NO.:
DATE:

Rewards

	school	home
sunny	+5	-5
cloudy	+3	+1

Rainnard P. Montayre

Discount = $\gamma = 0.9$

$R_{school} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$ $P_{school} = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix}$

$R_{home} = \begin{bmatrix} -5 \\ 1 \end{bmatrix}$ $P_{home} = \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.7 \end{bmatrix}$

Step 1

1. r_{π} sunny $= 0.5 \cdot 5 + 0.5 \cdot (-5) = 0$
2. r_{π} cloudy $= 0.5 \cdot 3 + 0.5 \cdot 1 = 2$
3. r_{π} matrix $= \begin{bmatrix} 0 \\ 2 \end{bmatrix}$

Step 2 Compute the policy transformation matrix

Row 1 (sunny)

- $P_{\pi}(1,1) = 0.5 \cdot 0.8 + 0.5 \cdot 0.9 = 0.85$
- $P_{\pi}(1,2) = 0.5 \cdot 0.2 + 0.5 \cdot 0.1 = 0.15$

Row 2 (cloudy)

- $P_{\pi}(2,1) = 0.5 \cdot 0.4 + 0.5 \cdot 0.5 = 0.45$
- $P_{\pi}(2,2) = 0.5 \cdot 0.6 + 0.5 \cdot 0.7 = 0.65$

$P_{\pi} = \begin{bmatrix} 0.85 & 0.15 \\ 0.45 & 0.65 \end{bmatrix}$

Step 3 Write the Bellman expectation equation

V_{π} sunny $= V_1 = 0 + 0.9(0.85 V_1 + 0.15 V_2)$

$V_1 = 0 + 0.765 V_1 + 0.135 V_2$

V_{π} sunny $= 0.235 V_1 - 0.135 V_2 = 0$

V_{π} cloud $= V_2 = 2 + 0.9(0.45 V_1 + 0.65 V_2)$

$V_2 = 2 + 0.405 V_1 + 0.585 V_2$

V_{π} cloud $= -0.405 V_1 + 0.415 V_2 = 2$

NO.:
DATE:

Step 4 Write the Bellman expectation equation

V_π cloudy

$V_1 = 0.5744680\, V_2$

V_π cloudy $= 0.4\, v_1 + 0.4\, v_2 = 2$

$= -0.2326595745 + 0.4\, v_2 = 2$

V_π cloudy $= \boxed{V_2 = 5.379902589 2}$

V_π sunny $= 0.5744680\,(5.379902589 2) = \boxed{3.090581880 6}$

Step 5 Write the Bellman optimality equations

- For Sunny (V_1) using school

$V_*(\text{sunny}) = 5 + 0.9(0.8 v_1 + 0.2 v_2) = 5 + 0.72 v_1 + 0.18 v_2$

- For Cloudy (V_2) using school

$V_*(\text{cloudy}) = 3 + 0.9(0.4 v_1 + 0.6 v_2) = 3 + 0.36 v_1 + 0.54 v_2$

- Sunny

$v_1 = 5 + 0.72 v_1 + 0.18 v_2 \Rightarrow 0.28 v_1 - 0.18 v_2 = 5$

- Cloudy

$v_2 = 3 + 0.36 v_1 + 0.54 v_2 \Rightarrow 0.46 v_2 - 0.36 v_1 = 3$

Step 6.

sunny:

$0.28 v_1 = 5 + 0.18 v_2 \Rightarrow v_1 = \frac{5 + 0.18 v_2}{0.28}$

Using it on cloudy

$= 0.46 v_2 - 0.36\left(\frac{5 + 0.18 v_2}{0.28}\right) = 3$

$= 0.46 v_2 - 0.231428571 v_2 = 3$

$v_2 = 41.25 = V_*$ cloudy

Using it on sunny:

$= v_1 = \frac{5 + 0.18(41.25)}{0.28}$

$v_1 = 44.375 = V_*$ sunny

STEP 7:

$q(1, \text{school}) = 5 + 0.9(0.8 v_1 + 0.2 v_2) = 44.375$

$q(1, \text{stay}) = -5 + 0.9(0.9 v_1 + 0.1 v_2) = 8.216$

$q(2, \text{school}) = 3 + 0.9(0.4 v_1 + 0.6 v_2) = 41.25$

$q(2, \text{stay}) = 1 + 0.9(0.5 v_1 + 0.7 v_2) = -7.2981$

NO.:
DATE:

Step 5.2 Write the Bellman Optimality equations

* For Sunny (V_1) using stayhome

$V_*(\text{sunny}) = -5 + 0.9(0.9V_1 + 0.1V_2) = 0.19V_1 - 0.9V_2 = -5$

* For Cloudy (V_2) using stayhome

$V_*(\text{cloudy}) = 1 + 0.9(0.5V_1 + 0.7V_2) = 0.37V_2 - 0.45V_1 = 1$

Step 6.2

sunny:

$V_1 = \frac{-5 - 0.9V_2}{0.19}$

using it on cloudy

$= 0.37V_2 - (0.45(\frac{-5 - 0.9V_2}{0.19})) = 1$

$= 0.37V_2 - (\frac{-2.25 - 0.405V_2}{0.19}) = 1$

$= -2.25 - 0.405V_2 = 0.19 - 0.0703V_2$

$= -0.3347V_2 = 2.44$

$= V_2 = \boxed{-7.290110546 = V_* \text{cloudy}}$

Using it on sunny

$V_1 = \frac{-5 - 0.9(-7.290110546)}{0.19}$

$= \boxed{8.216313[illegible]11 = V_* \text{sunny}}$