Assignment 2 - MDP

Laksh Bhatia **604561**

Solution 1

Encoding a weighted directed acyclic graph(DAG) can be done as follows:
States should be nodes of the weighted DAG
LivingRewards should be the weights of the DAG for every edge
Terminal state has the highest reward
Noise should be 0 i.e. all transitions are deterministic.
Actions are the possible transitions from one node to the other
Discount factor should be high i.e > 0.6

Solution 2

Part A
Optimal policy

1	1	1	1	4
2	0	2	1	4
2	0	0	0	0
2	3	1	1	2
0	0	0	0	0

Here 0 represents that we it is either the terminal state or an obstacle i.e. wall

I don't think my solution gives the optimal value for all possible combinations of livingrewards,noise and discount factor.

Part B

1. Prefer close Exit, Risking Cliff — livingReward = 0, noise = 0 discount factor between 0.1 and 0.3 i.e. 0.1 , 0.2 , 0.3. Most optimal value is discount factor of 0.1.

2. Prefer close Exit, Avoiding Cliff — livingReward = 0, noise = 0.1 discount factor between 0.1 and 0.3 i.e. 0.1 , 0.2 , 0.3. Most optimal value is discount factor of 0.1.

3. Prefer distant Exit, Risking Cliff — livingReward = 0, noise = 0 discount factor between 0.4 and 0.99 i.e. 0.4 , 0.5 , 0.6 , 0.7 , 0.8 , 0.9 , 0.99. Most optimal value is discount factor of 0.99.

4. Prefer distant Exit, Avoiding Cliff — livingReward = 0, noise = 0.5 discount factor between 0.6 and 0.99 i.e. 0.6 , 0.7 , 0.8 , 0.9 , 0.99. Most optimal value is discount factor of 0.99.