NO.:
DATE:

Baldeo, John Vernon B.

COM221

| | Go to School | Stay at 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 $\pi$

(1) Find $r_\pi$ for sunny:

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

$= 2.5 + (-2.5)$

$= 0$

(2) Find $r_\pi$ for cloudy

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

$= 1.5 + 0.5$

$= 2$

(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):

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

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

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

③ Write the Bellman expectation equations

$V_\pi$ (sunny):

$V_1 = 0 + 0.9(0.85V_1 + 0.15V_2)$

$V_1 = 0.765V_1 + 0.135V_2$

$V_1 - 0.765V_1 - 0.135V_2$

$0.235V_1 - 0.135V_2 = 0$ [1]

③ Write the Bellman expectation equations

$V_\pi$ (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 $V_2$ (cloudy):

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

$0.235 v_1 = 0.135 v_2$

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

Substitute to [2]

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

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

$-0.1808 v_2 + 0.415 v_2 = 2$

$0.234 v_2 = 2$

$v_2 \approx 8.55$

$v_1 = \frac{0.135 (8.55)}{0.235}$

$v_1 \approx 4.91$

$V_2$ (sunny) $= 4.91$ $V_2$ (cloudy) $= 8.55$

⑤ Write the Bellman optimality equations:

Find Sunny ($V_1$) using [Go to School]:

$V_1$ (sunny) $= 5 + 0.9(0.8 v_1 + 0.2 v_2)$

$\Rightarrow 0.28 v_1 - 0.18 v_2 = 5$

Find Cloudy ($V_2$) using [Go to School]

$V_1$ cloudy $= 3 + 0.9(0.4 v_1 + 0.6 v_2)$

$-0.36 v_1 + 0.46 v_2 = 3$

⑥ Solve for $V_1$ (cloudy) : $V_1$ (sunny)

$-0.36 v_1 + 0.46 v_2 = 3$

$0.46 v_2 = 3 + 0.36 v_1$

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

$\Rightarrow 0.28 v_1 - 0.18 \left(\frac{3 + 0.36 v_1}{0.46}\right) = 5$

$V_1$ (sunny) [$v_1$] $= 44.375$

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

$V_1$ (cloudy) [$v_2$] $= 41.25$

⑦ Solve for $q_*$ :

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

$= 44.375$

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

$= 34.65625$

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

$= 41.25$

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

$= 38.90875$