# Effective robot-human skill transfer via mutual reinforcement learning

Sayanti Roy[1,*], Emily Kieson[1], Charles Abramson [1], Christopher Crick[1]

**1 Oklahoma State University, Stillwater, Oklahoma**

***sayanti.roy@okstate.edu**

## Abstract

Mutual information exchange between robots and humans are increasingly becoming important in collaborative task environments to accomplish complex tasks. In this paper we developed a novel approach called mutual reinforcement learning (MRL) where both the robot and human act as reinforcement learners in a skill transfer scenario over continuous communication and feedback. Here, we discovered a significant $p$-value $= 0.038$ to demonstrate the occurrence of skill transfer using Simpson's psychometric model. Thus this shared model has a collective impact, improves human cognition and helps in building a successful human-robot relationship where the robot acts as an expert, invokes the exploration/exploitation tradeoff to improve its own cognitive strategy to accomplish the task, and improves its understanding of the mental model of its human partner.

## Introduction

Empathy plays a vital role in social interaction in all stages of human life, and many contemporary researchers are working on empathetic robots that are designed to respond to human behavior and emotion with appropriate social cues. Empathy and adaptation may not be enough, however, since social responses are only one component of effective human-robot interactions. Instead, robot interactions that facilitate mutual learning with their human counterparts may prove more effective in a teaching environment [9] due to the ability to learn, adapt, and create reinforcement feedback tailored to the individual.

It has been shown that humans can be induced to respond to perceived empathy from robot and computer interactions. Research indicates that empathy increases rapport between humans and robots, which is important for user comfort [4]. However, while empathy is important for contextual comfort, it may not be the only component of a learning environment and does not indicate a human response for the robot. While scientists may have developed robots to mimic empathy that can be detected by participants, humans have yet to respond with equal attachment or empathy towards robots [3]. This means that while adaptive empathetic robots may build some rapport with humans, the communication usually runs only from the robot to the human, with the robot less sensitive to human reactions and failing to respond in ways that may be necessary for human learning. Interactivism and process-oriented robots have been challenging to design in the past, as they must balance environmental stimuli and feedback and adapt their inner processes to respond. Robots using socially-inspired reinforcement, including verbal and behavioral feedback [5] [6], have only shown modest results. This study suggested that a more targeted approach, tailored to the individual, would be better suited for future robot-human interactions. To better facilitate natural social interactions [1] and learning environments of humans, robots need to adapt and respond appropriately to each individual. Positive reinforcement increases learning in animals and promotes voluntary behaviors, but the reinforcement tools need to be specific and based on individual preferences and experiences. In this sense, if robot-human interactions are to use socially-derived reinforcements as teaching tools, researchers need to take into account not just human social interactions, but individual differences as well. This means that the robots need to be programmed with an understanding of individually specific approaches to interactions based on principles of learning. Robots must be able to adapt and respond in ways that are tailored based on the individual's unique responses. In our work robot and human both act as empathizers in the context of skill transfer scenario using MRL.

We present a novel approach using mutual reinforcement learning (MRL) where the action of the agent acts as the reward to the participant and vice versa [7]. In this paper, MRL is used in the context of skill transfer from robot to its human counterpart using positive reinforcers. To identify the success of MRL-guided skill transfer, we divided the subject population into three major groups where participants get no reinforcement, random reinforcement, or individually-tailored learned reinforcement respectively. To evaluate the behavior of the participants and the effectiveness of the model, we analyze their performance using Simpson's psychometric model [8] developed for skill transfer.

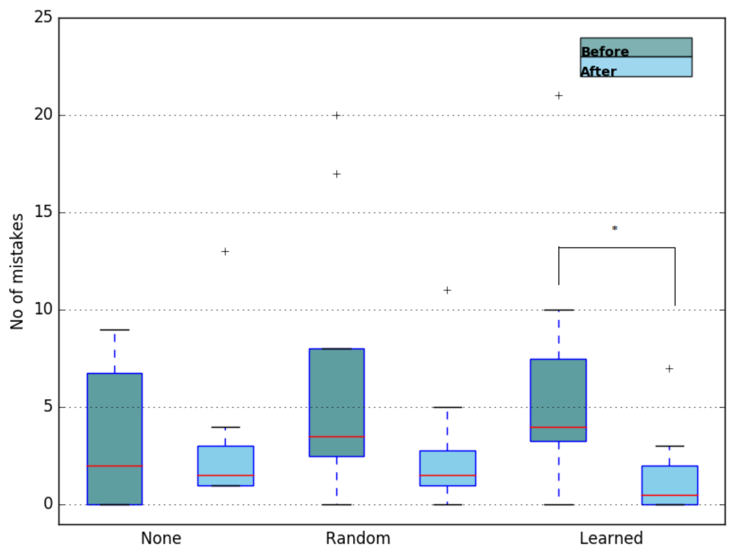


**Figure 1. Performance evaluation of participants before and after skill transfer using Baxter. * denotes $p$-value $<$ 0.05 (0.038 in this case).**

## Mutual Reinforcement Learning

MRL is a tuple $\{S, A(A'), T, R(R')\}$ where $S$ is a set of states; $A$ and $A'$ are sets of actions; $T$ is the set of state transition probabilities $P(s, a)$ upon taking action $a$ in state $s$; and $R$ and $R' : S \Rightarrow R(R')$ are the reward functions. Since in MRL, the action of the agent is the reward to the expert and vice versa, the tuple can be simplified as follows: Agent=$\{S, A', T, R\}$, Expert=$\{S, A, T, R'\}$ where if the agent excutes action $A'$, reward $R'$ is given out to the expert. This helps the expert to execute action $A$ using an exploration/exploitation strategy [2], which at the same time acts as a reward $R$ to the participant. If the action $A'$ is successful, then the robot realizes that the participant is fonder of reward $R$, which acts at the same time as reward $R'$ for the robot to understand its own performance or action $A$. Thus the response of the participant to the robot not only determines its cognitive orientation but also helps the robot to take better actions in future. Hence mutual reinforcement is based on psychological principles of social reinforcement and inclusion that improves skill transfer by adapting to the reward value systems of an individual learner.

## Experimental Procedure

$n$=34 (age $\mu$=19.69, $\sigma$=3.49, male=13, female=21, none=10, random=8, learned=10) participants were recruited for the experiments with Baxter. The task involved in the experiment was divided into two sections. In the first section, participants were taught to identify building blocks and were shown how to construct a pattern with them. These were evaluated by the robot, which, depending on the group, provided nothing but simple yes-or-no feedback or provided positive reinforcers (selected randomly or via MRL) at any point where the subject failed to reconstruct the pattern appropriately. Later, the subjects were asked to reconstruct the pattern twice, and were asked questions about the blocks (about which they had earlier been taught) without reinforcers to observe the skill transfer mechanism. The tasks were designed to cover perception, guided response, mechanism, adaptation etc. from Simpson's psychometric model to analyze skill transfer from expert to novice. At the end of each experiment, participants' opinions were probed with 5-point Likert scale questionnaires.

## Results

Fig. 1 illustrates the the experiment to observe the skill transfer scenario. We computed the number of mistakes in all of the cases to evaluate the performance of the participants. While the subjects were learning the task and the robot was engaged in teaching (with feedback in two of the experimental conditions), the number of mistakes made were, by group: none $\mu = 3.5$, $\sigma = 3.89$ , random $\mu = 6.62$, $\sigma = 7.5$, learned $\mu = 6.2$, $\sigma = 5.9$. After they were exposed to the teaching and reward feedback, they reconstructed the patterns without assistance. In this part of the experiment, the mistakes made by group were: none $\mu = 3.0$, $\sigma = 3.68$ , random $\mu = 2.87$, $\sigma = 3.6$, learned $\mu = 1.5$, $\sigma = 2.22$. All groups improved somewhat with practice, but the only group to improve to a statistically significant degree ($p = 0.038$) consisted of those individuals who received personalized feedback provided by the robot via MRL.

At the end of the experiment, those participants in the learned model group were asked to choose their preferred reinforcers. Baxter could correctly identify the preferred reinforcers in half of the cases (twice as effectively as a random baseline). Thus, mutual reinforcement learning allowed the robot successfully to identify the cognitive

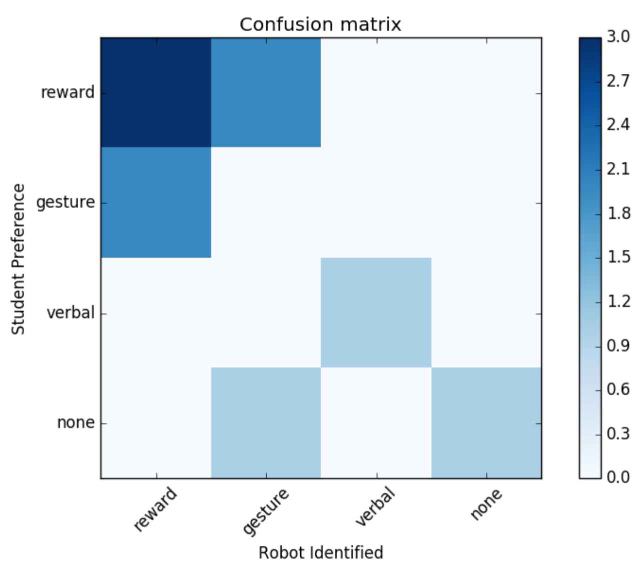


**Figure 2. Correlation of the reinforcer as preferred by the students and identified by the robot.**

orientation of the participants to a large extent. Fig. 2 denotes the correlation between the subjects' preferred reinforcers and the MRL-identified ones. During the task, since the number of interactions was limited, the robot did not have sufficient opportunity to engage in the exploitation aspect of the reinforcement learning, and thus its ability to identify preferred reinforcers was limited (but still reasonably successful). In these experiments, Baxter explored more than exploited, which impacted the types of reinforcers given out by the robot.

# Conclusion and Future Work

In order to effectively teach a skill, the instructor relies on the principles of learning theory and basic operant conditioning and positive reinforcement. Learning theory dictates, however, that the value of rewards in positive reinforcement are subjective in nature and highly dependent upon the individual. This means that a reward may be of high value to one individual and of no value to another. Hence feedback and positive reinforcement can be a powerful tool to encourage people and influence their thought processes. Positive reinforcement has a beneficial impact on human learning and behavior, but each individual is different. Hence if a robot can learn the appropriate reinforcement learning stategy and adapt accordingly, it will be more successful at training humans in complex tasks. Experiments like the ones detailed above lead to important future research questions, such as

- What necessary behaviorial changes are required to be adapted by the robots to become better trainers over time?

With a better understanding of these necessary changes in the robot's behaviour, the students will be better guided towards the correct actions.

- What responses robots should develop that will keep the learners thinking more about the task?

In the above experiment, the robot shows empathetic reactions when the participant fails at a task, but empathy is not enough. In future experiments, we will try to make the robots more responsive, critiquing their own performance in order to better guide novices by understanding their cognition. Currently, we only consider subjects' task performance and their level of engagement with the robots. In our future work, we would like to identify their emotions during the task, and incorporate emotionally appropriate responses into the robot's mutual reinforcement learning. Using MRL, we can not only improve the performance of the robot but also can effectively identify the mental models of the human participants involved in the experiment.

# Acknowledgments

This work was supported by NSF award #1527828 (NRI: Collaborative Goal and Policy Learning from Human Operators of Construction Co-Robots).

# References

1. A. G. Brooks, J. Gray, G. Hoffman, A. Lockerd, H. Lee, and C. Breazeal. Robot's play: interactive games with sociable machines. *Comput. Entertain.*, 2(3):10–10, 2004.

2. B. Clement, D. Roy, P.-Y. Oudeyer, and M. Lopes. Multi-armed bandits for intelligent tutoring systems. *Journal of Educational Data Mining*, 7(2):20–48, 2015.
3. V. Konok, B. Korcsok, Á. Miklósi, and M. Gácsi. Should we love robots?–the most liked qualities of companion dogs and how they can be implemented in social robots. *Computers in Human Behavior*, 80:132–142, 2018.
4. I. Leite, C. Martinho, and A. Paiva. Social robots for long-term interaction: a survey. *International Journal of Social Robotics*, 5(2):291–308, 2013.
5. S. Roy, C. Crick, E. Kieson, and C. Abramson. A reinforcement learning model for robots as teachers. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 294–299. IEEE, 2018.
6. S. Roy, E. Kieson, C. Abramson, and C. Crick. Using human reinforcement learning models to improve robots as teachers. In *Proceedings of the 13th ACM/IEEE Conference on Human-Robot Interaction (HRI)*, 2018.
7. S. Roy, E. Kieson, C. Abramson, and C. Crick. Mutual reinforcement learning with robot trainers. 2019.
8. E. Simpson. The psychomotor domain. *Washington DC: Gryphon House*, 1972.
9. A. L. Thomaz and C. Breazeal. Teachable robots: Understanding human behavior to build more effective robot learners. *Artificial Intelligence*, 172:716–737, 2008.