Automatic Design of Behaviors for Khepera IV Robots

Justin Manson

Professor : Alcherio Martinoli
Assistant(s) : Cyrill Baumann

With the growing use of mobile robotics in real-world situations, robots encounter more diverse challenges. A hierarchically structured controller can address this by comprising specialized sub-components and an arbitrator to choose the right sub-component for each situation. Currently, sub-components for tasks like obstacle avoidance are mainly programmed manually. Our work aims to automatically generate a library of basic skills which could then be leveraged using behavioral arbitrators for more complex scenarios.

We start from the results of a recent semester project and implement the Dynamics-Aware Unsupervised Discovery of Skills method (DADS). This method learns diverse skills by maximizing the mutual information between the skill variable and the state transitions it produces. The idea is that under each skill, the robot should produce distinctive and easily inferable state transitions.

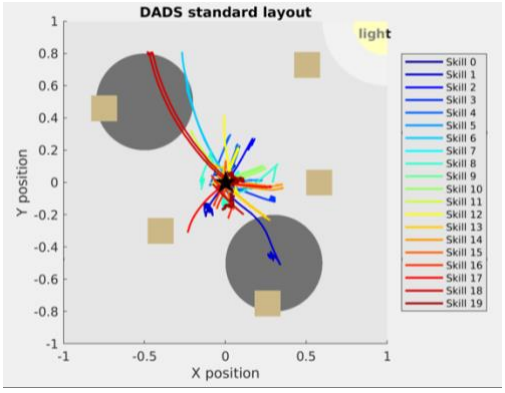

Figure 1: Paths taken at test time for twenty learned behaviors using DADS.

To the best of our knowledge, all mutual-information based skill discovery methods suffer from the “curse of dimensionality” issue. As the dimension of robot state increases (i.e. using more sensors), mutual information becomes a meaningless measure of diversity, as even slight changes in a robot’s position can yield a drastic change of state.

To address this limitation, we explore an alternative framework for skill discovery, using reward functions encoded by neural networks, trained iteratively to reward increasingly complex behaviors. This method leads to a DFS-like exploration of the environment. We refer to this method as 'neural rewards' (NR).

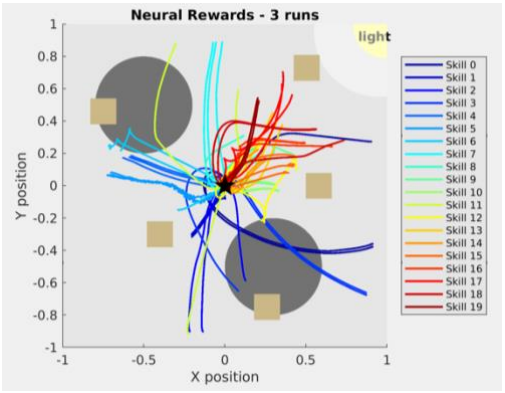

Figure 2: Paths taken at test time for twenty learned behaviors using NR.

We also develop our own skill discovery method, details of which will be published in a separate report. Our method is geared towards wheeled robots and yields highly diverse and dynamic skills, while keeping the library size small.

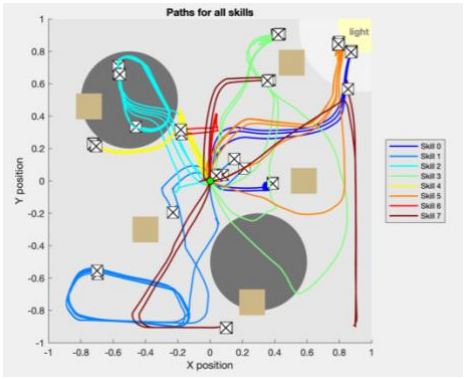

Figure 3: Paths taken at test time for eight learned behaviors using our own method.