

CS 7641 Machine Learning Spring 2017, Assignment 4

Markov Decision Processes

Numbers

The assignment is worth 15% of your final grade.

Why?

In some sense, we have spent the semester thinking about machine learning techniques for various forms of function approximation. It's now time to think about using what we've learned in order to allow an agent of some kind to act in the world more directly. This assignment asks you to consider the application of some of the techniques we've learned from reinforcement learning to making decisions.

The same ground rules apply for programming languages as with the previous assignments.

Read everything below carefully!

The Problems Given to You

You are being asked to explore Markov Decision Processes (MDPs):

1. Come up with two interesting MDPs. Explain why they are interesting. They don't need to be overly complicated or directly grounded in a real situation, but it will be worthwhile if your MDPs are inspired by some process you are interested in or are familiar with. It's ok to keep it somewhat simple. For the purposes of this assignment, though, make sure one has a "small" number of states, and the other has a "large" number of states.
2. Solve each MDP using value iteration as well as policy iteration. How many iterations does it take to converge? Which one converges faster? Why? Do they converge to the same answer? How did the number of states affect things, if at all?
3. Now pick your favorite reinforcement learning algorithm and use it to solve the two MDPs. How does it perform, especially in comparison to the cases above where you knew the model, rewards, and so on? What exploration strategies did you choose? Did some work better than others?

Coding Resources

[Brown-UMBC Reinforcement Learning and Planning \(BURLAP\) java code library](#)

What to Turn In

You must submit a tar or zip file named *yourgtaccount*.{zip,tar,tar.gz} that contains a single folder or directory named *yourgtaccount* that in turn contains:

- a file named *README.txt* that contains instructions for running your code
- your code
- a file named *yourgtaccount-analysis.pdf* that contains your writeup.
- any supporting files you need

The file *yourgtaccount-analysis.pdf* should contain:

- A description of your MDPs and why they are interesting.

- A discussion of your experiments.
- **Limit your analysis write-up to 10 pages.**

Grading Criteria

As always you are being graded on your analysis more than anything else.

Additional instructor's comments about your submission:

Problem Selection

- Good job in choosing the problems of different sizes. This helps in comparing the results of different algorithms. For example, the time taken to converge increases exponentially as the size of state space increases as you pointed out later in the assignment.

PI + VI

- Good job analyzing the results for VI and PI
- It is good to see that both the algorithms converged to the same answer
- Good job analyzing the change in size of state space wrt the time taken for convergence.

Q-learning

- Good job describing and implementing Q-learning
- Good job comparing the performance of Q-learning with PI and VI and describing how they are different.
- Good discussion on the exploration- exploitation trade-off. The values of the transition probabilities are needed for VI and PI. However, Q-learning learns the transition model by alternatively randomly exploring the learning space and finding the optimal policy.