

# Reinforcement Learning Lecture: Homework 10

Ngo Anh Vien

MLR, University of Stuttgart

June 16, 2016

## 1 Exercise 01

This homework implements three policy gradient algorithms. For each step of gradient update, you can use a fixed step-size  $\alpha = 0.07$  or RPROP. Other settings are similar to the previous homework.

1. Vanilla policy gradient
2. REINFORCE
3. G(PO)MDP

Results to report: the performance of three algorithms for 500 iterations of gradient updates. Report the results for both a fixed  $\alpha$  and the use of RPROP.