Personalised Behaviour Model for Autism Therapy

Michał Stolarz
Hochschule Bonn-Rhein-Sieg
Sankt Augustin, Germany
michal.stolarz@smail.inf.h-brs.de

Alex Mitrevski
Hochschule Bonn-Rhein-Sieg
Sankt Augustin, Germany
aleksandar.mitrevski@h-brs.de

Mohammad Wasil
Hochschule Bonn-Rhein-Sieg
Sankt Augustin, Germany
mohammad.wasil@h-brs.de

Paul G. Plöger
Hochschule Bonn-Rhein-Sieg
Sankt Augustin, Germany
paul.ploeger@h-brs.de

***Abstract*—In Robot-Assisted Therapy for children with Autism Spectrum Disorder, the therapists' workload is increased due to the necessity of controlling the robot manually. The solution for this problem is to increase the level of autonomy of the system, namely the robot should interpret and adapt to the behaviour of the child under therapy. The problem that we are adressing is to develop a behaviour model that will be used for the robot decision-making process, which will learn how to adequately react to certain child reactions. We propose the use of the reinforcement learning technique for this task, where feedback for learning is obtained from the therapist's evaluation of a robot's behaviour.**

I. INTRODUCTION

Children with autism find robots easier to communicate with than human therapists [1]. In the context of Robot-Assisted Therapy (RAT) [2], it is desirable to increase the autonomy of the system in order to reduce the workload of the therapist. In particular, the robot should interpret a child's behaviour and adapt its actions to the child's individual needs. Adaptation is possible if the robot actively learns a user model that can be integrated into the robot's decision-making algorithm [3].

One of the state-of-the-art solutions introduced the supervised autonomy system for RAT for children with ASD [4]. In this system, the robot produces actions according to therapeutic scripts defined by therapists, but when interactions do not go as planned, the robot choses appropriate actions on its own. Supervised autonomy means that, before executing any action, the robot requests the therapist for feedback, which is used to correct its behaviour model. However, the model used here is a neural network, whose output can be difficult to interpret. Moreover, the network has to be retrained every time the therapist's feedback is obtained, which would make this solution suboptimal for online real-time interactions [5], especially in case of long-time scenarios, as the more data is collected, the more time the learning process takes.

This work elaborates on an alternative solution, based on the Q-learning algorithm [6] [7] [8], to improve the interaction abilities of the QTrobot [9], used as a tutor in tablet-based therapeutic games. Section II formalizes the problem, describes certain design decisions, and presents approaches for learning and evaluation of the introduced model. Section III provides a conclusion and elaborates on the planned future work.

II. BEHAVIOUR MODEL

A. Problem formalization

To formalize the problem, the robot will be named an agent. The behaviour model for decision making will be represented as a deterministic Markov Decision Process that is defined as a tuple (S, A, P_a, R_a, γ). Each state $s \in S$ can be defined by the child's affective state, engagement, motivation, and game performance. Normally, the affect is modelled by three factors: valence, arousal and dominance [10]. Valence, which describes the positiveness of an emotion, is useful along with engagement [11], such that there are various methods of estimating them [12] [13]; however, autistic children usually have difficulties recognizing and expressing emotions [14], and using the affective state may result in a suboptimal behaviour model. As an alternative, an estimate of the child's motivation is used in [15], which is related to the speed of movements during a therapeutic game. Finally, the performance of the child can be defined by two variables: one indicates if the last move during the therapeutic game was correct or incorrect, while the second one exhibits the time since the last child's move [16]. In the end, it might be feasible to use only three variables: engagement, game performance, and motivation.

For a minimal example, the action space A consists of three actions a: encouragement, waving, and proposition [16]. When the child's motivation is low, the robot might perform encouragement in the form of motivating feedback. If the child is not really engaged, the robot can wave to catch the child's attention. When the child is not responding for a long time, the robot can invite the child to play the game (proposition). The agent in state s will move to a state s' after performing action a with probability $P_a(s, s')$. For such a transition, an agent receives an immediate reward $R_a(s, s')$ based on the therapist's feedback. Finally, γ is a discount factor.

B. Model learning

We propose to perform model learning with the tabular Q-Learning reinforcement learning technique as in [7], which is depicted in algorithm 1, for the purposes of easier interpretability. Here, $Q(s_t, a_t)$ is the value of a given entry in the Q-value table, s_t and s_{t+1} are the states before and after execution of the action a_t, respectively, $R_a(s_t)$ is the immediate reward after applying a_t in s_t, and α is a predefined learning rate. The variables defining the state s_t are in the range $[-1, 1]$ [15], discretised with resolution κ, so as to limit the state space and enable the evaluation described in subsection II-C. Before execution of a_t, the robot suggests it to the therapist. When the proposed action is not corrected, the agent receives $R_a(s_t) = \theta$. However, if the therapist selects a different action, the reward of $R_a(s_t) = \beta$ is received by the

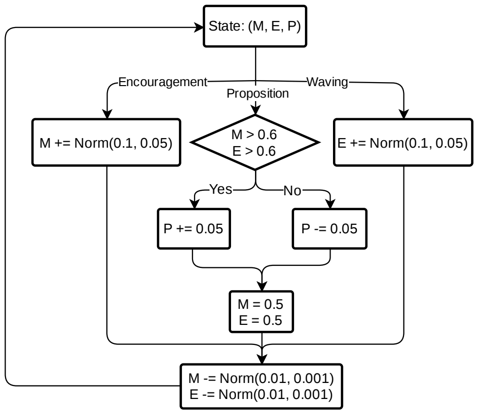

Fig. 1: Child model (based on [16])

agent. In this case, the reward is assigned to the action selected by the therapist and not the one proposed by the robot.

while $learning$ **do**
  $a_t = \max_a Q(s_t, a)$
  propose a_t to the therapist
  while $waiting\ for\ feedback\ (n\ seconds)$ **do**
    if $therapist\ selected\ another\ action$ **then**
      a_t = selected action
      reward, $R_a(s_t) = \beta$
    else
      reward, $R_a(s_t) = \theta$
    end
  end
  execute a_t, and transition to s_{t+1}
  $Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha(R_a(s_t) + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t))$
end

Algorithm 1: Model learning from the therapist's feedback (based on [7]).

C. Planned evaluation

Before conducting experiments with children, the proposed system will be tested using a rule-based child model [16] (Fig. 1), which includes processes which are dependent as well as independent on the robot behaviour. One example of independent child behaviour is decreasing engagement and motivation over time. As shown in Fig. 1, the child state is defined by three variables: motivation (M), engagement (E) and performance (P). In the end, the proposed algorithm should reduce the number of the required therapist interventions.

III. Conclusion

In this work, we are extending reinforcement learning used in the context of supervised autonomy [7] to the scenario of therapy for autistic children, where a neural network has previously been applied [4]. For future development, we need to select the most relevant variables to define the state space of our model, examine the convergence speed of the suggested algorithm, and we also want the robot to adapt the difficulty of the tasks during the intervention to each child. This problem requires modelling the full performance history of the child, which can be done in various ways [12] [17].

Acknowledgment

This work is conducted in the context of the MigrAVE project funded by the German Ministry of Education and Research (BMBF). We hereby thank our partners, Münster University of Applied Sciences (FHM) and the RFH – University of Applied Sciences, Cologne.

References

[1] B. Robins, K. Dautenhahn, and J. Dubowski, "Does appearance matter in the interaction of children with autism with a humanoid robot?," *Interaction Studies*, vol. 7, no. 3, pp. 479–512, 2006.

[2] I. Iacono, H. Lehmann, P. Marti, B. Robins, and K. Dautenhahn, "Robots as social mediators for children with Autism - A preliminary analysis comparing two different robotic platforms," in *IEEE Int. Conf. Development and Learning (ICDL)*, vol. 2, pp. 1–6, 2011.

[3] S. Rossi, F. Ferland, and A. Tapus, "User profiling and behavioral adaptation for HRI: A survey," *Pattern Recognition Letters*, vol. 99, pp. 3–12, 2017.

[4] P. G. Esteban *et al.*, "How to build a supervised autonomous system for robot-enhanced therapy for children with autism spectrum disorder," *Paladyn, Journal of Behavioral Robotics*, vol. 8, no. 1, pp. 18–38, 2017.

[5] E. Senft, *Teaching Robots Social Autonomy From In Situ Human Supervision*. PhD thesis, University of Plymouth, 2018.

[6] C. J. Watkins and P. Dayan, "Q-learning," *Machine learning*, vol. 8, no. 3-4, pp. 279–292, 1992.

[7] E. Senft, P. Baxter, J. Kennedy, S. Lemaignan, and T. Belpaeme, "Supervised autonomy for online learning in human-robot interaction," *Pattern Recognition Letters*, vol. 99, pp. 77–86, 2017.

[8] Y. Feng, Q. Jia, and W. Wei, "A control architecture of robot-assisted intervention for children with autism spectrum disorders," *Journal of Robotics*, vol. 2018, 2018.

[9] A. P. Costa, G. Steffgen, F. J. R. Lera, A. Nazarikhorram, and P. Ziafati, "Socially assistive robots for teaching emotional abilities to children with autism spectrum disorder," in *3rd Workshop on Child-Robot Interaction at HRI*, 2017.

[10] H.-L. Cao *et al.*, "A personalized and platform-independent behavior control system for social robots in therapy: development and applications," *IEEE Trans. Cognitive and Developmental Systems*, vol. 11, no. 3, pp. 334–346, 2018.

[11] G. Gordon *et al.*, "Affective personalization of a social robot tutor for children's second language skills," in *Proc. AAAI Conf. Artificial Intelligence*, vol. 30, 2016.

[12] S. Jain, B. Thiagarajan, Z. Shi, C. Clabaugh, and M. J. Matarić, "Modeling engagement in long-term, in-home socially assistive robot interventions for children with autism spectrum disorders," *Science Robotics*, vol. 5, no. 39, 2020.

[13] O. Rudovic, J. Lee, M. Dai, B. Schuller, and R. W. Picard, "Personalized machine learning for robot perception of affect and engagement in autism therapy," *Science Robotics*, vol. 3, no. 19, 2018.

[14] O. Rudovic, J. Lee, L. Mascarell-Maricic, B. W. Schuller, and R. W. Picard, "Measuring engagement in robot-assisted autism therapy: a cross-cultural study," *Frontiers in Robotics and AI*, vol. 4, p. 36, 2017.

[15] E. Senft, P. Baxter, J. Kennedy, and T. Belpaeme, "Sparc: Supervised progressively autonomous robot competencies," in *Int. Conf. Social Robotics ICSR*, pp. 603–612, Springer, 2015.

[16] E. Senft, P. Baxter, and T. Belpaeme, "Human-guided learning of social action selection for robot-assisted therapy," in *Machine Learning for Interactive Systems*, pp. 15–20, PMLR, 2015.

[17] P. Baxter, E. Ashurst, R. Read, J. Kennedy, and T. Belpaeme, "Robot education peers in a situated primary school study: Personalisation promotes child learning," *PloS one*, vol. 12, no. 5, p. e0178126, 2017.