CZ4046/CSC416:
INTELLGENT AGENTS

Assignment

Semester 2, 2019/2020

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

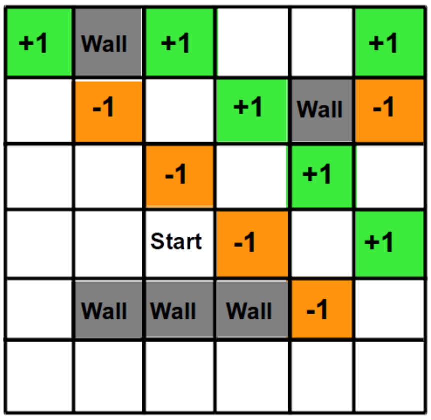
NANYANG TECHNOLOGICAL UNIVERSITY

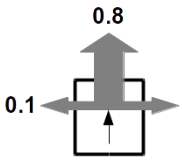
Assignment 1: Agent Decision Making

Deadline: Tuesday, March 31, 2020, 11:59:59PM

This assignment covers topics in Module 3: Agent Decision Making. Refer to Chapters 16 and 17 in the reference book "Artificial Intelligence: A Modern Approach" by S. Russell and P. Norvig. Prentice-Hall, third edition, 2010. For the ease of access, a softcopy of these two chapters can be found in the NTULearn system.

Consider the following maze environment, similar to the one in Section 17.1 of the above reference book. Write a Java program to solve the following questions.





The transition model is as follows: the intended outcome occurs with probability 0.8, and with probability 0.1 the agent moves at either right angle to the intended direction. If the move would make the agent walk into a wall, the agent stays in the same place as before. The rewards for the white squares are -0.04, for the green squares are +1, and for the brown squares are -1. Note that there are no terminal states; the agent's state sequence is infinite.

Part 1: Assuming the known transition model and reward function listed above, find the optimal policy and the utilities of all the (non-wall) states using both value iteration and policy iteration. Display the optimal policy and the utilities of all the states, and plot utility estimates as a function of the number of iterations as in Figure 17.5(a) in the above reference book (for value iteration, you should need no more than 50 iterations to get convergence). In this question, use a discount factor of 0.99. Below are some reference utility values (computed with a different discount factor) to help you get an idea if the trend of your answers is correct.

```
Reference utilities of states:
Coordinates are in (col,row) format with the top left corner being
(0,0).

(0,0): 18.538042
(0,1): 16.973925
```

```
(0,2): 15.645911
(0,3): 14.432489
(0,4): 13.401969
(0,5): 12.329421
(1,1): 14.585138
(1,2): 14.446991
(1,3): 13.511668
(1,5): 11.42872
(2,0): 16.593647
(2,1): 15.394961
(2,2): 13.399139
(2,3): 12.629236
(2,5): 10.583616
(3,0): 15.480907
(3,1): 15.827977
(3,2): 14.659492
(3,3): 12.626667
(3,5): 11.000803
(4,0): 15.37473
(4,2): 15.389912
(4,3): 14.222186
(4,4): 12.217787
(4,5): 11.860354
(5,0): 16.523903
(5,1): 14.189546
(5,2): 14.370738
(5,3): 15.058972
(5,4): 13.81445
(5,5): 12.731918
```

Part 2 (Bonus Questions): Design a more complicated maze environment of your own and re-run the algorithms designed for Part 1 on it. How does the number of states and the complexity of the environment affect convergence? How complex can you make the environment and still be able to learn the right policy?

Submission Instructions

The assignment should be submitted by uploading a single ZIP file (with name format of **lastname_firstname_matriculation.zip**) at the course site in NTULearn system. The link will be provide later. Each submission must consist of the following two parts:

- A report in PDF format. The report should briefly describe your implemented solution and fully answer all the questions above. Remember: you will not get credit for any solutions you have obtained, but not included in the report. The name of the report file should be **lastname_firstname_matriculation.pdf**

- Your Java source code. The code should be well commented, and it should be easy to see the correspondence between what's in the code and what's in the report.

Assessment Weightage

Using method of value iteration for Part 1 (**40 marks**)

- Descriptions of implemented solutions (10 marks)
- Plot of optimal policy (10 marks)
- Utilities of all states (10 marks)
- Plot of utility estimates as a function of the number of iterations (10 marks)

Using method of policy iteration for Part 1 (**40 marks**)

- Descriptions of implemented solutions (10 marks)
- Plot of optimal policy (10 marks)
- Utilities of all states (10 marks)
- Plot of utility estimates as a function of the number of iterations (10 marks)

Source code for Part 1 (**20 marks**)

Part 2 bonus questions (**20 marks**)

- Answers of the questions in the report (15 marks)
- Source code (5 marks)

Deadline

The due date of the assignment is: **Tuesday, March 31, 2020, 11:59:59PM**

Note that **FIVE (5)** marks will be deducted for the delay submission of each calendar day.