Development of RobotHub: Integration of external system to group conversation system for older adults

S. Tokunaga, K. Fukumori, K. Tamura, K. Inoue, T. Kawahara, M. Otake-Matsuura

Purpose Towards an aging society, there becomes much more critical to conduct social activities such as conversations for older adults to keep their cognitive function in their lives (Livingstone et al. 2017). Along with this social background, our research group has conducted a group conversation experiment for older adults to verify cognitive training with a randomized controlled trial (Otake-Matsuura et al., 2021). The group conversation method is called Coimagination method, which is designed to have a talk in a well-balanced manner with a specialized system (named Fonopane) and moderator robot Bono-05. Because of the COVID-19 situation, those face-to-face experiments have become challenging to conduct. Therefore, we need to extend system to overcome the above physical limitations by interfacing the system to a teleconferencing system such as Zoom. Active listening function is desirable in such a system, since active listening is an essential factor in addressing smooth conversation for older adults and providing rich human-computer interaction (Inoue et al., 2020). In general, integration of different systems has some technical challenges, such as different communication protocols. In addition, the existing active listening system has developed focusing a one-to-one style with humans and AI; hence, there are technical problems for integrating such a one-to-one style system into a multi-party group conversation system. In this paper, as a first step, we integrate system to encourage new fashion to facilitate group conversation for older adults with remote. **Method** We cope problems above with a step-by-step approach. As a first step, we designed to integrate the existing active listening systems which are used in an autonomous Android ERICA (Inoue et al., 2016). We have developed a system component (RobotHub) that allows each participant in a group conversation to activate multiple systems and link them to existing systems. The system runs on a cloud server and client devices with four components; FrontEndAudioExtraction, CloudAudioGateway, FonopaneRemoteBridge, and SystemDataMapper, as shown in Figure 1. Specifically, audio data is not obtained from the videoconferencing system directly, but from a client application, FrontEndAudioExtraction. The videoconferencing system interoperates with systems developed in the role of using only conventional functions, such as sharing output with screen and voice. FrontEndAudioExtraction is a system that converts the elderly's voices into power (namely, audio data) and then sends these audio data to the cloud server from their home. The system is designed to deploy on the client devices such as laptops or tablets for individual participants. It is designed easy enough to avoid cumbersome operation for people who are not good at devices, especially older adults. CloudAudioGateway is deployed on the cloud server to properly resupply the audio data sent from the client devices to another component inside the cloud server. FonopaneRemoteBridge converts the communication protocol with an existing system and an external system. SystemDataMapper is designed to assign correct relationships between microphones and participants. We also carefully designed to integrate systems to prevent with the sub-modules so that actions/events would not conflict with existing systems and external systems. In addition, we have designed and developed a system to migrate from one-to-one mapping style to one-to-multiple users. We also deployed our system on the cloud to provide experiments to release the physical restriction. **Results and Discussion** We have integrated the existing group conversation systems to external systems. We have preliminarily tested on laboratory members on Zoom (See Figure 2). RobotHub and other systems are now running on the cloud, so we can experiment anywhere as long as Zoom and FrontEndAudioExtraction are set up. RobotHub manages the utterance status of all participants. Therefore, even when N participants speak as they wish, the AI can detect and issue a response to the speaker who should be most likely to give a nodding response. We plan to use this mechanism in the future to conduct usability and subjective evaluation of older adults in conversational situations.

References

Livingston, G., Sommerlad, A., Orgeta, V., et al. (2017). Dementia prevention, intervention, and care. Lancet 390 (10113), 2673–2734. doi:10.1016/s0140-6736(17)31363-6

Otake-Matsuura M., Tokunaga S., Watanabe K., et al. (2021). Cognitive Intervention Through Photo-Integrated Conversation Moderated by Robots (PICMOR) Program: A Randomized Controlled Trial, Frontiers in Robotics and AI, https://doi.org/10.3389/frobt.2021.633076

Inoue, K., Lala, D., Yamamoto, K. et al. (2020). An attentive listening system with android ERICA: Comparison of autonomous and WOZ interactions. In Proc. SIGdial Meeting Discourse and Dialogue, pp.118--127

Inoue, K., Milhorat, P., Lala, D., Zhao, T. et al. Talking with ERICA, an autonomous android. In Proc. 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue, pp.212–215, 2016

Keywords: group moderated robot, robots, home-based experiments

Address: RIKEN, Center for Advanced Intelligence Project, Japan
Email: seiki.tokunaga@riken.jp

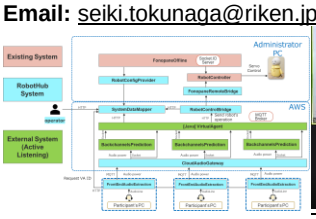

Figure 1. System Architecture of RobotHub

Figure 2. Screenshot of System Use-case on Zoom

RobotHub system and the external system determine whether to let the robot nod based on the participants' speaking status.