A Practical Reinforcement Learning System for Training Robots Through Human Interaction

Adam White

In this poster we present a research program based on developing a practical reinforcement learning system that can learn from interaction with a human. Typically, reinforcement learning research focuses on the convergence of algorithms with a particular set of parameters on a particular task--a hands-off, non-interactive approach to learning. Our research leverages human interaction and guidance to speed up learning where the tasks are defined by human reward.

Our research is motivated by an interactive robot teaching problem that is not often considered in machine learning and robotics research. Consider a commercial robot that can be taught to perform various tasks that its owner deems important: collecting garbage from the floor or locating its charger. The owner is a layman; they cannot modify the robot's code in any way. Instead the owner must help the robot learn each task through demonstrations, rewards and other communication primitives. Machine learning methods seem like a natural basis for such an interactive learning robot. Most work on programming by demonstration, apprenticeship learning and imitation learning, however, focuses on leveraging human input at the beginning of learning to achieve near-optimal performance on a task.

A variety of unique challenges arise when a human is allowed to interactively train a learning robot. Firstly, a teachable system's performance cannot be dependent on tuned parameters. Alternatively, the parameters could be used as a tool to the help teach the system: asking the robot to explore more, for example. Second, a practical learning system could provide the teacher with a human interpretable representation of the system's knowledge about the world, to help the teacher better guide the robot. This is challenging because the robot's internal state may be in terms of the robot's high frequency sensor readings and motor-voltage commands (subjective representation). Finally, the system should have the ability to continue learning after the teacher has discontinued rewarding and instructing the robot. Converting values learned during teaching, for example, into persistent reward could allow the robot to continue to improve its behavior without human guidance.

Our research will focus on human interaction studies: testing new learning algorithms, human-robot interfaces, training techniques and state representations on physical and simulated robots with human trainers. These studies will attempt to evaluate the overall helpfulness and practicality of our learning system: robustness to imperfect teaching, efficiency of the training process and the efficiency of the interface between the teacher and the robot. The learning system should optimize the training experience for the human.

Our poster will present preliminary results on interactively training a physical robot using several psychologically inspired training techniques including successive approximations, chaining and environmental shaping. The testbed for these experiments is the Critterbot: an omnidirectional robot that is outfitted with an unusually rich set of simple sensors, including sensors for touch, heat, motion, sound, vision, light, proximity, and detailed motor state.