2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 25-29, 2020, Las Vegas, NV, USA (Virtual)

Towards Conversational Interfaces and Visual Memory Representation for Social Robots helping the Elderly

Ifrah Idrees[1] and Stefanie Tellex[1]
[1]Dept. of Computer Science, Brown University, Providence, RI;

***Abstract*— Social robots have the potential to help the elderly through activities of daily living provided they can obtain enough situational awareness and build a rapport with their human partner. In the past, context-aware robots were able to gather information about the elderly's activities and the objects in the environments but lacked a natural conversational interface and a long-term memory representation for visual perception to support extended interactions. This paper provides proof of concept demonstrations of two use-cases of an autonomous mobile robot helping the elderly in the home. In the first, the robot helps the elder, especially with dementia, to find household objects, that they might have a hard time remembering. In the second, the robot engages in a dialog to guide the person through cooking a recipe. We are looking forward to participants in this workshop giving feedback on these use cases regarding the suitability of using a robot for such tasks and the potential failures of the proposed approach with the target user.**

I. INTRODUCTION

The older population, especially with dementia or cognitive impairment may have difficulty performing daily activities of living(ADLs) [15]. Robots can help the elderly live in their own homes and community more independently and safely by providing physical assistance or services as assistive social robots [8, 9]. Social robots have the potential to help by obtaining enough contextual-awareness and connecting with their human partner while having an engaging conversation.

Research in social robotics focuses on the social behaviors of robots in assisting the elderly [2, 6]. Systems have been proposed to engage the elderly for physical exercise, physiological therapy, and cognitive training [1]. However, less emphasis is placed on language grounding and more on using non-verbal cues. Social robots have also focused less on visual memory representation. Additionally, this work typically assumes that the robot does not have detailed information about the person's current actions, the semantics of their activities, and objects in the environment.

To address these problems, we propose an approach that exploits the enhanced perception capabilities made available from deep learning approaches to enable the robot to 1) guide the person through the activities of daily life based on being able to sense their current activity and 2) find objects in the environment based on its ability to perceive those objects and find those instances in its long-term memory. In particular, we deploy our prototypes on the social and interactive Kuri robot from Mayfield and hope to analyze the usefulness of these systems in assisting the elderly. [10].

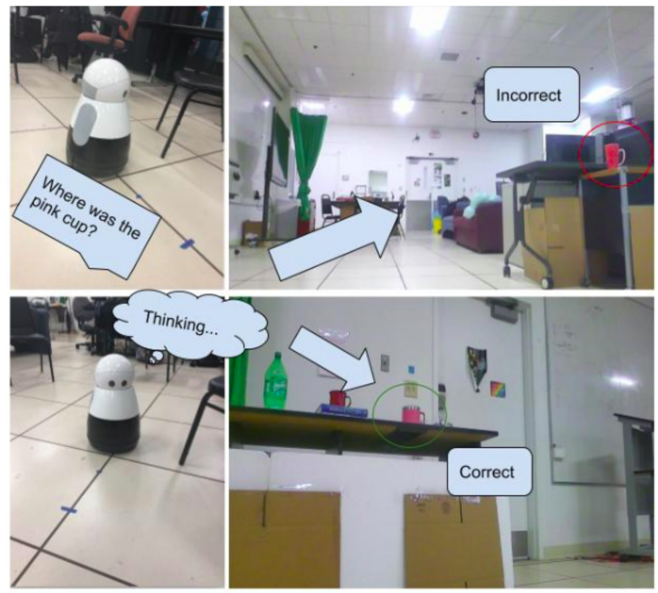

Fig. 1. Our mobile robot deployed with RoboMem monitors the environment for days. When asked about spatial-temporal queries about the objects, it has seen before; it returns a small set of candidate key frames, associated time intervals, and location in the map of the objects seen in the past. The first two candidate images are shown on the right. The top right image is the first ranked candidate image and does not contain our object of interest while the second-ranked image - bottom right correctly contains the object(pink cup)

In this paper, we present a proof of concept demonstration of these two prototypes. Both of these prototypes give a sense of what is possible. We are particularly looking forward to participants in this workshop, giving us feedback on these use cases.

II. USE CASES FOR OUR PROPOSED FRAMEWORK

We examine the following two use cases for situationally-aware home-service robots, which were decided on after careful consideration for the elderly's needs [11].

A. *Finding Objects*

A robot can help the elder find the lost objects that it has seen in the past by providing spatial-temporal information of the objects or navigating to the candidate places. For this task, we will allow the robot to roam the home environment of the user. We propose a pipeline for processing visual sensory input and a memory representation for a robot to perform correspondence of the object instances that the robot sees over time. Our prototype takes in sensory information - RGB, depth images, map, and pose estimates from the robot as it patrols the environment. Our algorithm then performs multi-tier processing using weakly supervised learning to identify unique object instances in the environment. These detections and their corresponding spatial-temporal information are then organized in a key-value database. This representation enables the robot to remember where the objects were and can be used to find the objects that the elderly might have lost over extended periods of time. A portion of this work has been published at MoRobAE - Mobile Robot Assistants for the Elderly workshop at ICRA 2019 [4]. Fig.1 shows the potential candidate images returned by our prototype as a response to a query of "Where is the pink cup?" asked by a human. We believe that this will be a useful application for a robotic assistant for the elderly and acknowledge the associated ethical concerns related to privacy. In the future work, we intend to get more insights into the effectiveness and concerns raised for this application from the stakeholders involved.

B. Engaging in dialogue with the elderly to guide them through cooking activity

The elderly with cognitive impairment have difficulty following and keeping track of the sequence of instructions involved in completing any activity [12]. To help an elder complete their cooking task safely; we propose a multi-modal dialog manager for the robot that engages with the elderly and guides the elderly through a cooking recipe while adapting to the actions of the elderly and the current state of the cooking recipe. The prototype's current implementation assumes to take as input binary results from inference models instead of performing inference itself. We plan to apply CNN's and change point detection algorithms to detect activities from a combination of sensors, if not all including RGB, depth camera (on the robot), and/or smart home devices such as a heat sensor or a situated camera on top of the stove. We implement a finite state machine (FSM) for a set of selected recipes such as boiling an egg, making avocado toast. The FSM integrates the inferences about the elder's action and the state of the recipe with the dialogue history and grounds the current user language input. An action policy is then learned to select a response and associated speech for the robot. Portions of this work were previously published at RoboDial, part of SIGDIAL 2020 [3]. Fig.2 shows a sample dialogue excerpt generated by the prototype invoking physical context to be inferred and checking the safety of the given recipe instruction. We believe that there is much potential for a sensor-grounded language model for robots that can carry out context-aware conversations to help the elderly through ADLs. We anticipate challenges such as hearing difficulties from the target user and hope to mitigate them in the future work.

III. Related work

This section will discuss the following lines of past research in robotics that make a technical contribution to the non-verbal front. All of these work pay less attention to the social implications of bringing robotics for elderly care outside the labs.

Robot	**Great, Next step in cooking is: Heat the pot on high heat and bring the water to a full rolling boil Do you want me to track to see if the activity is being completed safely or will you tell me when you are done with this step?**
User	**I will tell you when I am done**
Robot	**Yes you can surely complete the activity on your own, but for your safety let me monitor you and see when the activity gets completed**
User	**Hmm, okay you can monitor me**
Robot	**Sure, for your safety I will monitor you and see when the activity gets completed**
User	**Okay. I am proceeding with the step**
Robot	**Good job, you have completed the step safely.**

Fig. 2. Segment of Sample Dialogue - the pink represents the user, and blue represents the robot

Social Robotics: A great deal of research has been done on developing social behaviors for the robots for cognitive training and psychological therapy of the elderly, focusing on user motivation, engagement, and personalization [6, 1, 14]. These works focus on making communication better with non-verbal cues such as embodiment, gesture. An example is a robot exercise coach that measures its effectiveness on the elderly [2]. In most studies, the social robot is hard-coded with conversational cues, not focusing on a dialogue manager combining the physical and environmental context with learning an action policy to generate a natural language response. Pavez et al. [13] discusses the need of long-term memory for social robots with less focus on the application of finding lost objects.

Conversational Social Robots: Relevant examples include Mabu by Kidd et al. [5] and Koichiro et al. [7]. The former is a startup producing non-mobile robots that engage with the patients to address treatment adherence challenges. However, the user's input is not a natural language utterance; rather, users select input from provided prompts. On the other hand, the latter dialogue manager has a conversation interface to handle the elderly's loneliness but is not task-oriented or targeted to the ADL.

IV. Conclusion

In this paper, we show qualitative demonstrations of two use cases of a social robot helping the elderly in their everyday life. The first is to assist the elderly in finding the lost objects in their household, and the second one is to engage the elderly guiding them through cooking. We are particularly interested in getting feedback from the workshop participants about the possible failures of the proposed technology that they can foresee with the two presented use cases in mind. We anticipate ethical concerns of using robotic visual sensors as input and would like to learn potential alternatives for mitigating them. We are looking forward to getting insights from the healthcare robotics community that can be incorporated in future work for the long-term deployment of social robots at home to assist the elderly in their daily living activities.

REFERENCES

[1] Jordan Abdi, Ahmed Al-Hindawi, Tiffany Ng, and Marcela P Vizcaychipi. Scoping review on the use of socially assistive robot technology in elderly care. *BMJ open*, 8(2), 2018.

[2] Juan Fasola and Maja J Matarić. A socially assistive robot exercise coach for the elderly. *Journal of Human-Robot Interaction*, 2(2):3–32, 2013.

[3] Ifrah Idrees, Stefanie Tellex, and Momotaz Begum. Grounding human-robot dialogue in environment sensors to help elderly in activities of daily living.

[4] Ifrah Idrees, Steven P Reiss, and Stefanie Tellex. Robomem: Giving long term memory to robots. *arXiv preprint arXiv:2003.10553*, 2020.

[5] Cory Kidd, Devon Edwards, Gary Arnold, Brian Mirletz, and Brien Voorhees. Method and system for patient engagement. URL https://patents.google.com/patent/US10452816B2/en.

[6] Cory D Kidd and Cynthia Breazeal. Robots at home: Understanding long-term human-robot interaction. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3230–3235. IEEE, 2008.

[7] Yoshino Koichiro, Murase Yukitoshi, Lubis Nurul, Sugiyama Kyoshiro, Tanaka Hiroki, Sakti Sakriani, Takamichi Shinnosuke, and Nakamura Satoshi. Spoken dialogue robot for watching daily life of elderly people. In *International Workshop on Spoken Dialogue System Technology*, 2019.

[8] Mikaela Law, Craig Sutherland, Ho Seok Ahn, Bruce A MacDonald, Kathy Peri, Deborah L Johanson, Dina-Sara Vajsakovic, Ngaire Kerse, and Elizabeth Broadbent. Developing assistive robots for people with mild cognitive impairment and mild dementia: a qualitative study with older adults and experts in aged care. *BMJ open*, 9(9):e031937, 2019.

[9] Ester Martínez and Angel P. del Pobil. Personal robot assistants for elderly care: An overview. In *Personal Assistants: Emerging Computational Technologies*, pages 77–91. Springer, 01 2018. ISBN 978-3-319-62529-4.

[10] Robotics Mayfield. Meet kuri! the adorable home robot, 2018. URL https://www.heykuri.com/.

[11] C. Mucchiani, S. Sharma, M. Johnson, J. Sefcik, N. Vivio, J. Huang, P. Cacchione, M. Johnson, R. Rai, A. Canoso, T. Lau, and M. Yim. Evaluating older adults' interaction with a mobile assistive robot. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 840–847, Sep. 2017. doi: 10.1109/IROS.2017.8202246.

[12] The National Institute on Aging (NIA) website. Do memory problems always mean alzheimer's disease?, 2018. URL https://www.nia.nih.gov/health/do-memory-problems-always-mean-alzheimers-disease.

[13] Matías Pavez, Javier Ruiz-del-Solar, Victoria Amo, and Felix Meyer zu Driehausen. Towards long-term memory for social robots: Proposing a new challenge for the robocup@home league. *CoRR*, abs/1811.10758, 2018. URL http://arxiv.org/abs/1811.10758.

[14] Candace L. Sidner, Timothy Bickmore, Bahador Nooraie, Charles Rich, Lazlo Ring, Mahni Shayganfar, and Laura Vardoulakis. Creating new technologies for companionable agents to support isolated older adults. *ACM Trans. Interact. Intell. Syst.*, 8(3), July 2018. ISSN 2160-6455. doi: 10.1145/3213050. URL https://doi.org/10.1145/3213050.

[15] Healthcare Sova. How does dementia affect everyday life?, 2018. URL https://www.sovahealthcare.co.uk/blog/post/how-does-dementia-affect-everyday-life.