

Homework Assignment # 1
Due: Thursday, January 26, 2017, 11:59 p.m.
Total marks: 100

1 Questions from Chapter 2

Question 1. [10 MARKS]

Exercise 2.2 in Sutton and Barto. Show your steps; if you do not know what “show your steps” means please ask the instructor.

Question 2. [35 MARKS]

Programming: exercise 2.3 in Sutton and Barto. This will require you to implement three things:

1. a non-stationary 10-armed bandit
2. two learning agents (agent that uses sample average, and an agent that uses a constant step-size)
3. code to run the experiment (average over runs, etc.)

Please submitted yours plot(s) and all your code.

Question 3. [5 MARKS]

Exercise 2.4 in Sutton and Barto. Start off by explaining the big spike in early learning.

2 Questions from Chapter 3

Question 4. [10 MARKS]

Exercise 3.1 in Sutton and Barto.

Question 5. [5 MARKS]

Exercise 3.5 in Sutton and Barto.

Question 6. [5 MARKS]

Exercise 3.6 in Sutton and Barto.

Question 7. [5 MARKS]

Exercise 3.7 in Sutton and Barto.

Question 8. [5 MARKS]

Exercise 3.8 in Sutton and Barto.

Question 9. [5 MARKS]

Exercise 3.9 in Sutton and Barto.

Question 10. [5 MARKS]

Exercise 3.11 in Sutton and Barto.

Question 11. [10 MARKS]

Exercise 3.16 in Sutton and Barto.

Homework policies:

Your assignment will be submitted as a single pdf document and a zip file with code, on canvas. The questions must be typed; for example, in Latex, Microsoft Word, Lyx, etc. or must be written legibly and scanned. Images may be scanned and inserted into the document if it is too complicated to draw them properly. All code (if applicable) should be turned in when you submit your assignment. Use the RL-glue framework available on the course webpage (your code will be in c/c++), and any language of choice for plotting the results (learning curves).

Policy for late submission assignments: Unless there are legitimate circumstances, late assignments will be accepted up to 5 days after the due date and graded using the following rule:

on time: your score 1
1 day late: your score 0.9
2 days late: your score 0.7
3 days late: your score 0.5
4 days late: your score 0.3
5 days late: your score 0.1

For example, this means that if you submit 3 days late and get 80 points for your answers, your total number of points will be $80 \times 0.5 = 40$ points.

All assignments are individual work, no exceptions. All the sources used for problem solution must be acknowledged, e.g. web sites, books, research papers, personal communication with people, etc. You are expected to solve problems from scratch. That means, if you are asked to code something, do not find code online and modify it. Write it from scratch yourself. Academic honesty is taken seriously; for detailed information see Indiana University Code of Student Rights, Responsibilities, and Conduct.

Good luck!