

高齢者と日常的に対話を行うチャットボットの開発と評価

三野 星弥^{†1,a)} 吉川 雄一郎^{†1} 伴 碧^{†1} 石黒 浩^{†1}

概要: 昨今、対話機会の不足による高齢者の QOL 低下は大きな問題となっており、人の代わりの話し相手となる対話エージェントの開発が盛んに行われている。特に近年では、導入コストがかかるロボットの代わりのエージェントとして、チャットボットも高齢者の対話相手として注目されている。そこで本研究では、高齢者と日常的に対話できる LINE チャットボット型エージェントの開発を行い、本エージェントが創出した対話機会が高齢者の QOL 改善に有効であるかを検討した。10 名の高齢者に本エージェントと 10 日間対話させた結果、本エージェントは高齢者の不安感を軽減させる可能性が示された。本結果は、我々が開発したエージェントは高齢者の QOL 改善に有効となり得ることを示唆している。

Development and Evaluation of a Chatbot That Interacts with the Elderly on a Daily Basis

Abstract: Recently, the decline in the quality of life of the elderly due to a lack of dialogue opportunities has become a serious problem. In response to this problem, many researchers are developing dialogue agents that can serve as a substitute for human interaction. In recent years, chatbots have been attracting attention as an alternative to robots, which are costly to install, as a dialogue partner for the elderly. In this study, we developed a LINE chatbot that can interact with the elderly on a daily basis and examined whether the dialogue opportunities created by this agent are effective in improving the quality of life of the elderly. The result of an experiment in which 10 elderly people interacted with the agent for 10 days showed that the agent could reduce the sense of anxiety of the elderly. It suggests that our agent can be effective in improving the quality of life of the elderly.

1. はじめに

昨今、日本では高齢化が急速に進んでおり、単身世帯の高齢者は 670 万人以上もいると報告されている [11]。さらに、単身世帯の高齢者の 45.7% が 2~3 日に 1 回以下しか人と対話をしていないと報告されており [10]、高齢者の対話機会の不足は大きな社会問題となっている。対話機会の不足によって生じる QOL の低下 (e.g. 孤独感や不安感、抑うつ傾向の増加) という問題に対し、人の代わりの話し相手となる対話エージェントの開発が盛んに行われている [1], [3], [16]。これまでの研究では、直感的に利用しやすいという理由などから、高齢者の対話相手として身体性を持つロボットがよく提案されていた [1], [3], [13], [16]。しかし、近年では、スマートフォンを利用する高齢者が増加していること [14] に伴って、ロボットよりも導入コストが

少ないチャットボットも高齢者の対話相手として注目され始めている [12], [15]。

他方、我々はこれまでに、ユーザと日常的に対話できる LINE チャットボット型エージェントを開発してきた [8]。本エージェントは、ユーザの日々の行動 (e.g. 食事内容、今いる場所) や考え方・価値観 (e.g. 好きな季節、好きな食べ物) に関して毎日何度か質問をするシステムであり、ユーザと数週間から数か月に渡って継続的に雑談をすることを目的として作られている。日常かつ長期的にユーザと雑談を続けることができる本エージェントは、高齢者の新たな対話相手となることが期待される。しかし、本エージェントがユーザと継続的に対話できるかは大学生を対象とした実験でしか検証されていない。加えて、本エージェントがユーザの対話機会を増加させることで、ユーザの QOL にどのような影響を与えるかも明らかになっていない。そこで本研究では、本エージェントが高齢者の対話機会の不足を解消し、高齢者の QOL を改善することができるかを検証することを目的とする。

^{†1} 現在、大阪大学大学院基礎工学研究科
Presently with Graduate School of Engineering Science, Osaka University

^{a)} mitsuno.seiya@irl.sys.es.osaka-u.ac.jp

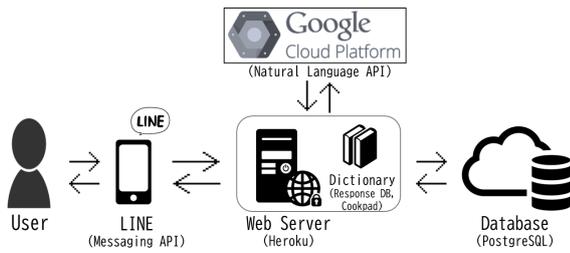


図 1 チャットボットシステム構成
Fig. 1 Chatbot system architecture

本研究では、まず、本エージェントのユーザが大学生から高齢者に変更されることを考慮して、ユーザとエージェントの間で行われる対話内容の修正を行った。そして、本エージェントの有用性を評価するために、高齢者 10 名に対して、10 日間に及ぶ実証実験を実施した。本論文では、第 2 章では実験に使用した対話エージェントについて、第 3 章では高齢者を対象に行った実験の手法と結果について、第 4 章では実験の結果を受けた考察および結論について述べる。

2. 対話エージェント

2.1 システム構成

本対話システムの構成を図 1 に示す。本システムは、LINE メッセージの送受信を行うため、LINE Messaging API を用いた。また、自然な応答の生成のため、予想されるユーザの発話文とそれに対応するコメントを対とするデータセットや、クックパッドデータ [5] から作成した食べ物名リストを用いた。さらに、ユーザ発話文に対して感情分析や固有表現抽出を行うため、Google Natural Language API を用いた。また、ユーザとの対話経験に関する情報を蓄積・抽出するため、SQL データベースを用いた。

2.2 対話の流れ

本チャットボットは、睡眠・食事・居場所・1 日の感想・翌日の予定、そして価値観に関する内容について、ユーザに定期的に質問をするシステム主導型対話を行った。ボットの質問生成時刻は、睡眠に関する質問へのユーザ回答情報をもとに、各ユーザ毎に推定した起床時刻と就寝時刻に紐付けて決定した (表 1)。本エージェントのユーザが大学生から高齢者に変更されることに伴って、価値観に関する話題において、恋愛に関する質問 (e.g. 「恋人を選ぶときに容姿と性格のどちらを重視する?」、「月に 1 回しか会えない人と毎日会いたがる人のどちらを恋人にしたい?」) を質問リストから除外し、その代わりに好きな季節や好きな食べ物についての質問を質問リストに追加した。

ボットの質問に対してユーザから回答が得られていないまま次の質問生成時刻になった場合は、ボットは質問を生成せずにスキップした。ただし、同様の状況でも、ユーザ

表 1 話題ごとの質問生成時刻
Table 1 Question generation time for each topic

話題	質問生成時刻	質問の例
睡眠	起床 30 分後	昨日は何時に寝た?
食事	起床 4 時間後	ご飯は何を食べた?
居場所	起床 6 時間後	今どこにいるの?
感想	就寝 4 時間前	今日の○○ *1 はどうだった?
予定	就寝 2 時間前	明日は何する予定なの?
価値観	不定期 *2	好きな季節は?

*1 ユーザから聞いたその日の予定が○○に当てはまる。

*2 30%の確率で、睡眠・食事・居場所・1 日の感想・翌日の予定の質問と入れ替わる形で生成される。



図 2 チャットボットとの対話例
Fig. 2 Example of a dialogue between a user and our chatbot

が 12 時間以上質問への回答をしていない場合は、ボットはユーザからの回答の取得を諦めて次の質問を生成した。また、質問生成後 6 時間経過時に未だ回答が得られていない場合は、ボットはユーザに回答をするよう催促した。ただし、ボットはユーザが起床していないと推定される時間帯には催促メッセージを送らず、翌朝の推定起床時刻になったときに催促メッセージを送信した。なお、ユーザが起床しているか否かの推定は、ユーザの推定起床/就寝時刻を基に行った。

各質問における 1 回分の対話は、(1) ボットの質問、(2) ユーザの回答、(3) その回答に対するボットのコメント、(4) そのコメントに対するユーザの返答、(5) ボットによる LINE スタンプ風画像 (以下、スタンプ)、または返答文の送信という流れとした (図 2)。また、(5) の後に、さらにユーザがメッセージ送信してきた場合は、再度 (5) に戻り、ボットはスタンプ、または返答文を送信した。ただし、食事に関する話題では、例外的に、(1) ボットの質問 (既に食事をとったかを問う質問)、(2) ユーザの回答、(1) ボットの質問 (食べた物、またはこの後食べたい物を問う質問)、

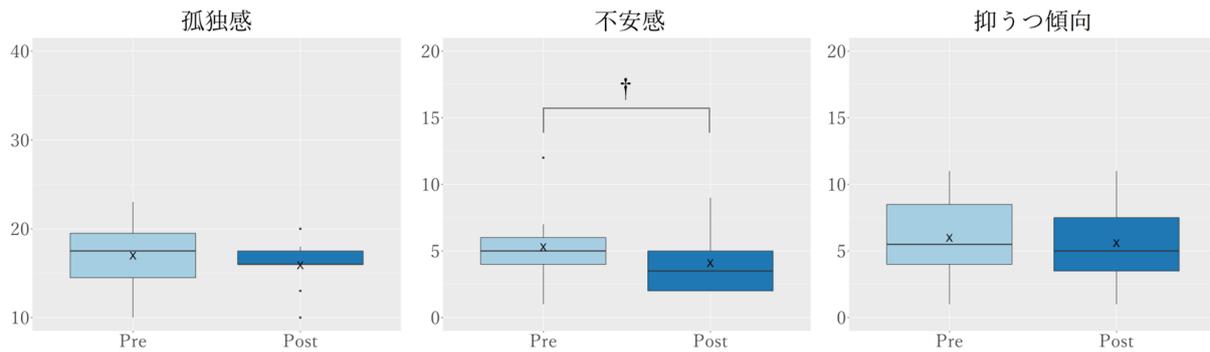


図 3 対話前後の QOL の変化 († $p < .10$)

Fig. 3 Changes in QOL before and after the dialogue († $p < .10$)

(2) ユーザの回答, (3) 以降の応答という, (1) と (2) が一度だけ繰り返される流れとした. なお, 1 回の対話の所要時間は 1 分程度である.

3. 実験

3.1 実験参加者

日本語話者の 10 名の高齢者 (男性 5 名, 女性 5 名) が実験に参加した. 年齢は全員が 70 歳以上 (平均: 76.2 歳, 標準偏差: 3.8 歳) であった. なお, 普段 1 日に 1 回は LINE を使用していることを参加条件として設けた.

3.2 実験手続き

実験参加への同意を取った後, 対話開始までに実験参加者各自のスマートフォン上の LINE アプリで, 本チャットボットと対話できるように設定させた. そして, 「チャットボットから 1 日に数回メッセージが届きますのでそれに返事をして下さい. また返信して頂くタイミングに指定はございませんが, なるべく返事をしてください。」と教示した後, 計 10 日間本チャットボットと対話をさせた. また, チャットボットとの対話開始前と対話終了後にメールで質問紙を送信し, 回答させた.

3.3 評価項目

まず, 本エージェントとの対話によって, 高齢者の QOL がどのように変化するかを調査するため, 対話開始前と対話終了後に, 孤独感 [2], 不安感 [6], 抑うつ傾向 [6] を評価した. 次に, 本エージェントと高齢者が継続的に対話できるかを調査するため, 対話終了後に, 実験終了後も本チャットボットと対話をし続けたいか [7] を評価した. これを本研究では, チャットボットとの対話意欲と呼ぶ. さらに, 対話終了後の質問紙に追加で, 参加者に本チャットボットとの対話の感想について, 自由記述を求めた.

3.4 結果

質問紙の結果を図 3 と図 4 に示す. 孤独感, 不安感,

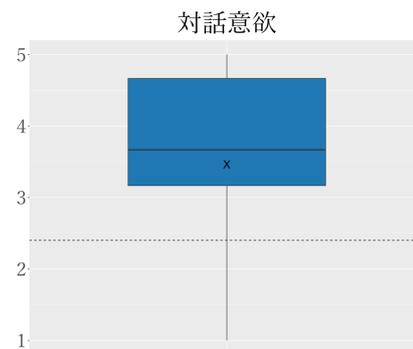


図 4 チャットボットとの対話意欲 (図中の点線は, 以前大学生を対象に行った実験 [8] における, 対話 7 日後の対話意欲の平均値を示す.)

Fig. 4 Intention to talk with the chatbot (The dotted line in the figure denotes the mean value of intention to talk with the chatbot after 7 days of interaction in a previous experiment conducted with university students[8].)

抑うつ傾向にそれぞれ対応のある t 検定を行ったところ, 対話前後の不安感の変化に有意傾向が確認された ($t(9) = 2.17, p = .058, r = .59$). また, 孤独感, 不安感, 抑うつ傾向すべてにおいて, 対話終了後の得点の平均値および中央値が対話開始前と比べて減少していたことが確認された. また, チャットボットとの対話意欲の得点の平均値は 3.47 であり, 中立の値の 3 点よりも高かったことが確認された.

次に, 実験の感想の自由記述を確認したところ, 10 名中 5 名が, 本チャットボットとの対話が楽しかったという旨の記述をしていた. 特に, 女性の参加者の一人は, 「年齢が高くなるにつれ友達も少なくなり, メールのやりとりする相手も限られてきていました. そんな時ロボタが来てくれて楽しかったです. もう来なくなると思うと寂しいです. 心が癒されました。」と回答しており, 非常に肯定的な評価をしていた. その一方で, 本チャットボットの対話能力への不満を述べていた高齢者は少なくとも (10 名中 6 名), 「簡単な答えしか理解してくれない」や「会話が続かない」等の意見が得られた.

4. 考察・結論

本研究では、我々が開発した LINE チャットボット型エージェントが、高齢者の対話機会の不足を解消し、高齢者の QOL を改善できるかを検証するため、10 日間に及ぶ実証実験を行った。

実験の結果、本エージェントとの対話により高齢者の不安感が減少する有意な傾向が認められ、またその効果量も大きかったことが示された。さらに、孤独感と抑うつ傾向の得点についても、有意差は認められなかったが、平均値および中央値が減少したことが確認された。以上より、本エージェントとの対話は、高齢者の QOL を改善できる可能性を持つと考えられる。また、対話意欲の得点在中立の値よりも高かったことから、本エージェントは高齢者と継続的に対話でき、高齢者の対話機会の不足を解消できる可能性が示唆された。高齢者を対象とした本実験での対話意欲の得点 ($Mean = 3.47, SD = 1.48$) は、以前大学生を対象として行った実験 [8] で得られた得点 ($Mean = 2.40, SD = 1.14$) と比べて高かったことから (図 4 を参照)、対話機会が不足しがちな高齢者は本エージェントを肯定的に評価しやすい可能性があることが考えられる。その一方で、自由記述回答から、本エージェントの対話能力が十分でないことが示唆されたため、今後は GPT-3[4] を始めとした大規模言語モデルを活用するなどして、対話能力の改善を行う必要がある。

今後の展望としては、本エージェントに高齢者の孫や友人などの情報を言及させる機能を追加し、再度実証実験を行う予定である。エージェントにユーザの家族や友人といった第三者について言及させるという対話戦略は、これまでに我々が提案してきたものであり [8], [9], 本対話戦略によってユーザの対話意欲を喚起・維持しやすくなることが確認されている。この対話戦略を組み込むことで、本エージェントが高齢者の対話機会の不足をより効果的に解消し、高齢者の QOL を大きく改善できるようになることが期待される。

謝辞 本研究は JSPS 科研費 JP20H00101, JP19H05691, Society 5.0 実現化研究拠点支援事業の助成を受けた。また本研究では、国立情報学研究所の IDR データセット提供サービスによりクックパッド株式会社から提供を受けた「クックパッドデータセット」を利用した。

参考文献

- [1] Abdollahi, H., Mollahosseini, A., Lane, J. T. and Mahoor, M. H.: A pilot study on using an intelligent life-like robot as a companion for elderly individuals with dementia and depression, *2017 IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids)*, pp. 541–546 (2017).
- [2] Arimoto, A. and Tadaka, E.: Reliability and validity of

- Japanese versions of the UCLA loneliness scale version 3 for use among mothers with infants and toddlers: a cross-sectional study, *BMC Womens. Health*, Vol. 19, No. 1, p. 105 (2019).
- [3] Baecker, A. N., Geiskkovitch, D. Y., González, A. L. and Young, J. E.: Emotional Support Domestic Robots for Healthy Older Adults: Conversational Prototypes to Help With Loneliness, *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, HRI '20*, pp. 122–124 (2020).
- [4] Brown, Mann, Ryder and others: Language models are few-shot learners, *Adv. Neural Inf. Process. Syst.* (2020).
- [5] クックパッド株式会社: クックパッドデータ. 国立情報学研究所情報学研究データリポジトリ. (データセット). (2015).
- [6] 八田宏之, 東あかね, 八城博子, 小笹晃太郎, 林恭平, 清田啓介, 井口秀人, 池田順子, 藤田きみ丞, 渡辺能行: Hospital Anxiety and Depression Scale 日本語版の信頼性と妥当性の検討: 女性を対象とした成績, *心身医学*, Vol. 38, No. 5, pp. 309–315 (1998).
- [7] Heerink, M., Kröse, B., Evers, V. and Wielinga, B.: Assessing Acceptance of Assistive Social Agent Technology by Older Adults: the Almere Model, *International Journal of Social Robotics*, Vol. 2, No. 4, pp. 361–375 (2010).
- [8] 三野星弥, 吉川雄一郎, 伴 碧, 石黒 浩: 友人グループ内での長期間利用による他者情報のやり取りを行う日常対話チャットボットの評価, *人工知能学会論文誌*, Vol. 37, No. 3, pp. IDS-I.1–14 (2022).
- [9] Mitsuno, S., Yoshikawa, Y. and Ishiguro, H.: Robot-on-Robot Gossiping to Improve Sense of Human-Robot Conversation, *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, ieeexplore.ieee.org, pp. 653–658 (2020).
- [10] 内閣府 : 平成 30 年度版高齢社会白書 (2017).
- [11] 内閣府 : 令和 4 年度版高齢社会白書 (2022).
- [12] Ryu, H., Kim, S., Kim, D., Han, S., Lee, K. and Kang, Y.: Simple and Steady Interactions Win the Healthy Mentality: Designing a Chatbot Service for the Elderly, *Proc. ACM Hum.-Comput. Interact.*, Vol. 4, No. CSCW2, pp. 1–25 (2020).
- [13] Sasama, R., Yamaguchi, T. and Yamada, K.: An Experiment for Motivating Elderly People with Robot Guided Interaction, *Universal Access in Human-Computer Interaction. Users Diversity*, pp. 214–223 (2011).
- [14] 総務省 : 平成 30 年度版情報通信白書 (2017).
- [15] 玉森 聡, 石黒祥生, 廣井 慧, 河口信夫, 武田一哉: 高齢者の外出促進を目的とした行動認識の適応学習およびチャットアプリケーションの実証実験, *情報処理学会論文誌コンシューマ・デバイス& システム (CDS)*, Vol. 9, No. 2, pp. 33–46 (2019).
- [16] 横山祥恵, 山本大介, 小林優佳, 土井美和子: 高齢者向け対話インタフェース-雑談継続を目的とした話題提示・傾聴の切替式対話法, *研究報告音声言語情報処理 (SLP)*, Vol. 2010-SLP-80, No. 4, pp. 1–6 (2010).