Adaptable Multimodal Interaction Framework for Robot-Assisted Cognitive Training

Aleksandar Taranović
Institut de Robòtica i Informàtica Industrial, CSIC-UPC
Barcelona, Spain
ataranovic@iri.upc.edu

Aleksandar Jevtić
Institut de Robòtica i Informàtica Industrial, CSIC-UPC
Barcelona, Spain
ajevtic@iri.upc.edu

Carme Torras
Institut de Robòtica i Informàtica Industrial, CSIC-UPC
Barcelona, Spain
torras@iri.upc.edu

ABSTRACT

The size of the population with cognitive impairment is increasing worldwide, and socially assistive robotics offers a solution to the growing demand for professional carers. Adaptation to users generates more natural, human-like behavior that may be crucial for a wider robot acceptance. The focus of this work is on robot-assisted cognitive training of the patients that suffer from mild cognitive impairment (MCI) or Alzheimer. We propose a framework that adjusts the level of robot assistance and the way the robot actions are executed, according to the user input. The actions can be performed using any of the following modalities: speech, gesture, and display, or their combination. The choice of modalities depends on the availability of the required resources. The memory state of the user was implemented as a Hidden Markov Model, and it was used to determine the level of robot assistance. A pilot user study was performed to evaluate the effects of the proposed framework on the quality of interaction with the robot.

KEYWORDS

Cognitive Training, Multimodal Interaction, Assistive Robotics

ACM Reference Format:

Aleksandar Taranović, Aleksandar Jevtić, and Carme Torras. 2018. Adaptable Multimodal Interaction Framework for Robot-Assisted Cognitive Training. In *HRI '18 Companion: 2018 ACM/IEEE International Conference on Human-Robot Interaction Companion, March 5–8, 2018, Chicago, IL, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3173386.3176911

1 INTRODUCTION AND RELEVANT WORK

The average life expectancy has increased significantly and consequently has the number of senior citizens who often suffer from reduced cognitive abilities, which affects their ability to live independently and makes them reliant on professional care. The focus of our work is the development of a socially assistive robot (SAR) for cognitive training that helps patients with mild cognitive impairment (MCI). A study shows that people prefer interacting with robots than with virtual agents when receiving healthcare instructions [2]. Moreover, some examples of SAR for cognitive training

HRI '18 Companion, March 5–8, 2018, Chicago, IL, USA

ACM ISBN 978-1-4503-5615-2/18/03.
https://doi.org/10.1145/3173386.3176911

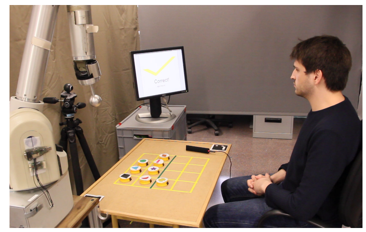

Figure 1: Sequential memory exercise setup

have already been proposed in literature [4]. In this work, we propose a sequential memory exercise scenario (Fig. 1). The user and scenario requirements were developed in collaboration with the medical staff from Fundació ACE, Barcelona Alzheimer Treatment & Research Center, to ensure the transferability of the results to the everyday practice with the patients. The robot's primary role is to help the user to perform the exercise correctly. Depending on the user's performance, the robot can adjust its action; for example, it can confirm a correct move or provide assistance in case of error. Besides, the robot has an algorithm to select and use only modalities with appropriate resources. We proposed a framework that integrates these mechanisms in order to produce adaptive robot behavior.

2 RESEARCH APPROACH AND METHODOLOGY

The proposed framework adapts to user input in two ways. First, it provides adaptive assistance using a Hidden Markov Model (HMM), and secondly, it executes actions using modalities with required resources. One of the major features of the framework (Fig. 2) is modality transfer, i.e., it uses modalities to perform the action depending on the availability of their respective resources [1]. The proposed sequential memory exercise consists in sorting the shapes on the board in a specified order. Because direct measuring of user's memory is impossible, we decided to estimate the probability that the user remembers a sequence using an HMM. This approach is inspired by the application of HMM in robotic tutoring systems [3]. In our sequential memory exercise, the unobservable variables correspond to probabilities that the user remembers a particular shape. The level of assistance that the robot provides depends on the estimated probability. User action is the observation, and it can be: making the correct guess, making a wrong guess, requesting help and reaching a time limit without providing any guess.

The robot can adjust the level of assistance. Besides confirming the correct guess, the robot can provide hints that guide the user towards the right solution. Thus, by not immediately providing the right answer, the robot stimulates user's effort to recall the shape

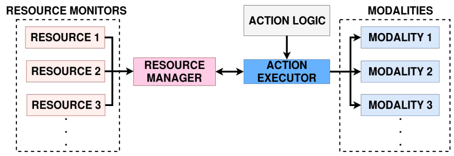

Figure 2: Resource-based modality selection

order. We defined three levels of robot assistance: high, medium and low. At the high level, the robot discloses the correct shape to the user. At the medium level, the robot shows two shapes of which one is correct. Finally, at the lowest level of assistance, the robot just indicates that the user made a mistake. This level of assistance is not considered when the user asks for assistance.

Robot actions are task-specific. In the proposed scenario, the actions are aimed at assisting the user to perform the correct move. Modalities represent the communication channels between the user and the robot, but also define a way robot actions are executed. Modalities require resources, which can be cognitive (user attention), or physical (board space, or speaking floor). For example, in case of user attention, the robot will not use the screen if the user is not looking at it. On the other hand, the robot will not point to a shape while the user's hand is inside the board space. The same goes for the speaking floor: the robot will not speak when the user is speaking. The possibility to freely define a resource required by a modality provides flexibility in the design of robotic systems. As shown in the diagram of the proposed resource-based modality selection algorithm in Fig. 2, the resource manager (RM) is informed by the resource monitors about the availability of their corresponding resources. RM informs the action executor (AE) about the state of all the resources, while AE communicates to RM requests for resources from all modalities. Action logic (AL) determines what action will be executed and when. In our scenario, AL is defined by the adaptive assistance and the exercise rules.

One important feature of the resource-based modality selection is the robot's ability to independently interrupt the execution of each modality if the required resource suddenly becomes unavailable. This gives the robot a human-like reactive behavior, such as stopping to speak when the user starts speaking or pulling back the arm when the user moves inside the board space to grasp a token. It is important to note that the proposed framework enables transfer of actions between the modalities according to the resources that are available to the robot, which ensures robustness in performing an action. Moreover, the framework is scalable because it allows adding new modalities and actions, and redefining their relations, i.e., what modalities will be used to perform each action. Even reprogramming the robot to perform a new task can be achieved by defining a task-specific set of modalities and associated actions, which allows the generalization of results.

3 RESULTS AND REMAINING WORK

At the beginning of the exercise, the robot displayed the shapes on the screen, one by one, in a random order. After that, the user could start to sort the shapes on the designated squares in the lower section of the board. The robot provided feedback after each move, either by informing the user that the move was correct or providing assistance in case of error. Additionally, participants were allowed to verbally request assistance. Both a request for assistance and an incorrect move were time-penalized; however, the penalty for an incorrect move was higher (30 s) than for a requested assistance (15 s). Thus, motivating the users to engage in the interaction with the robot. The available robot actions and associated modalities are shown in Table 1. High, medium and low level of assistance correspond to complete help, partial help, and wrong choice actions, respectively. Three modalities, speech, gesture, and display, relied on the availability of their corresponding resources, speaking floor, board space, and user attention, respectively.

Table 1: Robot actions and associated modalities

Action	Modality	Description
Wrong choice	Display	Show "Wrong" and a X mark
	Speaker	"Sorry, but that is wrong."
	Gesture	"Negative" gesture
Partial help	Display	Show the correct and a wrong shape
	Speaker	"It's either $Shape1 or $Shape2."
	Gesture	Point to correct and a wrong shape
Complete help	Display	Show the correct shape
	Speaker	"Correct shape is $Shape."
	Gesture	Point towards the correct shape
Correct choice	Display	Show "Correct" and a check mark
	Speaker	"Correct, next move."
	Gesture	Confirmatory gesture

The pilot study was performed with six participants, of ages between 24 and 32. The focus of the preliminary experiments was on the adaptive use of modalities, while the robot assistance was reduced to the high level. We compared the implementation of the proposed resource-based framework with the use of all the modalities regardless of the user input (baseline system). To evaluate the user behavior and interaction quality, we tracked their performance, usage of modalities and we asked the participants to fill in a Likert scale questionnaire. Users performed slightly better when interacting with the robot with the resource-based modality selection. It should be noted that the experiments were performed with healthy participants, and some participants found the exercise easy to do and minimally interacted with the robot.

The preliminary results will allow us to identify user profiles in order to develop a personalized interaction with the robot, for example, to learn about the user's preferred interaction modality. It will also allow us to propose personalized assistance strategies on how to improve individual user performance. In the future, we plan to perform studies with patients with MCI.

ACKNOWLEDGMENTS

This work was supported by the SOCRATES project funded from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 721619.

REFERENCES

[1] Crystal Chao and Andrea Thomaz. 2016. Timed Petri nets for fluent turn-taking over multimodal interaction resources in human-robot collaboration. *The International Journal of Robotics Research* 35, 11 (2016), 1330–1353.
[2] Jordan A. Mann, Bruce A. MacDonald, I.-Han Kuo, Xingyan Li, and Elizabeth Broadbent. 2015. People respond better to robots than computer tablets delivering healthcare instructions. *Computers in Human Behavior* 43 (2015), 112 – 117.
[3] Thorsten Schodde, Kirsten Bergmann, and Stefan Kopp. 2017. Adaptive Robot Language Tutoring Based on Bayesian Knowledge Tracing and Predictive Decision-Making. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI '17)*. 128–136.
[4] A. Tapus, C. Tapus, and M. J. Mataric. 2009. 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*. 924–929.