Learning to Provide Better Examples for Our Robots

Anca Dragan
Robotics Institute, Carnegie Mellon
adragan@cmu.edu

Siddhartha Srinivasa
Intel Labs Pittsburgh
siddhartha.srinivasa@intel.com

1. INTRODUCTION

Good coaches can put themselves in the shoes of their trainees, and provide them with useful examples. Simply watching a professional squash player during a game often does not provide mere beginners with enough information for them to improve their skills: their muscles are not as well formed and they are missing some of the elementary skills to build upon. When a squash expert is asked to coach a beginner, however, he will probably learn the beginner's limitations over time, and end up providing different, more suitable examples than originally.

We have reasons to expect that the same holds for robots: for example, humans are utter experts in manipulation tasks, while robots are novices with different kinematics and much less impressive basic capabilities for the teacher to build on. Therefore, when humans teach robots certain manipulation tasks, they might have to come up with a strategy for the robot to use that is different from the strategy they would apply themselves. For example, if the robot's hands are shaped differently (e.g. HERB, at Intel Labs Pittsburgh, shown in Figure 1), then it will not necessarily be able to open a fridge door the same way a human does. When the human is asked to teach the robot, they will likely not be aware of this at first, but when observing how the robot fails, they will strive to come up with a new way to approach the problem that is easier for the robot.

In our current work, we are exploring how humans adapt to the capabilities of the robot when teaching them such manipulation skills, and what interface is the most effective for this kind of adaptive teaching. We are looking at two types of interfaces: human tracking followed by mapping movement to the robot's kinematics, and direct control (in which the human can directly move or teleoperate the robot). Our future goal is to create a framework for learning by demonstration in which the expert is not automatically regarded as ideal, but is allowed time to provide exemplars that are directly applicable to the robot.

2. PREVIOUS WORK

Today's robots are still novices in a multitude of domains in which the average human is an expert by comparison. Learning from demonstration is thus a popular field because it strives towards transferring skills from humans to robots in a natural, effective and general way. There are two main teaching interfaces: the robot can track the human and map the example to its own action range (Atkeson[1], Schaal[7]), or the human can exert direct control of the robot. This can be achieved either by physically moving it (Hersch[4]), or by teleoperating, e.g. via joystick (Chernova[2], Grollamn[3]) or via a graphical interface, like the one proposed at HRI Pioneers 2010(Koening[6]). The advantage of the direct control type is that it is trivially mappable to the robot, filtering out unsuccessful examples automatically. However, it can also be a lot less intuitive for the human, thus preventing them from achieving the intended example.

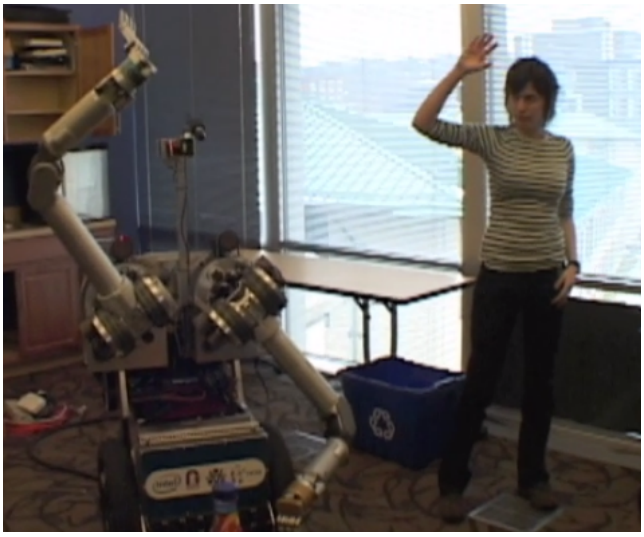

Figure 1: HERB imitates the motion of a human.

In a psychology study, Hinds[5] showed that it can be hard for experts to assess the capabilities of novices. If we extrapolate this to humans teaching robots, we can infer that the first exemplar a human might provide when demonstrating a task can very well be outside the robot's capabilities. Atkeson[1] found that in some cases, imitating the human directly is not feasible – unaware of what works for the robot, the human gives an example that is not well applicable. Another example is shown in Figure 2, when the expert is asked to demonstrate picking up the bottle, he chooses a trajectory that ends with a grasp from the top of the target object. However, HERB is not able to grasp in the same manner, thus the reaching trajectory not directly applicable: the robot can imitate it, yet at the end fail to accomplish its goal. This mismatch between expert and novice can happen in this case either because HERB was not yet taught how to grasp bottles from the top, or because HERB's hand is not adequate for that particular grasp.

Figure 2: The expert demonstrating how to reach for a bottle. Since HERB cannot grasp the bottle in the same way, the example is not particularly useful.

3. PILOT STUDY

Setup: In a pilot user study meant to evaluate a direct interface, 5 subjects were given 3 minutes to command a robot to successfully grasp an object in the absence of clutter. They could control it by planar motions of the end-effector using the iPhone as a joystick (a direct control interface) and hitting a grasp button when close enough to the target. The users could visualize the robot in a simulated environment (Figure 3).

Figure 3: The user attempting to teleoperate the robot using the iPhone.

Findings: The task has proven to be very challenging, with only one out of the 5 users successful. All users had problems with joint limits. Furthermore, the successful trajectory was unsatisfactory because it was very unsmooth (due to joint limits and planar hand motion). We expect that incorporating clutter into the scene (a situation in which the robot actually require assistance due to limited planning capabilities) would make successful demonstration even harder to attain. This indicates that the tracking interface (as opposed to the direct control one used in this study) has the potential to be more suitable for teaching, despite the difficulty in mapping the motion, provided the expert is given time to adapt and ensure that the examples are applicable to the robot.

4. PROPOSED EXPERIMENT

We propose an experiment targeted at the following questions:

1. *When it comes to teaching manipulation skills to a robot, does a human adapt over time to the limitations and constraints of the robot and the interface? In other words, does it take less time to successfully demonstrate how to accomplish a subsequent tasks than it does to accomplish the first one?*
2. *What type of demonstration interface is most effective for teaching manipulation skills?*

Interfaces: A first interface is a tracking one, in which a human demonstrates a trajectory to HERB, after which he/she observes the robot imitating the trajectory in a duplicate environment. The second interface is at the middle ground between tracking and direct control: the robot imitates the human's arm in real-time, which can be viewed as teleoperation. In case of failure, the human can reset the robot to the original configuration and attempt another demonstration. The third interface is direct control – the human moves the robot arm (with HERB in gravity compensation mode to allow for easy interaction). Again, the human can reset the robot.

Setup: Each subject is given basic information about an interface and 5 similar tasks for grasping a target object in a cluttered environment. We expect to see that the subsequent tasks become easier. We define a successful demonstration as one that yields a trajectory that the robot can re-execute without colliding and that results in a successful grasp. The teachers are told to stop when they are satisfied with the trajectory that robot executes, in terms of success and quality.

Measurable quantities:

- time required to get the first successful trajectory and the end trajectory (that the teacher is satisfied with), for each interface and task
- the cost of both trajectories, for each task (length and smoothness, defined as sum squared velocities)

Survey: The subjects will be asked to rate: 1) how fit the interface is for teaching, 2) how well they could get the robot to follow the movement they intended (ease of use), 3) the difficulty of each of the tasks and 4) how different the trajectory they originally had in mind is from the final trajectory (change in strategy) for each task.

5. DISCUSSION

We hope to find that humans do in fact adapt to the robot's capabilities (including the teaching interface), and to get an idea for what interface is more suitable for this sort of teaching. Ultimately, we want to show that learning from experience is facilitated by good examples that are adapted to the robot. We look forward to discussing this experimental setup at the HRI Pioneers Workshop, and developing new ideas about learning from adapted demonstration.

6. REFERENCES

[1] C. Atkeson and S.Schaal. Robot learning from demonstration. *International Conference on Machine Learning*, pages 11–73, 1997.

[2] S. Chernova and M. Veloso. ConïňĄdence-based policy learning from demonstration using gaussian mixture models. *International Conference on Autonomous Agents and Multiagent Systems*, 2007.

[3] D. H. Grollman and O. C. Jenkins. Learning robot soccer skills from demonstration. *International Conference on Development and Learning*, pages 276–281, 2007.

[4] M. Hersch, F. Guenter, S. Calinon, and A. Billard. Dynamical system modulation for robot learning via kinesthetic demonstrations. *IEEE Transactions on Robotics*, 24(6):1463–1467, 2008.

[5] P. Hinds. The curse of expertise: The effects of expertise and debiasing methods on prediction of novice performance. *Journal of Experimental Psychology: Applied*, 5(2):205–221, 1999.

[6] N. Koening, L. Takayama, and M. Mataric. Factors in learning from demonstrations. *HRI Young Pioneers Workshop*, 2010.

[7] S. Schaal, A. Ijspeert, and A. Billard. Computational approaches to motor learning by imitation. *Oxford University Press*, (1431):199–218, 2004.