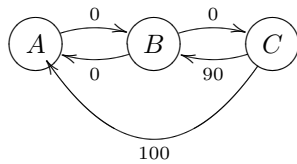


Homework-7

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

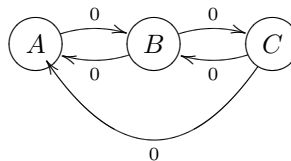
Guess the optimal policy for a discount rate of 0.9999999.

3.1

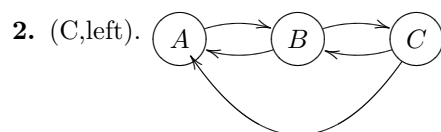
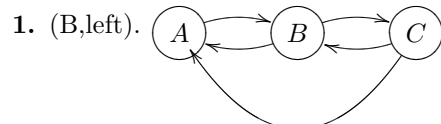
Compute V^* from your guess, then Q from V^* and then π^* from Q . Was your guess correct?

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:

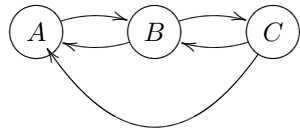


Produce the value of Q after the following pairs are considered one after the other:



.

3. (A,right).



4. (C,down).

