Reinforcement Learning

**Introduction**

Come up with 2 interesting MDPs.

- Explain why they are interesting

- Make sure 1 has small amount of states and the other has a large amount of states (200 is not large).

- No more than 1 of the MDPs should be a grid world problem

Solve each MDP using value iteration (VI)

- How many iterations does it take to converge?

Solve each MDP using policy iteration (PI)

- How many iterations does it take to converge?

Analysis

- How is convergence defined?

- Did they converge to the same answer?

- Which method converged faster?

- Why did one method converge faster?

- How did the number of states affect things?

Pick your favorite reinforcement learning algorithm and use it to solve the two MDPs.

- How does it perform?

- How does it compare to the two cases above which knew the model, rewards and so on?

- What exploration strategies did you choose?

- Did some exploration strategies work better than others?