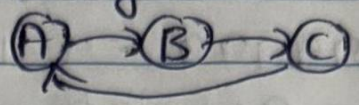
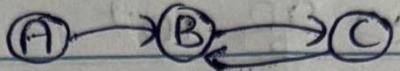
Homework - 6

Q1 1. Policy 1 

Policy 2 

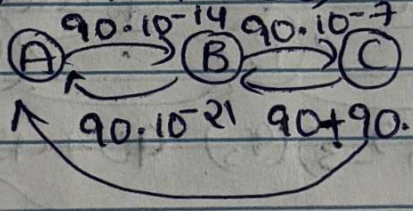
2. Policy 2

2.1 $r = 10^{-7}$

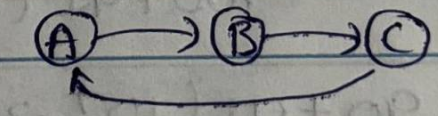
Policy 1 : 

$$V^*(C) = 90 + r \cdot 0 + r^2 \cdot 90 + r^3 \cdot 0 + \cdots$$
$$= 90[1 + r^2 + r^4 + \cdots]$$
$$= \frac{90}{1 - r^2}$$
$$\simeq 90.$$

$$V^*(B) = r \cdot V^*(C) = 90 \times 10^{-7}$$
$$V^*(A) = r \cdot V^*(B) = 90 \times 10^{-14}$$

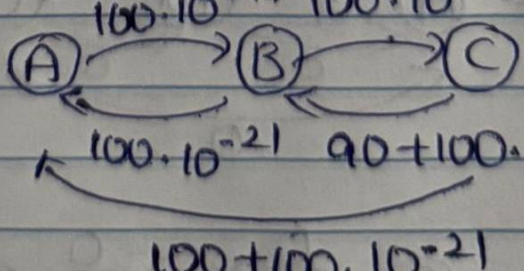
Q : 

$$10.0 + 90.10^{-21}$$

Policy 2: 

$$V(C) = 100 + r^3 100 + r^6 100 + \cdots$$
$$= \frac{100}{1 - r^3} = \frac{100}{1 - 10^{-21}} \simeq 100$$

$$V(B) = 100 \cdot 10^{-17}$$
$$V(A) = 100 \cdot 10^{-14}$$

$\therefore$ Q : 

$$100 + 100 \cdot 10^{-21}$$

$\therefore$ Policy 2 is the optimal policy

3. Policy 1

3.1 $\gamma = 0.9999999 = 1 - 10^{-7}$

$V^*(C) = 90 + 90\gamma^2 + 90\gamma^4 + \cdots$

$= \frac{90}{1-\gamma^2} = \frac{90}{1 - 1 - 10^{-14} + 2 \cdot 10^{-7}}$

$\simeq \frac{90}{2} \cdot 10^7$

$\simeq 45 \cdot 10^7$

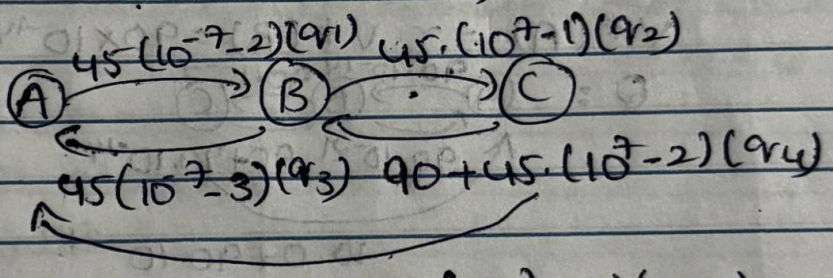
$V^*(B) = \gamma 45 \cdot 10^{-7} = (1 - 10^{-7}) 45 \cdot 10^7$

$= 45(10^7 - 1)$

$V^*(A) = \gamma \cdot 45(10^7 - 1)$

$= 45(10^7 - 1 - 1 + 10^{-7})$

$= 45(10^7 - 2)$



$q_4 = 90 + 45(10^7 - 2) = 45 \cdot 10^7$

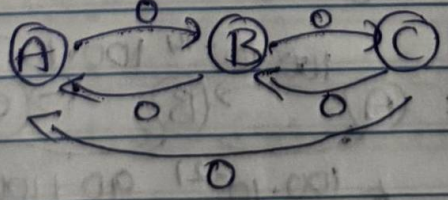
$q_5 = 100 + 45(10^7 - 3) = 100 + 45 \times 10^7 - 150 - 45 \times 10^7 - 50$

$\therefore q_4 > q_5$

$\therefore$ Policy 1 is better.

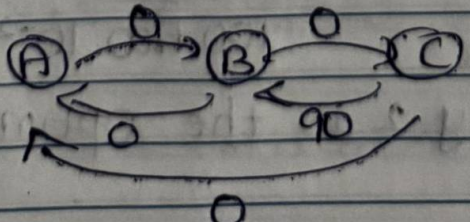
4.1 (B, left)

$\hat{Q}(S, a) =$



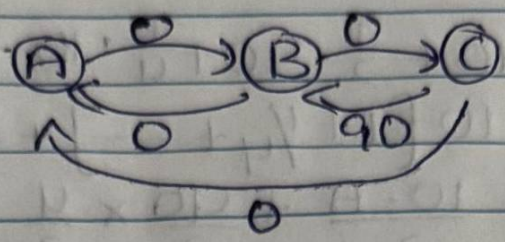
2. (C, left)

$\hat{Q}(S, a) =$



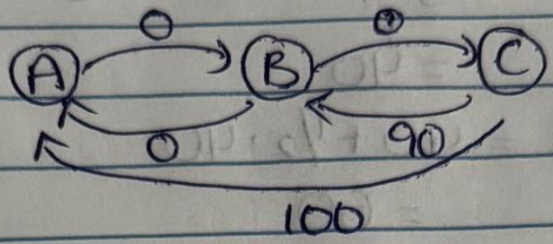
3. (A, right)

$\hat{Q}(S, a)$ =



4. (C, down)

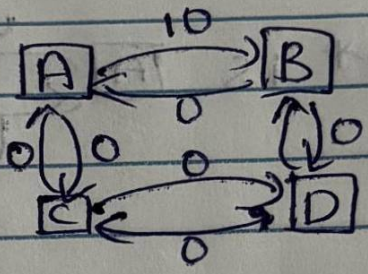
$\hat{Q}(S, a)$ =



Q2. PART-A

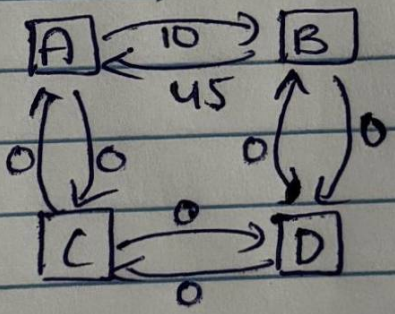
1. (A, right)

$\hat{Q}(S, a)$ =



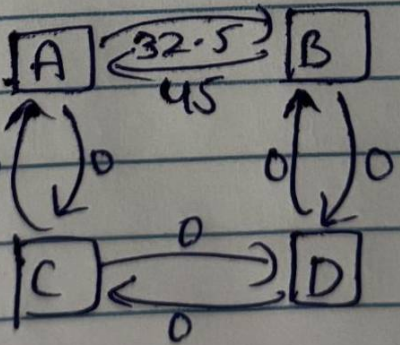
2. (B, left)

$\hat{Q}(S, a)$ =



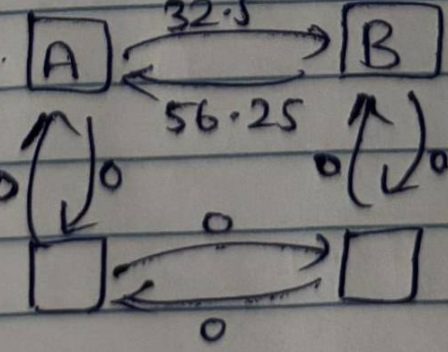
3. (A, right)

$\hat{Q}(S, a)$ =



4. (B, left)

$\hat{Q}(S, a)$ =



PART-B

1. $V_A^* = 10 + \frac{1}{2} \cdot 40 + \frac{1}{4} \cdot 10 + \frac{1}{8} \cdot 40 + \cdots$

$= 10\left[1 + 1/4 + \cdots\right] + \frac{40}{2}\left[1 + 1/4 + \cdots\right]$

$= \frac{10 \cdot 4}{3} + \frac{40}{2} \times \frac{4}{3}$

$= 40.$

$V_B^* = 40 + 1/2 \cdot 40$

$= 60$

∴ The optimal policy

∴ Q^* is