Learning Tasks and Skills Together from a Human Teacher

Baris Akgun, Kaushik Subramanian, Jaeeun Shim, Andrea L. Thomaz
Georgia Institute of Technology
College of Computing
801 Atlantic Dr., Atlanta, GA 30332
{bakgun, ksubrama, jaeeun.shim, athomaz}@gatech.edu

Introduction

Robot Learning from Demonstration (LfD) research deals with the challenges of enabling humans to teach robots novel skills and tasks (Argall et al. 2009). The practical importance of LfD is due to the fact that it is impossible to pre-program all the necessary skills and task knowledge that a robot might need during its life-cycle. This poses many interesting application areas for LfD ranging from houses to factory floors.

An important motivation for our research agenda is that in many of the practical LfD applications, the teacher will be an everyday end-user, not an expert in Machine Learning or robotics. Thus, our research explores the ways in which Machine Learning can exploit human social learning interactions—*Socially Guided Machine Learning* (SG-ML). We claim computational learning systems should make use of social learning principles because:

1. **It is better for the human:** A working hypothesis of this approach is that using characteristic aspects of human learning will be the most natural interface for everyday people. Several studies show that humans inherently and dynamically provide *social scaffolding* for learners. Greenfield describes studies of children learning language and girls learning to weave (Greenfield 1984), where the teachers engage in a complex process of dynamically adjusting their support of the learner based on skill level and success. The teachers are unconscious of the process or the method by which they are teaching. Thus, the partnership of social learning is an interaction that people fall into intuitively.
2. **It is better for the machine:** This point is generally less intuitive, but one way to think of it is that social interaction provides biases and constraints that simplify the problem for the machine. Social learning leads to a more efficient and robust machine learning process. We have several examples of this in our prior work (Thomaz and Breazeal 2008; Thomaz and Cakmak 2009; Cakmak et al. 2010; Zang et al. 2010).

We are interested both in novel algorithms and systems for LfD and in using everyday people as our benchmark for evaluation. As part of the AAAI 2011 LfD Challenge, we will be exhibiting our SG-ML approach to LfD.

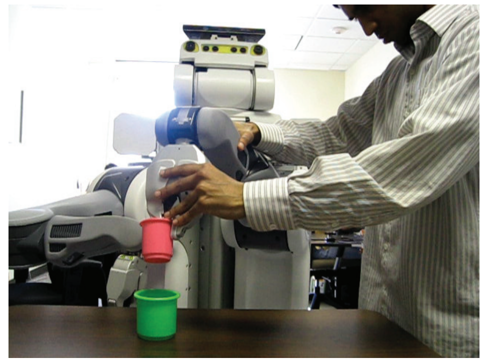

Figure 1: Interacting with the PR2 using kinesthetic demonstrations.

The domain of the AAAI 2011 LfD Challenge is food preparation tasks. Skills such as picking up, placing down, whisking, scooping, pouring and shaking are relevant to food preparation. Specific tasks could range from preparing cereal to making scrambled eggs.

We will be using the PR2 from Willow Garage as our robotic platform. The PR2 is a two arm mobile-manipulator with a variety of sensors. The arms have passive gravity compensation which we find desirable in human interaction scenarios. It uses the Robot Operating System (ROS) and has a developed codebase for various action and sensory capabilities.

Approach

We identify three parts of the LfD problem, low-level skill learning, high-level task learning and the interaction during demonstrations. The large body of literature have been concentrated on algorithms and representations related with either of the first two problems, while mostly neglecting the third. In our research we recognize that these problems are interrelated and tackle them together.

Task Learning

In prior work we have focused on the problem of task learning given a set of known primitive actions (Lockerd-Thomaz and Breazeal 2004; Breazeal and Thomaz 2008). In our task learning approach, the human instructs the robot on how to build a new task, through its components, from the set of known actions and tasks with speech and gesture. During task learning, the actions that the robot is asked to perform are monitored and the goals for each action as well as the overall task are inferred.

To infer a goal state for an action or task, the robot compares the world state before and after its execution. If this execution causes a change of state, this change is taken to be the goal. The learning problem for the robot is to identify what exactly about the state change was important for the task (was it the change to some particular object, a class of objects, etc.) To so do we expand a version space of all goal representations consistent with the current action/task example (Buchanan and Mitchell 1978). In a version space approach, there is a lattice of hypotheses consistent with the positive examples ordered from most specific to most general. Learning happens through a hypothesis elimination process as more examples of the concept are seen.

Skill Learning

In more recent work we have begun to also focus on the problem of skill learning. Skills can be regarded as low-level action primitives or policies that take the robot from one-state to another in a particular way. The low-level robot state can incorporate all or a subset of joint angles, end-effector (eef) pose, object-eef relations, basic sensory features (eef image position) and more. The two most common ways of representing skills are dynamic systems or mixture models.

We chose kinesthetic teaching as our skill demonstration method. In kinesthetic teaching, the human teacher guides the robot physically through the skill. Figure 1 provides an example of the kinesthetic interaction where the human is manually controlling the arm of the PR2 to demonstrate the required skill. We find that this method of interaction has significant advantages. It overcomes the correspondence problem often encountered in most LfD systems (Argall et al. 2009). The direct control of the robot's end-effectors eliminates the need for *record mapping* and *embodiment mapping*. Thus we can utilize the robot data directly and remove the effects of mapping errors. Furthermore, some of our recent work shows that humans are more comfortable using kinesthetic means when interacting with a robot than other methods (e.g. teleoperation). It provides them fine control over the demonstrated behavior that is not always available in other modes of interaction. We believe this aspect is essential for learning behaviors that are relevant in a kitchen environment.

Leveraging an Iterative Human Interaction

Finally, our approach recognizes that in a natural social learning interaction, both task and skill learning happen interactively within an ongoing social dialog.

In prior work, we have specifically shown the benefits of interactively collecting LfD examples as opposed to batch processing (Zang et al. 2010). This allows the human teachers to iteratively see the learned task model, and results in them having a better understanding of what kinds of examples or trajectories are most needed by the learner. We find that interaction helps people build the appropriate mental models of the learning process, and we see that they change their teaching strategies to improve learning.

In addition to the benefits of interaction, we will also be exploring the benefits of co-mingling task and skill learning within a single interaction. For example, the human can teach several low-level skills that they think will be necessary for a particular high-level food preparation task. Then, they begin teaching the high-level task by instructing the robot with a verbal dialog as to which skills to perform. In doing so, they may realize that one of the skills they taught needs modification, so they can switch to skill teaching and provide an additional demonstration of the skill, in the context of this high-level task. Afterwards, the high-level task instruction dialog continues.

Conclusion

We are interested in developing LfD systems that are tailored to be used by everyday people. As part of the AAAI 2011 LfD Challenge, we will demonstrate some of our most recent SG-ML work, in which the PR2 robot learns both low-level skills and high-level tasks through an ongoing social dialog with a human partner.

References

Argall, B.; Chernova, S.; Browning, B.; and Veloso, M. 2009. A survey of robot learning from demonstration. *Robotics and Autonomous Systems* 57(5):469–483.

Breazeal, C., and Thomaz, A. L. 2008. Learning from human teachers with socially guided exploration. In *IEEE International Conference on Robotics and Automation*.

Buchanan, B. G., and Mitchell, T. M. 1978. Model-directed learning of production rules. In Waterman, and Hayes-Roth., eds., *Pattern-directed inference systems*. New York: Academic Press.

Cakmak, M.; DePalma, N.; Arriaga, R.; and Thomaz, A. L. 2010. Exploiting social partners in robot learning. *Autonomous Robots* 29.

Greenfield, P. M. 1984. Theory of the teacher in learning activities of everyday life. In Rogoff, B., and Lave, J., eds., *Everyday cognition: its development in social context*. Cambridge, MA: Harvard University Press.

Lockerd-Thomaz, A., and Breazeal, C. 2004. Tutelage and socially guided robot learning. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

Thomaz, A. L., and Breazeal, C. 2008. Teachable robots: Understanding human teaching behavior to build more effective robot learners. *Artificial Intelligence Journal* 172:716–737.

Thomaz, A. L., and Cakmak, M. 2009. Learning about objects with human teachers. In *International Conference on Human-Robot Interaction (HRI)*.

Zang, P.; Tian, R.; Thomaz, A. L.; and Isbell, C. 2010. Batch versus interactive learning by demonstration. In *International Conference on Development and Learning (ICDL 2010)*.