Fortune Raphael C. Bermudez

Exercise 3
Reinforcement Learning

Step 1

1.) Find the r_π for sunny = $\boxed{r_\pi = 0}$

$r_\pi = 0.5 \times (5) + 0.5 \times (-5) = 0$

2.) Find the r_π for cloudy = $\boxed{r_\pi = 2}$

$r_\pi = 0.5 \times (3) + 0.5 \times (1) = 2$

3.) Find the r_π matrix = $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$

Step 2

Row 1 (Sunny)

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

Row 2 (Cloudy)

- $P_\pi (2,1) = 0.5 \times 0.4 + 0.5 \times 0.3 = 0.35$
- $P_\pi (2,2) = 0.5 \times 0.6 + 0.5 \times 0.7 = 0.65$

4.) Find the P_π matrix = $\begin{bmatrix} 0.85 & 0.15 \\ 0.35 & 0.65 \end{bmatrix}$

Step 3

5.) Find v_1 =

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

$v_1 = 0 + 0.765 v_1 + 0.135 v_2$

$v_1 - 0.765 v_1 - 0.135 v_2 = 0$

$0.235 v_1 - 0.135 v_2 = 0$

6.) Find v_2 =

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

$v_2 = 2 + 0.315 v_1 + 0.585 v_2$

$v_2 - 0.315 v_1 - 0.585 v_2 = 2$

$-0.315 v_1 + 0.415 v_2 = 2$

Step 4

7.) V_{π} (cloudy) = ?

$0.235 v_1 - 0.135 v_2 = 0$

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

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

$-0.315 v_1 + 0.415 v_2 = 2$

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

$-0.181 v_2 + 0.415 v_2 = 2$

$(0.415 - 0.181) v_2 = 2$

$\frac{0.234 v_2}{0.234} = \frac{2}{0.234}$

$\boxed{v_2 = 8.547} \rightarrow \boxed{V_{\pi} \text{ (cloudy)} = 8.547}$

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

$v_1 = \frac{0.135 (8.547)}{0.235} = \frac{1.154}{0.235} = 4.911$

$\boxed{V_{\pi} \text{ (sunny)} = 4.911}$

Step 5

Find Sunny (v_1) using Go to School

9.) $v_*(\text{sunny}) = ?$

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

$$v_1 = 5 + 0.72v_1 + 0.18v_2$$

$$v_1 - 0.72v_1 - 0.18v_2 = 5$$

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

Find Cloudy (v_2) using Go to School

10.) $v_*(\text{cloudy}) = ?$

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

$$v_2 = 3 + 0.36v_1 + 0.54v_2$$

$$v_2 - 0.36v_1 - 0.54v_2 = 3$$

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

Step 6

From the sunny equation:

$$0.28v_1 - 0.18v_2 = 5$$

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

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

Using the equation from cloudy:

$$-0.36v_1 + 0.46v_2 = 3$$

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

$$-0.36 \times \left(\frac{5}{0.28}\right) = -6.429$$

$$-0.36 \times \left(\frac{0.18v_2}{0.28}\right) = -0.231v_2$$

$$-6.429 - 0.231v_2 + 0.46v_2 = 3$$

$$-6.429 + (0.46 - 0.231)v_2 = 3$$

$$-6.429 + 0.229v_2 = 3$$

$$0.229v_2 = 3 + 6.429$$

$$\frac{0.229v_2}{0.229} = \frac{9.429}{0.229}$$

12.) $\boxed{v_2 = 41.175}$ $\boxed{v_*(\text{cloudy}) = 41.175}$

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

$$v_1 = \frac{5 + 0.18(41.175)}{0.28} = \frac{5 + 7.4115}{0.28} = \frac{12.4115}{0.28}$$

$\boxed{v_1 = 44.327}$ 11.) $\boxed{v_*(\text{sunny}) = 44.327}$

Step 7

13.) $q(1, \text{School}) = 5 + 0.9(0.8v_1 + 0.2v_2) = 44.327$

14.) $q(1, \text{Home}) = -5 + 0.9(0.9v_1 + 0.1v_2) = -38.265$

15.) $q(2, \text{School}) = 3 + 0.9(0.4v_1 + 0.6v_2) = 41.175$

16.) $q(2, \text{Home}) = 1 + 0.9(0.3v_1 + 0.7v_2) = -25.227$

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

$v_1 = -5 + 0.81v_1 + 0.09v_2$

$v_1 - 0.81v_1 - 0.09v_2 = -5$

$\boxed{0.19v_1 - 0.09v_2 = -5}$

$q(2, \text{Home}) = 1 + 0.9(0.3v_1 + 0.7v_2)$

$v_2 = 1 + 0.27v_1 + 0.63v_2$

$v_2 - 0.27v_1 - 0.63v_2 = 1$

$\boxed{-0.27v_1 + 0.37v_2 = 1}$

$0.19v_1 - 0.09v_2 = -5$

$\frac{0.19v_1}{0.19} = \frac{-5 + 0.09v_2}{0.19}$

$v_1 = \frac{-5 + 0.09v_2}{0.19}$

$$-0.27v_1 + 0.37v_2 = 1$$

$$-0.27\left(\frac{-5+0.09v_2}{0.19}\right) + 0.37v_2 = 1$$

$$-0.27\left(\frac{-5}{0.19}\right) = 7.105$$

$$-0.27\left(\frac{0.09v_2}{0.19}\right) = -0.128$$

$$7.105 - 0.128v_2 + 0.37v_2 = 1$$

$$7.105 + (0.37 - 0.128)v_2 = 1$$

$$7.105 + 0.242v_2 = 1$$

$$0.242v_2 = 1 - 7.105$$

$$\frac{0.242v_2}{0.242} = \frac{-6.105}{0.242}$$

$$\boxed{v_2 = -25.227}$$

$$V_1 = \frac{-5 + 0.09v_2}{0.19}$$

$$V_1 = \frac{-5 + 0.09(-25.227)}{0.19} = \frac{-5 - 2.27043}{0.19} = \frac{-7.27043}{0.19}$$

$$\boxed{V_1 = -38.265}$$