Learning Modular Neural Network Policies for Multi-Task and Multi-Robot Transfer

Coline Devin Abhishek Gupta

October 2016

1 Introduction

This paper talks about an application of Transfer Learning (It is the learning of a new task through the transfer of knowledge from a related task that has already been learned) and how it can help enabling zero-shot learning (Being able to solve a task despite not having received any training examples of that task). This reduces the importance of collecting real-world data for Reinforcement Learning problems (RL). RL has been able to automate a wide variety of robotic skills, but learning each new skill requires considerable real-world data collection. In this work the authors use Deep RL to train neural network policies which has some benefits but it worsen the challenge of data collection. Transfer learning is used to mitigate this problem by enabling transfer of information from one skill to another and even from one robot to another.

2 Theoretical Details

The work aims at teaching multiple robots to do multiple tasks (the tasks for all the robots are same) using deep RL. Authors shows that that neural network policies can be decomposed into ”task-specific” and “robot-specific” modules, where the task-specific modules are shared across robots, and the robot-specific modules are shared across all tasks on that robot. This allows for sharing of task information, such as perception, between robots and sharing robot information, such as dynamics and kinematics, between tasks. The decomposition to train mix-and-match modules that can solve new robot-task combinations that were not seen during training. Using a neural network architecture, they demonstrate the effectiveness of the method for enabling zero-shot generalization.

3 Experiments Results

To experimentally evaluate the network, tests are conducted on a number of simulated environments. For each experiment multiple robots and tasks are
used. The results showed that network allows them to train on a subset of possible worlds (a specific robot doing an specific task) in a universe (all possible combinations of robots and tasks) and achieve fast or zero-shot learning for an unseen world.

The evaluation is done against the baseline of training a separate policy network for each world instantiation, several simulated environments were made using the MuJoCo physics simulator. The model is evaluate on tasks that involve discontinuous contacts, moving and grasping objects, and processing raw images from simulated cameras. Further details and videos can be found [here].

The results showed that the task-specific modules are robot-invariant, and the robot-specific modules are task-invariant. This invariance allows the composition of modules to perform tasks well for robot-task combinations that have never been seen before. In some cases, previously untrained combinations might generalize immediately to the new task, while in other cases, the composition of previously trained modules for a new previously unseen task can serve as a very good initialization for speeding up learning.

4 Future Scope

A limitations of the current work is that, by utilizing the standard RL algorithms, the method requires different task-robot combinations to be trained simultaneously. A promising direction for future work will be to combine the approach with more traditional, sequential methods for transfer learning, such that the same robot can learn multiple tasks in sequence, and still benefit from modular networks. This would enable combined lifelong and multirobot learning.