Exercise # 3

Jayce Ann Colocado CCRNFLRL Com 221

Step 1: Compute state-wise average reward under the policy π

1.) Find r_π for sunny = ?

$r_\pi = 0.5 \times (5) + 0.5 \times (-5) = 2.5 + -2.5 = 0$

2.) Find r_π for cloudy = ?

$r_\pi = 0.5 \times (3) + 0.5 \times (1) = 1.5 + 0.5 = 2$

3.) Find r_π matrix = ?

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

Step 2: Compute the policy transition matrix

Row 1 (sunny):

* $P_\pi(1,1) = 0.5 \times 0.8 + 0.5 \times 0.9 = 0.4 + 0.45 = 0.85$
* $P_\pi(1,2) = 0.5 \times 0.2 + 0.5 \times 0.1 = 0.1 + 0.05 = 0.15$

Row 2 (cloudy)

* $P_\pi(2,1) = 0.5 \times 0.4 + 0.5 \times 0.3 = 0.2 + 0.15 = 0.35$
* $P_\pi(2,2) = 0.5 \times 0.6 + 0.5 \times 0.7 = 0.3 + 0.35 = 0.65$

4.) Find $P_\pi = \begin{bmatrix} 0.85 & 0.15 \\ 0.35 & 0.65 \end{bmatrix}$

Exercise # 3

Joyce Anne Colocado CCRNFLRL Con 221

Step 3: Write the Bellman expectation v_π (cloudy)

Sunny

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

$$v_1 = 3 + 0.765 v_1 + 0.135 v_2$$

$$v_1 - 0.765 v_1 - 0.135 v_2 = 3$$

5.) find $v_1 - 0.235 v_1 - 0.135 v_2 = 3$

Cloudy

$$v_2 = 2 + 0.9(0.35 v_1 + 0.15 v_2)$$

$$v_2 = 2 + 0.315 v_1 + 0.585 v_2$$

$$v_2 - 0.315 v_1 - 0.585 v_2 = 2$$

6.) find $v_2 = -0.315 v_1 + 0.415 v_2 = 2$

Step 4: Solve for v_π (cloudy)

$$0.235 v_1 - 0.135 v_2 = 3$$

$$\frac{0.235 v_1}{0.235} = \frac{3 + 0.135 v_2}{0.235}$$

$$v_1 = \frac{3 + 0.135 v_2}{0.235}$$

$$-0.315\left(\frac{3 + 0.135 v_2}{0.235}\right) + 0.415 v_2 = 2$$

$$\left(-0.315 \times \frac{3}{0.235}\right)\left(-0.315 \times \frac{0.135 v_2}{0.235}\right) + 0.415 v_2 = 2$$

$$(-0.315 \times 12.766)(-0.315 \times 0.574) + 0.415 v_2 = 2$$

$$-4.021 - 0.181 v_2 + 0.415 v_2 = 2$$

$$-4.021 + 0.234 v_2 = 2$$

$$0.234 v_2 = 2 + 4.021$$

7. v_π (cloudy) $= \frac{2 + 4.021}{0.234} = \frac{6.021}{0.234} = 25.731$

Exercise # 3

Joyce Anne Colorado CCRNFLRL Com 221

Step 4: Solve for π (sunny)

$$V_1 = \frac{3 + 0.135\,(25.731)}{0.235}$$

$$V_1 = \frac{3 + 3.474}{0.235}$$

$$V_1 = \frac{6.474}{0.235}$$

8.) V_π (sunny) = 27.549

Step 5: Write the Bellman optimality equations

Sunny

$$V_1 = 5 + 0.72 V_1 + 0.18 V_2$$

$$V_1 = -0.72 V_1 - 0.18 V_2 = 5$$

9.) V_* (sunny) = $0.28 V_1 - 0.18 V_2 = 5$

Cloudy

$$V_2 = 5 + 0.36 V_1 + 0.54 V_2$$

$$V_2 = -0.36 V_1 - 0.54 V_2 = 5$$

10.) V_* (Cloudy) = $-0.36 + 0.46 = 5$

Exercise # 3

Joyce Anne Colocado CCRNFLRL Con 221

Step 6: Solve for

$$0.28v_1 - 0.18v_2 = 5$$

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

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

Cloudy Equation

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

$$-0.36 \times \frac{5}{0.28} = -6.429$$

$$-0.36 \times \frac{0.18}{0.28} = -0.231v_2$$

$$-6.429 - 0.231v_2 + 0.46v_2 = 5$$

$$-6.429 + (0.46 - 0.231) = 5$$

$$-6.429 + 0.229v_2 = 5$$

$$0.229v_2 = 5 + 6.429 = 11.429$$

$$0.229v_2 = 11.429$$

11.) v_* (cloudy) $= \frac{11.429}{0.229} = 49.908$

Solve for v_* (sunny)

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

$$v_1 = \frac{5 + 0.18 \times 49.908}{0.28}$$

$$= \frac{5 + 8.983}{0.28}$$

$$= \frac{13.983}{0.28}$$

12.) v_* (sunny) $= 49.939$

Exercise #3

Joyce Ann Colocado CERNFLRL Com 221

Step 7 : Solve for q_*

13.) $q(1, School) = 5 + 0.9(0.8_{r_1} + 0.2_{r_2}) = 5.9$

14.) $q(1, Home) = -5 + 0.9(0.9_{r_1} + 0.1_{r_2}) = -4.1$

15.) $q(2, School) = 3 + 0.9(0.4_{r_1} + 0.6_{r_2}) = 3.9$

16.) $q(2, Home) = 1 + 0.9(0.3_{r_1} + 0.7_{r_2}) = 1.9$