Distributed Learning for Controlling Modular Robots

Paulina Varshavskaya, Daniela Rus and Leslie Kaelbling

What: Recently there has been an important research effort into modular, distributed robotics and in particular, self-reconfiguring robotics [2, 5, 8]. Issues with designing controllers for such systems range from constructing motor control primitives to ensuring cooperation between modules. For simpler tasks, such as locomotion in one direction, hand design is easy. However, as modular robots are tried in more complex domains or at more intricate tasks, designing effective and efficient controllers becomes a problem. We propose to have such robots learn their behaviors instead. We research the strategies and algorithms for learning controllers for self-reconfigurable robotic systems, in which each element has some computational and motor power, focusing in particular on applying reinforcement learning techniques. More generally, we are interested in developing distributed reinforcement learning algorithms that can be used to control distributed teams of robots and sensors.

Why: Reinforcement learning is especially appropriate whenever the problem involves agents interacting with an environment [7]. There are many standard techniques for learning optimal policies for agents in worlds which are fully observable Markov Decision Processes (MDPs). However, very few learning algorithms have been tried on physical robots, in particular because the real world, or even a simple subset of the real world, is at best a partially observable MDP. Where multiple robots, or robotic systems controlled by a distributed number of agents are concerned, there have been relatively few attempts at learning their behaviors [3, 4]. The vast field of research in multiagent learning (for example, [1]) focuses mainly on competitive games. Cooperative multiagent learning usually does not involve physical coupling. Self-reconfigurable robotic agents, on the other hand, must learn to cooperate and function while physically connected to their peers (as in simulation in figure 1). We would like to develop reinforcement learning approaches which can be applicable to such a setup of physically coupled robots with reconfigurable modular hardware and distributed processing.

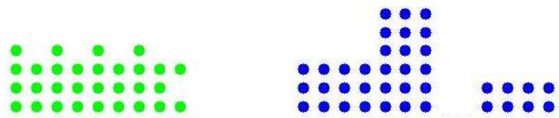

Figure 1: 2D simulation of self-reconfigurable motion. The modules (green) are approaching an obstacle course (blue).

The motivation for research in such distributed robotics is manifold. Modular systems can be more redundant, and therefore more fault-tolerant. Reconfigurable modular systems can be more versatile in locomotion and in function, reshaping themselves to achieve particular goals. Self-reconfigurable systems can do so autonomously, reducing the need for human operator intervention. In addition, self-assembling and self-reconfiguring robots with distributed computational abilities are a valuable platform for studies in artificial life.

Applying machine learning to the field of self-reconfigurable modular robotics will provide more adaptive machines. In addition to being able to automatically reshape for particular tasks in environments where human operation is undesirable, costly or hazardous, such systems could adapt to changes in the environment, and learn to perform new tasks.

How: Several things need to be done to address our problem fully. We compare a number of standard reinforcement learning techniques on our problem, both to see what solutions can be reused and to identify the hard parts of the problem. We will also design new algorithms specifically to address the setup of distributed cooperative agents controlling physically coupled robotic hardware. Finally we will need to build a physical platform for our learning experiments.

Progress: Currently, we have created a simulation world in which we conduct comparison tests (figure 1). The robot learns a simple locomotion task from local rewards that can be sensed by each module. We have implemented a local policy search algorithm, a modified version of GAPS with a lookup table [6], in each robotic module. We are running experiments in informed selection of starting points for searching policy space. One potential benefit of a modular system is the ability to add more components over time as the necessity arises. In our experiment, we start with only two modules and add more of them to the

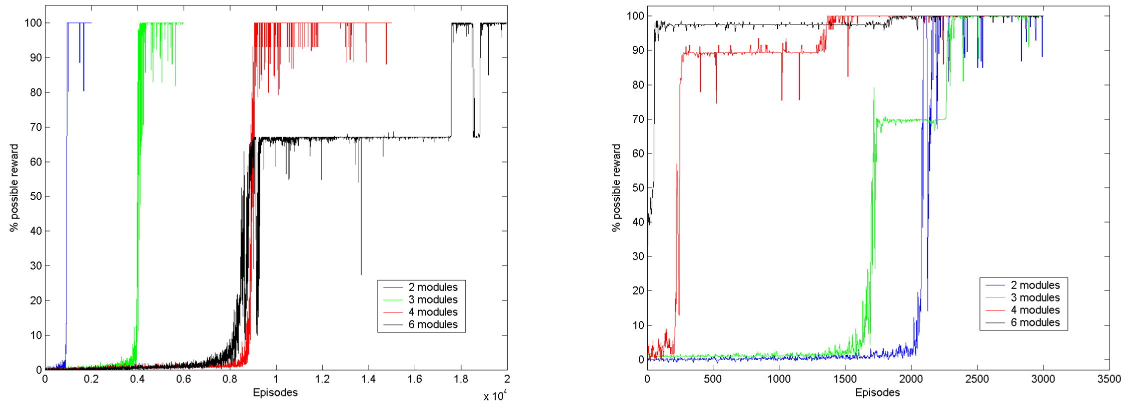

Figure 2: Results of learning using policy search with tied parameters of a simpler no-obstacle course. On the left, the starting point is always random. On the right, the starting point for the next number of modules is the learned policy of the lesser number of modules.

robot as the learning continues (figure 2). We are also experimenting with techniques designed for fully observable worlds such as Q-Learning, as well as with parameter tying between modules.

We are also working on learning to move over obstacles and perform turns. However, locomotion is not our sole aim. We are trying to further the understanding of the self-reconfigurable robotics domain – its theory, algorithms and applications.

Future: Our long-term goals are to develop distributed learning algorithms grounded in physical experiments that allow groups of autonomous agents to collectively learn, adapting to changing environments and new tasks. We plan to develop several learning techniques for distributed systems of robots, which ultimately can be used for automated coordination in any system composed of distributed hardware with processing capabilities. We would like to aim for functionality in self-reconfigurable robots - locomotion and reshaping could be used for exploration of unknown environments, rescue missions, robotic assembly or self-assembly. We plan to implement and test all our algorithms on physical hardware.

Research Support: This research is supported in part by NSF award 0225446 and by project Oxygen.

References:

[1] Michael Huhns and Gerhard Weiss, editors. *Machine Learning Journal, Special Issue on Multiagent Learning*, volume 33. 1998.

[2] Keith Kotay, Daniela Rus, Marsette Vona, and Craig McGray. The self-reconfiguring robotic molecule. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 424–431, Leuven, Belgium, 1998.

[3] Pattie Maes and Rodney A. Brooks. Learning to coordinate behaviors. In *Proceedings of the National Conference on Artificial Intelligence*, pages 796–802, 1990.

[4] Maja J. Matarić. Reinforcement learning in the multi-robot domain. *Autonomous Robots*, 4(1):73–83, March 1997.

[5] Satoshi Murata, Eiichi Yoshida, Harushisa Kurokawa, Kohji Tomita, and Shigeru Kokaji. Self-repairing mechanical systems. *Autonomous Robots*, 10:7–21, 2001.

[6] Leonid Peshkin. *Reinforcement Learning by Policy Search*. PhD thesis, Brown University, November 2001.

[7] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning*. a Bradford Book, the MIT Press, 1998.

[8] Mark Yim, David G. Duff, and Kimon D. Roufas. Polybot: a modular reconfigurable robot. In *Proceedings of the IEEE International Conference on Robotics and Automation*, 2000.