Leveraging Social Networks to Motivate Humans to Train Agents

(Extended Abstract)

Guangliang Li
University of Amsterdam
Amsterdam, Netherlands
g.li@uva.nl

Hayley Hung
Delft University of Technology
Delft, Netherlands
h.hung@tudelft.nl

Shimon Whiteson
University of Amsterdam
Amsterdam, Netherlands
s.a.whiteson@uva.nl

W. Bradley Knox
Massachusetts Institute of Technology
Cambridge, USA
bradknox@mit.edu

ABSTRACT

Learning from rewards generated by a human trainer observing the agent in action has been demonstrated to be an effective method for humans to teach an agent to perform challenging tasks. However, how to make the agent learn most efficiently from these kinds of human reward is still under-addressed. In this paper, we investigate the effect of providing social-network-based feedback intended to engender trainer competitiveness, focusing on its impact on the trainer's behavior. The results of our user study with 85 subjects show that the agent's social feedback can induce the trainer to train longer and give more feedback. Furthermore, the agent's performance was much better when social-competitive feedback was provided. The results also show that making the feedback active further increases the amount of time trainers spend training but does not further improve agent performance.

Categories and Subject Descriptors

I.2.6 [**Artificial Intelligence**]: Learning

Keywords

reinforcement learning; human-agent interaction; social agent; learning from human reward

1. INTRODUCTION

Autonomous agents have the potential to play a transformative role in many aspects of society in the near future. However, for agents to realize their transformative potential, they need to be able to efficiently learn how to perform challenging tasks from humans who, although experts in the tasks they are teaching, may have little expertise in autonomous agents or computer programming. Therefore, there is a great need for new methods that facilitate the interaction between human teachers and learning agents.

The feedback that the human provides during such interaction can take many forms, e.g., reward and punishment [5, 2], advice [4], guidance [6], or critiques [1]. Within them, learning from rewards generated by a human trainer observing the agent in action promises to be a powerful method for non-expert users in autonomous agents to teach the agent to perform challenging tasks. However, how to make the agent learn most efficiently from such human trainers is still under-addressed. Intuitively, when learning from humans, the agent's performance depends critically on the efficiency of the interaction between the agent and human trainer. It also depends on the information within the feedback provided by the human trainer. Therefore, we consider how the interaction between the trainer and the agent should be designed to reduce the human's effort or cost to train the agent to perform a task well. Previous work [3] showed that the way that the agent interacts with the trainer can greatly affect the trainer's engagement and the agent's performance and that the interaction between the agent and the trainer should ideally be *bi-directional*. In this paper, we seek to build on this past work by investigating how to improve its sophistication and efficacy, proposing a new *Socio-competitive* training interface. We use *TAMER* [2] as our foundation, which is one approach allowing the agent learn from human-generated rewards that reflect the human trainer's judgement of the quality of the agent's actions. A TAMER agent learns from this feedback by creating a predictive model of the human trainer's feedback and myopically choosing the action at each time step that it predicts will receive the highest feedback value.

2. SOCIO-COMPETITIVE TAMER

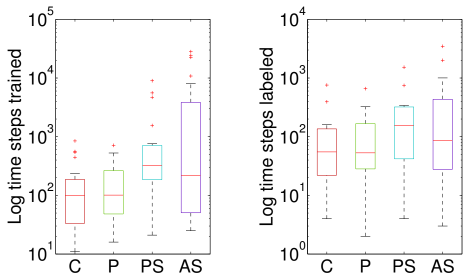
We have four conditions used in our experiment: the *control condition*, the original TAMER interface [2]; the *performance condition*, which shows the agent's history of performance to the trainer [3]; the *passive social condition* and *active social condition*, the novel conditions implemented with our socio-competitive training interface.

In the passive social condition, in addition to receiving feedback about how his/her agent is performing, the trainer

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*

now also sees a leaderboard that compares his/her agent's performance to that of his/her Facebook friends as well as all others using the Facebook app.

While both interfaces in previous work [3] and the social extension mentioned above are bi-directional, the agent's role is passive: it merely displays feedback for the trainer, which the trainer can choose to look at or ignore. To address this limitation, the agent in the active social condition *actively* provides feedback to the trainer using Facebook *notifications*, which are messages sent to Facebook users while they are not using the app that update the trainers on their performance relative to other trainers.



a b

Figure 1: Boxplots across the four conditions of (a) total time steps trained by subjects and (b) between-subject distribution of the total number of time steps that were labeled with feedback. C: control; P: performance; PS: passive social; AS: active social.

3. EMPERIMENTAL RESULTS

To evaluate our Socio-competitive training interface, we conducted a user study with 85 subjects on Facebook.

3.1 Engagement and Performance

As shown in Figure 1a and b, in terms of total number of time steps trained and the number of time steps with feedback, trainers in both passive and active social conditions trained longer than in the other conditions, while the active social condition resulted in a longer mean training time than the passive social condition but a lower median in training time. Thus, combined with the significant results of Mann Whitney U test between conditions, our results show that the social conditions positively affected training time and the quantity of time steps with feedback, which is consistent with our hypotheses.

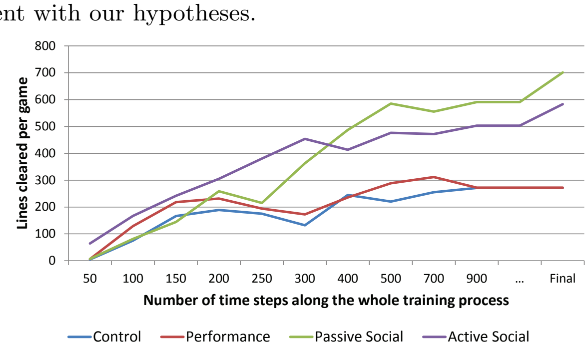


Figure 2: Mean offline performance.

The results in Figure 2 suggests that the social conditions improved the performance of the agent, consistent with our hypothesis that the increased engagement would lead to improved agent performance (For the calculation of the performance values in Figure 2, refer to [3]). Surprisingly, however, the active social condition, did not outperform the passive social condition, despite inducing more training and feedback. Further analysis is required to understand why this performance discrepancy between *active social* condition and *passive social* condition happened.

3.2 Influence of Social Information

To measure the extent to which social information influenced the trainers, we tried to measure how often they looked at the leaderboard, using the count of their mouseovers as a proxy metric. Our data shows that more than a half of the participants in the *passive* and *active* social conditions moved their mouse pointer over the leaderboard tabs at least once. With Pearson's correlation test, we also observed that for both conditions, the number of tab mouseovers correlates with the number of time steps trained ($r = 0.60, p \approx 0.006$ and $r = 0.89, p \approx 0$ for passive social and active social conditions respectively) and the trained agents' final offline performances ($r = 0.72, p \approx 0.0004$ and $r = 0.67, p \approx 0.0002$ for passive social and active social conditions respectively). In the active social condition, the number of notifications the trainer received correlates with the time steps trained ($r = 0.18, p = 0.39$) and the number of time steps with feedback ($r = 0.41, p = 0.04$).

4. CONCLUSION

By integrating agent training with an online social network via our Socio-competitive TAMER interface, this paper investigated the influence of social feedback on human training and the resulting agent performance. The results of our user study showed that the agent's social feedback can induce the trainer to train substantially longer and to give more feedback. Further, social feedback improved the agent's learned performance. The results also show that making the feedback active further increases the amount of time trainers spend training but does not further improve agent performance.

5. REFERENCES

[1] B. Argall, B. Browning, and M. Veloso. Learning by demonstration with critique from a human teacher. *HRI*, 2007.

[2] W. Knox and P. Stone. Interactively shaping agents via human reinforcement: The TAMER framework. *International Conference on Knowledge Capture*, 2009.

[3] G. Li, H. Hung, S. Whiteson, and W. B. Knox. Using informative behavior to increase engagement in the tamer framework. *AAMAS*, 2013.

[4] R. Maclin, J. Shavlik, L. Torrey, T. Walker, and E. Wild. Giving advice about preferred actions to reinforcement learners via knowledge-based kernel regression. In *National Conference on Artificial Intelligence*. AAAI Press, 2005.

[5] P. Pilarski, M. Dawson, T. Degris, F. Fahimi, J. Carey, and R. Sutton. Online human training of a myoelectric prosthesis controller via actor-critic reinforcement learning. *International Conference on Rehabilitation Robotics*, 2011.

[6] A. Thomaz and C. Breazeal. Reinforcement learning with human teachers: Evidence of feedback and guidance with implications for learning performance. *Proc. of the National Conference on AI*, 2006.