Original Manuscript

RATE

Journal of Rehabilitation and Assistive Technologies Engineering
Volume 9: 1–9

DOI: 10.1177/20556683221105768
journals.sagepub.com/home/jrt
SAGE

A simulated experiment to explore robotic dialogue strategies for people with dementia

Fengpei Yuan[1], Amir Sadovnik[2], Ran Zhang[3], Devin Casenhiser[4], Eun Jin Paek[4] and Xiaopeng Zhao[1]

Abstract

Introduction: Persons with dementia (PwDs) often show symptoms of repetitive questioning, which brings great burdens on caregivers. Conversational robots hold promise of helping cope with PwDs' repetitive behavior. This paper develops an adaptive conversation strategy to answer PwDs' repetitive questions, follow up with new questions to distract PwDs from repetitive behavior, and stimulate their conversation and cognition.

Methods: We propose a general reinforcement learning model to interact with PwDs with repetitive questioning. Q-learning is exploited to learn adaptive conversation strategy (from the perspectives of rate and difficulty level of follow-up questions) for four simulated PwDs. A demonstration is presented using a humanoid robot.

Results: The designed Q-learning model performs better than random action selection model. The RL-based conversation strategy is adaptive to PwDs with different cognitive capabilities and engagement levels. In the demonstration, the robot can answer a user's repetitive questions and further come up with a follow-up question to engage the user in continuous conversations.

Conclusions: The designed Q-learning model demonstrates noteworthy effectiveness in adaptive action selection. This may provide some insights towards developing conversational social robots to cope with repetitive questioning by PwDs and increase their quality of life.

Keywords

Conversational robot, Alzheimer's dementia, human-robot interaction, reinforcement learning, adaptive robot behavior

Date received: 14 December 2021; accepted: 22 May 2022

Introduction

According to the World Alzheimer Report 2018,[1] there were 50 million people living with Alzheimer's Disease and Alzheimer's Disease Related Dementias (AD/ADRD) in 2018 around the world, with one new case of dementia every 3 s. Alzheimer's disease is the most common form of dementia, contributing to 60 – 70% of cases.[2] Due to memory impairment, persons with AD/ADRD (PwDs) often show behavior of repetitive questioning, which can be very frustrating, tedious and exhausting to their

[1]Department of Mechanical, University of Tennessee, Knoxville, TN, USA
[2]Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN, USA
[3]Department of Electrical and Computer Engineering, Miami University, Oxford, OH, USA
[4]Department of Audiology and Speech Pathology, University of Tennessee Health Science Center, Knoxville, TN, USA

Corresponding author:
Xiaopeng Zhao, Department of Mechanical, Aerospace, and Biomedical Engineering, University of Tennessee, 1506 Middle Dr, Knoxville, TN 37916, USA.
Email: xzhao9@utk.edu

caregivers.[3,4] Worse yet, caregivers often do not receive sufficient training on how to communicate with people with dementia and do not have the time to learn the best communication methods to create a strong relationship.[5,6] Especially in senior homes, many PwDs do not receive meaningful daily interactions.[7] Social robots, an emerging common assistive technology to support dementia care,[8–10] could play the role of conversational companion engaging in conversations with PwDs. For example, a very recent late-breaking report[11] proposed proactive robotic listeners to encourage more responses from PwDs. The robot will ask a follow-up question related to the most recent topic or start a topic introduction if a PwD shows silence during conversation. In the situation when a PwD shows repetitive questioning, a conversational robot can be employed to answer those repetitive questions, with the advantages of high repeatability and no complaints and no fatigue.[12] Meaningful daily communication with a robot, as indicated in previous studies, may help reduce symptoms of AD/ADRD and improve quality of life and independence of people with AD/ADRD.[5,6,13–15] It has been found that having a robot to talk to can decrease the feelings of loneliness.[16–18]

On the other hand, adaptive robot interactions are necessary to provide a comfortable and effective interactions with target users,[19] i.e. PwDs in our study. An adaptive dialogue system would facilitate meaningful, effective communication and a more trusting relationship between the PwD and robot.[6,13] For example, Rudzicz et al.[20] confirms that the entire communication system will be more effective if the individual's mental state is taken into account. The adaptive (or autonomous) behaviors in robots have been investigated using the technique of Wizard of Oz (WoZ), where the robot is usually controlled by a human operator.[21,22] However, WoZ has been demonstrated to not be a sustainable technique in long term.[23] WoZ may be sufficient for narrow task domains and very specific user interactions, but it is limited in terms of flexibility and adaptivity to different individual users.[24] Particularly regarding the population of people with AD/ADRD, each individual may have a different personality, preference and cognitive abilities,[25] and show time-varying behaviors, emotions (e.g. behavioral and psychological symptoms of dementia, BPSD),[26] and personality[27] in both short and long term, as well as time-decreasing cognitive capabilities.

The limitations of WoZ can be compensated by using adaptive algorithms such as reinforcement learning (RL), which enables robots to learn from the interaction with the environment (e.g. users) and makes it possible to adapt and optimize robotic policies to different individual users. RL allows to integrate developer-defined rewards that better mimic the goal of conversational robot and to model the long-term influence of a generated response in an ongoing dialogue with PwDs.[28] It has been applied in some studies for dialogue management. For example, Cuayáhuitl[29] used deep reinforcement learning to perform action selection from raw text for the context of restaurant. The state space was defined as word-based features, and the action space included 35 dialogue actions in response to users' intentions. In another study[30] of using RL to learn a conversation strategy for autonomous robotic dialogue system for PwDs, the authors designed state space as the robot's internal motivation (closely associated with user's motivation) and previously selected action. Their action space was represented by three types of robot's action, including short response (simple agreement/encouragement), long response (question) and topic change. The results showed the robot was capable of maintaining conversations with seniors for at least 20 min.

Regarding the specific context of repetitive questioning in PwDs, the conversational robot would be expected to not only answer those repetitive questions for PwDs, but also further communicate with PwDs to distract their attention from the repetitive behaviors and stimulate their cognitive activities. Asking appropriate questions proactively by the robot can be a good approach to start conversations with PwDs.[11] However, the question difficulty level (e.g. closed-ended vs open-ended questions) must be taken into careful consideration during communication with PwDs.[31] Inspired by previous relevant research,[32,33] asking questions with optimal difficulty level with respect to PwD's cognitive capability may help engage PwD and keep them interested in interacting with the robot, and maximally stimulate their cognitive activities.

Our long-term goal is to build a cost-effective and adaptive conversational robot, as a complement of the caregivers, to cope with repetitive questioning in PwDs by answering their questions and following up with new questions to distract them from the repetitive behaviors. Inspired by previous studies,[32,33] asking questions with optimal difficulty level geared to a PwD's cognitive capability may keep them engaged and interested in interacting with the robot, and consequently, stimulate their cognitive activities to the most extent. In this paper, we propose a general RL framework for this problem (i.e. repetitive questioning), to explore an adaptive conversation strategy for a robot by tuning the frequency and difficulty level of asking a follow-up question when communicating with PwDs. To the best of our knowledge, this work is the first to learn adaptive conversation strategy for social robots to specifically cope with PwDs' repetitive questioning. The strategy is adaptive to PwDs with different cognitive capabilities and engagement levels. We expect this general framework would serve as a useful benchmark on the design of the adaptive conversation strategy for the robots when additional sensing modalities are available.

The remainder of the paper is organized as follows. The Section Method introduces the methodology, including the definition of Markov Decision Process, the modelling of the persons with AD/ADRD, and experiments. The simulation results and a demonstration are presented in the Section Results and discussed in the section of Discussion. A discussion of limitations in this study as well as the future work are listed in Section Limitation and future work. The last section concludes the paper.

Method

The technique of reinforcement learning is used to learn from the PwD-robot conversation (Figure 1) and investigate the optimal policy for a robot. To offset the stress of PwD's repetitively questioning on a caregiver, the robot should be able to always answer the questions asked by the PwD, propose some follow-up questions to distract PwD from the repetitive behaviors, and also to stimulate their cognitive activities. However, if a follow-up question is too challenging, difficult or complicated to the PwD, the PwD may not respond to it at all. Therefore, the robot should also be able to identify if the follow-up question is too difficult for the PwD, and adapt the difficulty level to users with different cognitive capabilities. Therefore, the optimal policy maps the PwD-robot interaction/dialogue to the robot's follow-up rate and difficulty level of follow-up question, so that there will be frequent conversation between the PwD and the robot, together with the brain activities in people with AD/ADRD being maximally stimulated.

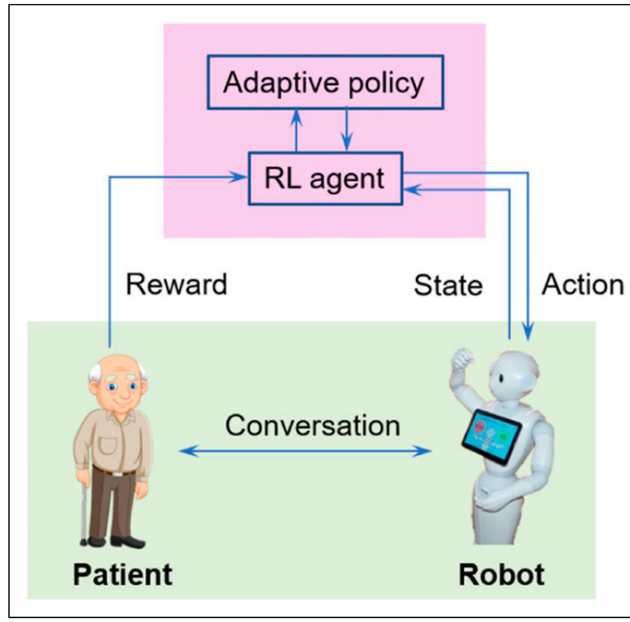

Figure 1. A schematic framework demonstrates the adaptive strategy of PwD-robot dialogue.

Definition of reinforcement learning

According to the aforementioned goals for PwD-robot dialogue, we define the key elements of Markov decision process (MDP) model as follows[34]:

- State space: A state is defined according to a user's situation during a conversational context. There are five potential states: the user asking a question, denoted by *Q* or *Question*, the user's question being simply answered by robot without following-up, denoted by *NF*, the user providing relevant response to the robot's follow-up question, denoted by *RR*, the user providing irrelevant response to the robot's follow-up question, denoted by *IR*, and the user providing no response to robot's follow-up question, denoted by *NR*. Noticeably, underlying the states *RR*, *IR*, and *NR*, there are always processes of the user's question being answered by the robot and further the robot asking a follow-up question.
- Action: An action includes two elements: the follow-up rate (i.e. the probability of the robot asking a question after answering PwD's question) for the robot and the difficulty level of a follow-up question. For simplification, currently the optional follow-up rate is 0.1, 0.4, 0.7 or 1.0. And there are three question difficulty levels: easy, moderately difficult, and difficult. Table 1 shows examples of follow-up questions with different difficulty levels.
- Reward: According to our goals for a social robot in the situation of repetitive questioning, i.e. answering PwDs' repetitive questions, asking follow-up questions to distract PwDs from repetitive questioning and to promote daily conversation, and avoiding too difficult follow-up questions towards PwDs, we build our immediate reward function as following,

$$Reward = \begin{cases} 0, & \text{if } Q \rightarrow \text{NF} \\ +1 \times \text{Question Difficulty}, & \text{if } Q \rightarrow \text{RR} \\ +0.5, & \text{if } Q \rightarrow \text{IR} \\ -0.2 \times \text{Question Difficulty}, & \text{if } Q \rightarrow \text{NR} \end{cases}$$

where the variable *QuestionDifficulty* can be 1, 2, or 3, separately corresponding to easy, moderately difficult, and difficult follow-up questions. The expression *Q*, *NF*, *RR*, *IR*, and *NR* represent the possible states in the MDP model. The symbol → indicates the state transition. Although meaning conversation (i.e. relevant responses from PwD to robot's questions) is the most suggested, irrelevant responses are still meaningful to PwDs.[31] Thus, a slightly positive reward (i.e. +0.5) was assigned to $Q \rightarrow IR$.

Table 1. Examples of follow-up questions by a robot.

Question difficulty	Example
Easy	"Would you like some tea?"
Moderate	"What would you like to drink?"
Difficult	"What do you think about this tea?"

The PwD-robot dialogue is a complicated interactive task, which can be continuing (i.e. the left diagram in Figure 2) or episodic tasks.[34] In this paper, we start our simulated exploration study with the most simple modelling of situation, one-step episodic task, as illustrated in the right diagram of Figure 2. Each episodic task starts by a PwD asking a repetitive question, that is, the state $s = Q$, and ends with a terminal state, i.e. $s = NF$, RR, IR, or NR.

Simulated person with Alzheimer's disease and Alzheimer's disease related dementias

Because the real-world samples of PwD-robot conversational interaction is costly in terms of time and labor and might be related to ethical issues, we start from simulated PwD-robot dialogue using RL modelling, which hopefully provides us insight into real-world PwD-robot dialogue in next step. A simulated individual with AD/ADRD during the PwD-robot conversation is characterized by their response to the robot's follow-up question. The response can be categorized into three types of response: relevant, irrelevant, and no response to the follow-up question. Therefore, an individual is characterized by the relevant response rate P_{Rresp} and irrelevant response rate P_{IRresp}. Notice here the rate of no response P_{Nresp} is dependent on P_{Rresp} and P_{IRresp}. The sum of these three variables is always 1.

We assume that an individual's relevant and irrelevant response rates are influenced by the individual cognitive capability and their engagement, as well as the question difficulty level. Engagement of a person with AD/ADRD is defined as the act of being occupied or involved with an external stimulus,[35] i.e. the conversational robot in our case. The engagement of an individual with AD/ADRD can be influenced by robot attributes and the individual's attributes (e.g. personality and preference).[36] There are three basic rules to create models of PwDs. Firstly, with the same engagement level, a person with a lower cognitive capability will show a lower relevant response probability. Also, we expect a person with lower cognitive capability will have a higher irrelevant response probability except the most severe PwD who will have extremely low response (both relevant and irrelevant) to all questions. Secondly, we assume that an individual with a lower engagement level will have a lower response rate (i.e. the sum of P_{Rresp} and P_{IRresp})

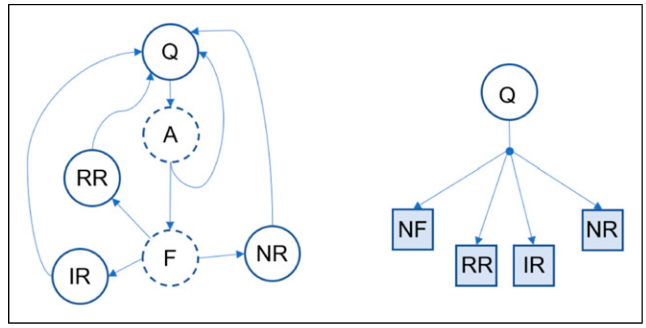

Figure 2. The Markov decision process diagram in one episode of PwD-robot dialogue interaction.

and thus a greater no response rate, P_{Nresp}. The specific effect on P_{Rresp} and P_{IRresp} is dependent on individuals. Therefore, the effects on individual response rate P_{Rresp} and P_{IRresp} are random in our simulation. Thirdly, given a more difficult follow-up question, the same user with the same engagement is expected to show a lower relevant response rate P_{Rresp} and a higher no response rate P_{Nresp}. Naturally, it takes people with AD/ADRD more cognitive workload to answer a more difficult follow-up question. For example, PwDs only need to answer "yes" or "no" to the easy question in Table 1. Comparatively, to answer the difficult question in this table, PwDs need to think more to understand the question and express their opinions.

During our simulation, there are four basic simulated users, each with three different levels of engagement, i.e. high, medium, and low. The parameters of the simulated User $1-4$ with different engagement levels are listed in Tables 2–5. User 1, 2, 3 and 4, correspond to older adults without cognitive impairment, with mild cognitive impairment, moderate dementia, and severe dementia, respectively.

Experiments

On the basis of the aforementioned definition for the three key elements (state, action, and reward function) of MDP, we train our reinforcement learning model for our four simulated users with three different engagement levels. We apply the technique of Off-policy Q-learning to solve the MDP problem, with the ϵ-greedy policy ($\epsilon = 0.1$), constant learning rate $\alpha = 0.03$, and discount factor, $\gamma = 0.95$. An episode starts by a PwD asking a repetitive question and ends with a terminal state (NF, RR, IR, or NR), as illustrated on the right of Figure 2. We evaluate the performance of our model with the metrics proposed by,[13] including the average return per epoch, the starting state value V (*Question*) in each epoch, and the sum of Q-value updates during each epoch. Here, due to our definition of one-step episodic task, we only need to consider the starting state value in the starting state, *Question* (i.e. the PwD asking repetitive

Table 2. Parameters of simulated User 1 without cognitive impairment.

Engagement	Question difficulty	P_{Rresp}	P_{IRresp}	P_{Nresp}
High	Easy	1	0	0
	Moderate	1	0	0
	Difficult	1	0	0
Medium	Easy	0.95	0	0.05
	Moderate	0.92	0	0.08
	Difficult	0.90	0	0.10
Low	Easy	0.90	0	0.10
	Moderate	0.88	0	0.12
	Difficult	0.85	0	0.15

Table 3. Parameters of simulated User 2 with mild cognitive impairment.

Engagement	Question difficulty	P_{Rresp}	P_{IRresp}	P_{Nresp}
High	Easy	0.9	0.1	0
	Moderate	0.86	0.14	0
	Difficult	0.82	0.18	0
Medium	Easy	0.83	0.11	0.06
	Moderate	0.75	0.15	0.10
	Difficult	0.68	0.20	0.12
Low	Easy	0.75	0.14	0.11
	Moderate	0.65	0.16	0.19
	Difficult	0.50	0.18	0.32

Table 4. Parameters of simulated User 3 with moderate dementia.

Engagement	Question difficulty	P_{Rresp}	P_{IRresp}	P_{Nresp}
High	Easy	0.70	0.20	0.10
	Moderate	0.63	0.22	0.15
	Difficult	0.50	0.23	0.27
Medium	Easy	0.60	0.21	0.19
	Moderate	0.50	0.25	0.25
	Difficult	0.30	0.20	0.50
Low	Easy	0.35	0.15	0.50
	Moderate	0.20	0.13	0.67
	Difficult	0.08	0.10	0.82

question). The start state value V (*Question*) indicates the expected return. Each epoch is composed of 150 episodes. Furthermore, we analyze the performance of Q-learning by comparing the learning results using randomized action selection.

Results

Based on the evaluation metrics, we observe that in all cases (i.e. User $1-4$ with different engagement level in Tables 2–5) our RL model converges within 30 epochs. Therefore, we

Table 5. Parameters of simulated User 4 with severe dementia.

Engagement	Question difficulty	P_{Rresp}	P_{IRresp}	P_{Nresp}
High	Easy	0.04	0.08	0.88
	Moderate	0.02	0.08	0.90
	Difficult	0.01	0.04	0.95
Medium	Easy	0.02	0.05	0.93
	Moderate	0.01	0.04	0.95
	Difficult	0.005	0.02	0.975
Low	Easy	0.01	0.04	0.95
	Moderate	0	0.02	0.98
	Difficult	0	0.01	0.99

only show the learning processes of the model during epoch $1-100$ in Figure 3. In this figure, the first, second, and third column correspond to the learning process (i.e. average return) of all users with high, medium, and low engagement, respectively. The solid and dashed curves with the same color in each sub-figure represent the learning results for the same user using Q-learning and random action selection, respectively. A user with a higher engagement level is associated with greater average return, compared to the same user with lower engagement level. With the same engagement level, the converged average return in User 1 and 2 is obviously greater than that in User 3 and 4. The optimal policies learned by the RL agent for different users with different engagement levels are listed in Table 6.

Demonstration of the dialogue policy

For purpose of demonstration, we implement the optimal dialogue policies obtained in Table 6 using a humanoid robot, Pepper,[37] and show the interaction between Pepper and a participant to simulate scenarios of repetitive questioning. Here, a researcher plays the role of a PwD with repetitive questioning behaviors. Pepper is 1.2-m tall and has 17 joints to support expressive and appealing gestures and other body movements. Thanks to its attractive appearance and capability of multimodal interaction (e.g. verbal communication, body movement, and eye contact), the Pepper robot is perceived to be acceptable, appealing, and engaging to human users in daily activities.[38–40] Pepper is equipped with a tablet in its chest, which can be used to display dialogues and pictures. In addition, the robot has a range of sensors, which can be used to evaluate the human user's response relevance, engagement level, and emotions.

We use Pepper's software development kit (SDK), QiSDK, to implement the conversational policy (as listed in Table 6) so that Pepper is able to respond to the repetitive questions as well as to come up with follow-up questions to distract users from their repetitive questioning behaviors. The robot's body movement, gestures and eye contact are used to support the interaction and engage the human user in

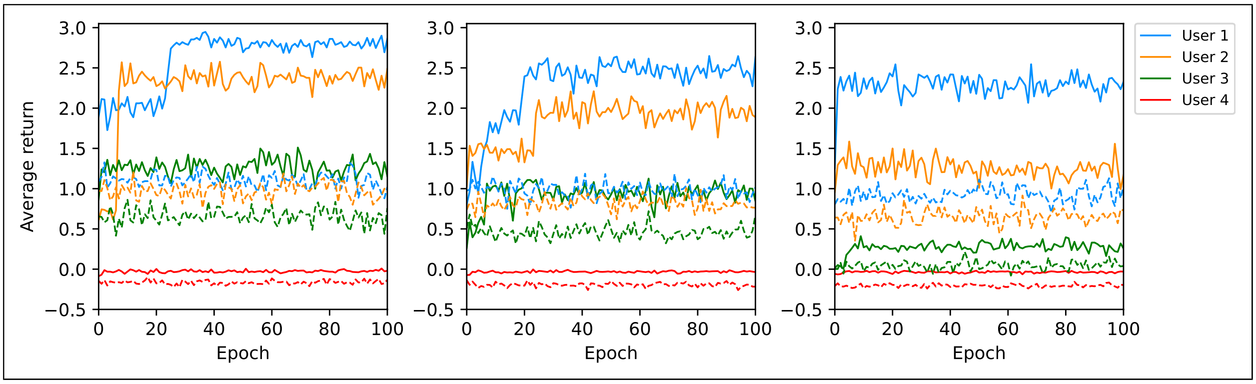

Figure 3. The learning results of average return for User 1 − 4 with high (first column), medium (second column), and low (third column) engagement. In each sub-figure, the solid and dashed curves represent the learning results by Q-learning and random action selection model, respectively.

Table 6. Optimal policy suggested by the Q-learning.

Engagement	User 1	User 2	User 3	User 4
High	[1.0, *D*]	[1.0, *D*]	[1.0, *D*] or [1.0, *M*]	[0.1, *E*]
Medium	[1.0, *D*]	[1.0, *D*]	[1.0, *M*]	[0.1, *E*]
Low	[1.0, *D*]	[1.0, *D*] or [1.0, *M*]	[1.0, *E*]	[0.1, *E*]

Note User 1, 2, 3 and 4 represent a person without cognitive impairment, with mild cognitive impairment, moderate dementia and severe dementia. *D* = difficult follow-up questions; *M* = moderately difficult follow-up questions; *E* = easy follow-up questions.

the interaction. For example, during the conversation, the robot Pepper will always make eye contact with the human user. All the verbal conversation between the user and the robot is displayed in Pepper's tablet. A list of repetitive questions are incorporated into the robot using the Application Programming Interface (API), *QiChatbot* and *topic*, through which we can define how the robot will verbally respond to a human user's verbal input (e.g. repetitive questions) and come up with a follow-up question for the user.

Figure 4 shows a screenshot of the verbal interaction between the user and the robot Pepper. The robot is coded with the policy of [1.0, *M*] in Table 6. During the scenario, when the robot Pepper detects a repetitive question (e.g. "What time is the party?") from the user, the robot first answers the question with the pre-defined answer (e.g. "It is at 7p.m.") and then comes up with a follow-up question for the user.

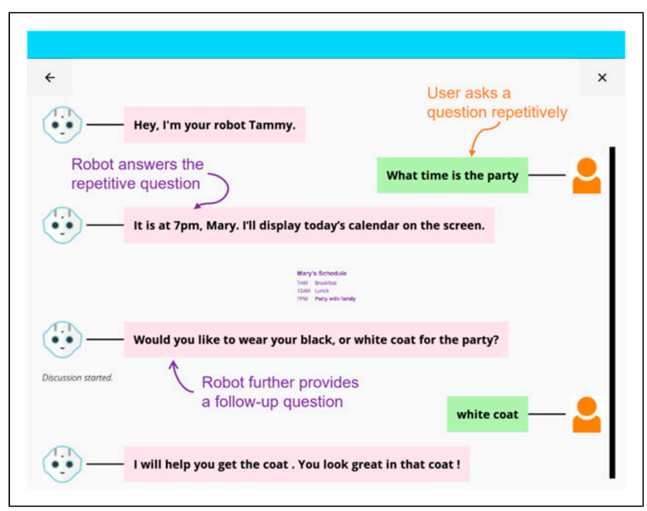

Figure 4. Verbal interaction between a human user (text in green) and the social robot Pepper (text in pink) in a scenario of repetitive questioning.

Discussion

From Figure 3, we can see that the implemented RL agent is able to learn the best policy within 30 epochs. The solid average return curves for all users learned by Q-learning are greater than dashed curves learned by random action selection, which indicates that Q-learning here is helpful for action selection. The learning processes here converge with spikes, which is due to the simulation of stochastic response rate (Tables 2–5).

In Figure 3, as the learning curves converge, the average return for an individual with higher cognitive capability (e.g. User 1 and 2) are greater than an individual with lower cognitive capability (e.g. User 1 and 2). This makes sense considering that an individual without or with mild cognitive impairment is more likely to answer a question. Comparing the optimal policies learned by our RL agent for four types of user with the same engagement level, for example, the row of medium engagement, the agent is able to learn the best policy for PwDs with different cognitive capabilities. More specifically, the optimal policy (i.e. follow-up rate and question difficulty) for a user with higher cognitive capability (e.g. User 1 and 2) is always asking the difficult follow-up question. However, towards users with lower cognitive capability, User 3 and 4, the RL agent separately adapts the policy to always following up with moderate-level difficult question, and following up very rarely also with easy questions.

Comparing the learning results for the same user but with different engagement level, for example, the green curve representing User 3 in the three sub-figures in Figure 3, the average return obtained per epoch decreases as the user's engagement decreases, which makes sense because an individual with lower engagement is expected to less likely join activities and conversations. Moreover, the column of User 3 in Table 6 shows that the question difficulty needs to decrease as an adaptation to individual's decreasing engagement level. This indicates that, although merely observing relevance of user's response, the RL agent is able to detect the change of latent variable (i.e. user's engagement level) and adaptively adjust the question difficulty level accordingly.

Limitation and future work

There are some limitations in the study. First, a state in our MDP only considers the relevance of user's response to a robot's question. Although the current MDP seems to adapt the policy when the latent variables (e.g. user's cognitive capability and engagement level) changes, the inclusion of more variables associated with users, e.g. PwD's cognitive and affective states, may facilitate better learning performance. Previous studies has showed that a PwD's engagement level can be read using sensing technologies such as camera,[41] heart-rate sensor,[19] and passive brain-computer interfaces (BCIs, e.g. electroencephalography).[42] In the future, we will take into consideration one or some PwD's cognitive and affective states as additional dimension(s) of state space, by integrating sensing technologies in a cost-effective way (e.g. in terms of computational complexity).

Second, in the case of demonstration, we manually coded the optimal policy suggested by RL as well as the list of follow-up questions in the robot, which is an impediment to the development of adaptive conversational robot from long term. In the future, we will integrate the robot Pepper and the RL model into one framework, where the input will be raw data of user's states and the output will be Pepper automatically performing an action suggested by RL. We will also work on the list of follow-up questions, which could be appropriate to users' environment and social context.

Third, it is simplified that an individual's cognitive capability and engagement level is consistent during the whole PwD-robot dialogue, which is usually not true in real world considering the intra-individual variability and disease progression in people with AD/ADRD. In future, we will conduct more research on this direction. We will collect real-world repetitive questions from PwDs who show time-varying cognitive capability and time-varying engagement in either the short and long term and adapt our RL model accordingly. Then we will investigate the performance of the RL model for those PwDs in both the short and long term in real world.

Conclusions

In this paper, we have developed a general RL framework, to learn the adaptive conversation strategy for a robot to cope with the problem of repetitive questioning by persons with dementia. The model allows the robot's dialogue system to adjust an appropriate rate of asking a follow-up question and the question difficulty level, given the robot only perceives PwDs' spoken conversation (e.g. repeating a question and relevance of response). We have also presented a demonstration of the optimal dialogue policy, where a physically embodied social robot interacts with a human user with repetitive questioning behaviors. During the scenario, the robot is able to respond to the user's repetitive questions with appropriate answers and further come up with a follow-up question. This study may allow a conversational social robot to help caregivers with PwDs' repetitive questioning behaviors. Moreover, the design of follow-up question may also distract PwDs from repetitive behaviors and stimulate PwDs' brain activities through the conversation with different difficulty level. In the future, the MDP model will be improved by leveraging the multimodal sensing technologies. And more work is needed to train the RL model with users who has time-decreasing cognitive capabilities and time-varying engagement level. Also, future study is needed to integrate the RL agent with the robot Pepper and investigate the performance with PwDs in real world.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Institute of Health under the grant number R01AG077003.

ORCID iD

Fengpei Yuan https://orcid.org/0000-0002-5653-5552

References

1. Patterson C. *World Alzheimer Report 2018*. London: Alzheimer's Disease International (ADI), September 2018.
2. World Health Organization. Dementia. URL https://www.who.int/news-room/fact-sheets/detail/dementia (2020, accessed 24 November 2021.
3. Reeve E, Molin P, Hui A, et al. Exploration of verbal repetition in people with dementia using an online symptom-tracking tool. *Int Psychogeriatrics* 2017; 29(6): 959–966.
4. Hamdy RC, Lewis JV, Copeland R, et al. Repetitive questioning exasperates caregivers. *Gerontol Geriatric Medicine* 2018; 4: 2333721417738915.
5. Kuwamura K, Nishio S and Sato S. Can we talk through a robot as if face-to-face? Long-term fieldwork using teleoperated robot for seniors with alzheimer's disease. *Front Psychology* 2016; 7: 1066.
6. Pou-Prom C, Raimondo S and Rudzicz F. A conversational robot for older adults with alzheimer's disease. *ACM Trans Human-Robot Interaction* 2020; 9(3): 1–25.
7. Paudel A, Resnick B and Galik E. The quality of interactions between staff and residents with cognitive impairment in nursing homes. *Am J Alzheimer's Dis Other Dementias®* 2020; 35: 1533317519863259.
8. Kruse CS, Fohn J, Umunnakwe G, et al. Evaluating the facilitators, barriers, and medical outcomes commensurate with the use of assistive technology to support people with dementia: a systematic review literature. *Healthcare* 2020; 8: 278. Multidisciplinary Digital Publishing Institute.
9. Woods D, Yuan F, Jao YL, et al. Social robots for older adults with dementia: a narrative review on challenges & future directions. In: *Social Robotics*. Cham: Springer, pp. 411–420.
10. Liu Z, Collier PN, Wang C et al. A demonstration of human-robot communication based on multiskilled language-image analysis. In: CHASE 2021: IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies. Washington D.C. 16–17 December 2021.
11. Li Y, Lai C, Lala D et al. Alzheimer's dementia detection through spontaneous dialogue with proactive robotic listeners. In: Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction, Sapporo, Hokkaido, Japan, 7–10 March 2022, pp. 875–879.
12. Taheri AR, Alemi M, Meghdari A, et al. Clinical application of humanoid robots in playing imitation games for autistic children in iran. *Proced - Soc Behav Sci* 2015; 176: 898–906. DOI: 10.1016/j.sbspro.2015.01.556
13. Tsiakas K, Dagioglou M, Karkaletsis V, et al. Adaptive robot assisted therapy using interactive reinforcement learning. In: *Social Robotics*. Cham: Springer, pp. 11–21.
14. Yamazaki R, Kase H, Nishio S et al. A conversational robotic approach to dementia symptoms: measuring its effect on older adults. In Proceedings of the 7th International Conference on Human-Agent Interaction, Kyoto, Japan, 6–10 October 2019, pp. 110–117.
15. Shibata T. Therapeutic seal robot as biofeedback medical device: qualitative and quantitative evaluations of robot therapy in dementia care. *Proc IEEE* 2012; 100(8): 2527–2538.
16. Wang RH, Sudhama A, Begum M, et al. Robots to assist daily activities: views of older adults with alzheimer's disease and their caregivers. *Int Psychogeriatrics* 2017; 29(1): 67–79.
17. Loveys K, Fricchione G, Kolappa K, et al. Reducing patient loneliness with artificial agents: design insights from evolutionary neuropsychiatry. *J Medical Internet Research* 2019; 21(7): e13664.
18. Pirhonen J, Tiilikainen E, Pekkarinen S, et al. Can robots tackle late-life loneliness? Scanning of future opportunities and challenges in assisted living facilities. *Futures* 2020; 124: 102640.
19. Cruz-Sandoval D, Favela J, Parra M et al. Towards an adaptive conversational robot using biosignals. In Proceedings of the 7th Mexican Conference on Human-Computer Interaction, Merida, Mexico, 29–31 October 2018, pp. 1–6.
20. Rudzicz F, Wang R, Begum M, et al. Speech interaction with personal assistive robots supporting aging at home for individuals with alzheimer's disease. *ACM Trans Accessible Comput* 2015; 7(2): 1–22.
21. Cruz-Sandoval D and Favela J. Semi-autonomous conversational robot to deal with problematic behaviors from people with dementia. In: International conference on ubiquitous computing and ambient intelligence, Philadelphia, PA, USA, 7–10 November 2017, Springer, pp. 677–688.
22. Cruz-Sandoval D, Favela J and Sandoval EB. Strategies to facilitate the acceptance of a social robot by people with dementia. In: Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, Chicago, IL, USA, 5–8 March 2018, pp. 95–96.
23. Esteban P, Hernández García D, Lee HR et al. Social robots in therapy: Focusing on autonomy and ethical challenges. In: Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, Chicago, IL, USA, 5 March 2018, pp. 391–392.
24. Hemminahaus J and Kopp S. Towards adaptive social behavior generation for assistive robots using reinforcement learning. In: 2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE, Vienna, Austria, 6–9 March 2017, pp. 332–340.
25. Kobayashi M, Kosugi A, Takagi H, et al. Effects of age-related cognitive decline on elderly user interactions with

voice-based dialogue systems. In: IFIP Conference on Human-Computer Interaction, Paphos, Cyprus, 2–6 September 2019, Springer, pp. 53–74.
26. Baharudin AD, Din NC, Subramaniam P, et al. The associations between behavioral-psychological symptoms of dementia (bpsd) and coping strategy, burden of care and personality style among low-income caregivers of patients with dementia. *BMC Public Health* 2019; 19(4): 447.
27. Islam M, Mazumder M, Schwabe-Warf D, et al. Personality changes with dementia from the informant perspective: new data and meta-analysis. *J Am Med Directors Assoc* 2019; 20(2): 131–137.
28. Li J, Monroe W, Ritter A, et al. *Deep Reinforcement Learning for Dialogue Generation*. arXiv preprint arXiv:160601541 2016.
29. Cuayáhuitl H. Simpleds: a simple deep reinforcement learning dialogue system. In: *Dialogues with Social Robots*. Singapore: Springer, 2017, pp. 109–118.
30. Magyar J, Kobayashi M, Nishio S et al. Autonomous robotic dialogue system with reinforcement learning for elderlies with dementia. In: 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC). Bari, Italy, 6–9 October 2019, pp. 3416–3421.
31. Alzheimer's Association. Communication and alzheimer's. URL https://www.alz.org/help-support/caregiving/daily-care/communications, Accessed 24 November 2021.
32. Csikszentmihalyi M. *Beyond Boredom and Anxiety*. San Francisco, CA: Jossey-Bass, 2000.
33. Tapus A, Tapus C and Mataric MJ. The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia. In: 2009 IEEE international conference on rehabilitation robotics, Kyoto, Japan, 23–26 June 2009, pp. 924–929.
34. Sutton RS and Barto AG. *Reinforcement Learning: An Introduction*. MA: MIT press, 2018.
35. Cohen-Mansfield J, Dakheel-Ali M and Marx MS. Engagement in persons with dementia: the concept and its measurement. *Am Journal Geriatric Psychiatry* 2009; 17(4): 299–307.
36. Cruz-Sandoval D, Morales-Tellez A, Sandoval EB et al. A social robot as therapy facilitator in interventions to deal with dementia-related behavioral symptoms. In: Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, Cambridge, UK, 23–26 March 2020, pp. 161–169.
37. Pandey AK and Gelin R. A mass-produced sociable humanoid robot: pepper: the first machine of its kind. *IEEE Robotics Automation Mag* 2018; 25(3): 40–48.
38. Yuan F, Klavon E, Liu Z, et al. A systematic review of robotic rehabilitation for cognitive training. *Front Robotics AI* 2021; 8: 105.
39. Yuan F, Blackburn J, Condon C et al. A social robot-based psycho-educational program to enhance alzheimer's caregiver health. In: CHASE 2021: IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies, Washington D.C, 16–17 December 2021.
40. Schüssler S, Zuschnegg J, Paletta L, et al. The effects of a humanoid socially assistive robot versus tablet training on psychosocial and physical outcomes of persons with dementia: protocol for a mixed methods study. *JMIR Research Protocols* 2020; 9(2): e14927.
41. Chen H, Park HW and Breazeal C. Teaching and learning with children: impact of reciprocal peer learning with a social robot on children's learning and emotive engagement. *Comput Edu* 2020; 150: 103836.
42. Alimardani M and Hiraki K. Passive brain-computer interfaces for enhanced human-robot interaction. *Front Robotics AI* 2020; 7: 125.