SPARC: Supervised Progressively Autonomous Robot Competencies

Emmanuel Senft, Paul Baxter, James Kennedy, and Tony Belpaeme

Centre for Robotics and Neural Systems, Plymouth University, United Kingdom
{emmanuel.senft, paul.baxter, james.kennedy, tony.belpaeme}@plymouth.ac.uk

Abstract. The Wizard-of-Oz robot control methodology is widely used and typically places a high burden of effort and attention on the human supervisor to ensure appropriate robot behaviour, which may distract from other aspects of the task engaged in. We propose that this load can be reduced by enabling the robot to learn online from the guidance of the supervisor to become progressively more autonomous: Supervised Progressively Autonomous Robot Competencies (SPARC). Applying this concept to the domain of Robot Assisted Therapy (RAT) for children with Autistic Spectrum Disorder, a novel methodology is employed to assess the effect of a learning robot on the workload of the human supervisor. A user study shows that controlling a learning robot enables supervisors to achieve similar task performance as with a non-learning robot, but with both fewer interventions and a reduced perception of workload. These results demonstrate the utility of the SPARC concept and its potential effectiveness to reduce load on human WoZ supervisors.

1 Introduction

Over the last two decades, an increasing amount of research has been conducted to explore Robot Assisted Therapy (RAT). Using robots in therapies for children with Autism Spectrum Disorder (ASD) has revealed promising results [5, 10, 11]. The Wizard-of-Oz (WoZ) paradigm is typically used for this application, and others, where the robots are not autonomous but tele-operated. Many motivating factors for moving away from WoZ in RAT have been put forward [8, 13]. In particular, autonomous behaviour facilitates repetition of the robot behaviour and decreases the workload on therapists, freeing them to pay attention to other aspects of the interaction. It is the intention of our research to facilitate this shift to robot autonomy.

As the optimal robot behaviour is unlikely to be known in advance (be it in a therapeutic or indeed other domain), and with adaptability during and between the different interactions being generally desirable, it is necessary to provide the robot with learning capabilities. In the context of RAT, by using the knowledge of a therapist, the learning can be guided so that it is faster and safer, especially as the robot cannot use random exploration to acquire knowledge about its environment when interacting with children with ASD in case of negative therapeutic and/or clinical outcomes. We propose an approach taking inspiration from the Learning from Demonstration and online learning literature, and call it SPARC: Supervised Progressively Autonomous Robot Competencies. In SPARC, a therapist guides the robot in the early stages of the interaction, and progressively, the robot learns an action policy adapted to the particular therapeutic session [12]. Assuming the effective learning of the robot in this context, the therapist can allow the robot to behave increasingly autonomously, whilst maintaining oversight. Although not reducing the attentional requirements, this would reduce the physical interventions to direct the robot behaviour required by the therapist. Thus, by proposing and executing good actions, SPARC can reduce the therapists' workload.

A RAT scenario typically involves three parties: the patient, a robot, and the human therapist. In this context, the therapist does not interact with the patient directly, but rather through the robot. The therapist could therefore be described as playing the role of a robot supervisor. The focus of this paper is not on a new learning algorithm, but rather on the interaction between the robot and the therapist (supervisor), and the role that robot autonomy can play in this relationship. Specifically, as an initial validation of the principle, we seek to assess whether the SPARC concept can feasibly result in a reduction in workload for the supervisor, even given different strategies used by different individuals. A user study employing a novel methodology is conducted (section 3), demonstrating that progressive robot autonomy does indeed result in lower supervisor workload (section 4). This outcome provides support for the proposed approach and motivates further development efforts in the domain of RAT.

2 Related work

A number of research groups have studied the use of robot in therapy for children with ASD, which allowed children to express previously unseen social behaviour for example [9, 10]. Two primary methods have been used for these investigations: using an autonomous robot following preprogrammed rules [6, 14], or using the WoZ paradigm, allowing more flexibility in the robot's reaction. As noted in [8, 13], using WoZ allows testing and prototyping of interaction scenarios, but researchers should consider moving away from it to achieve more scalability, more repeatability, and to allow the use of robots without increasing the workload on therapists. Complex behaviour is required for a therapeutic robot, thereby making learning a desirable feature for future, more autonomous, RAT. As therapists possess the knowledge required to make appropriate decisions in different contexts, Learning from Demonstration [1] provides a useful starting point. Recently, Knox *et al.* proposed the Learning from Wizard paradigm in [7]. The robot is first controlled by a human operator as in a WoZ scenario, and after a number of interactions, batch learning is applied on the previous interaction data to obtain autonomous behaviour.

A fixed action policy of this type is however not desirable for RAT as children may not be consistent between interactions, and thus online learning is required

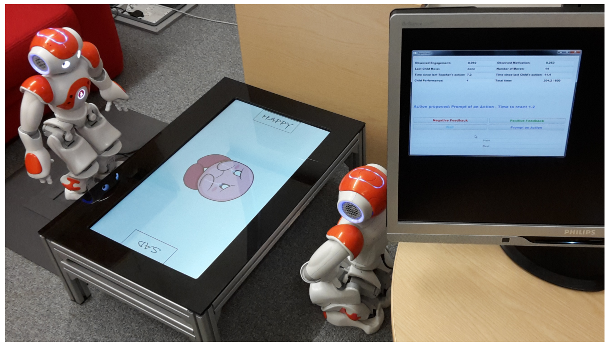

Fig. 1. Setup used for the user study from the perspective of the human supervisor. The *child-robot* (left) stands across the touchscreen (centre-left) from the *wizarded-robot* (centre-right). The supervisor can oversee the actions of the *wizarded-robot* through the GUI and intervene if necessary (right).

to provide the robot with the adaptability necessary to update its action policy depending on the current circumstances. Several experimenters in HRI have studied active learning: a robot actively questions a human teacher in order to request data points or demonstration for an uncertain scenario. A study exploring the type of questions that a robot could ask and the human reactions can be seen in [3], and Chernova and Veloso propose a progressive learning algorithm where a robot can estimate the confidence in its action decision in a fixed environment [4]: if the confidence is too low, a demonstration from a human teacher is required to complete the task.

However, an important element missing from the current literature is online learning for interaction. The robot needs to be able to progressively create an action policy, and update it later if necessary, to reach a more complex interaction behaviour. This paper explores how supervised progressive learning can be used in an interaction scenario and introduces a novel methodology to test this technique.

3 Assessing the effect of a progressively autonomous robot on supervisor workload

The focus of the present study is to assess whether the application of the SPARC concept to RAT results in a decrease in workload for the human supervisor. Two types of robot controller are employed to determine the presence and magnitude of this effect: a robot that learns from the actions of the supervisor to progressively improve its behaviour (*learning* controller), and a robot that only generates random actions (*non-learning* controller).

The methodology used in this paper is based on a real scenario for RAT for children with ASD based on the Applied Behaviour Analysis therapy framework. The aim of the therapy is to help the child to develop/practice their social skills: the task we focus on here is emotion recognition. This scenario involves a child playing a categorisation game with a robot on a mediating touchscreen device [2]. Images of faces or drawings are shown to the child, and she has to categorise them by moving the image to one side or the other depending on whether the picture shown denotes happiness or sadness (e.g. fig. 1). The human supervisor is physically present and guides the robot using the Wizard of Oz paradigm, but does not interact with the child directly.

In our proposed system, the basic interaction structure following the SPARC concept is as follows: the robot suggests an action to the supervisor, the supervisor agrees or disagrees with this suggestion (providing an alternative if disagreeing), the robot executes the action, and then both robot and supervisor observe the outcome. Over time, it is possible for the robot to learn an appropriate strategy based on observations of the child and oversight from the supervisor, with the supervisor still maintaining overall control if necessary.

Given the focus on human supervisor workload, it is necessary to provide a consistent experimental environment across both conditions in which the task, setup, and interaction partner is kept constant. A minimal model of child behaviour is therefore used to stand in for a real child. A second robot is employed in the interaction to embody this child model: we term this the *child-robot*. The robot being directly guided by the human supervisor is termed the *wizarded-robot* (fig. 1).

3.1 Child model

The purpose of the child model is not to realistically model a child (with or without autism), but to provide a means of expressing some of the behaviours we observed in our interactions with children in a repeatable manner. The child-robot possesses an internal model encompassing an *engagement* level and a *motivation* level, together forming the *state* of the child. The engagement represents how often the child-robot will make categorisation moves and the motivation gives the probability of success of the categorisation moves. Bound to the range $[-1, 1]$, these states are influenced by the behaviour of the wizarded-robot, and will asymptotically decay to zero without any actions from the wizarded-robot. These two states are not directly accessed by either the supervisor or the wizarded-robot, but can be observed through behaviour expressed by the child-robot: low engagement will make the robot look away from the touchscreen, and the speed of the categorisation moves is related to the motivation (to which gaussian noise was added). There is thus incomplete/unreliable information available to both the wizarded-robot and the supervisor, making the task non-trivial.

The influence of the wizarded-robot behaviour on the levels of engagement and motivation are described below (section 3.2). In addition to this, if a state is already high and an action from the wizarded-robot further increases it, then there is a chance that this level will sharply decrease, as an analogue of child-robot *frustration*. When this happens, the child-robot will indicate this frustration verbally (uttering one of eight predefined strings). The reason this mechanism is required is that it prevents a straightforward engagement and motivation maximisation strategy, thus better approximating the real situation, and requiring a more complex strategy to be employed by the supervisor.

3.2 Wizarded-robot control

The wizarded-robot is controlled through a Graphical User Interface (GUI) and has access to multiple variables characterising the state of the interaction. The wizarded-robot has a set of four actions, which each have a button in the GUI:

- Prompt an Action: Encourage the child-robot to do an action.
- Positive Feedback: Congratulate the child-robot on making a good classification.
- Negative Feedback: Supportive feedback for an incorrect classification.
- Wait: Do nothing for this action opportunity, wait for the next one.

The impact of the action on the child-robot depends on the internal state and the type of the last child-robot move: good, bad, or done (meaning that feedback has already been given for the last move and supplementary feedback is not necessary). A prompt always increases the engagement, a wait has no effect on the child-robot's state, and the impact of positive and negative feedback depends on the previous child-robot move. Congruous feedback (positive feedback for correct moves; negative feedback for incorrect moves) results in an increase in motivation, but incongruous feedback can decrease both the motivation and the engagement of the child-robot. The supervisor therefore has to use congruous feedback and prompts, whilst being careful not to use them too often, to prevent the child-robot becoming frustrated. A 'good' strategy would keep the engagement and motivation high, leading to an increase in performance of the child-robot in the categorisation task.

Through the GUI, the supervisor has access to *observed states* (noisy estimations of the child-robot state), and information about the interaction history: number of moves, child-robot performance, time since last child-robot and wizarded-robot actions, type of the last child-robot move, and elapsed time. However the supervisor can not control the wizarded-robot directly, actions can only be executed only at specific times triggered by the wizarded-robot. Two seconds after each child-robot action, or if nothing happens in the interaction for five seconds, the wizarded-robot proposes an action to the supervisor by displaying the action's name and a countdown before execution. Only after this proposition has been done can the supervisor provide feedback to the wizarded-robot. If the supervisor does nothing in the following three seconds, the action proposed by the wizarded-robot is executed. This mechanism allows the supervisor to passively accept a suggestion made by the wizarded-robot or actively make an *intervention* by selecting a different action and forcing the wizarded-robot to execute it.

3.3 Learning algorithm

The two robot controllers used for the study were a learning controller and a non-learning random action selection controller. The learning algorithm used was a Multi-Layer Perceptron, trained with back propagation (five input, six hidden and four output nodes): after each new decision from the supervisor, the network was fully retrained with all the previous state-action pairs and the new one.

3.4 Participants

In WoZ scenarios, the wizard is typically a technically competent person with previous experience controlling robots. As such, to maintain consistency with the target user group, the participants for this study (assuming the role of the supervisor) are taken from a robotics research group. Ten participants were used (7M/3F, age M=29.3, 21 to 44, SD=4.8 years).

3.5 Hypotheses

To evaluate the validity of our method and the influence of such an approach, four hypotheses were devised:

- H1 A 'good' supervisor (i.e. keeping the motivation and engagement of the child-robot high) will lead to a better child-robot performance.
- H2 When interacting with a new system, humans will progressively build a personal strategy that they will use in subsequent interactions.
- H3 Reducing the number of interventions required from a supervisor will reduce their perceived workload.
- H4 Using a learning wizarded-robot allows the supervisor to achieve similar performance with fewer interventions when compared to the same scenario with a non-learning wizarded-robot.

3.6 Interaction Protocol

Each participant experienced both robot controllers, with the order changed between participants to control for any ordering effects. In *Condition LN* the participants first interact with the learning wizarded-robot, and then with the non-learning one; in *Condition NL* the participants first interact with the non-learning wizarded-robot, and then the learning robot. Participants were randomly assigned to one of the two conditions.

The interactions took place on a university campus in a dedicated experiment room. Two Aldebaran Nao robots were used; one robot had a label indicating that it was the *Child-Robot*. The robots face each other with a touchscreen between them, and participants assuming the role of the supervisor sit at a desk to the side of the wizarded-robot, with a screen and a mouse to interact with the wizarded-robot (fig. 1). The participants were able to see the screen and the child-robot.

A document explaining the interaction scenario was provided to participants. After the information had been read, a 30s video presenting the GUI in use was shown to familiarise them with it, without biasing them towards any particular intervention strategy. The participant then clicked a button to start the first interaction which lasted for 10 minutes. The experimenter was sat in the room outside of the participants' field of view. After the end of the first interaction, a post-interaction questionnaire was administered. The same protocol was applied in the second part of the experiment with another post-interaction questionnaire following. Finally, a questionnaire asking the participants to explicitly compare the two conditions was administered.

4 Results

4.1 Interaction data

The state of the child and the interaction values were logged at each step of the interaction (at 5Hz). All of the human actions were recorded: acceptance of the wizarded-robot's suggestion, selection of another action (intervention), and the states of the child-robot (motivation, engagement and performance) at this step. From this the intervention ratio was derived: the number of times a user chose a different action to the one proposed by the wizarded-robot, divided by the total number of executed actions. On average, after a first exploration phase, where the participant discovers the system, the learning robot robot has an intervention ratio lower than the non learning one (fig. 2, left)

The performance indicates the number of good categorisations executed by the child-robot minus the number of bad categorisations. A strong positive correlation (Pearson's r=0.79) was found between the average child-robot motivation and engagement and its performance.

In both conditions, the average performance in the second interaction (M_{LN-2} =38, 95% CI [36.2, 39.8], M_{NL-2}=34.8, 95% CI [30.8, 38.8]) was higher than in the first one (M_{LN-1}=29.4, 95% CI [25.3, 33.5], M_{NL-1}=24.3, 95% CI [19.4, 29.4]; Fig. 2 *left*). The 95% *Confidence Interval of the Difference of the Mean* (CIDM) for the L-NL condition is [4.1, 13.1] and for the NL-L condition is [4.0, 16.8]. However, the performance is similar when only the interaction order (first or second) is considered. The participants performed slightly better in the LN condition, but the CIDM includes zero in both cases (95% CIDM_1 [-1.5, 11.5], 95% CIDM_2 [-1.2, 7.6]). In the condition L-NL, the intervention ratio increased between the learning and non learning condition (M_{LN-1}=0.31, 95% CI [0.20, 0.42] to M_{LN-2}=0.68, 95% CI [0.66, 0.70], CIDM_{LN}=[0.26, 0.48]). But in the NL condition, the intervention ratio is almost identical between the two interactions but slightly lower for the learning case (M_{NL-1}=0.50, 95% CI [0.44, 0.57] to M_{NL-2}=0.46, 95% CI [0.40, 0.51], CIDM_{NL} [-0.03, 0.13]). This shows that when the wizarded-robot learned, a similar performance is attained as without learning, but the number of interventions required to achieve this is lower.

4.2 Questionnaire data

The post-interaction questionnaires evaluated the participant's perception of the child-robot's learning and performance, the quality of suggestions made by the

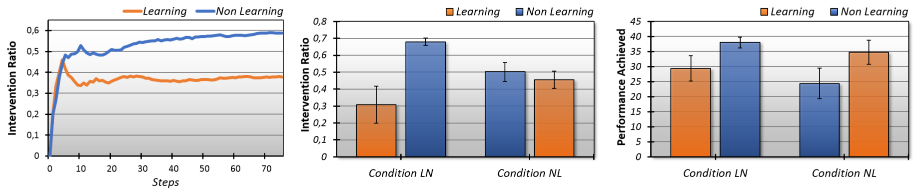

Fig. 2. (Left) evolution of intervention ratio over time for the learning and non learning cases. Intervention ratio (centre) and final performance (right) for the two conditions and the two interactions (*errors bars show 95% CI*). In condition LN participants started wizarding a robot which learns their interaction style, followed by a non-learning robot; in condition NL participants started with a non-learning robot, followed by a learning robot. Results show that a learning robot reduces the workload of the wizard, but performs equally well as a non-learning robot that needs wizarding at all times.

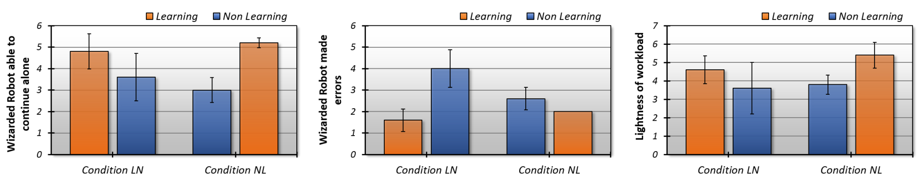

Fig. 3. Questionnaire responses (*mean and 95% CI*): increased confidence in the learning wizarded-robot over the non-learning version is apparent, as is a lower perceived workload.

wizarded-robot, and the experienced workload. All responses used seven point Likert scales.

Across the four possible interactions, the rating of the child-robot's learning was similar (M=5.25, 95% CI [4.8, 5.7]). The same effect was observed for the evaluation of the child performance (M=4.75, 95% CI [4.3, 5.2]). As the child-robot was using the same interaction model in all four conditions, this result is expected.

Participants report the wizarded-robot as more suited to operate unsupervised in the learning than in the non learning condition (M_{LN-1}=4.8, M_{LN-2}=3.6, M_{NL-1}=3, M_{NL-2}=5.2 ; CIDM for LN condition [-0.2, 2.6], CIDM for the NL condition [1.6, 2.8]).

Similarly, a trend was found showing that learning wizarded-robot is perceived as making fewer errors than the non-learning robot (M_{LN-1}=1.6, M_{LN-2}=4.0, M_{NL-1}=2.6, M_{NL-2}=2 ; CIDM for LN condition [1.3, 3.4], CIDM for the NL condition [0.1, 1.1]).

The participants tended to rate the workload as lighter when interacting with the learning robot, and this effect is much more prominent when the participants interacted with the non-learning robot first (M_{LN-1}=4.6, M_{LN-2}=3.6, M_{NL-1}=3.8, M_{NL-2}=5.4 ; CIDM for LN condition [-0.6, 2.6], CIDM for the NL condition [0.7, 2.5]).

5 Discussion

Strong support for H1 (a good supervisor leads to a better child performance) was found, a correlation between the average states (engagement and motivation) and the final performance for all of the 10 participants was observed (r=0.79). We could expect a similar effect when working with real children, but measuring these values would be a challenge.

The results also provide support for H2 (supervisors create personal strategies): all the participants performed better in the second interaction than in the first one. This suggests that participants developed a strategy when interacting with the system in the first interaction, and were able to use it to increase their performance in the second interaction. Looking in more detail at the interaction logs, it is possible to see that different people used different strategies.

H3 (reducing the number of interventions will reduce the perceived workload) is partially supported: the results show a trend for participants to rate the workload as lighter when interacting with the learning robot, and another trend between using a learning robot and the intervention ratio. However, when considering the difference of workload rating and intervention ratios between the two interactions, a positive correlation is only found for the LN condition, which could be accounted for by the initial steep learning curve for the study participants. Nevertheless, regardless of the order of the interactions, the learning robot consistently received higher ratings for lightness of workload (fig. 3).

Finally, H4 (using learning keeps similar performance, but decreases interventions) is supported: interacting with a learning robot results in a similar performance than interacting with a non-learning robot, whilst requiring fewer active interventions from the supervisor. This has real world utility, it frees some time for the supervisor, to allow her to focus on other aspects of the intervention, e.g. analysing the child's behaviour rather than focusing on the robot control.

It should be noted that the actual learning algorithm used in this study is only of incidental importance, and that certain features of the supervisor's strategies may be better approximated with alternative methods – of importance for the present work is the presence of learning at all. Future work will assess what the most appropriate machine learning approach is given the observed features of supervisor strategy from this study.

In conclusion, this paper proposed the SPARC concept (Supervised Progressively Autonomous Robot Competencies). Based on a suggestion/intervention system, this approach allows online learning for interactive scenarios, thus increasing autonomy and reducing the demands on the supervisor. Results showed that interacting with a learning robot allowed participants to achieve a similar performance as interacting with a non-learning robot, but requiring fewer interventions to attain this result. This suggests that while there is always adaptation in the interaction (leading to similar child-robot performance given the two wizarded-robot controllers), the presence of learning shifts this burden of adaptivity onto the wizarded-robot rather than on the human. This indicates that a learning robot could allow the therapist to focus more on the child than on the robot, with improved therapeutic outcomes as potential result.

Acknowledgements

This research was funded by the EU FP7 DREAM project (grant no. 611391).

References

1. Argall, B.D., Chernova, S., Veloso, M., Browning, B.: A survey of robot learning from demonstration. Robotics and autonomous systems 57(5), 469–483 (2009)
2. Baxter, P., Wood, R., Belpaeme, T.: A touchscreen-based sandtray to facilitate, mediate and contextualise human-robot social interaction. In: Human-Robot Interaction (HRI), 2012 7th ACM/IEEE International Conference on. pp. 105–106. IEEE (2012)
3. Cakmak, M., Thomaz, A.L.: Designing robot learners that ask good questions. In: Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction. pp. 17–24. ACM (2012)
4. Chernova, S., Veloso, M.: Interactive policy learning through confidence-based autonomy. Journal of Artificial Intelligence Research 34(1), 1 (2009)
5. Dautenhahn, K.: Robots as social actors: Aurora and the case of autism. In: Proc. CT99, The Third International Cognitive Technology Conference, August, San Francisco. vol. 359, p. 374 (1999)
6. Feil-Seifer, D., Mataric, M.: B3IA: A control architecture for autonomous robot-assisted behavior intervention for children with Autism Spectrum Disorders. RO-MAN 2008. (2008), http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4600687
7. Knox, W.B., Spaulding, S., Breazeal, C.: Learning social interaction from the wizard: A proposal. In: Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence (2014)
8. Riek, L.: Wizard of Oz Studies in HRI: A Systematic Review and New Reporting Guidelines. Journal of Human-Robot Interaction 1(1), 119–136 (Aug 2012)
9. Robins, B., Dautenhahn, K., Boekhorst, R.T., Billard, A.: Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills? Universal Access in the Information Society 4(2), 105–120 (Jul 2005)
10. Robins, B., Dautenhahn, K., Dickerson, P.: From Isolation to Communication: A Case Study Evaluation of Robot Assisted Play for Children with Autism with a Minimally Expressive Humanoid Robot. Conferences on Advances in Computer-Human Interactions pp. 205–211 (Feb 2009)
11. Scassellati, B., Admoni, H., Mataric, M.: Robots for use in autism research. Annual review of biomedical engineering 14, 275–294 (2012)
12. Senft, E., Baxter, P., Belpaeme, T.: Human-guided learning of social action selection for robot-assisted therapy (2015), in press for 4th Workshop on Machine Learning for Interactive Systems
13. Thill, S., Pop, C.A., Belpaeme, T., Ziemke, T., Vanderborght, B.: Robot-assisted therapy for autism spectrum disorders with (partially) autonomous control: Challenges and outlook. Paladyn 3(4), 209–217 (2012)
14. Wainer, J., Dautenhahn, K., Robins, B., Amirabdollahian, F.: Collaborating with kaspar: Using an autonomous humanoid robot to foster cooperative dyadic play among children with autism. In: Humanoid Robots (Humanoids), 2010. pp. 631–638. IEEE (2010)