Teaching Robots how to Interact with Humans

Emmanuel Senft
Centre for Robotics and Neural Systems
Plymouth University
Plymouth (UK)
emmanuel.senft@plymouth.ac.uk
http://www.tech.plym.ac.uk/SoCCE/CRNS/staff/esenft/

***Abstract*—To have meaningful social interaction with humans, robots have to be able to improve their behaviour over time by learning from humans. My research is focused on how a robot can learn how to interact with humans while interacting with someone and being supervised by a teacher.**

I. Introduction

Social interactions are complex environments governed by a large number of social norms and where all participants have a high number of expections on the others. As it seems impossible to provide directly a robot with all the knowledge required to behave in society, we argue that robots need to be able to learn how to interact socially. Furthermore, as the context of interaction and the partners can evolve over times, robots have to continuously update their behaviour. The problem of improving an action policy over time by interacting in an environment is well represented by Reinforcement Learning. Powerful tools exist to improve an agent's behaviour as it is evolving in an environment. By maximising a reward function evaluating its actions, the agent can improve its behaviour over time.

However classic methods in Reinforcement Learning rely on exploration, often random, to acquire information on the environment and generally require many iterations to converge to an appropriate action policy. When interacting with humans in the real world, acquiring data can take a prohibitive amount time, and the risk of exploration can be high: failing to satisfy expectations from human partners can annoy them or even lead to the end of the interaction. To tackle this issue, we propose to make use of a human supervisor to guide the robot's behaviour, preventing it to make mistakes and demonstrating it a correct action policy.

II. Proposition

A. Supervised Autonomy

Robots cannot have optimal behaviours, especially in the early stages of learning. To address this drawback, we propose the concept of Supervised Autonomy: a robot is mostly autonomous, but supervised by a human. Before executing an action, the robot proposes this action to a supervisor who is able to correct it before execution, ensuring that only actions validated by a human are executed. This provides robustness against suboptimal action policy or sensory errors.

This validation of action is combined with an automatic execution after a short delay: for all correct actions, the supervisor does not have to act and can let the robot being autonomous, only in cases where the proposed action is incorrect, the supervisor has to step in.

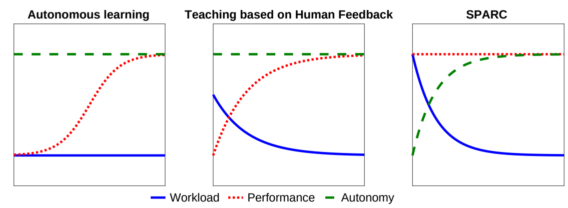

Fig. 1. Comparison of workload on human teacher, performance of the robot and autonomy for three frameworks of learning: autonomous learning, learning with human feedback and SPARC

B. SPARC: Supervised Progressively Autonomous Robot Competencies

When the robot is interacting in a Supervised Autonomous fashion, it can also use the human inputs to improve its behaviour over time. Combined with algorithms from the Reinforcement Learning field or other algorithms directly mapping inputs to actions, this approach ensures that the behaviour of the robot is constantly correct and the frequency of intervention from the supervisor decreases as the robot learns. We called this method: SPARC for Supervised Progressively Autonomous Robot Competencies.

Other methods using humans to teach a robot how to interact often use the teacher only to label actions from the robot or to provide corrective behaviours once the robot has made a mistake or have the robot request a demonstration from the human when the robot's confidence in the output of the learning algorithm is too low. On the other hand, SPARC aims to give control over the robot's action to a teacher to ensure a high performing action policy at every step of the interaction. Figure 1 show an illustrative comparison between SPARC and other learning methods. By giving control on *every* robot's action to the teacher, SPARC is the only method allowing the robot to have an efficient behaviour even in early stages of the learning while reducing the number of actions from the teacher over time.

III. Experiments

SPARC has been tested in two explorative studies and so far has demonstrated promising results.

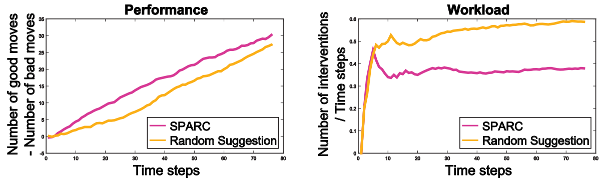

Fig. 2. Comparison of workload on human teacher and performance of the robot for two control scenario: SPARC and a random suggestion equivalent to a Wizard of Oz control.

A. Initial Eploration with Direct Mapping

The first study [1] involved 10 participants teaching a teacher robot how to provide feedback to a learner in a classification task. For the sake of the study, the learner was modeled by a second robot whose behaviour was impacted by the teacher robot's action. This study used a simple feed-forward neural network to map the state to action. As shown in Figure 2, compared to a Wizard of Oz interaction (where the human has to select manually all the actions executed by the robot), SPARC can achieve a similar performance while requiring a lower number of inputs.

B. Combining Supervised Autonomy with Reinforcement Learning

The second study [2] used Sophie's Kitchen, the environment proposed by Thomaz and Breazeal in [3] to explore how a human supervisor can teach a robot to complete a task in a deterministic and discrete environment. The learning algorithm is from the Reinforcement Learning field, the agent receives rewards from the environment or the human teacher and learns how to maximise these rewards to reach an efficient action policy and complete the task. This study involved 40 participants teaching the robot how to complete the task in two conditions. In the Interactive Reinforcement Learning condition, the teacher can provide feedback and limited guidance to the robot. In the second condition, SPARC, the teacher is informed of the action the robot is about to execute and can correct it before execution. With SPARC, the rewards are implicit, each action executed by the robot is rewarded with $+1$ as it received the approval from the teacher either explicitly (if selected) or implicitly (if letting be executed).

As shown in Figure 3, during the first interaction with an algorithm, participants reached a higher performance with SPARC compared to IRL, while requirering less inputs, less time and reaching less failure states during the teaching process.

IV. Discussion and Future Work

These two studies showed that SPARC has promises as interacting principles to teach a robot how to interact in its environment while not suffering of classical drawbacks of methods from Reinforcement Learning (poor performance at the start of the interaction and necessity of large number of samples). However, these studies were not teaching social behaviours. We are designing a new study to explore how to teach a robot how to interact with children in an educative game.

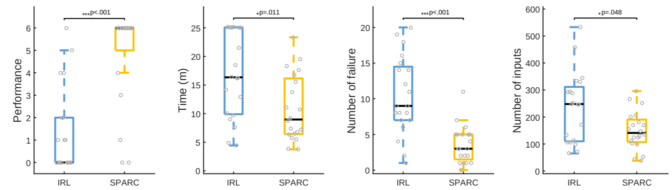

Fig. 3. Comparison of performance, teaching time, number of failure during teaching and number of inputs providing for the conditions: IRL and SPARC.

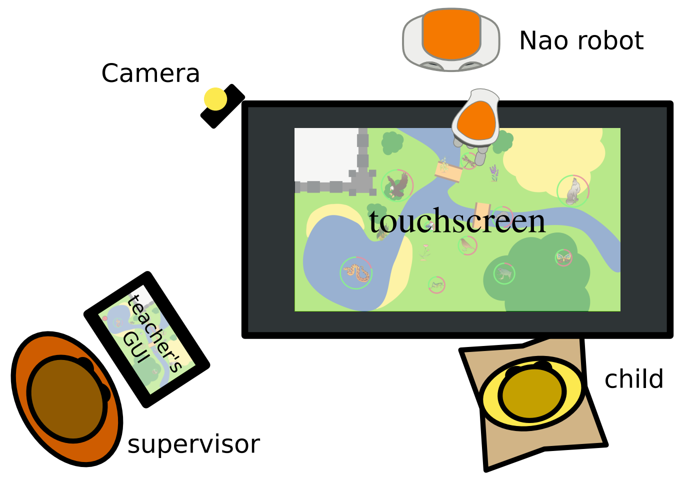

Fig. 4. Setup for the future study to teach a robot to interact with children.

Children will interact with a touchscreen on a game designed to teach food web: interconnected food chains present in nature. The game has 10 animals and 3 plants and the children will learn what each animal is eating. The robot will be used to provide feedback, advices and encouragements to the child. And this teaching behaviour will be taught by an adult using a graphical user interface on a tablet (see Figure 4).

The state will combine both features from the games such are the relative positions between the animals or their status and more social features, such as where the child is looking. And the actions will be game-related as moving animals. Or more social actions such as drawing the attention of the child to certain elements of the game or providing feeback related the last action executed by the child.

References

[1] E. Senft, P. Baxter, J. Kennedy, and T. Belpaeme, "Sparc: Supervised progressively autonomous robot competencies," in *International Conference on Social Robotics*, 2015, pp. 603–612.

[2] E. Senft, P. Baxter, J. Kennedy, S. Lemaignan, and T. Belpaeme, "Supervised autonomy for online learning in human-robot interaction," *Pattern Recognition Letters*, 2017.

[3] A. L. Thomaz and C. Breazeal, "Teachable robots: Understanding human teaching behavior to build more effective robot learners," *Artificial Intelligence*, vol. 172, no. 6, pp. 716–737, 2008.