

CPSC422 Spring 2007

Assignment 1

Due: 2:00pm, Thursday 25 January 2007.

The aim of this assignment is to learn about MDPs. You are to write code for the simple game at:

`http://www.cs.ubc.ca/spider/poole/demos/rl/sGame.html`

You are free to use any of the code for that applet. You can also refer to and use any of the code for the value iteration applet at:

`http://www.cs.ubc.ca/spider/poole/demos/mdp/vi.html`

Please read and post to the bulletin board in the course WebCT site for hints on how to modify the controllers.

You can either do this assignment by yourself or with someone else (i.e., in pairs).

Question 1

Write a controller for the simple game (modifying only SameController.java) that uses:

- (a) value iteration
- (b) asynchronous value iteration
- (c) modified policy iteration

If you are working by yourself, then you only need to do one of these. If you are working in a pair, you need to do all three.

Question 2

In this question, you are to think about the effect of changing the discount factor on your implementation.

- (a) How does changing the discount factor affect the rate of convergence?

- (b) How does changing the discount factor affect the policy found?
- (c) What is an appropriate discount factor for this game?

Write full sentences, and justify your answers with specific examples from running your code.

Question 3

[Note that this question is worth marks, so don't forget to do it.]

- (a) For each question and each part in this assignment, say how long you spent on it. Was this reasonable? What did you learn?
- (b) If there was more than one person in your group, say what each person did.
- (c) Tell us every person you discussed this with (including the TA and fellow students), and every external source (e.g., web site, research paper, book) you referenced to do this assignment.