

Deep Reinforcement Learning Homework 5-a: Exploration

Ex 1-3, PointMass:



This figure shows the performance of : no exploration; exemplar density count estimation; histogram state counting; RBF density count estimation. We can see that without an exploration bonus, the curve stagnates pretty early on, and peaks at ~70 avg return. RBF and Exemplar density models both rise very fast, but RBF starts falling off for some reason. Histogram takes more time to peak, but finally reaches the same level of performance as Exemplar density estimation.

EX 4, RoboschoolHalfCheetah-v1:



With half-cheetah, all experiments were too noisy to make any conclusions.