The RLAI Robotic Simulator

Marc G. Bellemare MGBELLEMARE@CS.UALBERTA.CA
Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada

Recent advances in Reinforcement Learning have focused on developing scalable, flexible algorithms. In this abstract, we propose a robotic simulator which acts as a rich environment for investigating Artificial Intelligence research challenges. Our research laboratory, the Reinforcement Learning and Artificial Intelligence (RLAI) laboratory at the University of Alberta, has been building a robotic platform for testing a variety of learning problems. This robotic platform, the Critterbot, has a wide array of sensors and is equipped with a holonomic drive. In parallel with the hardware robot, we have been developing a software suite, the RLAI Robotic Simulator. This simulator is the focus of this abstract.

1. A Research Tool

The RLAI Robotic Simulator is a flexible toolkit for building robotic environments. The simulator is designed to be able to describe a variety of situations, although the currently implemented components are influenced by the Critterbot platform. The software is written in Java and is an object-oriented, modular framework which users can easily customize. It is compatible with RL-Glue, a standard interface for reinforcement learning (Tanner & White, 2009), but may also be used as a standalone research tool. The simulator aims to satisfy a growing need in the RL community for large testbeds. As such, the emphasis has been put on flexibility over optimization. For example, it is easy for the user to quickly set up a Reinforcement Learning task by using the RL-Glue version of the simulator to provide a reward function and a description of the specific environment. It is also straightforward for the user to implement additional simulator components not already included in the simulator distribution.

2. A Challenging Framework

The simulator offers the Reinforcement Learning (RL) community many exciting research challenges. It provides large scale continuous environments with a rich action space. Physical interactions with various objects in the world yield dynamical environments. Additionally, there are currently more than ten different sensors implemented, including light, distance and bump sensors. There is also documentation on how to add new sensors and physicalities.

Agents may be designed with both short-term and long-term control in mind. In general, successful algorithms will have to learn to control the robot at multiple time scales. The two principal factors challenging traditional one-time-scale algorithms are the simulator's high operation frequency (100Hz) and the versatility of the platform, which allows complex tasks to be designed by the users.

Dealing with subjective observations - rather than objective state information - is a challenge in RL, partly due to the well-known complexity of partially observable domains (Cassandra et al., 1994). The simulator, being focused on generating robot-centric sensory information, provides such subjective observations. Using a robot-centric state has been studied in the context of learning policies in agent space (Konidaris & Barto, 2007) and discovery meaningful structure in robotic sensory data (Grollman et al., 2006).

The simulator is also meant to act as a lifelong learning research platform. An example of a lifelong learning agent is one which seek to indefinitely accumulate knowledge about its environment (Utgoff & Stracuzzi, 2002). Questions such as "What should an agent learning on a long term scale do?" may be answered by setting up complex experiments in the simulator. Because of the diversity of sensory information available via the simulated robot, discovery of its environment and its underlying structure may directly be addressed. Studying how best to transfer learned skills, or to learn transferable skills, is also an advantage of using the simulator over more traditional RL tasks.

3. A Set of RL Problems

The RLAI Robotic Simulator allows users to craft RL environments suited to answer their specific questions. At one end of the spectrum, it is possible to recover the objective information about the simulator's state - the location of objects of interest, for example. This is in sharp contrast with dealing with an actual robot, where complex calibration would be necessary to obtain such information. At the other end of the spectrum lies the completely subjective data provided by the robot's sensors. As such, the simulator can be used to create learning problems ranging from simple to extremely challenging.

Besides the choice of objective and subjective information, the simulator itself can be used to describe a variety of environments with little additional effort. Figure 1 shows a visualization of a particular simulator environment, the Garbage Collection environment. In this task, the robot is required to collect objects of various shapes and place them in the basket. The agent is rewarded for successfully moving objects to the basket.

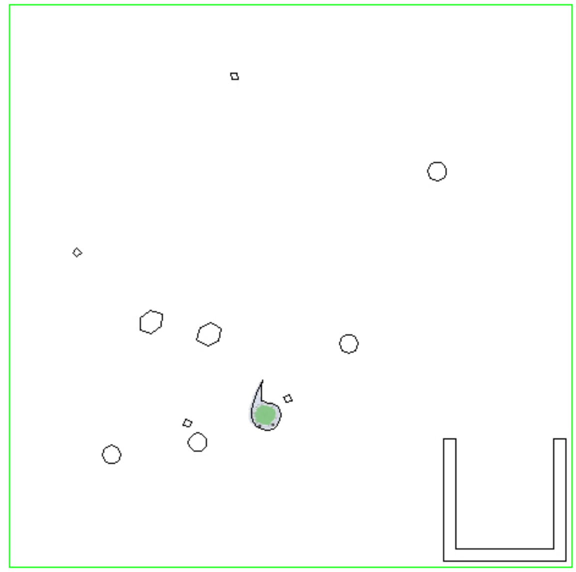

Figure 1. An example RLAI Robotic Simulator environment. Here, the Critterbot agent is tasked with collecting objects and bringing them to the bottom-right enclosure.

References

Cassandra, A. R., Kaelbling, L. P., & Littman, M. L. (1994). Acting optimally in partially observable stochastic domains. *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI).*

Grollman, D. H., Jenkins, O. C., & Wood, F. (2006). Discovering natural kinds of robot sensory experiences in unstructured environments. *Journal of Field Robotics*, *23*, 1077–1089.

Konidaris, G., & Barto, A. (2007). Building portable options: Skill transfer in reinforcement learning. *Proceedings of the 20th International Joint Conference on Artificial Intelligence* (pp. 6–12).

Tanner, B., & White, A. (2009). Rl-glue homepage. `http://glue.rl-community.org`.

Utgoff, P. E., & Stracuzzi, D. J. (2002). Many-layered learning. *Neural Computation*, *14*, 2497–2529.