Towards Active Learning for Socially Assistive Robots

Adriana Tapus Maja Matarić Department of Computer Science University of Southern California Interaction Lab Los Angeles, CA 90089 adriana.tapus@ieee.org; mataric@usc.edu

Introduction: The recent trend toward developing a new generation of robots capable of operating in human-centered environments, interacting with people, and participating and helping us in our daily lives, has introduced the need for robotic systems capable of learning to use their embodiment to communicate and to react to their users in a social and engaging way. Social robots that interact with humans have thus become an important focus of robotics research. Nevertheless, Human-Robot Interaction (HRI) for assistive applications is still in its infancy. In this work, the target user population is post-stroke patients. Stroke is the leading cause of serious, long-term disability among adults, with over 750,000 people suffering a new stroke each year in the US alone. Therefore, in this work we investigated the role of robot's active learning in the assistive therapy process. We tried to address the following research question: How should the behavior and encouragement of the therapist robot adapt as a function of the user's personality, profile, preferences, and task performance?

Methodology: Creating robotic systems capable of adapting their behavior to user personality, user preferences, and user profile in order to provide an engaging and motivating customized protocol, is a very difficult task, especially when working with vulnerable user populations. Various learning approaches for human-robot interaction have been proposed in the literature, but none include the user's profile, preferences, and/or personality. To the best of our knowledge, no work has yet tackled the issue of robot personality and behavior adaptation as a function of user personality in the assistive human-robot interaction context. In the work described here, we address those issues and propose a reinforcement-learning-based approach to robot behavior adaptation. In the learning approach, the robot incrementally adapts its behavior and thus its expressed personality as a function of the user's extroversion-introversion level and the amount of performed exercises, attempting to maximize that amount. We formulated the problem as policy gradient reinforcement learning (PGRL) and developed a learning algorithm that consists of the following steps: (a) parametrization of the behavior; (b) approximation of the gradient of the reward function in the parameter space; and (c) movement towards a local optimum. The main goal of our robot behavior adaptation system is to enable us to optimize on the fly the three main parameters (interaction distance/proxemics, speed, and verbal and paraverbal cues) that define the behavior (and thus personality) of the therapist robot, so as to adapt it to the user's personality and thus improve the user's task performance. Task performance is measured as the number of exercises performed in a given period of time; the learning system changes the robot's personality, expressed through the robot's behavior, in an attempt to maximize the task performance metric.

Experimental Results: Two experiments were designed to test the adaptability of the robot's behavior to the participant's personality and preferences. The

experimental task was a common object transfer task used in post-stroke rehabilitation and consisted of moving pencils from one bin on the left side of the participant to another bin on his/her right side. The bin on the right was on a scale in order to measure the user's task performance. The system monitored the number of exercises performed by the user. The robot used PGRL algorithm to adapt its behavior to match each participant's preferences in terms of therapy style, interaction distance, and movement speed. The learning algorithm was initialized with parameter values that were in the vicinity of what was thought to be acceptable for both extroverted and introverted individuals, based on our previous experiments [2] and psychology literature. The PGRL algorithm used in our experiments evaluated the performance of each policy over a period of 60 seconds. The reward function, which counted the number of exercises performed by the user in the last 15 seconds, was computed every second and the results over the 60 seconds "steady" period were averaged to provide the final evaluation for each policy. The result is a novel stroke rehabilitation tool that provides individualized and appropriately challenging/nurturing therapy style that measurably improves user task performance.

Discussion: Due to the large number of combinations of parameter values that have to be investigated during the adaptation phase the optimal policy might be obtained only after a period of time that exceeds our session of exercise (i.e., 15 minutes). However, we feel that this does not reduce the efficiency of our approach or the relevance of our results, as our research targets interaction with patients for an extended period of time and where many therapy sessions are required for complete rehabilitation. Thus, if the optimal policy is not reached during one therapy session the adaptation process can be extended over several sessions, with most of the interaction occurring with the optimal policy in place. In fact, this is very similar to real-life situations where therapists get to know patients over several therapy sessions and respond to their clues to provide a more efficient recovery environment.

Conclusions: We presented a non-contact therapist robot intended for monitoring, assisting, encouraging, and socially interacting with post-stroke users during rehabilitation exercises. The experimental results provide first evidence for the effectiveness of robot behavior adaptation to user personality and performance.

References

[1] Tapus, A., Tapus, C., and Matarić, M., J. (2008) User-Robot Personality Matching and Assistive Robot Behavior Adaptation for Post-Stroke Rehabilitation Therapy. *Journal on Intelligence Service Robotics* (to appear).

[2] Tapus, A., and Matarić, M., J. (2006) User personality matching with hands-off robot for post-stroke rehabilitation therapy. In *Proc. International Symposium on Experimental Robotics* (ISER'06), Rio de Janeiro, Brazil, July.