Enhancing Social Human-Robot Interaction with Deep Reinforcement Learning

Neziha Akalin, Andrey Kiselev, Annica Kristoffersson, Amy Loutfi
Örebro University, SE-701 82, Örebro, Sweden
neziha.akalin@oru.se, andrey.kiselev@oru.se, annica.kristoffersson@oru.se, amy.loutfi@oru.se

Abstract

This research aims to develop an autonomous social robot for elderly individuals. The robot will learn from the interaction and change its behaviors in order to enhance the interaction and improve user experience. For this purpose, we aim to use Deep Reinforcement Learning. The robot will observe the user's verbal and nonverbal social cues by using its camera and microphone, the reward will be positive valence and engagement of the user.

1 Introduction

In the effort to support elderly people in their domestic environments, to preserve their independence and to relieve the burden of caregivers, social robots have great potential. As the number of interactions with robots has been increasing, it is becoming important to understand how people perceive and feel about potential encounters with social robots. The manner in which the robot behaves during the interaction with a human may affect the human's perception, well-being, the sense of support and security, and willingness to interact. It has not yet been apparent how a robot should behave to achieve natural communication and to result in a safe and secure relationship between a robot and an elderly person.

In this study, the aim is to find an evaluation method for the quality of interaction with a focus on sense of safety and security in social human-robot interaction (sHRI), especially for elderly people. We aim to use Deep Reinforcement Learning (DRL) which will provide an adaptive system in which the robot learns through the interaction and adapts its behavior in order to obtain high quality interaction and keep its user feeling safe and secure.

2 Related Work

There are several recent works whose main focus is quality of interaction. [Castellano *et al.*, 2017] used machine learning methods for automatic estimation of quality of interaction by using game and social context based features where the scenario was playing chess with iCub. The features which were considered as dimensions of interaction quality were social engagement, help, friendship and presence. [Bensch *et al.*, 2017] whose aim was to understand the quality of interaction, indicated that interaction quality depends on static and dynamic properties of the involved humans, robots and environment. The authors also mentioned that the quality of interaction can be measured with a combination of performance metrics and to obtain high quality interaction, the robot's action should not depend only on the currently perceived data. Also the history, the robot's state, prior knowledge, and robot's general capabilities should be taken into account. On the other hand, robot behaviors also affect the quality of interaction. Another study [Wade *et al.*, 2011] presented quality of interaction during therapeutic sessions by investigating the role of various communication modalities during robot-guided motor task practice with post-stroke individuals (possible target group for our studies).

One of the dimensions of quality of interaction is sense of safety and security. It is not a well studied term though there are some similar terms in literature such as perceived safety [Bartneck *et al.*, 2009], psychological safety [Lasota *et al.*, 2014] and mental safety [Nonaka *et al.*, 2004]. Feeling safe and secure during the interaction is associated with comfort [Bartneck *et al.*, 2009; Lasota *et al.*, 2014], and emotions [Nonaka *et al.*, 2004].

3 Methodology Design

The aim of the approach presented in this paper is to enhance sHRI especially with elderly people. For this purpose, we aim to use deep reinforcement learning which is a revolutionary approach towards building autonomous systems. Deep Learning (DL) has the ability to perform automatic feature extraction from raw data and DRL introduces DL to approximate the optimal policy and/or optimal value functions [Arulkumaran *et al.*, 2017].

Most works using DRL so far were focused on video games, they achieved human level learning by using high dimensional visual data [Mnih *et al.*, 2015; Silver *et al.*, 2016]. Recently, however, a research group has begun to focus on the applicability of DRL in sHRI [Qureshi *et al.*, 2016; 2017; 2018].

In the current work, we aim to learn from interaction and adapt the robot's behavior. For that goal, we propose to use DRL where the input is the raw camera and microphone streams and the reward is valence and engagement of the user during the interaction to provide a customized behavior. The

10.21437/AI-MHRI.2018-12

targeted use cases for testing the system will include general elderly needs as well as games and entertainment. Some of these general elderly needs include: reminding them of taking the medication and medical analysis, encouraging for physical activities, providing contact with friends and family, giving advice about healthy eating and entertaining with games.

One of the experimental scenarios for the proposed method is summarized below:

Hans had a stroke two years ago, resulting in mobility and memory problems started after having had the stroke. He needs to train his muscles with physical exercises. He does physical exercises with the robot which also help his memory as he tries to remember the successive motions. If some days, he forgets to exercise, the robot approaches him, and suggests that he exercises. The robot observes him and changes the difficulty level of the exercises or suggests to stop/pause whenever it detects that Hans has difficulties or if pleasantness (valence) and engagement decrease dramatically between two exercises. The objective of the robot is to learn when to change difficulty level and pause or stop the exercising. The problem formulation for this example scenario is as follows:

States

The state space has three dimensions: (1) valence, (2) engagement, (3) current mode of the interaction (difficulty level, stopped, paused).

Actions

There are five actions that the robot can take: increasing the difficulty level, decreasing the difficulty level, not changing the difficulty level, pausing the exercising and stopping the exercising.

Reward

The reward will be positive valence and engagement of the user obtained from Affdex SDK [McDuff *et al.*, 2016] and OpenSmile [Eyben *et al.*, 2010]. Currently, we are conducting experiments to understand the importance each feature (valence and engagement) in this kind of scenario.

Components

The example scenario includes four primary components:

- A social robotic platform (currently we are using Pepper robot but the system will be able to use any Robot Operating System (ROS) compatible robotic platform)
- Affdex SDK [McDuff *et al.*, 2016] to analyze facial expressions in real time
- OpenSmile [Eyben *et al.*, 2010] to analyze voice in real time
- A Deep Reinforcement Learning architecture which gets input from camera and microphone and integrates affective information from Affdex and OpenSmile through ROS

The proposed DRL system will enhance the interaction by adapting the robot's behavior based on the user's pleasantness and engagement during the interaction. It will enable the robot to learn verbal and nonverbal parameters by using its camera and microphone and also adapt its behavior.

4 Future Work

This position paper presents the general outline of the planned work. Thus far, there are few studies using DRL in sHRI. The main challenges for our approach are the way in which data is collected, determining which behaviors to adapt and designing long enough user interactions to obtain enough data to be able use DL.

We plan to develop a robot which is capable of learning from interaction and adapting itself based on its observations and its user. This workshop will provide an opportunity to discuss with leading experts and gain a better understanding of the challenges for the proposed method.

References

[Arulkumaran *et al.*, 2017] Kai Arulkumaran, Marc Peter Deisenroth, Miles Brundage, and Anil Anthony Bharath. A brief survey of deep reinforcement learning. *arXiv preprint arXiv:1708.05866*, 2017.

[Bartneck *et al.*, 2009] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, 1(1):71–81, 2009.

[Bensch *et al.*, 2017] Suna Bensch, Aleksandar Jevtic, and Thomas Hellström. On interaction quality in human-robot interaction. In *ICAART 2017 Proceedings of the 9th International Conference on Agents and Artificial Intelligence, vol. 1*, pages 182–189. SciTePress, 2017.

[Castellano *et al.*, 2017] Ginevra Castellano, Iolanda Leite, and Ana Paiva. Detecting perceived quality of interaction with a robot using contextual features. *Autonomous Robots*, 41(5):1245–1261, 2017.

[Eyben *et al.*, 2010] Florian Eyben, Martin Wöllmer, and Björn Schuller. Opensmile: the munich versatile and fast open-source audio feature extractor. In *Proceedings of the 18th ACM international conference on Multimedia*, pages 1459–1462. ACM, 2010.

[Lasota *et al.*, 2014] Przemyslaw A Lasota, Gregory F Rossano, and Julie A Shah. Toward safe close-proximity human-robot interaction with standard industrial robots. In *Automation Science and Engineering (CASE), 2014 IEEE International Conference on*, pages 339–344. IEEE, 2014.

[McDuff *et al.*, 2016] Daniel McDuff, Abdelrahman Mahmoud, Mohammad Mavadati, May Amr, Jay Turcot, and Rana el Kaliouby. Affdex sdk: a cross-platform real-time multi-face expression recognition toolkit. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 3723–3726. ACM, 2016.

[Mnih *et al.*, 2015] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529, 2015.

[Nonaka *et al.*, 2004] Seri Nonaka, Kenji Inoue, Tatsuo Arai, and Yasushi Mae. Evaluation of human sense of security for coexisting robots using virtual reality. 1st report: evaluation of pick and place motion of humanoid robots. In *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, volume 3, pages 2770–2775. IEEE, 2004.

[Qureshi *et al.*, 2016] Ahmed Hussain Qureshi, Yutaka Nakamura, Yuichiro Yoshikawa, and Hiroshi Ishiguro. Robot gains social intelligence through multimodal deep reinforcement learning. In *Humanoid Robots (Humanoids), 2016 IEEE-RAS 16th International Conference on*, pages 745–751. IEEE, 2016.

[Qureshi *et al.*, 2017] Ahmed Hussain Qureshi, Yutaka Nakamura, Yuichiro Yoshikawa, and Hiroshi Ishiguro. Show, attend and interact: Perceivable human-robot social interaction through neural attention q-network. In *Robotics and Automation (ICRA), 2017 IEEE International Conference on*, pages 1639–1645. IEEE, 2017.

[Qureshi *et al.*, 2018] Ahmed Hussain Qureshi, Yutaka Nakamura, Yuichiro Yoshikawa, and Hiroshi Ishiguro. Intrinsically motivated reinforcement learning for human–robot interaction in the real-world. *Neural Networks*, 2018.

[Silver *et al.*, 2016] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

[Wade *et al.*, 2011] Eric Wade, Jonathan Dye, Ross Mead, and Maja J Matarić. Assessing the quality and quantity of social interaction in a socially assistive robot-guided therapeutic setting. In *Rehabilitation Robotics (ICORR), 2011 IEEE International Conference on*, pages 1–6. IEEE, 2011.