

Reinforcement Learning Lecture: Homework 11

Ngo Anh Vien

MLR, University of Stuttgart

June 24, 2016

1 Exercise 01

This homework implements the episodic natural actor-critic algorithm (slide 46). The setting is similar to the last exercise, with some following changes:

- the variance of the Gaussian controller changes to 0.1
- the maximum length of each episode changes to 500
- using RPROP for an adaptive step-size

Results to report: run the program with a various setting of the number of sampled trajectories from 50 to 2000 (from very few trajectories to many), report the result, and explain it.