Proposal of a Cloud-based Agent for Social Human-Robot Interaction that Learns from the Human Experimenters

G. Magyar and M. Vircikova
Center for Intelligent Technologies,
Department of Cybernetics and Artificial Intelligence,
Faculty of Electrical Engineering and Informatics,
Technical University of Kosice, Slovakia

{Gergely.Magyar, Maria.Vircikova}@tuke.sk

Abstract—**The field of human-robot interaction usually uses an experimental technique called Wizard of Oz where a human operator (the experimenter or a confederate) remotely controls the behavior of the system. Per contra, if robots are autonomous during the interaction, they have a limited pre-programmed set of behaviors. We propose to use reinforcement learning for adaptation of autonomous robotic behavior during the interaction and to benefit from the advantages that brings the field of cloud computing. The overall goal is to design robotic behaviors less boring and more effective and thus, to prepare robots for a long-term human-robot interaction.**

I. Introduction

Research outcomes in the field of human-robot interaction prove the effectiveness of using social robots in different roles. The experiments have moved from laboratories to different kinds of real-world environments. Social robots support young patients in hospital as they learn to manage a lifelong metabolic disorder (diabetes) as demonstrated by Baxter et al. in [1]. Robots proved that they can to motivate physical exercise for older adults as work by Fasola et al.[2]. Using robots in autism therapy is increasing as well, where among the pioneers are Scassellati et al.[3] and Dautenhahn & Billard [4]. Bhargava et.al.[5] present an artificial embodied intelligent tutoring system that is capable of empathic conversations with school pupils. Another examples of using robots as tutors gives Deshmukh et al. [6], Mubin et.al [7] and Castellano et al. [8].

These and other researchers in social human-robot interaction highlight the need to make robots ready for long-term human-robot interaction (HRI). Building long-term relationship with robots study Kasap et al. [9], Walters et al. [10] and Leite et al. [11] review the research on long-term interaction between users and social robots.

The Wizard of Oz technique or simply, a teleoperation during the HRI sessions supplies the perspective of human-centered robotics in which the robots need to react appropriately to human expectations and behavior. On the other hand, the human teleoperator cannot be presented always, controlling the robot in the scenarios that are supposed to be long-term. This paper proposes a system based on artificial intelligence for social robots that can learn how to choose the proper behavior of the robot. First of all, we design a Wizard of Oz interface, which is used by human observing the experiment to control the robot. Then we propose a reinforcement-based learning algorithm, which adopts the Wizard's knowledge. In addition, we show two real-world examples of placing Nao robots in elementary schools.

From the perspective of artificial intelligence, we believe that this approach leads to the increase the machine intelligent quotient(MIQ). The three-dimensional model of machine intelligence was proposed by Bien et al. in [12] as seen on Figure 1. Regardless of the classes of intelligent machines, they expect that autonomy and human–machine interaction are the common components of intelligent machines.

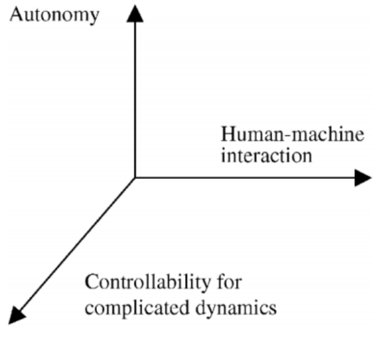

Fig. 1. Model space of machine intelligence for a large-scaled system as proposed by Bien et al. [12], where the controllability for complicated dynamics includes the behavior of the robot

978-1-4673-7939-7/15/$31.00©2015 IEEE

II. DESIGN OF THE CLOUD-BASED WIZARD OF OZ INTERFACE

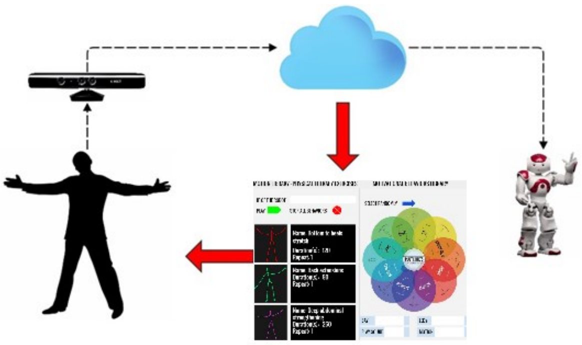

Fig. 2. Architecture of the system. User, Kinect, cloud environment, robot and GUI for the Wizard

Cloud Robotics, as defined by RoboEarth project [13] is an emerging field of robotics rooted in cloud computing, cloud storage, and other Internet technologies centered around the benefits of converged infrastructure and shared services. More on the topic of an adaptive cloud-based platform can be found in our previous work in [14].

The main goal of the proposed solution is to create a modular system that has a potential in different HRI scenarios able to serve as a common platform for researchers. It is cloud-based which means that it is always online and after connecting the robot to the system based on its IP address the users are ready to go. It allows robots to benefit from the powerful computational, storage, and communications resources of modern data centers. In addition, it removes overheads for maintenance and updates, and reduces dependence on custom middleware. From the HRI perspective we expect that the system can benefit from the knowledge of quasi unlimited number of human experts (Wizards).

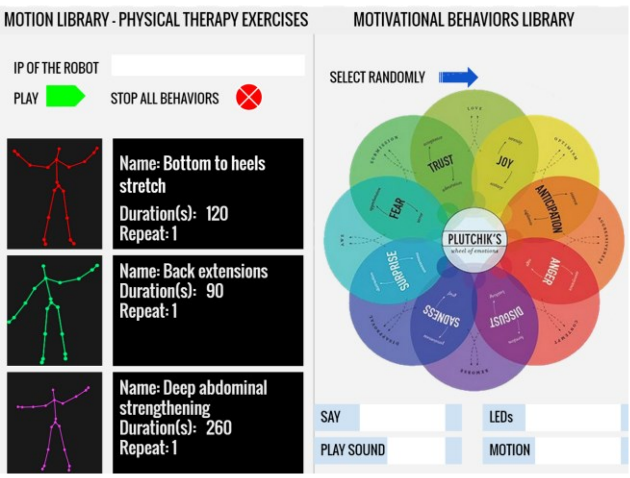

Fig. 3. Interface for the Wizard of Oz that can control the robot during the experiment

More in detail, our system consists of the following parts:

1) Motion Library

The Motion Library is a collection of motions for the humanoid robot that depend on the application. We show two examples:

1. 'Robot as a physiotherapist' scenario, in which it contains the physical therapy exercises and

2. 'Robot as a tutor of foreign language' in which it represents the motion-based gestures used by the robot when explaining the foreign language vocabulary.

The Wizard can choose the exercises/gestures from the database, number of repeats and set the order of execution. Another feature is recording new exercises with a Microsoft Kinect sensor, which enables the creation of more diverse rehabilitation sessions. This part of the system is dynamic and can be changed according to the given experiment.

2) Motivational behavior library

This one states for the emotional database that contains emotional expressions of the robot based on Plutchik's emotional model [15] (joy, satisfaction, anger, sadness, surprise, fear). The Wizard can also control the LED animations and the phrases said by the robot.

Next part of our research is to make an agent based on reinforcement learning. This states for a system that determines how to map situations to actions and also tries to maximize a numerical reward signal. In other words, how to set the teaching process, e.g. when to activate the motivational mode of the robot. The agent should be able to adapt according to the quality of a human-robot interaction. In this part of the work we were inspired by the work of Olsen and Goodrich [16]. We agree that a very important metric in measuring the autonomy of a robot with respect to some task (and corresponding task effectiveness metric) is the robot's neglect tolerance (NT). Neglect tolerance is a measure of how the robot's current task effectiveness declines over time when user neglects the robot. We can establish an acceptable minimum effectiveness threshold and using the characteristic neglect curve we can define the neglect tolerance as the time that can expire before the robot's effectiveness drops below the acceptable minimum.

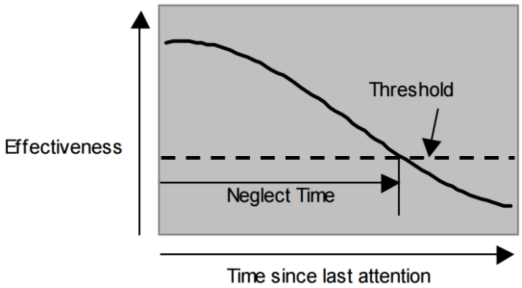

Fig. 4. Neglect time as proposed by Olsen & Goodrich [16]

III. Learning from the Wizard Proposal

In order to include human input to a machine learning process several approaches were designed, e.g. machine learns by observing human behavior, human explicitly directs action of the machine, human provides high-level evaluation, feedback or labels to a machine learner, etc.

According to Sutton and Barto[17], reinforcement learning is a learning method that determines how to map situations to actions and also tries to maximize a numerical reward signal. The actions performed by the agent are not defined explicitly, and they have to be discovered through exploration to get the most reward.

Traditional methods of reinforcement learning algorithms were used successfully in many application areas - however they were not primarily designed for learning from real-time social interaction from humans. This kind of learning addresses some additional challenges, such as dealing with limited human patience or ambiguous human input. The most important task of the designer of the algorithm is to ensure that the system learns the right thing at the right time [18]

Tapus et al. in [19] investigated the role of the robot's personality in the hands-off therapy process, focusing on the relationship between the level of extroversion-introversion of the robot and the user. They demonstrated a behavior adaptation system, using the Policy Gradient Reinforcement Learning. It is capable of adjusting its social interaction parameters (e.g., interaction distances/proxemics, speed, and vocal content) toward customized post-stroke rehabilitation therapy based on the user's personality traits and task performance.

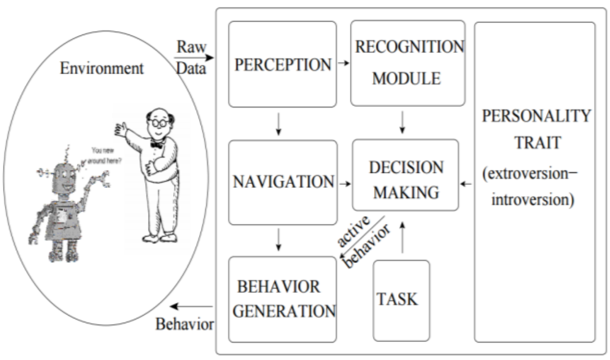

Fig. 5. HRI Information Processing using the Personality Model of the User as proposed by Tapus in [19]

In the case of learning from observation of the human behavior, the learning can occur implicitly or in other scenarios the human is explicitly teaching the machine a new skill. This type of approach was used in programming by example, learning by watching, and learning by demonstration [20] etc.

In other scenarios, the humans explicitly directed the action of the machine to provide them experience from which they can learn. In comparison with the above mentioned approach, this method is more interactive, but the human has to learn how he/she has to interact with the machine. The mechanism was used in various robot learning scenarios, such as learning navigation tasks by following a human demonstrator.

The third mechanism is used to influence the experience of the machine with higher level constructs such as giving feedback to a reinforcement learning agent or labeling examples in an active learning scenario.

In the above mentioned approaches, the main goal is to gain the learning performance of the machine through human inputs. Socially guided machine learning reframes the machine learning problem as an interaction between the human and the machine. In the case of the traditional supervised learning a human provides inputs to the machine learning algorithm that performs its task and provides output. Opposite to this is the setting of a social learning system, where the machine learning provides output but also interacts with the human teacher. Such a system is designed to learn efficiently from people with no experience in machine learning.

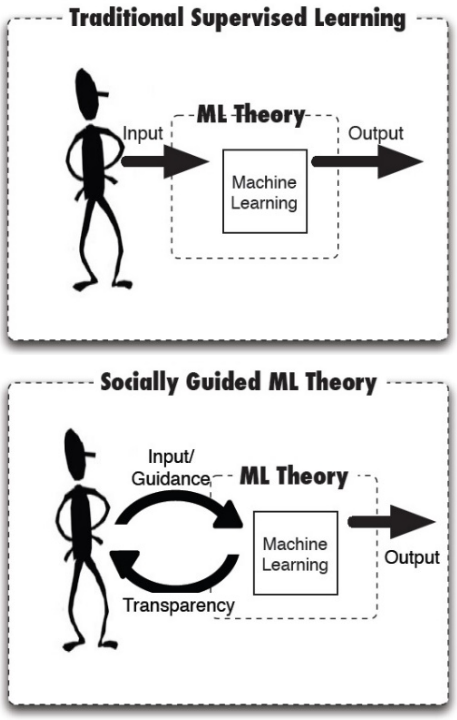

Fig. 6. Traditional supervised learning versus socially guided machine learning theory, [18]

This type of reinforcement learning is called interactive reinforcement learning and was successfully used in various scenarios, e.g. [22][23][24].

In our work, we aim to use this method for learning and adapting the robot's social behavior in order to move from the Wizard of Oz technique towards an autonomous mode. To achieve this, the learning has to consist of the following parts:

1. during the experiment the human operator (Wizard) controls the robot and based on the subject's non-verbal reactions (gestures, emotions) changes its behavior.
2. the states of the subject are labeled, and the associations between them and the actions of the Wizard are saved in the form of rules. In order to get a database with as much data as is possible we plan to use the crowdsourcing technique.
3. during the interaction the algorithm tries to approximate the policy - how the operator chooses its actions in different states - of the operator. After a given point, the algorithm is capable of replacing the Wizard, such as it was presented by Broekens in [24]. When in autonomous mode the robot senses the subject using Microsoft's Kinect.
4. in the first three phases the robot learns a pattern of social behavior, which can serve as a base when we want to create personalized behaviors.
5. to create a personalized social behavior we use interactive reinforcement learning algorithm where the reward is represented by the subject's states.

IV. Case Studies: Experiments in the Wild

We propose two scenarios for a social robot and following role: as a physiotherapist for rehabilitation and prevention of scoliosis and as a teacher of English language. This scenarios can also fuse play and rehabilitation techniques using a robotic design to induce child-robot interaction, in which the criteria was to make the educational process entertaining and effective for the children.

1) Robot as a physiotherapist

In the first scenario, our motivation was principally to motivate children to exercise, as there is a wide application potential. The problem of scoliosis in today's society is growing, and also children tend to become overweight (similar application using the same system can be for example the weight loss). In general, it is fundamental to ensure adequate motor skill development during childhood.

In collaboration with a professional physiotherapist, we programmed a set of exercises that can reduce the symptoms of spinal disorders. The exercises improve cosmetic appearance, reduce pain, improve breathing and function levels, reduce the existing curvature and in some cases avoid the need for scoliosis surgery.

Fig. 7. 'Robot as a physiotherapist' scenario

Fig. 8. 'Robot as a foreign language tutor' scenario

2) Robot as a tutor of foreign language

Using robots to support teaching and learning, from secondary school to undergraduate courses to graduate education, has become a popular research topic. We decided to design a simple scenario, in which children have to play a game with the robot - so it acts as a peer to learn new English words.

In both setup-ups, the robot remembers the correct answers/exercises and motivates children to continue with the "game". The motivational features consist of various affective gestures, sounds and verbal commands. The robot has a set of emotional expressions for basic emotions as surprise, anger, joy or sadness. However, these are selected randomly by the robot and after some time, children stop to pay attention. In the next step, we designed an interface for human operators, who can control the robot and this seems more intelligent and interesting for the children.

V. Discussion

Our research highlights the need of moving from the short-term human-robot interaction scenarios towards the applications where robots serve as long-term partners/peers/companions. The robot should not have a limited pre-programmed set of behaviors, but should be able to adapt according to the users' needs ad according to its role within an application. We propose an approach that combines the advantages of the Wizard of Oz technique (teleoperation) and autonomous robotics.

The paper shows an interface for the human experts (e.g. psychologists/therapists) using which they can control the robots used in the scenarios. The system is placed in the cloud environment, thus several experts can use it. The interface has two basic parts: a motion library, where the Wizard controls the motion of the robot in real-time and a so-called motivational behavior library that represents the emotional expressions of the robot. The key factor is the learning ability of the system – the robots connected to Cloud can adopt the behavior that the system learns from the Wizards. To achieve this, we use a reinforcement-based agent. We show two case studies: in one the robot acts as a physiotherapist and in the second one the robot has a role of a tutor of a foreign language.

We believe that this learning mechanism can serve for solving two research problems in HRI. Firstly it creates a learning environment where the operators' work can be replaced by a semi-or fully autonomous behavior of the robot. Secondly, the learned social behaviors can be personalized so robots would be more successful in long-term HRI. In future, we plan to conduct a long term experiment to confirm our hypothesis that the adaptation of social behavior makes the cooperation between robots and human users more enjoyable and effective.

Acknowledgment

This publication is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the EREDF.

References

[1] P. Baxter et al., "Long-term human-robot interaction with young users," in IEEE/ACM Human-Robot Interaction 2011 Conference (Robots with Children Workshop), 2011.

[2] J. Fasola, and M. J. Mataric., "Using socially assistive human–robot interaction to motivate physical exercise for older adults," in Proc. of the IEEE 100.8, 2012, pp. 2512-2526.

[3] B. Scassellati, H. Admoni, and M. Mataric, "Robots for use in autism research," in Annual Review of Biomedical Engineering 14, 2012, pp. 275-294.

[4] K. Dautenhahn, and A. Billard, "Games Children with Autism Can Play with Robota", in Universal Access and Assistive Technology: Proceedings of the Cambridge Workshop on UA and AT'02. Springer Science & Business Media, 2013, p. 179.

[5] Bhargava, S. et al., "Demonstration of the Emote Wizard of Oz Interface for Empathic Robotic Tutors," in Proceedings of SIGdial, 2013.

[6] A. Deshmukh, et al., "Empathic Robotic Tutors: Map Guide", in Proceedings of the Tenth Annual ACM/IEEE Int Con on HRI. ACM, 2005, p. 311.

[7] O. Mubin, et al., "A review of the applicability of robots in education", in Journal of Technology in Education and Learning 1, 2013.

[8] G. Castellano, et al., "Towards empathic virtual and robotic tutors", in Artificial Intelligence in Education. Springer Berlin Heidelberg, 2013, pp. 733-736.

[9] K., Zerrin, and N. Magnenat-Thalmann, "Building long-term relationships with virtual and robotic characters: the role of remembering", in The Visual Com., 2012, pp. 87-97.

[10] M. L. Walters, et al, "A long-term HRI proxemic study", in RO-MAN, 2011 IEEE. IEEE, 2011, pp. 137-142.

[11] I. Leite, C. Martinho, and A. Paiva. "Social robots for long-term interaction: a survey," in International Journal of Social Robotics 5.2, 2013, pp. 291-308

[12] Z. Bien, et al., "Machine intelligence quotient: its measurements and applications," in Fuzzy sets and systems 127.1, 2002, pp. 3-16.

[13] RoboEarth Project Webiste. http://roboearth.org/cloud_robotics/ [viewed on 12.08.2015].

[14] G. Magyar, et al. "Towards adaptive cloud-based platform for robotic assistants in education", in Applied Machine Intelligence and Informatics (SAMI), 2014 IEEE 12th Int. Sym. on. IEEE, 2014, pp. 285-289.

[15] R. Plutchik, "The emotions", in University Press of America, 1991.

[16] D. R. Olsen and M. Goodrich, "Metrics for Evaluating HRI", in perMIS, 2003

[17] R. S. Sutton and A. G. Barto,, "Reinforcement Learning: An Introduction," in Cambridge, MA, London, England: The MIT Press, 1998.

[18] A. L. Thomaz, "Socially Guided Machine Learning, PhD Thesis", 2006.

[19] A. Tapus, et al., "User-Robot Personality Matching and Robot Behavior Adaptation for Post-Stroke Rehabilitation Therapy, " in Intelligent Service Robotics, Springer Verlag (Germany), 2008, 1 (2), pp.169-183

[20] A. L. Thomaz and C. Breazeal, C., "Reinforcement learning with human teachers: Evidence of feedback and guidance with implications for learning performance", in AAAI (Vol. 6), 2006, pp. 1000-1005.

[21] U. Kartoun, H. Stern and Y. Edan, "A Human-Robot Collaborative Reinforcement Learning Algorithm," in Journal of Intelligent Robotic Systems, vol. 60, no. 2, 2010, pp. 217-239.

[22] N. Mitsunaga, et al.: "Adapting Robot Behavior for HRI", in IEEE Trans. on Robotics, vol. 24, no. 4, 2008, pp. 911-916.

[23] H. B. Suay and S. Chernova Sonia, "Effect of Human Guidance and State Space Size on Interactive Reinforcement Learning," in Proceedings of the 20th IEEE Int Conf on Robot and Human Interactive Communication, 2011, pp. 1-6.

[24] J Broekens, "Emotion and Reinforcement: Affective Facial Expressions Facilitate Robot Learning", in Artificial Intelligence for Human Computing, 2007, pp. 113-132.