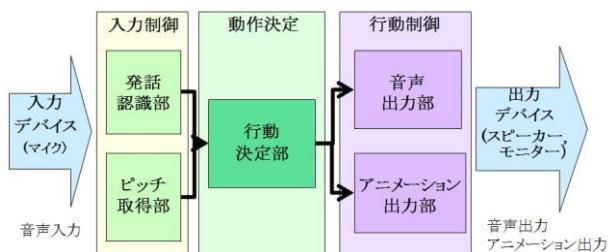


図 2: システム構成図

4 評価実験

4.1 実験実施

各被験者には以下の 3 種類の会話エージェントと会話をしてもらった。

- (1) 協調的エージェント: 相手の発話音声が続いている場合は、相槌や頷きを返すことにより相手の発話に反応し、6 秒以上発話音声がない場合に次の質問に移る。
- (2) 非協調的エージェント: 相手の発話に関わらず、10 秒間隔で質問を発話する。相手の発話中には相槌や頷きは行わない。
- (3) モノログエージェント: 相手の発話の有無にかかわらず、10 秒間隔で一方向的に発話をする。

(1) は会話のテンポがよくなり、会話が弾む。(2) は会話にはなるが、エージェントの質問によって被験者の発話が遮られ、円滑な会話にはならない。(3) は会話は成立せず、一方向的に話すエージェントに対し、被験者は退屈に感じると予想される。

この 3 条件により、(1)(2)間では相槌の有無と質問のタイミングによる差異が、(2)(3)間ではコンテンツの差による差異が現れるのではないかと考えた。対話時間は各 7 分間の計 21 分間である。予備実験同様、耳の聞こえづらい被験者にはイヤホンを使用していた。被験者は千葉労災病院の物忘れ外来に通院中の認

知症患者男性 6 名、女性 7 名、合計 13 人である。MMSE の値は 11~29、平均値 22.6 で比較的軽度の方が多かった。被験者の正面と横からビデオデータ撮影すると同時に別マイクにより音声データを収録した。図 3 に評価実験の様子を示す。



図 3: 評価実験の様子

4.3 実験結果

実験の結果、協調的エージェントに対する平均発話量は 88.9 秒で、非協調的エージェントに対する発話量 92.8 秒とあまり差はなかった。一方、モノログエージェントに対して被験者が語りかけることは少なく、発話量も少なくなっている (40.1 秒)。従って、エージェントが協調的に語りかけることより、発話量が増えるという結果は得られなかった。一方、ユーザの発話とエージェントの発話がオーバーラップする回数を計測したところ、協調的エージェントでは平均 1.15 回、非協調的エージェントでは、4.08 回であり、t 検定の結果統計的に有意な差が得られた ($t = -3.37061$, $p < 0.005$)。この結果から、発話量には差はないが、協調的エージェントと会話をする方が話しやすかったのではないかと推測される。実験中の被験者の反応を観察すると、興味のある話題とそうでない話題とで、発話量が大きく異なることが分かった。今後は、ユーザの性別や趣味等によって興味を持ちそうな話題をあらかじめ選別することが必要であると考えられる。また、単純に質問するのではなく、エージェントが自分の話も交えながら質問する方が、より相手からの発話を引き出しやすこともわかった。

5 まとめ

認知症患者の精神安定、介護者の負担軽減を目指して、認知症患者支援用の会話エージェントを実装した。今後は会話エージェントを利用した認知症患者のためのコミュニケーション支援モデルの確立に向け、収集したビデオデータを使用し各条件と表情の変化の関係性等詳細なデータ分析を進める予定である。また、エージェントの質問内容も相手の興味に合わせて変更する等のコンテンツの工夫も進めていく予定である。

謝辞: 本研究の実験に協力してくださった被験者の皆様、千葉労災病院の皆様に深謝いたします。

参考文献

- [1] Jonathan, et al.; "Virtual Rapport." Lecture Notes in Computer Science, 2006,
- [2] W. Tsukahara and N. Ward; Responding to Subtle, Fleeting Changes in the User's Internal State. CHI: Conference on Human Factors in Computer Systems (ISBN: 1-58113-327-8), ACM, p.p. 77-84, 2001.
- [3] 山岡 忠夫 Home; 音階 & リズム測定ソフト Pitch&RhythmMonitor (prmonitor) <http://hp.vector.co.jp/authors/VA046927/>