SLHAP: Simultaneous Learning of Hierarchy and Primitives

Anahita Mohseni-Kabir, Changshuo Li, Victoria Wu, Daniel Miller, Benjamin Hylak, Sonia Chernova, Dmitry Berenson, Candace Sidner, and Charles Rich

Worcester Polytechnic Institute
100 Institute Rd, Worcester, MA 01609
rich@wpi.edu

In robot learning from demonstration (LfD), a human teaches a robot how to perform a task by executing the task himself. For complex tasks, such as the tire rotation shown in the accompanying video,[1] this involves learning at two levels: the robot needs to learn the motion *primitives* and also how these primitives are combined into a *hierarchy* of steps to achieve the complete task. These two kinds of LfD have traditionally been studied separately. The contribution of this work is a novel human-robot interaction paradigm, called SLHAP (for simultaneous learning of hierarchy and primitives), in which these two kinds of LfD are interleaved in a way that is natural for a human teacher.

We have implemented a SLHAP proof of concept system in which an autonomous robot learns from a human teacher through a mixture of narration, in which the human speaks the name of a primitive when he executes it, and dialogue, in which the human answers the robot's questions about how to group primitives into subtasks. The human's motions are also tracked using a Vicon motion capture system.

Learning Task Hierarchy

The robot uses the techniques described in [2] to interactively learn a hierarchical task network (HTN) for a simple form of tire rotation. The robot asks the human questions based on two heuristics that group actions that (1) use the same object, e.g., picking up and then hanging a tire, or (2) are repeated on multiple objects of the same type, e.g., unscrewing the nuts on three studs of a hub. The human is also asked to provide a name for each new subtask, so that it can be used later in the interaction.

Learning Primitives

Learning task primitives is a two-step process. First, the section of motion data corresponding to each primitive action is identified using the techniques described in [3]. Second, the task space region (TSR) motion planning constraints defining each primitive action are learned from the motion data using the techniques described in [1].

In the first learning step, the human narrations provide a very rough estimate of the beginning and end of each primitive action, which is then refined using motif-based pattern recognition. The most representative instance of each type of primitive action is passed to the TSR learning step.

Learning the TSR constraints for a primitive is more powerful than learning an abstracted motion trajectory, because the TSR constraints allows the action to be used in more varied contexts, such as with different obstacles.

Limitations of Proof of Concept System

- Speech recognition and understanding is not general-purpose; we use a push-to-talk button operated off-screen and a predefined grammar for the human utterances.
- Movement of the robot base is not autonomous; base is controlled by offscreen joystick
- Learning primitives is not real-time; the primitives that the robot executed in the video were previously learned offline from similar motion capture data in which each primitive was demonstrated in isolation.
- Four (of eight total) primitive action types were not learned by demonstration; however, these actions (picking up and putting down a tire or a nut) do not need TSR learning, because they are only constrained at their endpoints.

This work is supported in part by the Office of Naval Research under grant N00014-13-1-0735.

References

[1] Changshuo Li and Dmitry Berenson. "Learning Object Orientation Constraints and Guiding Constraints for Narrow Passages from One Demonstration". In: *Int. Symp. on Experimental Robotics*. 2016.

[2] Anahita Mohseni-Kabir et al. "Interactive Hierarchical Task Learning from a Single Demonstration". In: *Proc. ACM/IEEE Int. Conf. on Human-Robot Interaction*. 2015.

[3] Anahita Mohseni-Kabir et al. "What's in a Primitive? Identifying Reusable Motion Trajectories in Narrated Demonstrations". In: *IEEE Int. Symp. on Robot and Human Interactive Communication*. 2016.

[1] https://youtu.be/GXjoybXFD70

HRI '17 Companion March 06-09, 2017, Vienna, Austria

ACM ISBN 978-1-4503-4885-0/17/03.

DOI: http://dx.doi.org/10.1145/3029798.3036641