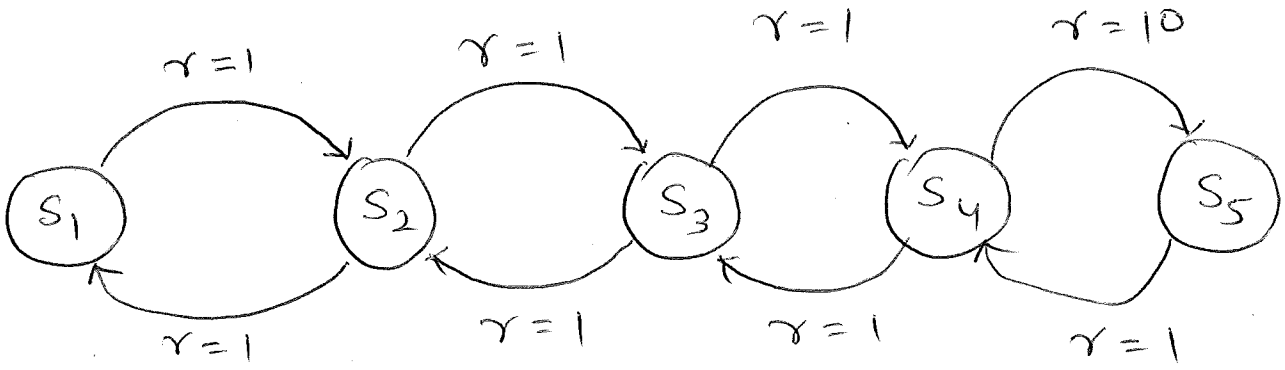
Sample Questions:

1. Consider the following MDP



5 states

Two actions {LEFT, RIGHT}

Rewards are on the edges

Assume a discount factor $\gamma = 0.7$

(a) What is the optimal policy of this MDP ?

(b) What is the optimal value function of state S_5, i.e, $V^*(S_5)$?

(c) Consider running Q-learning.

All $Q(s,a)$ values are initialized to 0

Learning rate $\alpha = 0.4$

Greedy exploration policy (i.e, always choose action with maximum Q value).

Break ties by choosing LEFT

If the robot starts in state S_2, what are the first 15 state-action pairs?

2. Value Iteration and Policy Iteration algorithms can be seen as two equivalent views of solving the policy optimization problem.

What is the corresponding equivalence??

3. The Sample Complexity of passive learning algorithms is exponentially better than Active learning algorithms.

(True / False)

4. How does Bagging affect the bias and variance of base learner?

5. ML algorithms cannot be compared based on time-complexity analysis.

(True/False)

6. AdaBoost will always give ZEROS training error if run for a large number of iterations

(True/False)

7. Why is it possible to exploit dynamic programming for computing the optimal value function of an MDP?