Socially-Aware Reinforcement Learning for Personalized Human-Robot Interaction

Doctoral Consortium

Hannes Ritschel
Human-Centered Multimedia, Augsburg University
Augsburg, Germany
ritschel@hcm-lab.de

ABSTRACT

This research in the context of Human-Robot Interaction explores how to tailor the robot behavior to the human's individual preferences in real-time. Algorithmically, Reinforcement Learning is the method of choice as it allows the robot to explore and learn autonomously. Instead of relying on task-related data, the proposed approach is primarily based on human social signals, which occur all the time and provide valuable information which cannot be extracted from the task itself. Including social signal data in the Reinforcement Learning framework enables us to adapt robot behavior depending on the current user behaviors without additional interaction.

KEYWORDS

Social Robotics; Human-Robot Interaction; Social Signals; Reinforcement Learning

ACM Reference Format:
Hannes Ritschel. 2018. Socially-Aware Reinforcement Learning for Personalized Human-Robot Interaction. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018*, IFAAMAS, 3 pages.

1 INTRODUCTION

Creating customized and individualized interaction is one important challenge in Human-Robot Interaction (HRI). Social robots with adaptive behavior based on Reinforcement Learning (RL) are used in different contexts, including long-term adaptation with focus on empathic supportive strategies [2], affective behavior of a tutoring robot [1], in the assistive domain for post-stroke rehabilitation [10] or children with autism [4]. In general, the goal is to tailor the robot's behavior to the human's preferences or needs in order to maintain user engagement and solve the task most efficiently. This requires to get feedback from the user to evaluate whether the robot's behavior is expedient or not. Often, this data comes from task-based information, e.g. the number of solved exercises.

This work explores how social signals can be integrated in the adaptation process. More specifically, it examines how they can be used to implement adaptation in combination with RL. As a result, findings about the possibilities and limits of this kind of adaptation process and how social signals can be used most effectively in selected HRI scenarios, are expected.

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden.

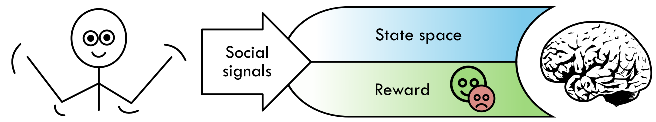

Figure 1: Options for using social signals with Reinforcement Learning.

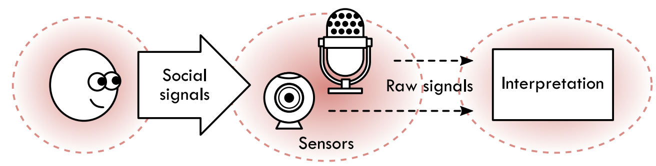

Figure 2: Different types of noise occuring when using social signals.

2 OVERVIEW

Generally, there are two options how to incorporate human social signals in the RL process (see figure 1). First of all, by extending the state space with social signal data, the agent is able to learn how to react in which situation, depending on the human's behaviors or state expressed via social signals (e.g. when the user shows low or high engagement or different emotions). In addition, this data is also an option to calculate the reward signal to give the robot feedback whether its actions are expedient or not. As an example, a decrease in user engagement may be undesirable and interpreted as a negative reward.

Combining social signals with RL is a promising approach with both advantages and challenges. One challenge is the problem of different kinds of noise (see figure 2). Since the human user can be interpreted as a nondeterministic environment, her or his reactions in terms of social signals do not need to be correlated with the actions that the robot executes and can vary from time to time, e.g. due to external influences. Moreover, the sensing hardware itself is subject to physical restrictions which limit the signals which can be perceived (e.g. camera field of view or resolution, etc.). When processing this data, interpreting the raw data often relies on machine learning itself and the result can only be an approximation of the actual user's behaviors. As a consequence, the data received for learning are noisy and allow to draw conclusions with regard to the user's intentions, needs or preferences only to a certain degree. Finally, the human's reaction to the robot's actions and behavior may vary from time to time as preferences may change, too.

Another challenge results from the fact that RL is based on trial and error. Initially, when the robot starts learning, it has to explore random actions as long as no initial knowledge is provided. However, learning "from scratch" in an HRI scenario should be avoided. Otherwise, the human would be confused or frustrated, when robot behavior is not consistent or transparent. One possible approach is the use of Wizard-of-Oz experiments to generate an initial policy, which is done e.g. for dialog systems [6]. Another option is to conduct a preliminary study to identify a robot behavior which is preferred by most of the participants. In both cases, the initial knowledge can serve as a reasonable starting behavior which then can be refined with real-time learning during interaction. Also connected to this issue is the question of how to model the learning process, specifically in which interaction context adaptation makes sense and what are appropriate actions for the RL process.

One advantage is the omnipresence of human social signals while interacting with the robot. Instead of asking the user directly, adaptation can be realized in the background based on the user's nonverbal signals. This is the most unobtrusive source of information available for adaptation. Moreover, since social signals occur frequently, they also provide the opportunity to optimize problem modelling by defining actions which are as short as possible to make the system learn faster.

Since human interaction takes much time and is very expensive from the machine learning perspective, simulations are the means to an end to compare the performance of different learning algorithms and to find an preliminary parameter setting. Simulations are performed before conducting a study to answer the question of whether the selected algorithm is suitable for dealing with potentially sparse and noisy social signals in nondeterministic environments. So far, learning algorithms with discrete state spaces, such as Q-Learning and SARSA [9], have been explored (see below). Since social signals occur continuously with different intensity, algorithms for continuous state spaces are of particular interest and will be examined in the future [11].

Apart from the algorithmic point of view, the question arises of which social signals can be used in which interaction context and how they can be used for learning. This is also the point where algorithmic simulations reach their limits and data from corpora of human-robot interaction has to be inspected. Different interaction scenarios serve as a basis to investigate these questions and to obtain new insights on social adaptation in the context of robot learning. A systematic overview of application scenarios and algorithms used for adaptive social robots and/or agents, as well as an overview of algorithms suitable for realizing adaptation, is essential, especially with regard to the challenges mentioned above.

3 COMPLETED WORK

So far, our research has focused on using human engagement in a storytelling scenario using Natural Language Generation (NLG) and RL to keep the user interested in the interaction [7, 8]. A *Reeti* robot acts as a story teller describing the main characters in the book "Alice in Wonderland". In parallel, it learns from feedback based on human engagement which personality it should express in terms of extraversion/introversion. While personality is generally a long-term trait, humans are able to a certain extent to portray a particular personality trait if required, e.g. to advance their projects [3]. Similar behavior could be used by social robots, e.g. to restore human attention and engagement. Figure 3 outlines the process, representing the general approach:

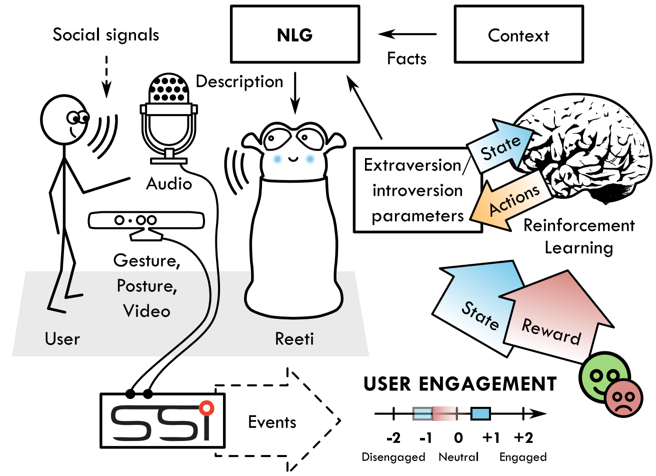

Figure 3: Manipulating robot extraversion based on social signals and RL.

(1) Sense human social signals and interpret the raw data. In our scenario, a Microsoft Kinect 2 sensor provides posture and video data which is used to estimate human engagement based on a dynamic bayesian network.
(2) Incorporate social signals in the learning process. Here, the current user engagement is part of the state space. In addition, the change of user engagement over time serves as reward (when increasing) or punishment (when decreasing) for RL. The combination of using engagement both as reward and as feature in the state space allows the robot to learn the optimal behavior depending on the current user engagement.
(3) RL manipulates the robot's behavior, driven by the reward based on social signals. This closes the loop between learnt robot behavior and human social signals/reactions. In our scenario, the robot's policy determines its expressed personality, which is reflected in utterances generated by NLG, following the approach by Mairesse and Walker [5].

4 FUTURE WORK

Current work complements the system described above with appropriate gaze behavior to include another modality to make the robot more expressive. We will evaluate whether the integration of natural language with dynamic gaze behavior will lead to a robot's behavior that is identified as extravert/introvert. Subsequently, we will investigate whether using engagement is a sensible option for such a scenario or not. The next experiments will explore algorithms based on continuous state spaces as well as learning humor preferences from vocal laughs and smile [11].

REFERENCES

[1] Goren Gordon, Samuel Spaulding, Jacqueline Kory Westlund, Jin Joo Lee, Luke Plummer, Marayna Martinez, Madhurima Das, and Cynthia Breazeal. 2016. Affective Personalization of a Social Robot Tutor for Children's Second Language Skills. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA.*, Dale Schuurmans and Michael P. Wellman (Eds.). AAAI Press, 3951–3957.

[2] Iolanda Leite, André Pereira, Ginevra Castellano, Samuel Mascarenhas, Carlos Martinho, and Ana Paiva. 2011. Modelling Empathy in Social Robotic Companions. In *Advances in User Modeling - UMAP 2011 Workshops, Girona, Spain, July 11-15, 2011, Revised Selected Papers (Lecture Notes in Computer Science)*, Liliana Ardissono and Tsvi Kuflik (Eds.), Vol. 7138. Springer, 135–147.

[3] Brian R. Little. 2008. Personal Projects and Free Traits: Personality and Motivation Reconsidered. *Social and Personality Psychology Compass* 2, 3 (2008), 1235–1254.

[4] Changchun Liu, Karla Conn, Nilanjan Sarkar, and Wendy Stone. 2008. Online Affect Detection and Robot Behavior Adaptation for Intervention of Children With Autism. *IEEE Trans. Robotics* 24, 4 (2008), 883–896.

[5] François Mairesse and Marilyn A. Walker. 2011. Controlling User Perceptions of Linguistic Style: Trainable Generation of Personality Traits. *Computational Linguistics* 37, 3 (2011), 455–488.

[6] Verena Rieser and Oliver Lemon. 2011. *Reinforcement Learning for Adaptive Dialogue Systems - A Data-driven Methodology for Dialogue Management and Natural Language Generation.* Springer.

[7] Hannes Ritschel and Elisabeth André. 2017. Real-Time Robot Personality Adaptation Based on Reinforcement Learning and Social Signals. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI '17)*. ACM, New York, NY, USA, 265–266.

[8] Hannes Ritschel, Tobias Baur, and Elisabeth André. 2017. Adapting a Robot's Linguistic Style Based on Socially-Aware Reinforcement Learning. In *Proceedings of the 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 378–384.

[9] Richard S. Sutton and Andrew G. Barto. 1998. *Reinforcement learning - an introduction.* MIT Press.

[10] Adriana Tapus, Cristian Tapus, and Maja J. Mataric. 2008. User - robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics* 1, 2 (2008), 169–183.

[11] Klaus Weber, Hannes Ritschel, Florian Lingenfelser, and Elisabeth André. 2018. Real-Time Adaptation of a Robotic Joke Teller based on Human Social Signals. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018)*. International Foundation for Autonomous Agents and Multiagent Systems.