Real-time robot personality adaptation based on reinforcement learning and social signals

Hannes Ritschel, Elisabeth André

Angaben zur Veröffentlichung / Publication details:

Ritschel, Hannes, 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, Vienna, Austria, March 06 - 09, 2017, edited by Bilge Mutlu, Manfred Tscheligi, Astrid Weiss, and James E. Young, 265–66. New York, NY: ACM Press. https://doi.org/10.1145/3029798.3038381.

Real-Time Robot Personality Adaptation based on Reinforcement Learning and Social Signals

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

Elisabeth André
Human Centered Multimedia
Augsburg University
Augsburg, Germany
andre@hcm-lab.de

ABSTRACT

Recent findings in Human-Robot Interaction (HRI) indicate that the adaptation of a robot's behaviors to the human's personality profile makes interaction more engaging, but also that it depends on the task context whether a similar or opposing robot personality is preferred. This late breaking report presents our ongoing work on an approach using Reinforcement Learning and social signals for figuring out and adapting to the human preferences, i.e. desired personality profile. Our scenario involves a "Reeti" robot in the role of a story teller talking about the main characters in the novel "Alice's Adventures in Wonderland" by generating descriptions with varying degree of introversion/extraversion. The learning process is running in real-time during the interaction and allows for simultaneous adaptation without explicitly asking the user about its preferences.

Keywords

Adaptation; Reinforcement Learning; Personality; introversion/extraversion; Social Robotics; Dialog; Social Signals

1. BACKGROUND AND PROTOTYPE

In HRI, implementing compelling personality in a social robot makes interaction more interesting, desirable and encourages the establishment of a relationship between user and robot [3]. Moreover, recent research points out that humans prefer robots with similar personality to their own and that the adaptation of a robot's personality to the human's profile makes interaction more engaging [1]. Other results indicate that it depends on the task context whether a similar or opposing personality is preferred [4]. In any case the existence of an adaptation process is important to address the user's needs and preferences, to customize and keep interaction engaging [7].

When implementing a new HRI application, we don't know which type of personality the robot should express. Therefore, we aim to develop an approach for figuring out and adapting to the user's personality preferences automatically without sticking to the similarity or complementary attraction principle. In our human-robot dialog prototype a "Reeti" robot describes and presents facts about the main characters in the novel "Alice's Adventures in Wonderland". Reeti's personality is expressed via linguistic style: utterances are not predefined, but generated automatically with varying amount of introversion/extraversion (one of the "Big Five" personality dimensions [6]). An adaptation process based on RL and social signals controls and optimizes the robot's personality to keep the user engaged.

DOI: http://dx.doi.org/10.1145/3029798.3038381

2. ADAPTATION PROCESS

Implementing adaptation in an HRI scenario requires to get information about the user's actual needs or preferences. We do not want to explicitly ask the user and thus cannot get this data directly. Moreover, depending on the task, we do not know whether the robot's personality should match these preferences or not. Therefore, our approach relies on social signals which are available during interaction anyway: we use the Social Signals Interpretation (SSI) framework [8] for estimating the user's current level of engagement based on the operationalization in [2]. Engagement serves us as an indication of whether Reeti's personality engages the user or not. Based on this estimation, RL is our algorithmic method of choice for manipulating the robot's personality and learning about the user's needs and preferences in a given scenario.

Figure 1 illustrates the interaction and adaptation process. SSI estimates the current level of user engagement E_t at time t based on the user's gestures, posture, video and audio. This value is discretized in an integer interval $[-2; +2]$ ("not engaged at all" to "very engaged").

The RL process uses this information for learning about the relationship between the robot's expressed personality and the user's engagement. Therefore, it manipulates the robot's current extraversion level X, which is in the integer interval $[-2; +2]$ ("very introverted" to "very extraverted"). Three actions allow to increase or decrease X by 1 as well as to leave X unchanged.

The state space includes the robot's current X value as well as the last sensed level of user engagement E_{t-1}. This allows to learn how to maximize engagement by increasing or decreasing X. In order to find the optimal value, the RL process uses the change of engagement $\Delta E_t = E_t - E_{t-1}$ between two sequent points in time $t-1$ and t as reward signal. For $\Delta E_t = 0$, the robot gets a small reward $+0.5$ for preventing a decrease of engagement.

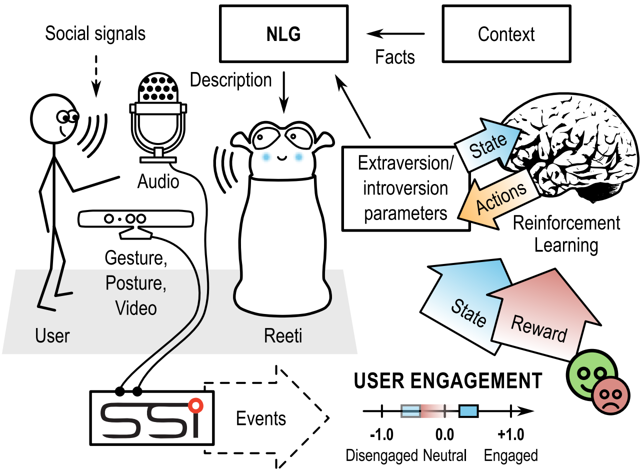

Figure 1: Prototype dialog scenario.

Introverted: *"The rabbit is sort of nervous."*

Neutral: *"The rabbit has a short tail. While he is nervous, he is quick. He has no name. He is ... he is the first character Alice encounters. He has no name."*

Extraverted: *"Did you say the rabbit? Ok, he is confident, on the other hand he is nervous. His tail is really short, also the rabbit has no name, mate. He is the first character Alice encounters and he is darn quick, isn't he?"*

Figure 2: Examples of generated descriptions.

In each time/learning step, one description of a story character is generated and presented by the robot. We rely on a generation approach inspired by Personage [5] to dynamically generate utterances with varying amount of extraversion. Facts are transformed into descriptions by taking a variety of style parameters from [5] into concern. Based on the extraversion level X, parameters like verbosity, restatements, etc., are set with some variation to generate appropriate utterances (for generated examples see figure 2). Generated texts are presented by the robot's Text-To-Speech (TTS) module to the human user.

3. SIMULATION AND RESULTS

Since our prototype is not completely finished at the time of this writing, we conducted initial experiments based on Q-Learning with ϵ-greedy exploration to simulate the learning process. We use exploration rate $\epsilon = 0.2$ (high enough for handling noise) and learning rate $\alpha = 0.5$ (low enough to not eliminate all previous knowledge in case of noise). The simulated user's engagement increases when the robot's personality matches the actual preferences and decreases otherwise. Since such a deterministic user behavior is far from being realistic, noise simulates random changes in user engagement as well as deviations of the sensed from the real E_t value in each learning step. Figure 3 plots the averaged reward for 30 trials. In each trial, the robot presents 30 descriptions (30 learning steps) to a simulated user with random personality preferences, the robot starts with neutral extraversion ($X = 0$) and an empty Q-Learning table. The initial random seed is the same for each noise level experiment.

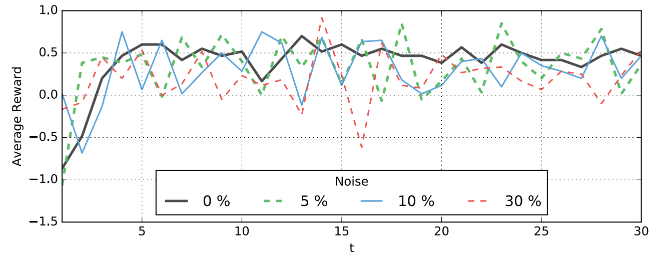

Figure 3: Initial Q-Learning results.

Without noise, learning is obviously quite robust. The average reward approaches 0.5: when the robot's extraversion level X equals to the user's preference, it learns to change X not anymore. Negative rewards can be attributed to exploration. Increasing noise leads to negative rewards occuring more frequently.

These results lead us to the conclusion that the adaptation process is able to learn correctly in the simulation. After completing the implementation of the prototype, we will focus on live experiments and optimization of the RL process with regard to sensor noise and exploration.

Acknowledgment

This research was funded by the Bavarian State Ministry for Education, Science and the Arts (STMWFK) as part of the ForGenderCare research association.

4. REFERENCES

[1] A. Aly and A. Tapus. Towards an intelligent system for generating an adapted verbal and nonverbal combined behavior in human–robot interaction. *Autonomous Robots*, 40(2):193–209, 2015.

[2] T. Baur, D. Schiller, and E. André. *Modeling User's Social Attitude in a Conversational System*, pages 181–199. Human–Computer Interaction Series. Springer International Publishing, 2016.

[3] C. Breazeal. *Designing sociable robots*. MIT press, 2004.

[4] M. Joosse, M. Lohse, J. G. Perez, and V. Evers. What you do is who you are: The role of task context in perceived social robot personality. *2013 IEEE International Conference on Robotics and Automation*, pages 2134–2139, 2013.

[5] F. Mairesse and M. A. Walker. Controlling user perceptions of linguistic style: Trainable generation of personality traits. *Computational Linguistics*, 37(3):455–488, 2011.

[6] R. R. McCrae and P. T. Costa. The five-factor theory of personality. *Handbook of personality: Theory and research*, 3:159–181, 2008.

[7] A. Tapus, M. Mataric, and B. Scassellati. Socially assistive robotics [grand challenges of robotics]. *IEEE Robotics & Automation Magazine*, 14(1):35–42, 2007.

[8] J. Wagner, F. Lingenfelser, T. Baur, I. Damian, F. Kistler, E. André, J. Wagner, F. Lingenfelser, T. Baur, I. Damian, F. Kistler, and E. André. The social signal interpretation (SSI) framework. *Proceedings of the 21st ACM international conference on Multimedia - MM '13*, pages 831–834, 2013.