

CSE 574 Homework 3

Zackary Crosley

September 28, 2018

Problem 1. Markov Decision Processes: Consider the world shown in the figure. Assume that 80% of the time the agent moves in its intended direction and 10% of the time it moves in each of the two right angles to that direction. Implement value iteration for each r value below. Discount rewards with a discount factor 0.99. Show the policy obtained in each case and explain why that policy was the intuitive output.

$$v^*(s) = R(s) + \gamma \max_a \sum_{s'} p(s, a, s') v^*(s') \text{ for } s = (i, j) \text{ for } i, j \in [1, 2, 3] \quad (1)$$

All values below were generated by the python file value_iteration.py provided.

1. r = 100

	189	78	10
Iteration 1	78	-1.99	6.7
	-1.99	-1.99	-1.99
	365.1	235.3	10
Iteration 2	255.3	69.1	14.4
	60.58	-1.99	3.55

2. r = -3

	-4.2	6.722	10
Iteration 1	-1.99	-1.99	6.722
	-1.99	-1.99	-1.99
	0.257	14.4	10
Iteration 2	-1.99	4.25	14.4
	-1.99	-1.99	3.55

3. r = 0

	-0.10	6.722	10
Iteration 1	-1.20	-1.99	6.722
	-1.99	-1.99	-1.99
	4.65	14.4	10
Iteration 2	-1.20	4.25	14.4
	-1.99	-1.99	3.55

4. r = 3

	5.574	6.722	10
Iteration 1	1.178	-1.99	6.722
	-1.99	-1.99	-1.99
	10.35	14.4	10
Iteration 2	4.96	4.45	14.4
	-0.41	-1.99	3.55

Problem 2. Value Iteration: For the zero-sum, turn-taking game in Homework 2, let $R(s)$ be the reward for A in state s . Let $-R(s)$ be the reward for B with A in state s . Let $U_A(s)$ be the utility of state s when it's A's turn to move in s , and $U_B(s)$ the same for B.

1. Write down Bellman equations defining $U_A(s)$ and $U_B(s)$.

$$v^*(s) = R(s) + \gamma \max_a \sum_{s'} p(s, a, s') v^*(s') \quad (2)$$

$$U_A(s) = R(s) + \gamma \max_{a_A} \sum_{s'} p(s' | a_A, s) - U_B(s') = R(s) + \gamma \max_{a_A} -U_B(s') \quad (3)$$

$$U_B(s) = -R(s) + \gamma \min_{a_B} \sum_{s'} p(s' | a_B, s) U_A(s') = -R(s) + \gamma \min_{a_B} U_A(s') \quad (4)$$

$p(s, a, s')$ can be omitted since the probability of moving to intended next state is always 1. Since we only have one s' , the $\sum_{s'}$ may also be omitted. Since unspecified, γ will be set to 0.9.

2. Explain how to do two player value iteration with these conditions and define suitable termination criteria. States are tuples $(A = [1 - 4], B = [1 - 4])$ where $A \neq B$. May be shortened to simply numbers, so $(A = 1, B = 2) \leftrightarrow (1, 2)$

Cycle value iterations between positive value iteration from the goal of A and negative value iteration from the goal of B. This is similar to the minimax algorithm in the way it cycles back and forth. Iteration can begin in states that lead to goal nodes, such as $(A=3, B=2)$.

3. Draw the state space showing moves by A in solid lines and moves by B in dashed lines. Mark each state with $R(s)$. Arrange states s_A s_B on a grid, using s_A and s_B as coordinates.

See Figure 1.

4. Apply two-player value iteration to solve the game, and derive the optimal policy.

As can be seen in Figure 1, not all states give one or both A and B a choice in move. Any item in position 1 or 4 can only move in one direction, and half those states are win conditions. Only B has a decision in movement in 1A3B; only A has a decision in movement in 2A4B; both have a decision in 2A3B and 3A2B. Thus, these four states are the only ones which require value iteration. Note that the valid movements for the restricted states (excluding the terminal states XA1B and 4AXB, where no movement is valid) are: A right in 1A2B, 1A3B, 1A4B and A left in 3A4B; B left in 1A4B, 2A4B, 3A4B and B right in 1A2B.

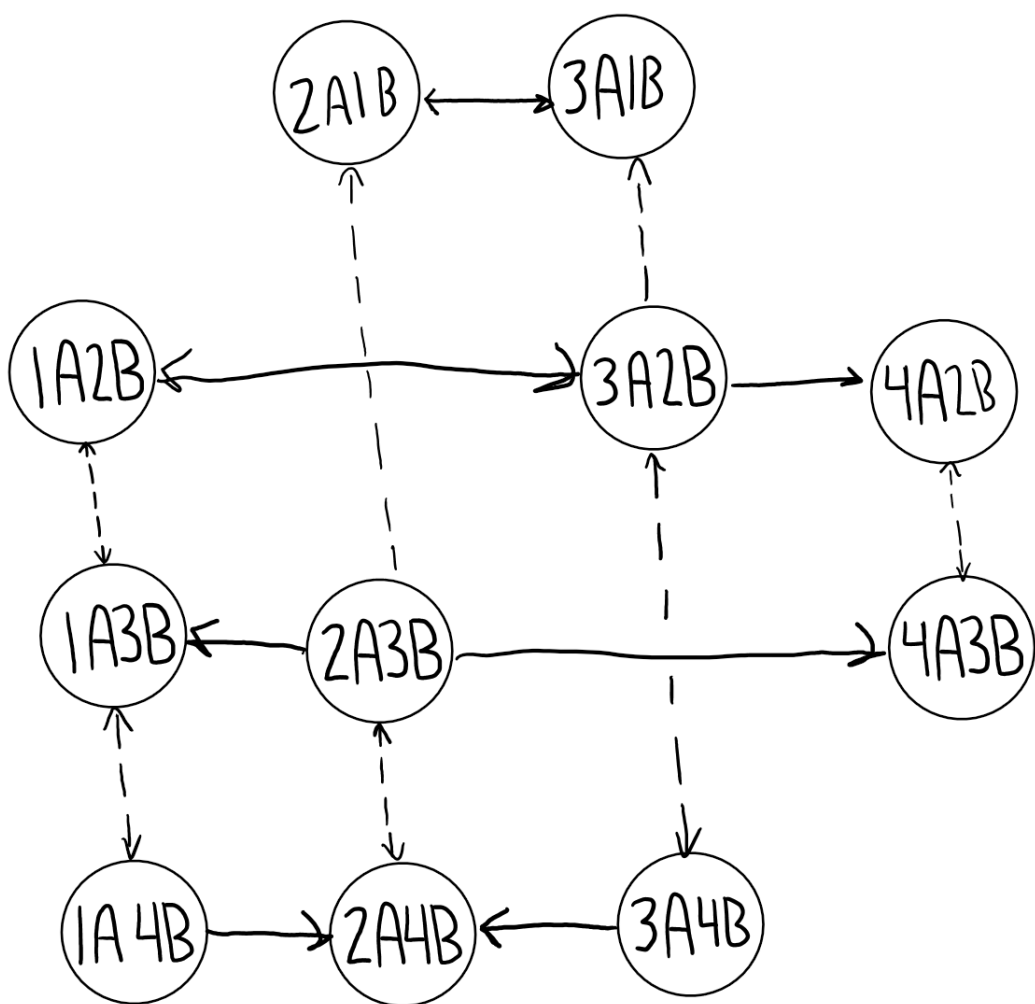


Figure 1: Problem 2c Game State Graph.

$$R(N, 1) = -1; R(4, M) = 1; R(N, M) = 0 \text{ where } N \neq 4 \text{ and } M \neq 1 \quad (5)$$

(A = 2, B = 3)

$$U_A(s) = R(s) + \gamma \max_{a_A} -U_B(s') \quad (6)$$

$$U_A(s) = R((2, 3)) + \gamma \max_{a_A} [-U_B((4, 3)), -U_B((1, 3))] \quad (7)$$

$$U_A(s) = 0 + 0.9 \max_{a_A} [1, x < 1] = 0.9 \text{ (Right)} \quad (8)$$

$$U_B(s) = -R(s) + \gamma \min_{a_B} U_A(s') \quad (9)$$

$$U_B(s) = -R((2, 3)) + 0.9 \min_{a_B} [U_A((2, 1), U_A(2, 4))] \quad (10)$$

$$U_B(s) = 0 + 0.9 \min_{a_B} [-1, x > -1] = -0.9 \text{ (Left)} \quad (11)$$

(A = 3, B = 2)

$$U_A(s) = R((3, 2)) + \gamma \max_{a_A} [-U_B((4, 2)), -U_B((1, 2))] \quad (12)$$

$$U_A(s) = 0 + 0.9 \max_{a_A} [1, x < 1] = 0.9 \text{ (Right)} \quad (13)$$

$$U_B(s) = -R((3, 2)) + 0.9 \min_{a_B} [U_A((3, 1), U_A(3, 4))] \quad (14)$$

$$U_B(s) = 0 + 0.9 \min_{a_B} [-1, x > -1] = -0.9 \text{ (Left)} \quad (15)$$

(A = 2, B = 4)

$$U_A(s) = R((2, 4)) + \gamma \max_{a_A} [-U_B((3, 4)), -U_B((1, 4))] \quad (16)$$

$$U_A(s) = R((2, 4)) + \gamma \max_{a_A} \left[- \left(-0 + \gamma \max_{a_A} 0.9 \right), - \left(-0 + \gamma \max_{a_A} [U_A((1, 3))] \right) \right] \quad (17)$$

$$U_A(s) = R((2, 4)) + \gamma \max_{a_A} [-0.81, 0] = 0 \text{ (Left)} \quad (18)$$

(A = 1, B = 3)

$$U_B(s) = -R((1, 3)) + 0.9 \min_{a_B} [U_A((1, 2), U_A(1, 4))] \quad (19)$$

$$U_B(s) = -R((1, 3)) + 0.9 \min_{a_B} \left[\left(R((1, 2)) + \gamma \max_{a_A} [-U_B((3, 2))] \right), \left(R((1, 4)) + \gamma \max_{a_A} [-U_B((2, 4))] \right) \right] \quad (20)$$

$$U_B(s) = 0.9 \min_{a_B} [(0.9^2), (0)] \quad (21)$$

$$U_B(s) = 0 \text{ (Right)} \quad (22)$$

Therefore, the optimal policy for A is to move right and B to move left unless that direction is blocked or allows the opponent to jump.

Problem 3. Suppose we have a machine that is either running or broken down. If it runs throughout one week, it makes a gross profit of \$100. If it fails, the profit is \$0. If it is running at the start of the week and we perform preventive maintenance, the chance it will fail is 0.4. If we do not perform maintenance, the probability is 0.7. Maintenance costs \$20. When the machine is broken, it may be repaired for \$40 with a probability of failure of 0.4 or replaced at a cost of \$150 which is guaranteed to run through its first week. Find the optimal repair, replacement, and maintenance policy for a four week period. Start with a new machine at time 0.

States: Broken (\$0), Running (\$100)

Actions:

Maintain = (Cost: -\$20, 0.4 S+1 = Broken, 0.6 S+1 = Running)

NoMaintain = (Cost: \$0, 0.7 S+1 = Broken, 0.3 S+1 = Running)

Repair = (Cost: -\$40, 0.4 S+1 = Broken, 0.6 S+1 = Running)

Replace = (Cost: -\$150, 1 S+1 = Running)

$$J_N(x_N) = g_N(x_N) \quad (23)$$

$$J_k(x_k) = \min_{\mu \in \mu} E_{w_k} [g_k(x_k, \mu_k, w_k) + J_{k+1}(f_k(x_{k+1}, \mu_{k+1}, w_{k+1}))] \quad (24)$$

Week 4

$$\begin{aligned} v^*(Running) &= \max_a [-20 + (0.4 * 0 + 0.6 * 100), 0 + (0.7 * 0 + 0.3 * 100)] \\ &= \max_a [-20 + 60, 0 + 30] \\ &= 40 \text{ (Maintain)} \end{aligned} \quad (25)$$

$$\begin{aligned} v^*(Broken) &= \max_a [-40 + (0.4 \times 0 + 0.6 \times 100), -150 + (1 \times 100)] \\ &= \max_a [-40 + 60, 0 - 50] \\ &= 20 \text{ (Repair)} \end{aligned} \quad (26)$$

Week 3

$$\begin{aligned} v^*(Running) &= \max_a [-20 + (0.4 \times [0 + 20] + 0.6 \times [100 + 40]), 0 \\ &\quad + (0.7 \times [0 + 20] + 0.3 \times [100 + 40])] \\ &= \max_a [-20 + 8 + 84, 14 + 42] \\ &= 72 \text{ (Maintain)} \end{aligned} \quad (27)$$

$$\begin{aligned} v^*(Broken) &= \max_a [-40 \\ &\quad + (0.4 \times [0 + 20] + 0.6 \times [100 + 40]), (-150 + 1 \times [100 + 40])] \\ &= \max_a [-40 + 8 + 84, -10] \\ &= 54 \text{ (Repair)} \end{aligned} \quad (28)$$

Week 2

$$\begin{aligned} v^*(Running) &= \max_a [-20 + (0.4 \times [0 + 54] + 0.6 \times [100 + 72]), 0 \\ &\quad + (0.7 \times [0 + 54] + 0.3 \times [100 + 72])] \\ &= \max_a [-20 + 21.6 + 103.2, 37.8 + 51.4] \\ &= 104.8 \text{ (Maintain)} \end{aligned} \quad (29)$$

$$\begin{aligned}
v^*(Broken) &= \max_a [-40 \\
&\quad + (0.4 \times [0 + 54] + 0.6 \times [100 + 72]), (-150 + 1 \times [100 + 72])] \\
&= \max_a [-40 + 21.6 + 103.2, 22] \\
&= 84.8 \text{ (Repair)}
\end{aligned} \tag{30}$$

Week 1

$$\begin{aligned}
v^*(Running) &= \max_a [-20 + (0.4 \times [0 + 84.8] + 0.6 \times [100 + 104.8]), 0 \\
&\quad + (0.7 \times [0 + 84.8] + 0.3 \times [100 + 104.8])] \\
&= \max_a [-20 + 33.92 + 122.88, 59.36 + 61.44] \\
&= 136.8 \text{ (Maintain)}
\end{aligned} \tag{31}$$

$$\begin{aligned}
v^*(Broken) &= \max_a [-40 + (0.4 \times [0 + 84.8] + 0.6 \times [100 + 104.8]), (-150 + 1 \\
&\quad \times [100 + 104.8])] \\
&= \max_a [-40 + 33.92 + 122.88, 54.8] \\
&= 116.8 \text{ (Repair)}
\end{aligned} \tag{32}$$

Therefore, the optimal policy π^* is to maintain each week and repair if the machine breaks.

Problem 4. Dice Blackjack Game - Players bust if they exceed seven they bust. Highest score wins. In event of tie, player two wins. First player rolls until they stop, followed by second. Second must roll until at least value 4. Determine stopping strategy for first player when the second player's initial roll is a 3. Hint: Let $N = 6$ and consider 15 states: 1 busted, 2-8 stopped at sum 1-7, 9-15 at sum 1-7 but still rolling.

$$P(win|x^i) = \max_{\mu} P(win|x^i) \tag{33}$$

$$P(win|x^1) = 0 \text{ (busted)} \tag{34}$$

Since player 2 must roll to get a 4, starts at 3, and wins on tie, stopped on 4 or less player 1 only wins if player 2 busts.

$$P(win|x^2) = P(win|x^3) = P(win|x^4) = P(win|x^5) = \frac{2}{6} \text{ (Probability p2 busted)} \tag{35}$$

Player 2 rolls exactly once since he is at 3, stops once he makes or exceeds 4.

$$P(win|x^6) = P(roll = 1) + p(roll = 5) + p(roll = 6) = \frac{1}{2} \tag{36}$$

$$P(win|x^7) = 1 - (P(roll = 3) + p(roll = 4)) = \frac{4}{6} \tag{37}$$

$$P(win|x^8) = 1 - p(roll = 4) = \frac{5}{6} \tag{38}$$

$$P(win|x^i) \text{ for } i \text{ in } [8, 15] = \tag{39}$$

$$J_N(x_N) = g_N(x_N) \quad (40)$$

$$J_k(x_k) = \min_{\mu \in \mu} E_{w_k} [g_k(x_k, \mu_k, w_k) + J_{k+1}(f_k(x_{k+1}, \mu_{k+1}, w_{k+1}))] \quad (41)$$

$$J_N(x^i) = p(\text{win}|x^i) - (1 - p(\text{win}|x^i)) \quad (42)$$

Roll 7

$$J_N(x^1) = 0 \text{ (No available actions)} \quad (43)$$

$$J_N(x^{15}) = \max_{\mu} \left[\text{Stop}: 1 \times \frac{5}{6}, \text{Roll}: 1 \times 0 \right] = \frac{5}{6} \text{ (Stop)} \quad (44)$$

Roll 6

$$J_N(x^1) = 0 \text{ (No available actions)} \quad (45)$$

$$J_N(x^{14}) = \max_{\mu} \left[\text{Stop}: 1 \times \frac{4}{6}, \text{Roll}: \frac{5}{6} \times 0 + \frac{1}{6} \times \frac{5}{6} \right] = \max_{\mu} \left[\text{Stop}: \frac{4}{6}, \text{Roll}: \frac{5}{36} \right] = \frac{4}{6} \text{ (Stop)} \quad (46)$$

$$J_N(x^{15}) = \max_{\mu} \left[\text{Stop}: 1 \times \frac{5}{6}, \text{Roll}: 1 \times 0 \right] = \frac{5}{6} \text{ (Stop)} \quad (47)$$

Roll 5

$$J_N(x^1) = 0 \text{ (No available actions)} \quad (48)$$

$$J_N(x^{13}) = \max_{\mu} \left[\text{Stop}: 1 \times \frac{3}{6}, \text{Roll}: \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{5}{6} + \frac{4}{6} \times 0 \right] = \max_{\mu} \left[\text{Stop}: \frac{3}{6}, \text{Roll}: \frac{9}{36} \right] = \frac{1}{2} \text{ (Stop)} \quad (49)$$

$$J_N(x^{14}) = \max_{\mu} \left[\text{Stop}: 1 \times \frac{4}{6}, \text{Roll}: \frac{5}{6} \times 0 + \frac{1}{6} \times \frac{5}{6} \right] = \frac{4}{6} \text{ (Stop)} \quad (50)$$

$$J_N(x^{15}) = \max_{\mu} \left[\text{Stop}: 1 \times \frac{5}{6}, \text{Roll}: 1 \times 0 \right] = \frac{5}{6} \text{ (Stop)} \quad (51)$$

Roll 4

$$J_N(x^1) = 0 \text{ (No available actions)} \quad (52)$$

$$J_N(x^{12}) = \max_{\mu} \left[\text{Stop}: 1 \times \frac{2}{6}, \text{Roll}: \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \frac{1}{2} + \frac{3}{6} \times 0 \right] = \max_{\mu} \left[\text{Stop}: \frac{2}{6}, \text{Roll}: \frac{12}{36} \right] = \frac{2}{6} \text{ (Either)} \quad (53)$$

$$J_N(x^{13}) = \max_{\mu} \left[\text{Stop}: 1 \times \frac{3}{6}, \text{Roll}: \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{5}{6} + \frac{4}{6} \times 0 \right] = \frac{1}{2} \text{ (Stop)} \quad (54)$$

$$J_N(x^{14}) = \max_{\mu} \left[\text{Stop}: 1 \times \frac{4}{6}, \text{Roll}: \frac{5}{6} \times 0 + \frac{1}{6} \times \frac{5}{6} \right] = \frac{4}{6} \text{ (Stop)} \quad (55)$$

$$J_N(x^{15}) = \max_{\mu} \left[\text{Stop}: 1 \times \frac{5}{6}, \text{Roll}: 1 \times 0 \right] = \frac{5}{6} \text{ (Stop)} \quad (56)$$

Roll 3

$$J_N(x^1) = 0 \text{ (No available actions)} \quad (57)$$

$$J_N(x^{11}) = \max_{\mu} \left[\text{Stop}: \frac{2}{6}, \text{Roll}: \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{1}{2} + \frac{1}{6} \times \frac{2}{6} \right] = \max_{\mu} \left[\text{Stop}: \frac{2}{6}, \text{Roll}: \frac{14}{36} \right] = \frac{14}{36} \text{ (Roll)} \quad (58)$$

$$J_N(x^{12}) = \max_{\mu} \left[\text{Stop}: 1 \times \frac{2}{6}, \text{Roll}: \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \frac{1}{2} + \frac{3}{6} \times 0 \right] = \frac{2}{6} \text{ (Either)} \quad (59)$$

$$J_N(x^{13}) = \max_{\mu} \left[\text{Stop: } 1 \times \frac{3}{6}, \text{Roll: } \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{5}{6} + \frac{4}{6} \times 0 \right] = \frac{1}{2} \text{ (Stop)} \quad (60)$$

$$J_N(x^{14}) = \max_{\mu} \left[\text{Stop: } 1 \times \frac{4}{6}, \text{Roll: } \frac{5}{6} \times 0 + \frac{1}{6} \times \frac{5}{6} \right] = \frac{4}{6} \text{ (Stop)} \quad (61)$$

$$J_N(x^{15}) = \max_{\mu} \left[\text{Stop: } 1 \times \frac{5}{6}, \text{Roll: } 1 \times 0 \right] = \frac{5}{6} \text{ (Stop)} \quad (62)$$

Roll 2

$$J_N(x^1) = 0 \text{ (No available actions)} \quad (63)$$

$$J_N(x^{10}) = \max_{\mu} \left[\text{Stop: } \frac{2}{6}, \text{Roll: } \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{1}{2} + 2 \times \left(\frac{1}{6} \times \frac{2}{6} \right) \right] = \max_{\mu} \left[\text{Stop: } \frac{2}{6}, \text{Roll: } \frac{16}{36} \right] = \frac{16}{36} \text{ (Roll)} \quad (64)$$

$$J_N(x^{11}) = \max_{\mu} \left[\text{Stop: } \frac{2}{6}, \text{Roll: } \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{1}{2} + \frac{1}{6} \times \frac{2}{6} \right] = \frac{14}{36} \text{ (Roll)} \quad (65)$$

$$J_N(x^{12}) = \max_{\mu} \left[\text{Stop: } 1 \times \frac{2}{6}, \text{Roll: } \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \frac{1}{2} + \frac{3}{6} \times 0 \right] = \frac{2}{6} \text{ (Either)} \quad (66)$$

$$J_N(x^{13}) = \max_{\mu} \left[\text{Stop: } 1 \times \frac{3}{6}, \text{Roll: } \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{5}{6} + \frac{4}{6} \times 0 \right] = \frac{1}{2} \text{ (Stop)} \quad (67)$$

$$J_N(x^{14}) = \max_{\mu} \left[\text{Stop: } 1 \times \frac{4}{6}, \text{Roll: } \frac{5}{6} \times 0 + \frac{1}{6} \times \frac{5}{6} \right] = \frac{4}{6} \text{ (Stop)} \quad (68)$$

$$J_N(x^{15}) = \max_{\mu} \left[\text{Stop: } 1 \times \frac{5}{6}, \text{Roll: } 1 \times 0 \right] = \frac{5}{6} \text{ (Stop)} \quad (69)$$

Roll 1

$$J_N(x^1) = 0 \text{ (No available actions)} \quad (70)$$

$$J_N(x^9) = \max_{\mu} \left[\text{Stop: } \frac{2}{6}, \text{Roll: } \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{1}{2} + 3 \times \left(\frac{1}{6} \times \frac{2}{6} \right) \right] = \max_{\mu} \left[\text{Stop: } \frac{2}{6}, \text{Roll: } \frac{18}{36} \right] = \frac{18}{36} \text{ (Roll)} \quad (71)$$

$$J_N(x^{10}) = \max_{\mu} \left[\text{Stop: } \frac{2}{6}, \text{Roll: } \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{1}{2} + 2 \times \left(\frac{1}{6} \times \frac{2}{6} \right) \right] = \frac{16}{36} \text{ (Roll)} \quad (72)$$

$$J_N(x^{11}) = \max_{\mu} \left[\text{Stop: } \frac{2}{6}, \text{Roll: } \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{1}{2} + \frac{1}{6} \times \frac{2}{6} \right] = \frac{14}{36} \text{ (Roll)} \quad (73)$$

$$J_N(x^{12}) = \max_{\mu} \left[\text{Stop: } 1 \times \frac{2}{6}, \text{Roll: } \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \frac{1}{2} + \frac{3}{6} \times 0 \right] = \frac{2}{6} \text{ (Either)} \quad (74)$$

$$J_N(x^{13}) = \max_{\mu} \left[\text{Stop: } 1 \times \frac{3}{6}, \text{Roll: } \frac{1}{6} \times \frac{4}{6} + \frac{1}{6} \times \frac{5}{6} + \frac{4}{6} \times 0 \right] = \frac{1}{2} \text{ (Stop)} \quad (75)$$

$$J_N(x^{14}) = \max_{\mu} \left[\text{Stop: } 1 \times \frac{4}{6}, \text{Roll: } \frac{5}{6} \times 0 + \frac{1}{6} \times \frac{5}{6} \right] = \frac{4}{6} \text{ (Stop)} \quad (76)$$

$$J_N(x^{15}) = \max_{\mu} \left[\text{Stop: } 1 \times \frac{5}{6}, \text{Roll: } 1 \times 0 \right] = \frac{5}{6} \text{ (Stop)} \quad (77)$$

Therefore the optimal policy is to keep rolling below four and stop above four. At four, either stopping or continuing to roll is equally likely to result in a win.

Problem 5. We are given six coins, one of which is counterfeit and is known to have different weight than the rest. Construct a strategy to find the counterfeit coin using a two-pan scale in a minimum average number of tries. Hint: There are two initial decisions that make sense: (1) test two of the coins against two others, and (2) test one of the coins against one other.

Obvious strategy is to first compare two coins against two coins. If these are the same weight, then we know the oddly weighted coin is in the remaining two. The first four coins can be marked or separated, depending on what works best. If they are not the same, the unweighed pair of coins must be the normally weighted. Weigh one of the pairs of two coins from the first step against the remaining two. If they are the same, then one of the other two coins must be differently weighted. Else, if they are different, then the pair of coins from the original comparison must contain the oddly weighted coin. We are now down to two coins, one of which must be differently weighted, in at most two comparisons. Weigh one of the remaining coins against a coin which is of known normal weight. If this weight is unequal, then this coin is the counterfeit one. Else, the last coin is. This policy finds the counterfeit coin in two cases, in the best case, and three in the worst.

Problem 6. In the framework of the basic problem, consider the case where the cost has the multiplicative form:

$$E_{w_{k=0,1,\dots,N-1}} \{g_N(x_N) \times g_{N-1}(x_{N-1}, u_{N-1}, w_{N-1} \dots g_0(x_0, u_0, w_0)\} \quad (78)$$

Develop a DP-like algorithm for this problem assuming that $g_k(x_k, u_k, w_k) \geq 0$ for all x_k, u_k, w_k , and k .

```
# Assumes x, u, and w are lists of the values for x, u, and w
# Also assumes function 'g' already exists and handles missing parameters
# shown in function.
E(N, x, u, w):
    ret g(N, x[N]) * g_recursive(N-1, x, u, w)

g_recursive(N, x, u, w):
    ret g(N, x[N], u[N], w[N]) * g_recursive(N-1, x, u, w)
```

Problem 7. $\neg \setminus ({}^{\vee}) _ /$