Teaching Robots: Mastering Tasks with Reinforcement Learning

Popular science summary of the Master Thesis [1]
Teodor Åstrand, June 2024

Combining machine learning and robotics has led to a new era of intelligent robots. Robots would previously rely on a human to guide their decisions, however, the integration with machine learning unveils new opportunities. By combining reinforcement learning with robotics tasks and utilizing technologies such as deep neural networks, a robot can gain the ability to learn from and adapt to dynamic tasks.

This thesis employed a reinforcement learning agent as the brain of a seven-axis robot arm and focuses on teaching agents in a simulated robot environment with reward functions as guidance. The purpose was to gain insight into how an agent controlling a robot can learn to complete a given task by itself, without a human specifying the robot's path and movements.

The agent's view of a robot task can be interpreted as a game where it strives to maximize its accumulated reward. The rewards were therefore designed to encourage the agent to control the arm toward completing a task. As an example, large rewards were given to the agent if it managed to grasp a targeted object.

Agents were trained for a variety of tasks, where they directly commanded the robot joints to manipulate the robot's movement. This differs from traditional approaches for controlling robot arms where inverse kinematics and path planning are common strategies. The robot tasks were designed to explore how to teach primitive robot actions such as pushing, grasping, and avoiding an object. Thereafter agents were taught to combine and improve upon their strategies to carry out more complex tasks involving extracting an object through grasping while not disturbing other surrounding objects. The purpose was to gain insight into how agents could be taught to isolate a target object, by extracting surrounding objects.

Developing agents capable of learning robot strategies that can be carried over from simulation to a real process was a central objective during the task design and agent training. The thesis found that agents could easily learn strategies to complete a task in unconventional and undesired ways. For example, for a task, where the objective is to remove a cube from a zone, the agent could develop strategies involving sweeping swiftly across the zone and knocking the cube away.

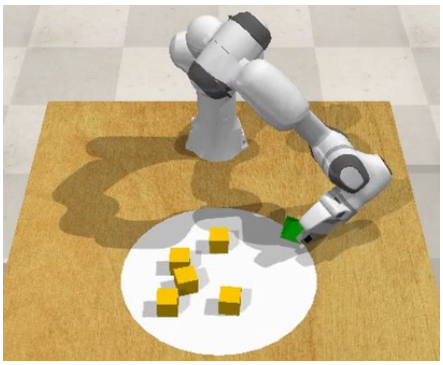

Figure 1: A simulated robot environment in which a reinforcement learning agent can learn to complete robotics tasks.

Rewards were designed to counter undesired behavior and favor good approaches to solving robot tasks. This resulted in finding methods to teach good, grasping approaches and more reliable cube extraction techniques. Furthermore, the thesis explored different ways to communicate a robot task environment to an agent and investigated how an agent can be brought from a simulated robot to its real-life counterpart. Strategies to achieve this were found and encountered problems for the simulation to real process were discussed.

[1] Teodor Åstrand. *"Robot Reinforcement Learning for Object Isolation"*. Master's thesis, Dept. of Automatic Control, LTH, Lund University, 2024. Available at https://lup.lub.lu.se/student-papers/search/.