Facilitating Safe Adaptation of Learning Agents using Interactive Reinforcement Learning

Konstantinos Tsiakas
[1]Computer Science and Engineering Dept,
University of Texas at Arlington, USA
[2]Institute of Informatics and Telecommunications,
National Centre for Scientific Research 'Demokritos', Greece
konstantinos.tsiakas@mavs.uta.edu

IUI'16 Companion, March 07-10, 2016, Sonoma, CA, USA
ACM 978-1-4503-4140-0/16/03.
http://dx.doi.org/10.1145/2876456.2876457

Abstract
In this paper, we propose a learning framework for the adaptation of an interactive agent to a new user. We focus on applications where safety and personalization are essential, as Rehabilitation Systems and Robot Assisted Therapy. We argue that interactive learning methods can be utilised and combined into the Reinforcement Learning framework, aiming at a safe and tailored interaction.

Author Keywords
Interactive Reinforcement Learning, Interaction Management, Adaptation, Learning Agents

ACM Classification Keywords
I.2.6 [Artificial Intelligence]: Learning.

Context and Motivation
Interactive Learning Agents are entities that learn through continuous interaction with their environment (world, humans, other agents). A significant attribute of these agents is the adaptability of their behavior towards a goal, in a dynamic and stochastic environment. as when a human end-user is involved in the interaction [6].

Interactive agents have been successfully employed to Robot Assisted Therapy and Computer Aided Training systems, with applications to physical and cognitive

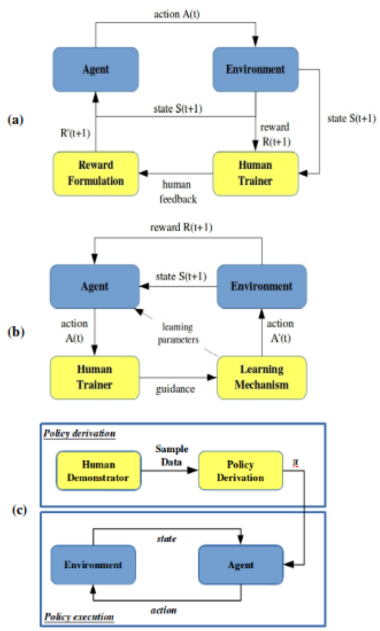

Figure 1: IRL approaches
(a) Learning from Feedback,
(b) Learning from Guidance,
(c) Learning from Demonstration

rehabilitation, vocational training and assistive robotics.

Two major aspects of such applications are *safety* and *personalization*. An intelligent interactive agent should be adaptive to the user needs, preferences and abilities [10]. Moreover, it should be robust to possibly inappropriate actions, that may lead to to an ineffective and harmful interaction [5]. Such applications may also support the participation of a *secondary user*, who participates as an external supervisor or observer, as the system interacts with the *primary user* (e.g. patient and therapist).

Background

In my research, I investigate how interactive learning methods can be used to facilitate the adaptation of an interactive agent to new users, following the Reinforcement Learning framework. In the next section, we briefly present applications of RL for modeling the interaction management of an interactive system.

Reinforcement Learning for Interaction Management

Interaction management is the problem of deciding on what to do in a given context, knowing that this context will be influenced by the decision [11].

Reinforcement Learning (RL) provides an appropriate framework for interaction optimization and has been successfully applied to model the interaction management of Adaptive Dialogue Systems [12],[4], Intelligent Tutoring Systems [2],[9] and Recommender Systems [8], considering the interaction to be a sequential decision making process. sharel;atex A key challenge of applying RL to interactive systems is ensuring a safe interaction while adapting the agent's behavior to each specific user, especially in sensitive environments, where exploration-based learning is not desirable. In the next section, we briefly present Interactive Reinforcement Learning, which studies how human interaction can change the agent learning process [14].

Interactive Reinforcement Learning

Interactive Reinforcement Learning (IRL) is a variation of RL that studies how a human can be integrated in the agent learning process. There are three approaches, based on how the human trainer intervenes with the learning process, as we show in Fig. 1.

Learning from Feedback treats the human feedback as a reinforcement signal after the executed action [7]. *Learning from Guidance* allows human intervention to the selected action before execution, proposing (corrective) actions [3] and *Learning from Demonstration* uses human demonstration samples to approximate a policy based on which the agent will interact with the environment [1].

These techniques refer to interactive systems, where the human trainer is not the primary user, but a secondary user that supervises the agent learning. We argue that both primary and secondary users can be integrated to the adaptation of the interactive agent, combining properly the different IRL techniques.

Statement of Thesis

In my research, I investigate how IRL techniques can be used for adaptation, exploiting the expertise of a secondary user that guides and supervises the interaction, when needed, as well as the *implicit* feedback that can be provided by the primary user in the form of an *affective signal* (facial expressions, speech, body posture, gestures).

We focus on applications as Rehabilitation Systems [15] and Robot Assisted Therapy [13]. Such systems must be able to assist the user on their task, while ensuring a safe and tailored interaction.

The contribution of this research will be a learning framework for interactive agents, following the RL paradigm, used to facilitate the agent's adaptation to a new user. We follow a *supervised autonomy* approach [13], where the system interacts with the primary user autonomously, while a secondary user intervenes with corrective information, when needed. We show the proposed framework in Fig. 2.

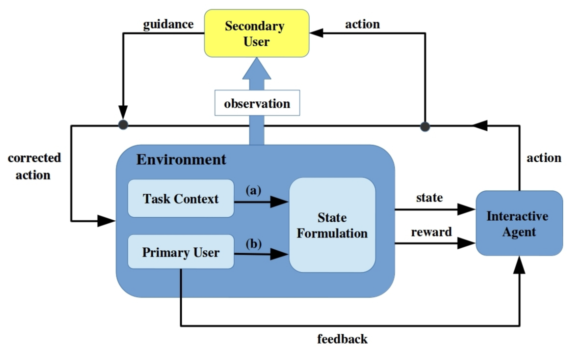

Figure 2: The proposed framework

The agent interacts with the primary user during a task execution (e.g. rehabilitation exercises). The system acts primarily autonomously, proposing actions based on its policy. The secondary user observes the interaction and intervenes with corrective information when needed. The primary user provides feedback, in an explicit way, that can be used for tailoring the interaction to their needs and preferences.

Both *communication channels* (guidance and feedback) must be properly integrated to the agent learning mechanism. The goal is to learn a progressively optimal policy, taking advantage of the information captured by both channels, during the adaptation process.

Open Questions

In this section, we address research questions that arise from the integration of IRL techniques to the adaptation process of an interactive agent.

How to combine guidance and feedback properly? Guidance and feedback can be seen as two different communication channels that need to be integrated properly to the RL mechanism. We need to develop appropriate methods to handle each channel separately, but also techniques that combine both channels properly into the RL framework.

How to consider and handle co-adaptation? An important aspect we need to consider is *co-adaptation*. Co-adaptation refers to the fact that as the system adapts to the users (primary, secondary), the users adapt their behaviors to the system, as well. Considering this, the amount of interventions (of the secondary user) should decrease over time. Similarly, the feedback of the primary user should be handled differently, as the agent learns.

How to simulate primary and secondary users? Using real users can be really time-consuming and infeasible. We need to find proper simulation techniques to test the framework with simulated users (primary and secondary), considering co-adaptation and the role of each user to the interaction.

Evaluation Metrics. Since we follow the RL framework, we will evaluate the adaptation using RL evaluation metrics, as learning performance and convergence speed. We also need to define proper evaluation metrics of the interface, regarding basic HCI concepts.

Conclusion

In this paper, we proposed a learning framework used for the adaptation of an interactive agent. We argue that integrating IRL techniques can facilitate a safe adaptation of the agent behavior to a new user, including a primary and secondary user in the learning process.

Acknowledgements

This material is based upon work supported by NSF under award numbers CNS 1035913, CNS 1338118. Moreover, this work is also supported by the educational program of NCSR "Demokritos" in collaboration with the University of Texas at Arlington.

References

[1] Argall, B. D., Chernova, S., Veloso, M., and Browning, B. A survey of robot learning from demonstration. *Rob.Aut.Sys 57*, 5 (2009), 469–483.

[2] Chi, M., VanLehn, K., Litman, D., and Jordan, P. An evaluation of pedagogical tutorial tactics for a natural language tutoring system: A reinforcement learning approach. *Int. Jour. of AI in Ed. 21*, 1-2 (2011), 83–113.

[3] Cruz, F., Twiefel, J., Magg, S., Weber, C., and Wermter, S. Interactive reinforcement learning through speech guidance in a domestic scenario. In *Proc. IJCNN*, IEEE (2015), 1–8.

[4] Cuayáhuitl, H., Kruijff-Korbayová, I., and Dethlefs, N. Hierarchical dialogue policy learning using flexible state transitions and linear function approximation. In *COLING* (2012), 95–102.

[5] Giullian, N., Ricks, D., Atherton, A., Colton, M., Goodrich, M., and Brinton, B. Detailed requirements for robots in autism therapy. In *Proc SMC 2010*, 2595–2602.

[6] Janarthanam, S., and Lemon, O. Adaptive generation in dialogue systems using dynamic user modeling. *Computational Linguistics 40*, 4 (2014), 883–920.

[7] Knox, W. B., and Stone, P. Reinforcement learning from simultaneous human and mdp reward. In *Proc. AAMAS* (2012), 475–482.

[8] Mahmood, T., Mujtaba, G., and Venturini, A. Dynamic personalization in conversational recommender systems. *Inf.Sys and e-Business Management 12*, 2 (2014), 213–238.

[9] Martin, K. N., and Arroyo, I. Agentx: Using reinforcement learning to improve the effectiveness of intelligent tutoring systems. In *Intelligent Tutoring Systems*, Springer (2004), 564–572.

[10] Picard, R. *Affective Computing*. MIT Press, 1997.

[11] Pietquin, O., and Lopes, M. Machine learning for interactive systems: Challenges and future trends.

[12] Rieser, V., and Lemon, O. *Reinforcement learning for adaptive dialogue systems: a data-driven methodology for dialogue management and natural language generation*. Springer Science & Business Media, 2011.

[13] Thill, S., Pop, C. A., Belpaeme, T., Ziemke, T., and Vanderborght, B. Robot-assisted therapy for autism spectrum disorders with (partially) autonomous control: Challenges and outlook. *J.of Beh. Rob. 3*, 4 (2012), 209–217.

[14] Thomaz, A. L., Hoffman, G., and Breazeal, C. Real-time interactive reinforcement learning for robots. In *Proc. AAAI* (2005).

[15] Tsiakas, K., Huber, M., and Makedon, F. A multimodal adaptive session manager for physical rehabilitation exercising. In *Proc. PETRA*, ACM (2015), 33.