

Autonomous Agents Assignment 2

Tobias Stahl
10528199

Spyros Michaelides
10523316

Ioannis Giounous Aivalis
10524851

Francesco Stablum
6200982

September 30, 2013

1 Introduction

2 Q-Learning

The initial exercise of this lab assignment is to implement Q-learning the temporal-difference (TD) learning algorithm. This algorithm is to be used by the predator agent to catch the prey.

2.1 Summary

Q-learning is an off-policy TD control algorithm. An off-policy TD algorithm is one in which the estimated value functions can be updated using hypothetical actions, without having actually executed the actions themselves. Using this approach the algorithm can separate exploration from control, meaning the agent could learn through from the environment without necessarily having had the explicit experience. Q-learning

References

- [1] <http://http://www.cse.unsw.edu.au/cs9417ml/RL1/tdlearning.html>
- [2] <http://webdocs.cs.ualberta.ca/sutton/book/ebook/node65.html>