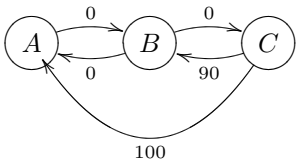
Homework-6

Question 1



1

Guess two good policies for the above. Draw them as a copy of the above diagram with arrows corresponding to the policy actions.

2

Guess the optimal policy for a discount rate of 0.0000001.

2.1

Compute V^* from your guess, then Q from V^* and then π^* from Q. Was your guess correct?

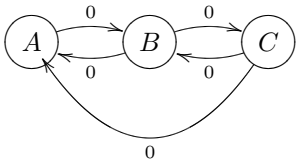
3

Guess the optimal policy for a discount rate of 0.9999999.

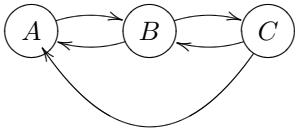
3.1

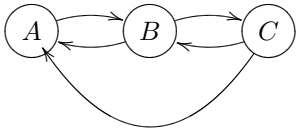
Compute V^* from your guess, then Q from V^* and then π^* from Q. Was your guess correct?

4

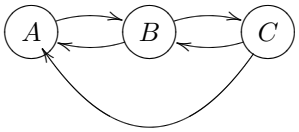
Run the Q learning algorithm for this problem with a discount rate of 1/2. Start with $\hat{Q}(s, a) = 0$, which can be described by the following diagram: 

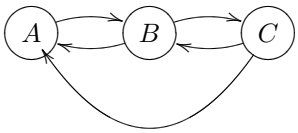
Produce the value of Q after the following pairs are considered one after the other:

1. (B,left). 

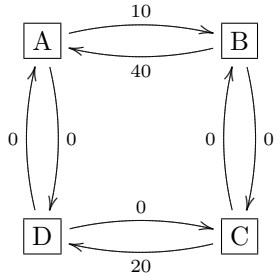
2. (C,left). 

.

3. (A,right). 

4. (C,down). 

Question 2

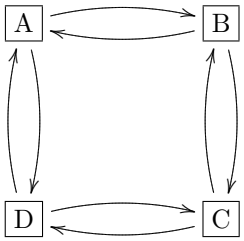


The following questions are related to the reinforcement online Q-learning algorithm applied to the above directed graph. The weights indicate rewards. Whenever needed use a discount rate of 0.5.

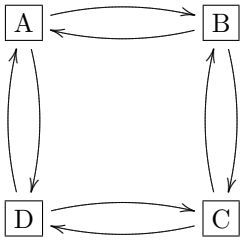
Part A

Run 4 iterations of the Q learning algorithm on this problem with a discount rate of 0.5. Start with $\hat{Q}(s,a) = 0$ for all states and actions, and compute the value of $\hat{Q}$ after the given actions are considered one after the other.

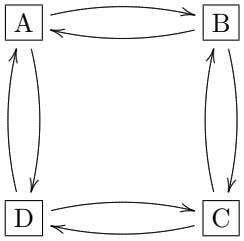
1. (A,right).



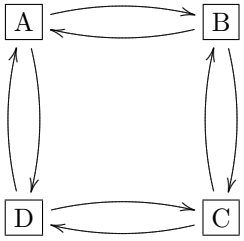
2. (B,left), after 1.



3. (A,right), after 2.



4. (B,left), after 3.



Part B

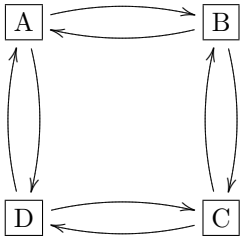
Starting with $\hat{Q}(s,a) = 0$ for all states and actions, the online Q-learning algorithm is applied repeatedly to the following actions

$$Q(B, \text{left}),\ Q(A, \text{right})$$

infinitely many times. (There are no other actions besides those.) Do you expect the algorithm to converge to fixed values of $\hat{Q}$ or keep changing?

Answer: Select 1 or 2.

1. The algorithm converges to $Q =$



(fill in the values of Q after convergence).

2. The algorithm will not converge. The values of $\hat{Q}$ will keep changing, and will not be converging to fixed values.

Question 3

Consider the problem of clustering the $m = 4$ points below into $k = 2$ clusters.

1	(0,0)
2	(4,0)
3	(5,1)
4	(6,0)

Part 1

Suppose the points 3 and 4 are selected using the Lloyd technique for computing initial means.

1. What is the initial clustering computed by the k-means algorithm?

Answer:

2. What clustering is computed after the first iteration of the k-means algorithm?

Answer:

3. What clustering is computed after the second iteration of the k-means algorithm?

Answer:

4. What clustering is computed after the third iteration of the k-means algorithm?

Answer:

5. What clustering is computed after the fourth iteration of the k-means algorithm?

Answer:

6. What clustering is computed after the fifth iteration of the k-means algorithm?

Answer:

Question 4

Consider the problem of clustering the $m = 6$ points below into $k = 2$ clusters.

1	(1,1)
2	(1,2)
3	(1,0)
4	(4,1)
5	(4,2)
6	(4,0)

Part 1

Here we will be using k-means.

1. What clstering is obtained by k-means if the initial points selected by the Lloyd technique are points 3 and 5? What is the corresponding quantization error?

Answer:

$$C(1) = \ , \ C(2) = \ , \ C(3) = \ , \ C(4) = \ , \ C(5) = \ , \ C(6) = \ , \quad E =$$

2. What clstering is obtained by k-means if the initial points selected by the Lloyd technique are points 4 and 6? What is the corresponding quantization error?

Answer:

$$C(1) = \ , \ C(2) = \ , \ C(3) = \ , \ C(4) = \ , \ C(5) = \ , \ C(6) = \ , \quad E =$$

Part 2

Here we will be using k-means++. Suppose the first point selected (at random) by the algorithm is Point 1.

1. Complete the following table for the squared distances of all points from Point 1:

1	$d =$
2	$d =$
3	$d =$
4	$d =$
5	$d =$
6	$d =$

2. Compute the probability of selecting each one of the points as the second point.

1	$p =$
2	$p =$
3	$p =$
4	$p =$
5	$p =$
6	$p =$

3. Suppose the algorithm selects the point with the largest probability. What is the clstering obtained by k-means++? What is the corresponding quantization error?

Answer:

$$C(1) = \ , \ C(2) = \ , \ C(3) = \ , \ C(4) = \ , \ C(5) = \ , \ C(6) = \ , \qquad E =$$

4. Will your answer change if the algorithm selects the point with second largest probability?

Answer:

5. Will your answer change if the algorithm selects the point with third largest probability?

Answer:

6. What is the probability that the algorithm selects one of the three points with the largest probability?

Answer: