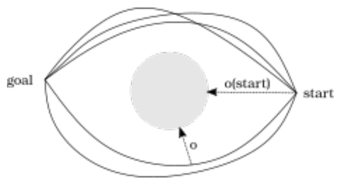
Robot Learning
Weekly Exercise 4

Marc Toussaint & Wolfgang Hönig
Learning & Intelligent Systems Lab, Intelligent Multi-Robot Coordination Lab, TU Berlin
Marchstr. 23, 10587 Berlin, Germany

Summer 2024

1 Multi-Modal Distributions

Consider a circular single integrator robot with 2D single integrator dynamics ($q = (x, y)$, $u = (v_x, v_y)$, $\dot{q} = f(q, u) = u = (v_x, v_y)$). The robot is equipped with a LIDAR and processes the resulting point cloud to get observation $o = (d_x, d_y)$, i.e., a vector pointing to the closest boundary point of any obstacle, see the figure for some examples (dotted lines). From experts, we obtained five example trajectories for a given scenario with a single circular obstacle in the middle, see the figure for these trajectory (black lines). Our goal is to learn a policy that directly maps observations to controls ($\pi : o \mapsto u$).



a) Discretize the observation into 8 parts (4 directional ranges and 2 magnitude ranges). For each of these possible input ranges, we "learn" the optimal action assuming an MSE loss. Draw the resulting action vectors (one for each possible observation) qualitatively.

b) Use the learned policy from a) and draw the resulting solution trajectory qualitatively.

c) Now consider that we learn a Gaussian Mixture Model (GMM) with two modes per discretized observation instead. Draw the resulting action distributions (one for each observation) qualitatively.

d) Explain how you can use the learned policy from c). Draw the resulting solution trajectory distribution qualitatively.

e) What changes if we do not discretize the observation? Explain what possible policy function approximators you might use, what learning algorithms are applicable, and what the expected outcomes compared to b) and d) are.

2 Application Problems

This is a discussion question, where there is no single correct solution.

Consider the following robotics problems, where our goal is to use imitation learning to mimic the outcome of an expert effectively and efficiently. For each problem, define i) what exactly should be learned, ii) how to gather the data, iii) and how to train the model.

a) A model-car that is racing on a fixed off-road track. The car has an IMU and RGBD camera. The expert is an external high-performance computer running nonlinear MPC with access to a tracking system that can accurately estimate the pose of the car. The goal to race using on-board sensors, only.

In addition to the general questions above: What needs to change if iv) the track is not fixed?

b) A robotic manipulator with two endeffectors tasked with folding laundry. The robot has force/torque sensors on each joint and a overhead RGBD camera. The experts are humans with direct line-of-sight (but without access to the on-board sensors).

3 Mountain Car Imitation Learning

This is a coding exercise. Please bring your laptop and connect to the HDMI in the tutorial to show your results. (If you upload a pdf, just include a screenshot of results in the pdf.)

a) Option1: DAgger for the mountain car (would need to have a different expert generator, e.g., human-in-the-loop) Option2: Generative model to learn a distribution (e.g., CVAE, diffusion); should have a second expert policy (e.g., first accelerating the other direction) to show multi-modal distribution learning Option3: Combination of the two options.