Reinforcement Learning

Jonathan Somer

April 28, 2018

Reinforcement Learning In Python Course

1   Introduction

SAR:

- Start in state S_t

- Apply action A_t

- Get reward R_{t+1}

SAS:

- Start in state S_t

- Apply action A_t

- Move to state S_{t+1}

2   Return of the Multi-Armed Bandit

Epsilon-Greedy Explore-Exploit:

Algorithm 1 Epsilon-Greedy Explore-Exploit

1: **for** turn **do**
2:     draw a random p $\in$ [0,1]
3:     **if** p $< \epsilon$ **then**
4:         explore()
5:     **else**
6:         exploit()
7:     **end if**
8: **end for**

Efficient Mean Update:

$$\bar{X}_N = \frac{N-1}{N}\bar{X}_{N-1} + \frac{1}{N}X_N$$