Learning Behavior-Grounded Tool Affordances with Generalization Accross Different Tools

Jivko Sinapov
Department of Computer Science
Iowa State University
jsinapov@cs.iastate.edu

Alexander Stoytchev
Department of Computer Science
Iowa State University
alex@cs.iastate.edu

I. Introduction

In previous work, Stoytchev [1] introduced a method for representing the affordances of tools by grounding the representation in the behavioral and perceptual repertoire of the robot. In this representation the affordances of the tool are expressed in concrete terms (e.g., behaviors and observable outcomes) that are directly available to the robot.

This work extends that model by introducing a framework which allows the robot to learn a compact predictive model for the affordances of a tool. Experiments are conducted that highlight the model's generalization properties. The robot is evaluated on how well it can use knowledge ackquired from familiar tools in order to predict the affordances of novel ones. The robot is also tested on how well it can handle familiar tools whose size has been changed after the training stage.

II. Experiments and Results

All experiments were performed using the open-source dynamic robot simulator BREVE. The robot is a simulated arm with 6 degrees of freedom and a gripper attached to the wrist. Six different tools: T-stick, L-stick, L-hook, Stick, T-hook, and Paddle were used in the experiments. The robot is capable of grasping each tool and sliding it in the horizontal plane in any direction in order to affect the position of a small cylindrical puck. The task of the robot is to predict the displacement of the puck as it performs an action with the tool (see Figure 1).

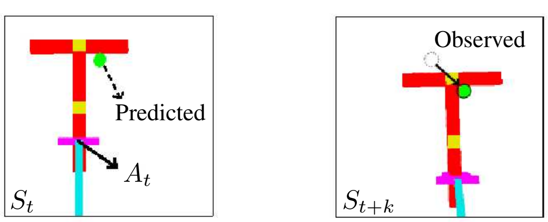

Fig. 1. A sample trial with the T-stick tool. At the start of the trial (left), the robot makes a prediction regarding the outcome of its tool action, A_t, based on the current sensory input, S_t. Once movement of the puck is detected, the robot is able to verify its prediction by comparing it with the observed outcome (right).

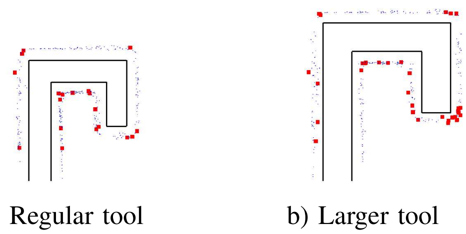

Fig. 2. Visualization of the prediction errors made for the regular L-hook tool (a) and its enlarged version (b). In both cases, the model is trained on the regular-sized tool. Each point in the two plots represents the puck's starting position relative to the tool during some particular learning trial. The points represented by the large squares indicate cases in which the predicted and the observed displacement vectors of the puck differ by more than $20°$.

The robot collects sample data during a behvaioral babbling training stage. The robot is allowed to use multiple frames of reference, i.e., it can focus on the tool, the puck, its gripper, or the center of the camera image, when building a model and making predictions.

The robot was tested on predicting the affordances of familiar tools, novel tools, and larger versions of familiar tools. Generalization across new tools is best achieved when the novel tool shares similar features with the tool on which the robot was trained. Figure 2 shows a visualization of the prediction errors of the model for an L-hook tool. The likelihood of prediction error is higher if the puck is located near one of the tool's corners and lower if the puck is near one of the tool's smooth surfaces. The plot also shows that the robot is capable of using the model trained on the regular-sized L-hook even after the size of the tool has been increased. MPEG movies of the behavior babbling training stage and details regarding the learning methods used to construct the models are available at: (http://www.cs.iastate.edu/~jsinapov/PosterRSS/).

References

[1] A. Stoytchev, "Behavior-Grounded Representation of Tool Affordances," In Proceedings of IEEE International Conference on Robotics and Automation (ICRA), 2005.