

# An HRI Approach to Learning from Demonstration

**Baris Akgun, Kalesha Bullard, Vivian Chu and Andrea L. Thomaz**

School of Interactive Computing  
Georgia Institute of Technology, Atlanta, GA 30332  
bakgun@gatech.edu

## Introduction

The goal of this research is to enable robots to learn new things from everyday people. For years, the AI and Robotics community has sought to enable robots to efficiently learn new skills from a knowledgeable human trainer, and prior work has focused on several important technical problems. This vast amount of research in the field of robot Learning by Demonstration has by and large only been evaluated with expert humans, typically the system's designer. Thus, neglecting a key point that this interaction takes place within a social structure that can guide and constrain the learning problem. We believe that addressing this point will be essential for developing systems that can learn from everyday people that are not experts in Machine Learning or Robotics.

Our work focuses on new research questions involved in letting robots learn from everyday human partners (e.g., What kind of input do people want to provide a machine learner? How does their mental model of the learning process affect this input? What interfaces and interaction mechanisms can help people provide better input from a machine learning perspective?) Often our research begins with an investigation into the feasibility of a particular machine learning interaction, which leads to a series of research questions around re-designing both the interaction and the algorithm to better suit learning with end-users. We believe this equal focus on both the Machine Learning and the HRI contributions are key to making progress toward the goal of machines learning from humans.

In this abstract we briefly overview four different projects that highlight our HRI approach to the problem of Learning from Demonstration.

## Keyframe-based Learning from Demonstration

In the typical LfD teaching interaction, each demonstration is an entire state *trajectory*, that is, the teacher provides a continuous uninterrupted demonstration of the skill to the robot. Alternatively, we proposed learning from much more sparse *keyframe* trajectories, which are sequential points that the teacher demonstrates to the robot. In a series of user

studies, we evaluated keyframe demonstrations against trajectory demonstrations from an HRI perspective, revealing a set of advantages and disadvantages for each (Akgun et al. 2012). Some users' trajectory demonstrations were inconsistent and noisy making learning difficult or impossible. However, all the users were able to utilize keyframe demonstrations to teach successful skills. On the other hand, keyframes cannot be directly used to provide dynamics information. We have proposed *hybrid demonstrations*, which can have both trajectory or keyframe segments, to combine the advantages of both types of demonstrations. An example LfD interaction can be seen in Fig. 1a.

We have also observed that users concentrate on achieving the goal of the demonstrated skill rather than on consistent demonstrations of how to achieve it. We interpret this as people being task-oriented. This fits well with keyframe demonstrations. Users can use them to highlight salient parts of the skill, which is difficult to extract from trajectories.

## Learning Action and Task Models Simultaneously

Based on our observations about the users being task-oriented, we propose to learn both action and task models simultaneously. The action corresponds to how the robot moves and the task corresponds to the desired change of state associated with the demonstration. This is most relevant for object oriented tasks, for example closing a box. The action component would be the robot's end-effector getting under the lid moving up and then moving down on top the box. The task component would be to the state change of a previously opened box being closed.

The action model is responsible for executing the skill. Hence, the action model is learned using the data from robot's joint space, end-effector space or the control space. The task model is responsible for monitoring the execution success. Since we are interested in object-oriented skills, the task models are learned from object data, extracted from an RGBD camera. Both types of data are gathered during the same set of demonstrations.

## Goal Learning from Demonstration

In addition to learning primitive actions, we also work on learning high level tasks comprised of sequences of these

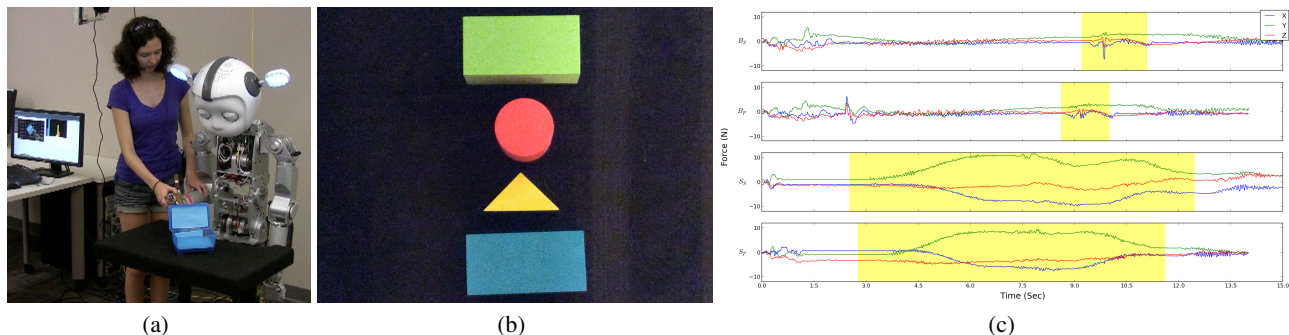


Figure 1: (a) A kinesthetic interaction for skill learning. The user guides the robot’s arm to demonstrate the skill. Both action data and object data are collected. (b) A valid goal state for the *Make Sandwich* task, composed of four *pick-and-place(object, location)* actions: place the top slice of bread, meat, cheese, and bottom slice of bread (c) Affordance learning: force data captured during closing the box (label *B*) and scooping beans tasks (label *S*). Subscripts *S* and *F* represent success and failure. Highlighted portions are when the robot touched the object.

actions. We are interested in object-directed tasks, whereby the goal is achieved by some configuration of objects. Our prior work focuses on learning a generalized model of the task goal; We explored the use of active learning to introduce transparency into the learning process and refine the goal representation based on human feedback (Chao, Cakmak, and Thomaz 2010) and examined the transfer of knowledge from known tasks in order to learn novel tasks faster (Chao, Cakmak, and Thomaz 2011).

Our current work extends this learning of task goals to learning the goal representation for the entire sequence of actions. Fig. 1b shows a goal state for the *make sandwich* task abstraction. It is composed of four pick and place actions on a top slice of bread, meat, cheese, and a bottom slice of bread. During a demonstration, the teacher provides the sequence of actions, while an image of the robot’s workspace is recorded, before and after each action. Each action is represented by *preconditions* and *postconditions*. Since each action is object-directed, preconditions characterize the object to be selected (*e.g.* color, shape, size) and postconditions characterize placement of the selected object (*e.g.* location). The hypothesis for preconditions consists of conjunctions of discrete object attribute values, and the hypothesis postconditions is a two-dimensional Gaussian distribution of the object’s goal location. The preconditions are learned by pruning the object attributes and the postconditions are learned by fitting the distribution. To meet the preconditions, an object must possess all of the relevant values. To meet the postconditions, the object must be placed within some variance of the distribution.

Teachers will have limited time and patience in a realistic scenario, resulting in a limited set of demonstrations. With such a limited data set, the resulting learned models will be uncertain. We are currently developing an active learning framework to address this.

### Affordance Learning from Demonstration

In the two previous examples, only visual features were considered in the expected perceptual state change over an action. To improve the execution of skills and tasks, we want

to incorporate haptic information into the robot’s model of a manipulation skill. Specifically, we are interested in understanding how force/torque (F/T) data and LfD can play a role in affordance learning. Affordances, defined as “action possibilities” by psychologist J.J. Gibson (Gibson 1977), look at the relationship between agents, actions, and the environment. For example a robot arm (agent) tapping (action) a ball on a flat surface (environment) has the affordance of rolling. Once learned, affordances can be used for many tasks such as planning for task execution (Krüger et al. 2011) and imitation learning (Montesano et al. 2008).

Research in this area has traditionally focused on visual information and building controllers to execute the motion primitives. By using LfD, we can seed motion primitives for the robot to execute rather than creating a controller for each primitive. Furthermore, prior work in this area (Thomaz and Cakmak 2009) demonstrate that humans are particularly good at providing rare environment configurations that can improve and speed up the affordance learning process.

Initial results where we collected F/T information from a robot arm performing a series of tasks demonstrated from a human operator show that LfD is a viable solution for exploration for affordance learning. The robot successfully performed five unique tasks (*e.g.* tightly closing a box or scooping beans) that traditionally would have taken an expert weeks to tune a controller for. The data collected from the F/T sensors can be seen in Fig. 1c and show promise for integrating haptic information to develop a richer model of affordances using both visual and F/T information.

### Conclusion

Our research aim is to develop interactions and algorithms for learning from naïve human teachers through demonstration. In all of the four aforementioned projects, we take an HRI approach to the problem of LfD. In our work, we have introduced ways to get better demonstrations from users with keyframes and active learning for both low level and high level tasks. We leverage the user behavior with learning task models during low level skill demonstrations and to learn otherwise difficult to discover affordances.

## References

- Akgun, B.; Cakmak, M.; Yoo, J. W.; and Thomaz, A. L. 2012. Trajectories and keyframes for kinesthetic teaching: A human-robot interaction perspective. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, HRI '12, 391–398. New York, NY, USA: ACM.
- Chao, C.; Cakmak, M.; and Thomaz, A. L. 2010. Transparent active learning for robots. In *Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on*, 317–324. IEEE.
- Chao, C.; Cakmak, M.; and Thomaz, A. L. 2011. Towards grounding concepts for transfer in goal learning from demonstration. In *Development and Learning (ICDL), 2011 IEEE International Conference on*, volume 2, 1–6. IEEE.
- Gibson, J. 1977. The concept of affordances. *Perceiving, acting, and knowing* 67–82.
- Krüger, N.; Geib, C.; Piater, J.; Petrick, R.; Steedman, M.; Wörgötter, F.; Ude, A.; Asfour, T.; Kraft, D.; Omrčen, D.; et al. 2011. Object–action complexes: Grounded abstractions of sensory–motor processes. *Robotics and Autonomous Systems* 59(10):740–757.
- Montesano, L.; Lopes, M.; Bernardino, A.; and Santos-Victor, J. 2008. Learning object affordances: From sensory–motor coordination to imitation. *Robotics, IEEE Transactions on* 24(1):15–26.
- Thomaz, A. L., and Cakmak, M. 2009. Learning about objects with human teachers. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, 15–22. ACM.