Autonomous Robot Skill Acquisition

George Konidaris
Autonomous Learning Laboratory
Department of Computer Science
University of Massachusetts Amherst
gdk@cs.umass.edu

Introduction

A credible theory of intelligence must provide a satisfactory account of the wide spectrum of learning abilities displayed by humans and some other animals. One of the most impressive of these is *skill acquisition*—the ability to create new skills, refine them through practice, and apply them in new task contexts. This ability to retain and refine solutions to important subproblems and then employ them later lies at the heart of two important aspects of human intelligence: the ability to perpetually improve at difficult control tasks through practice, and the ability to solve increasingly difficult problems.

The design and coordination of independent specialized skill units (often called behaviors) is fundamental to modern robotics (Arkin 1998). However, a robot that must act in a complex environment over an extended period of time should do more than just use existing skills: it should also learn new skills that increase its capabilities and facilitate later problem solving. Although robots exist that can learn a skill given a reward function and hand-engineered state space, none exist today that display truly autonomous skill acquisition.

Reinforcement learning (Sutton & Barto 1998) is a well developed theoretical framework for learning control policies in stochastic environments. The recent development of the options framework (Sutton, Precup, & Singh 1999) provides a theoretical basis for hierarchical reinforcement learning, including principled methods that tackle exactly the problems faced by an agent that must acquire and use new skills. I propose to adapt and apply the options framework to the robot skill learning problem, to create a robot capable of autonomously acquiring novel skills.

Approach

Previous robot learning research has considered skill learning largely in isolation. In contrast, I frame the skill acquisition problem in the context of an autonomous robot control architecture, as follows:

1. *The robot has an environmental context.* The robot is in the process of solving some task set in an environment that is unknown, but that is drawn from a class of environments whose shared properties are known.
2. *The robot has a task context.* The robot is equipped with a set of controllers sufficient for it to (possibly inefficiently or inconsistently) solve the task, and employs a learning or planning algorithm to sequence these controllers to do so.
3. *The robot has a control context.* While solving the task, the robot decides that something that it has just done should be isolated and learned as a new skill.

In this context skill acquisition serves two main purposes. First, a skill makes a subgoal prominent (by making it reachable using a single decision) in the planning or learning algorithm. Second, learning can make skill execution more efficient or consistent over time.

This setting has two important implications not commonly present in the robot learning literature. First, the robot is already capable of sequencing its existing controllers to achieve the goal of the target skill. In particular, it has a successful trajectory sample that it can use to decide how to learn that skill and which can be used to initialize the skill policy so that it does not need to learn from scratch.

Second, the robot is free to choose the extent of the learned skill. In particular, it may choose to learn *multiple* component skills to achieve a given subgoal, rather than a single monolithic skill.

Most prior research on robot skill learning has focused on the rapid learning of a skill given a hand-engineered state space and a prespecified reward function expressing the skill's goals. For skill acquisition to be truly autonomous we must remove these human design elements, and develop algorithms to determine *in what space a new skill should be learned*, as well as to determine *when a new skill should be created and what it should do*.

Selecting a Space

Most useful robots have multiple sensors and actuators from which a very large number of features can be extracted, and real-time learning is extremely difficult in such large state spaces. Learning only becomes feasible when the learning problem is posed in the right space—one in which a small number of relevant (often heavily preprocessed) features are present, and a large number of irrelevant features are not. As a result, successful robot learning applications always make use of a hand-engineered state space that is carefully designed to be task-relevant.

This poses a problem: if a robot is to autonomously decide which skills to acquire, how should it determine which state space to use for each skill?

I propose that instead of having either a single large but general state space or a single small but problem-specific state space, an autonomous robot should be equipped with a *set* of potential state spaces in the form of sensorimotor couplings, from which it can *select* when it decides to learn a new skill. This allows knowledge of the robot and the class of environments it faces to guide the design of a set of sensorimotor couplings sufficient to cover any skill the robot may decide to acquire, without requiring knowledge of the specific attributes of any single skill.

The first contribution of my thesis is a *state space selection algorithm* that selects an appropriate sensorimotor coupling for learning, given a sample solution trajectory (or set of trajectories). The algorithm performs an incremental least-squares weighted Monte Carlo fit for each coupling, and returns both an error metric for selection and a policy fit to the sample trajectory so that learning does not have to start from scratch.

In order to evaluate this algorithm I have developed a physically realistic simulation of a mobile robot approaching a door, opening it and moving through the doorway; the algorithm successfully selects an appropriate sensorimotor coupling for each subskill in the task. A paper reporting the results of this work is currently under review.

Creating Modular Skills

Regardless of the state space used to learn a skill, once it has been created and learned it can be integrated into the robot's control system as a new option (Konidaris & Barto 2007). This naturally leads to the question of when a new skill should be created.

The skill creation problem is well established, and several algorithms for it already exist in the options framework literature. However, they all assume discrete domains where experience is inexpensive and are thus ill-suited to robotics. I propose a mechanism for efficient robot skill acquisition consisting of a pair of interleaved processes as the second contribution of this thesis.

First, a robot exploring an environment using an existing set of behaviors creates new skills using *skill chaining*. Certain classes of events are designated as salient by the robot's designer and used as skill creation triggers. Once a robot encounters a salient event and creates a skill to reach it, the robot then also designates entering a state where that skill can be executed as salient (and thereby a skill creation trigger). This should result in a chain of new skills leading to a salient event, each in an appropriate space and each individually retained for later use. It also redistributes salience: skills created along a skill chain to reach a salient event also become salient, wherever they are later executed by the robot.

Second, the robot synthesises new skills out of existing skills using *skill composition*. When a pair of skills are consistently sequentially executed, they are composed into a single skill that efficiently simultaneously achieves the goals of both using coarticulation (Rohanimanesh & Mahadevan 2002). This should result in efficient specialized composite skills for common tasks to complement more general component skills.

Skill Acquisition on a Real Mobile Robot

Finally, I propose implementing these methods on a real mobile robot platform (the uBot-5, developed at the University of Massachusetts Amherst Laboratory for Perceptual Robotics) to create a robot capable of autonomously acquiring novel skills.

References

Arkin, R. 1998. *Behavior-Based Robotics*. Cambridge, Massachusetts: MIT Press.

Konidaris, G., and Barto, A. 2007. Building portable options: Skill transfer in reinforcement learning. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence*.

Rohanimanesh, K., and Mahadevan, S. 2002. Learning to take concurrent actions. In *Proceedings of the 16th Annual Conference on Neural Information Processing Systems*, 1619–1626.

Sutton, R., and Barto, A. 1998. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Sutton, R.; Precup, D.; and Singh, S. 1999. Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence* 112(1-2):181–211.