NO. _____

DATE _____

Christian Joshua Alberto

| | School | Home |
|---|---|---|
| Sunny | +5 | -5 |
| Cloudy | +3 | +1 |

$R_{school} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$ $R_{Home} = \begin{bmatrix} -5 \\ +1 \end{bmatrix}$

State Transition Matrix

| | S | C |
|---|---|---|
| S | 0.9 | 0.1 |
| C | 0.3 | 0.7 |

$P_{stay\ school} = \begin{bmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{bmatrix}$

$P_{school} = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix}$

1. Compute State-wise average reward under the policy

1.1 Find $r_\pi$ for sunny

$r_\pi = (0.5)(5) + 0.5(-5)$

$= 2.5 + (-2.5)$

$= 0$

1.2 Find $r_\pi$ for cloudy

$r_\pi = 0.5(3) + 0.5(1)$

$= 1.5 + 0.5$

$= 2$

1.3 Find $r_\pi$ matrix:

$r_\pi = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$

2. Compute the policy transition matrix

Row 1 (sunny):

$P_\pi(1,1) = 0.5(0.8) + 0.5(0.9) = 0.85$

$P_\pi(1,2) = 0.5(0.2) + 0.5(0.1) = 0.15$

Row 2 (cloudy):

$P_\pi(2,1) = 0.5(0.4) + 0.5(0.3) = 0.35$

$P_\pi(2,2) = 0.5(0.6) + 0.5(0.7) = 0.65$

$P_\pi = \begin{bmatrix} 0.85 & 0.15 \\ 0.35 & 0.65 \end{bmatrix}$

2 Write the Bellman expectation equations

$V_\pi$ (sunny):

$V_2 = 0 + 0.9(0.85v_1 + 0.15v_2)$

$V_1 = 0.765v_1 + 0.135v_2$

$V_1 = 0.765v_1 - 0.135v_2$

$0.235v_1 - 0.135v_2 = 0$ [1]

$V_\pi$ (cloudy):

$V_2 = 2 + 0.9(0.35v_1 + 0.65v_2)$

$= 2 + 0.315v_1 + 0.585v_2$

$V_2 - 0.315v_2 - 0.585v_2$

$-0.315v_1 + 0.415v_2 = 2$ [2]

4. Solve for $V_\pi$ (cloudy)

$0.235v_1 - 0.135v_2 = 0$ (1)

$0.235v_1 = 135v_2$

$v_1 = \frac{0.135v_2}{0.235}$

$-0.315v_1 + 0.415v_2 = 2$ (2)

$-0.315\left(\frac{0.135v_2}{0.235}\right) + 0.415v_2 = 2$

$0.1808v_2 + 0.415v_2 = 2$

$0.734v_2 = 2$

$\boxed{v_2 \approx 8.55}$

$v_1 = \frac{0.135(1.55)}{0.235}$

$\boxed{v_1 \approx 4.91}$

$V_\pi$ (sunny) = 21.91 $V_\pi$ (cloudy) = 8.55

5. Write the Bellman optimality equations

Find Sunny ($v_1$) using Go to school

$V_*$ (sunny) $= 5 + 0.9(0.8v_1 + 0.2v_2)$

$\boxed{= 0.28v_1 - 0.18v_2 = 5}$

Find cloudy $v_2$ using Go to school

$V_*$ (cloudy) $= 3 + 0.9(0.4v_1 + 0.6v_2)$

$\boxed{-0.36v_1 + 0.46v_2 = 3}$

6. Solve for $V_*$ (cloudy) + $V_*$ (Sunny)

$-0.36v_1 + 0.46v_2 = 3$

$0.46v_2 = 3 + 0.36v_1$

$\boxed{v_2 = \frac{3 + 0.36v_1}{0.46}}$

$= 0.2v_1 - 0.1\left(\frac{3 + 0.36v_1}{0.46}\right) = 5$

$\boxed{V_*$ (sunny) $[v_1] = 44.375}$

$v_2 = \frac{3 + 0.36(44.375)}{0.46}$

$\boxed{V_*$ (cloudy) $[v_2] = 41.25}$

7. Solve for q(a)

$q(1, \text{school}) = 5 + 0.9(0.8(44.375) + 0.2(41.25))$

$\boxed{= 44.375}$

$q(1, \text{home}) = -5 + 0.9(0.9(44.375) + 0.1(41.25))$

$\boxed{= 34.65625}$

$q(2, \text{school}) = 3 + 0.9(0.4(44.375) + 0.6(41.25))$

$\boxed{= 41.25}$

$q(2, \text{home}) = 1 + 0.9(0.3(44.375) + 0.7(41.25))$

$\boxed{= 38.96875}$