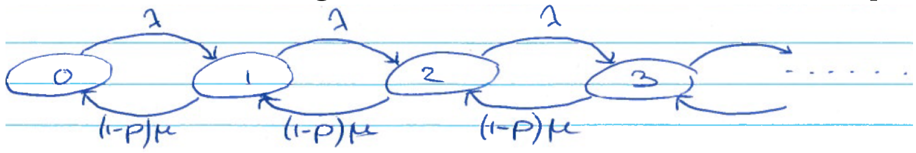
Answers Example Exam 2 Stochastic Models M8

1. **Consider the following variant of the M/M/1 queueing system. Upon completion of a customer's service, the customer again joins the (end of the) queue with probability p. Each customer continues to do so after each service until he or she has left the system. All service times are mutually independent exponentially distributed variables. Also, the probability p is independent of the number of services a customer has already completed.**

   (a) **Draw the transition diagram for the number of customers in this queueing system.**

   (b) **Give the balance equations and solve these.**

   Local balance equations:

   $\lambda P_i = (1-p)\mu P_{i+1}$

   $\sum_{i=0}^{\infty} P_i = 1$

   Solving these:

   $P_1 = \frac{\lambda}{(1-p)\mu} P_0$

   $P_2 = \frac{\lambda}{(1-p)\mu} P_1 = \left(\frac{\lambda}{(1-p)\mu}\right)^2 P_0$

   $P_3 = \frac{\lambda}{(1-p)\mu} P_2 = \left(\frac{\lambda}{(1-p)\mu}\right)^3 P_0$

   $P_i = \left(\frac{\lambda}{(1-p)\mu}\right)^i P_0$

   $\sum_{i=0}^{\infty} \left(\frac{\lambda}{(1-p)\mu}\right)^i P_0 = 1 \Rightarrow \frac{1}{1-\frac{\lambda}{(1-p)\mu}} P_0 = 1$

   $P_0 = 1 - \frac{\lambda}{(1-p)\mu}$

   (c) **Give the stability condition for this problem.**
   $\lambda < (1-p)\mu$

   (d) **What is the probability that a customer undergoes exactly k services?**
   $(1-p)p^{k-1}$, $k \geq 1$

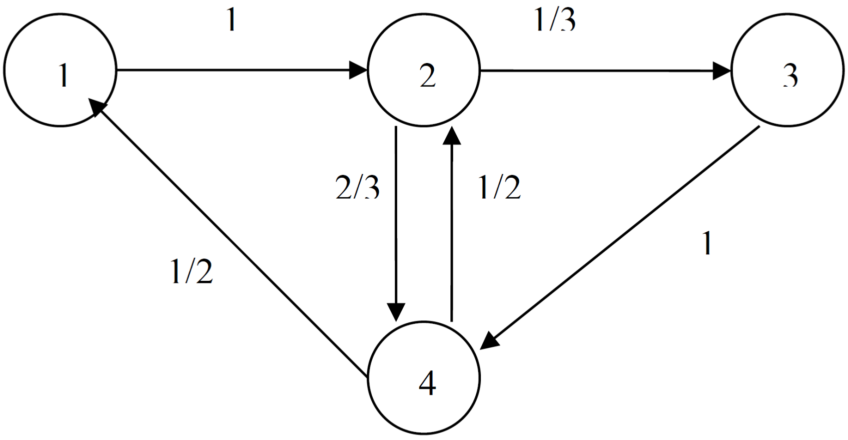
   (e) **What is the mean number of customers in the queue?**
   $E(L^q) = \frac{\tilde{\rho}^2}{1-\tilde{\rho}}$ with $\tilde{\rho} = \frac{\lambda}{(1-p)\mu}$.

   (f) **What is the utilization of the server?**
   $1 - P_0 = \frac{\lambda}{(1-p)\mu}$

   (g) **Which classical M/M/1 queueing system has the same average number of customers in the queue as the system considered in this problem?**
   The M/M/1 queuing system with arrival rate λ and service rate $(1-p)\mu$.

2. **Consider the closed network in the figure underneath. The numbers at the arrows give the transition probabilities for a customer leaving the queue to route to a subsequent queue. Every station contains a single server, and all arriving customers may enter the**

station. Service is in order of arrival. The service times have an exponential distribution with: $\mu_1 = 4, \mu_2 = 3, \mu_3 = 2, \mu_4 = 1$.



(a) **Give the joint stationary distribution for the number of customers in the four stations for** $m = 1, 2,$ **and** 3 **(**m**=total number of customers in the network).**
Use Buzen's algorithm:
The stationary distribution follows from:

$$\begin{aligned}\pi_1 &= \frac{1}{2}\pi_4\\ \pi_2 &= \pi_1 + \frac{1}{2}\pi_4\\ \pi_3 &= \frac{1}{3}\pi_2\\ \pi_4 &= \frac{2}{3}\pi_2 + \pi_3\\ \sum_i \pi_i &= 1\end{aligned}$$

So $\pi_1 = \frac{3}{17}$, $\pi_2 = \frac{6}{17}$, $\pi_3 = \frac{2}{17}$, $\pi_4 = \frac{6}{17}$. Also, $\rho_i = \frac{\pi_i}{\mu_i}$ so $\rho_1 = \frac{3}{68}$, $\rho_2 = \frac{2}{17}$, $\rho_3 = \frac{1}{17}$, $\rho_4 = \frac{6}{17}$.
In Buzen's algorithm, we have $C_1(k) = \rho_1^k$, which gives:
$C_1(0) = \rho_1^0 = \left(\frac{3}{68}\right)^0 = 1$
$C_1(1) = \rho_1^1 = \left(\frac{3}{68}\right)^1 = \frac{3}{68}$
$C_1(2) = \rho_1^2 = \left(\frac{3}{68}\right)^2 = \frac{9}{4624}$
$C_1(3) = \rho_1^3 = \left(\frac{3}{68}\right)^3 = \frac{27}{314432}$
For $i > 1$, we have $C_i(0) = 1$ and $C_i(k) = C_{i-1}(k) + \rho_i C_i(k-1)$, which gives:
$C_2(0) = 1$
$C_2(1) = C_1(1) + \rho_2 C_2(0) = \frac{11}{68}$
$C_2(2) = C_1(2) + \rho_2 C_2(1) = \frac{97}{4624}$
$C_2(3) = C_1(3) + \rho_2 C_2(2) = \frac{803}{314432}$
$C_3(0) = 1$
$C_3(1) = C_2(1) + \rho_3 C_3(0) = \frac{15}{68}$
$C_3(2) = C_2(2) + \rho_3 C_3(1) = \frac{157}{4624}$
$C_3(3) = C_2(3) + \rho_3 C_3(2) = \frac{1431}{314432}$
$C_4(0) = 1$
$C_4(1) = C_3(1) + \rho_4 C_4(0) = \frac{39}{68}$
$C_4(2) = C_3(2) + \rho_4 C_4(1) = \frac{1093}{4624}$
$C_4(3) = C_3(3) + \rho_4 C_4(2) = \frac{27663}{314432}$
Now, the joint stationary distribution for the number of customers in the four stations follows from

$$\Pi_m(\mathbf{n}) = \Pi_m(n_1, n_2, n_3, n_4) = \frac{\rho_1^{n_1}\rho_2^{n_2}\rho_3^{n_3}\rho_4^{n_4}}{C_4(m)}$$

Thus, we get:
$\Pi_1(\mathbf{n}) = \frac{68}{39} \cdot \left(\frac{3}{68}\right)^{n_1} \left(\frac{2}{17}\right)^{n_2} \left(\frac{1}{17}\right)^{n_3} \left(\frac{6}{17}\right)^{n_4}$
$\Pi_2(\mathbf{n}) = \frac{4624}{1093} \cdot \left(\frac{3}{68}\right)^{n_1} \left(\frac{2}{17}\right)^{n_2} \left(\frac{1}{17}\right)^{n_3} \left(\frac{6}{17}\right)^{n_4}$
$\Pi_3(\mathbf{n}) = \frac{314432}{27663} \cdot \left(\frac{3}{68}\right)^{n_1} \left(\frac{2}{17}\right)^{n_2} \left(\frac{1}{17}\right)^{n_3} \left(\frac{6}{17}\right)^{n_4}$

(b) **Use Mean Value Analysis to obtain the average number of customers and the average sojourn time in the four queues for $m = 1, 2$, and 3.**
$m = 1$:
$F_1(i) = \frac{1}{\mu_i}$, so the average sojourn time in the four queues is
$F_1(1) = \frac{1}{4},\ F_1(2) = \frac{1}{3},\ F_1(3) = \frac{1}{2},\ F_1(4) = 1.$
$1 = m = \sum_i L_1(i) = \lambda_1 \sum_i \pi_i F_1(i) = \lambda_1(\frac{3}{17} \cdot \frac{1}{4} + \frac{6}{17} \cdot \frac{1}{3} + \frac{2}{17} \cdot \frac{1}{2} + \frac{6}{17} \cdot 1) = \frac{39}{68}\lambda_1$
So, $\lambda_1 = \frac{68}{39}$.
Now $\lambda_1(i) = \lambda_1 \cdot \pi_i$, so $\lambda_1(1) = \frac{4}{13},\ \lambda_1(2) = \frac{8}{13},\ \lambda_1(3) = \frac{8}{39},\ \lambda_1(4) = \frac{8}{13}$.
Also, $L_1(i) = \lambda_1(i) \cdot F_1(i)$, so the average number of customers in the four queues is
$L_1(1) = \frac{1}{13},\ L_1(2) = \frac{8}{39},\ L_1(3) = \frac{4}{39},\ L_1(4) = \frac{8}{13}$.

$m = 2$:
$F_2(i) = \frac{1}{\mu_i} + L_1(i)\frac{1}{\mu_i}$, so the average sojourn time in the four queues is
$F_2(1) = \frac{7}{26},\ F_2(2) = \frac{47}{117},\ F_2(3) = \frac{43}{78},\ F_2(4) = \frac{21}{13}$.
$2 = m = \sum_i L_2(i) = \lambda_2 \sum_i \pi_i F_2(i) = \lambda_2(\frac{3}{17} \cdot \frac{7}{26} + \frac{6}{17} \cdot \frac{47}{117} + \frac{2}{17} \cdot \frac{43}{78} + \frac{6}{17} \cdot \frac{21}{13}) = \frac{1093}{1326}\lambda_2$
So, $\lambda_2 = \frac{2652}{1093}$.
Now $\lambda_2(i) = \lambda_2 \cdot \pi_i$, so $\lambda_2(1) = \frac{468}{1093},\ \lambda_2(2) = \frac{936}{1093},\ \lambda_2(3) = \frac{312}{1093},\ \lambda_2(4) = \frac{8936}{1093}$.
Also, $L_2(i) = \lambda_2(i) \cdot F_2(i)$, so the average number of customers in the four queues is
$L_2(1) = \frac{126}{1093},\ L_2(2) = \frac{376}{1093},\ L_2(3) = \frac{172}{1093},\ L_2(4) = \frac{1512}{1093}$.

$m = 3$:
$F_3(i) = \frac{1}{\mu_i} + L_2(i)\frac{1}{\mu_i}$, so the average sojourn time in the four queues is
$F_3(1) = \frac{1219}{4372},\ F_3(2) = \frac{1469}{3279},\ F_3(3) = \frac{1265}{2186},\ F_3(4) = \frac{2605}{1093}$.
$3 = m = \sum_i L_3(i) = \lambda_3 \sum_i \pi_i F_3(i) = \lambda_3(\frac{3}{17} \cdot \frac{1219}{4372} + \frac{6}{17} \cdot \frac{1469}{3279} + \frac{2}{17} \cdot \frac{1265}{2186} + \frac{6}{17} \cdot \frac{2605}{1093}) = 1.11658\lambda_3$
So, $\lambda_3 = 2.68677$.
Now $\lambda_3(i) = \lambda_3 \cdot \pi_i$, so $\lambda_3(1) = 0.47414,\ \lambda_3(2) = 0.94827,\ \lambda_3(3) = 0.31609,\ \lambda_3(4) = 0.94827$.
Also, $L_3(i) = \lambda_3(i) \cdot F_3(i)$, so the average number of customers in the four queues is
$L_3(1) = 0.13220,\ L_3(2) = 0.42483,\ L_3(3) = 0.18292,\ L_3(4) = 2.26006$.

(c) **For $m = 1$, determine the average number of transitions for a customer to return for the first time to station 1.**
For $m = 1$, our closed network is an ergodic Markov chain, in which we have to determine the *mean first passage time* from station 1 to station 1, m_{11}.

$$m_{ii} = \frac{1}{\pi_i} \qquad \text{so} \qquad m_{11} = \frac{1}{\pi_1} = \frac{17}{3}$$

Hence, it takes on average $\frac{17}{3}$ transitions for a customer to return for the first time to station 1.

3. **G. Ambler has €10,000 available for a second hand car, but would like to buy a fast car that costs €25,000. He needs the money for that car quickly, and would like to increase his capital to €25,000 via a gambling game. To this end, he can play a game in which he is allowed to toss an imperfect (with probability 0.4 for heads) coin three times. For each toss he may bet each amount (in multiples of €1000, and the amount should be in his possession). He will win the amount (i.e., receives twice the amount of the bet) when he tosses heads, and lose his betted amount when he tosses tails. Use stochastic dynamic programming to determine a strategy that maximizes the probability of reaching €25,000 after three tosses.**

(a) **Determine the phases n, states i, decisions d, and optimal value function $f_n(i)$ for this stochastic dynamic programming problem.**
Phases n: the tosses: $n = 1, 2, 3$.

States i: the amount of money G. Ambler has (in multiples of €1000).
Decisions d: the amount of money G. Ambler is going to bet: $0 \leq d \leq i$.
Optimal value function: $f_n(i)$ = the maximal probability that G. Ambler will have €25,000 at the end of the game if he owns capital i at toss n.

(b) **Give the recurrence relations for the optimal value function.**

$$f_4(i) = \begin{cases} 1 & \text{if } i \geq 25, \\ 0 & \text{if } i < 25. \end{cases}$$

For $n < 4$:

$$f_n(i) = \max_{0 \leq d_n \leq i} \{r(i, d_n) + 0.4 \cdot f_{n+1}(i + d_n) + 0.6 \cdot f_{n+1}(i - d_n)\},$$

where

$$\begin{aligned} r(i, 0) &= 1 \text{ for } i \geq 25, \\ r(i, d_n) &= 0 \text{ otherwise,} \end{aligned}$$

and d_n should be integer (G. Ambler may only bet in multiples of €1000).
(We aim to find $f_1(10)$.)

(c) **Determine the optimal policy, and describe in words what this policy does. What is the expected probability of success?**
$n = 3$:
At toss 3, there are three options:

- $i \geq 25$: G. Ambler has already reached his goal, so he does not have to bet ($d_3 = 0$).
- $13 \leq i < 25$: by betting $d_3 \geq 25 - i$, G. Ambler can reach his goal with probability 0.4.
- $i < 13$: G. Ambler cannot reach 25 in one toss, and thus has already lost the game ($d_3 = 0$).

Hence, we get:

$$f_3(i) = \begin{cases} 1 & \text{if } i \geq 25, \\ \max_{0 \leq d_3 \leq i}\{0 + 0.4 \cdot f_4(i + d_3) + 0.6 \cdot f_4(i - d_3)\} = 0.4 \cdot 1 = 0.4 & \text{if } 13 \leq i < 25, \\ 0 & \text{if } i < 13. \end{cases}$$

$n = 2$:
At the start of toss 2, G. Ambler's maximal capital is 20, because he starts with 10 at toss 1. Four situations need to be distinguished at toss 2:

- $19 \leq i \leq 20$: bet $d_2 = 25 - i$ and get to 25 with probability 0.4. In case of a loss (prob. 0.6) there is still $i \geq 13$ left, so there is still a chance to win the game at toss 3.
- $13 \leq i < 19$: bet $d_2 = 25 - i$ and get to 25 with probability 0.4. In case of a loss (prob. 0.6), the capital left is $i < 13$, so there is no chance to win the game at toss 3.
- $7 \leq i < 13$: there is no chance of winning the game immediately, G. Ambler has to get to $i \geq 13$ to have a chance to win the game in toss 3. This can be done with probability 0.4, by betting $d_2 \geq 13 - i$. With probability 0.6, G. Ambler looses toss 2 and reaches game over immediately.
- $i < 7$: G. Ambler cannot get to 25 at all, so he has already lost the game (because the highest starting capital he could possibly reach for toss 3 is $i < 13$).

Hence, we get:

$$\begin{aligned} f_2(i) &= \max_{0 \leq d_2 \leq i} \{0 + 0.4 \cdot f_3(i + d_2) + 0.6 \cdot f_3(i - d_2)\} \\ &= \begin{cases} 0.4 \cdot 1 + 0.6 \cdot 0.4 = 0.64 & \text{if } 19 \leq i \leq 20, \\ 0.4 \cdot 1 + 0.6 \cdot 0 = 0.4 & \text{if } 13 \leq i < 19, \\ 0.4 \cdot 0.4 + 0.6 \cdot 0 = 0.16 & \text{if } 7 \leq i < 13, \\ 0 & \text{if } i < 7. \end{cases} \end{aligned}$$

$n = 1$:
At the start of toss 1, G. Ambler has 10, so $i = 10$. He can do the following:

- $d_1 \geq 9$ win: get to $19 \leq i_2 \leq 20$; loose: game over.
- $4 \leq d_1 < 9$ win: get to $13 \leq i_2 < 19$; loose: game over.

- $d_1 = 3$ win: get to $13 \le i_2 < 19$; loose: get to $7 \le i_2 < 13$.
- $0 \le d_1 < 3$ win: get to $7 \le i_2 < 13$; loose: $7 \le i_2 < 13$.

Hence, we get:

$$\begin{aligned} f_1(10) &= \max_{0\le d_1 \le 10}\{0 + 0.4 \cdot f_2(10+d_1) + 0.6 \cdot f_2(10-d_1)\} \\ &= \max\{0.4 \cdot 0.64 + 0.6 \cdot 0, \qquad d_1 \ge 9, \\ &\qquad 0.4 \cdot 0.4 + 0.6 \cdot 0, \qquad 4 \le d_1 < 9, \\ &\qquad 0.4 \cdot 0.4 + 0.6 \cdot 0.16, \qquad d_1 = 3, \\ &\qquad 0.4 \cdot 0.16 + 0.6 \cdot 0.16\} \qquad 0 \le d_1 < 3, \\ &= \max\{0.256, 0.160, 0.256, 0.160\} = 0.256. \end{aligned}$$

Concluding: the expected probability of success is 0.256. G. Ambler has two possible strategies for optimally playing the game:
Option 1: Bet ≥ 9 in toss 1, if winning, bet $25 - i$ in toss 2 (if loosing, there is not second toss). If winning toss 2, 25 has been reached, so don't bet in toss 3. If loosing toss 2, bet $25 - i$ in toss 3. (G. Ambler can win the game by either winning in toss 1 and toss 2, or by winning toss 1, loosing toss 2, and winning toss 3.)
Option 2: Bet 3 in toss 1. If winning, bet $25 - i$ in toss 2 and either win or loose the game immediately; there is no third toss. If loosing toss 1, bet $\ge 13 - i$ in toss 2. If loosing toss 2, the game is over. If winning toss 2, bet $\ge 25 - i$ in toss 3. (In this case, G. Ambler can win the game by either winning in toss 1 and toss 2, or by loosing toss 1 and winning tosses 2 and 3.)

4. **The inventory of a certain good is inspected periodically. If a replenishment order is placed of size $x > 0$ (integer), the ordering costs are $8 + 2x$. The delivery time is zero. The demand is stochastic and equals 1 or 2 with probability $\frac{1}{2}$. Demand in subsequent periods is independent. The size of a replenishment order must be such that (i) demand in a period is always satisfied, and (ii) the stock at the end of a period never exceeds 2. The holding costs in a period are 2 per unit of inventory remaining at the end of a period. The target is to minimize the expected discounted costs over an infinite horizon, using discount factor 0.8.**

   (a) **Model this problem as a Markov decision problem. What do you choose as states, decisions and optimal value function?**
   States i: a period's starting inventory (= the previous period's ending inventory), $i \in \{0, 1, 2\}$.
   Decisions: the size x of the order that is placed.
   Possible decisions: $D(0) = \{2, 3\},\ D(1) = \{1, 2\},\ D(2) = \{0, 1\}$.
   Optimal value function $V(i)$: the minimum expected discounted costs over an infinite horizon if the stock is i at the beginning.

   (b) **Determine the direct costs and transition probabilities for each state and decision.**
   Direct costs:
   $c(0,2) = \frac{1}{2} \cdot 0 \cdot 2 + \frac{1}{2} \cdot 1 \cdot 2 + 8 + 2 \cdot 2 = 13$
   $c(0,3) = \frac{1}{2} \cdot 1 \cdot 2 + \frac{1}{2} \cdot 2 \cdot 2 + 8 + 2 \cdot 3 = 17$
   $c(1,1) = \frac{1}{2} \cdot 0 \cdot 2 + \frac{1}{2} \cdot 1 \cdot 2 + 8 + 2 \cdot 1 = 11$
   $c(1,2) = \frac{1}{2} \cdot 1 \cdot 2 + \frac{1}{2} \cdot 2 \cdot 2 + 8 + 2 \cdot 2 = 15$
   $c(2,0) = \frac{1}{2} \cdot 0 \cdot 2 + \frac{1}{2} \cdot 1 \cdot 2 = 1$
   $c(2,1) = \frac{1}{2} \cdot 1 \cdot 2 + \frac{1}{2} \cdot 2 \cdot 2 + 8 + 2 \cdot 1 = 13$

   Transition probabilities:
   $p(0|0,2) = 0.5 \quad p(0|1,1) = 0.5 \quad p(0|2,0) = 0.5$
   $p(1|0,2) = 0.5 \quad p(1|1,1) = 0.5 \quad p(1|2,0) = 0.5$
   $p(1|0,3) = 0.5 \quad p(1|1,2) = 0.5 \quad p(1|2,1) = 0.5$
   $p(2|0,3) = 0.5 \quad p(2|1,2) = 0.5 \quad p(2|2,1) = 0.5$
   $p(j|i,x) = 0$ for all other combinations of i, j, and x

   (c) **Formulate the corresponding optimality equations.**
   $V(i) = \min_x\{c(i,x) + 0.8 \sum_j p(j|i,x)V(j)\}$

(d) **Carry out the initialization and two additional iterations of the value iteration algorithm. What approximation to the optimal values do you find? What is the policy that corresponds to these values?**

$$V_0(0) = \min \begin{cases} 13^* & (x=2) \\ 17 & (x=3) \end{cases}$$

$$V_0(1) = \min \begin{cases} 11^* & (x=1) \\ 15 & (x=2) \end{cases}$$

$$V_0(2) = \min \begin{cases} 1^* & (x=0) \\ 13 & (x=1) \end{cases}$$

$$V_1(0) = \min \begin{cases} 13 + 0.8(1/2V_0(1) + 1/2V_0(0)) = 22.6 & (x=2) \\ 17 + 0.8(1/2V_0(2) + 1/2V_0(1)) = 21.8^* & (x=3) \end{cases}$$

$$V_1(1) = \min \begin{cases} 11 + 0.8(1/2V_0(1) + 1/2V_0(0)) = 20.6 & (x=1) \\ 15 + 0.8(1/2V_0(2) + 1/2V_0(1)) = 19.8^* & (x=2) \end{cases}$$

$$V_1(2) = \min \begin{cases} 1 + 0.8(1/2V_0(1) + 1/2V_0(0)) = 10.6^* & (x=0) \\ 13 + 0.8(1/2V_0(2) + 1/2V_0(1)) = 17.8 & (x=1) \end{cases}$$

$$V_2(0) = \min \begin{cases} 13 + 0.8(1/2V_1(1) + 1/2V_1(0)) = 29.64 & (x=2) \\ 17 + 0.8(1/2V_1(2) + 1/2V_1(1)) = 29.16^* & (x=3) \end{cases}$$

$$V_2(1) = \min \begin{cases} 11 + 0.8(1/2V_1(1) + 1/2V_1(0)) = 27.64 & (x=1) \\ 15 + 0.8(1/2V_1(2) + 1/2V_1(1)) = 27.16^* & (x=2) \end{cases}$$

$$V_2(2) = \min \begin{cases} 1 + 0.8(1/2V_1(1) + 1/2V_1(0)) = 17.64^* & (x=0) \\ 13 + 0.8(1/2V_1(2) + 1/2V_1(1)) = 25.16 & (x=1) \end{cases}$$

The approximation to the optimal values is $V_2(0) = 29.16, V_2(1) = 27.16, V_2(2) = 25.16$ with corresponding policy $\delta(0) = 3,\ \delta(1) = 2,\ \delta(2) = 0$.

(e) **Choose an ordering policy, and use the policy iteration algorithm to investigate whether or not this policy is optimal.**
E.g., start with the ordering policy $\delta(0) = 3,\ \delta(1) = 2,\ \delta(2) = 0$. For this policy, the value determination equations are:

$$\begin{aligned} V_\delta(0) &= 17 + 0.8\,[0.5V_\delta(2) + 0.5V_\delta(1)] \\ V_\delta(1) &= 15 + 0.8\,[0.5V_\delta(2) + 0.5V_\delta(1)] \\ V_\delta(2) &= 1 + 0.8\,[0.5V_\delta(1) + 0.5V_\delta(0)] \end{aligned}$$

Solving these equations we obtain $V_\delta(0) = \frac{407}{7} = 58.14,\ V_\delta(1) = \frac{393}{7} = 56.14,\ V_\delta(2) = \frac{327}{7} = 46.71$. The policy improvement step of the policy iteration methods yields

$$T_\delta(0) = \min \begin{cases} 13 + 0.8[0.5V_\delta(1) + 0.5V_\delta(0)] = \frac{411}{7} = 58.71 & (x=2) \\ 17 + 0.8[0.5V_\delta(2) + 0.5V_\delta(1)] = \frac{407}{7} = 58.14^* & (x=3) \end{cases}$$

$$T_\delta(1) = \min \begin{cases} 11 + 0.8[0.5V_\delta(1) + 0.5V_\delta(0)] = \frac{397}{7} = 56.71 & (x=1) \\ 15 + 0.8[0.5V_\delta(2) + 0.5V_\delta(1)] = \frac{393}{7} = 56.14^* & (x=2) \end{cases}$$

$$T_\delta(2) = \min \begin{cases} 1 + 0.8[0.5V_\delta(1) + 0.5V_\delta(0)] = \frac{327}{7} = 46.71^* & (x = 0) \\ 13 + 0.8[0.5V_\delta(2) + 0.5V_\delta(1)] = \frac{379}{7} = 54.14 & (x = 1) \end{cases}$$

We have $T_\delta(i) = V_\delta(i)$ for all states i, so δ is an optimal stationary policy.

Suppose that, instead of minimizing the expected discounted costs, the objective would be to minimize the average costs.

(f) **Formulate an LP-model that could be used to determine an average optimal policy. Also, describe how the average optimal policy may be obtained from the optimal solution of the LP.**

$\min 13\pi_{0,2} + 17\pi_{0,3} + 11\pi_{1,1} + 15\pi_{1,2} + 1\pi_{2,0} + 13\pi_{2,1}$
s.t.
$\pi_{0,2} + \pi_{0,3} = 0.5\pi_{0,2} + 0.5\pi_{1,1} + 0.5\pi_{2,0}$
$\pi_{1,1} + \pi_{1,2} = 0.5\pi_{0,2} + 0.5\pi_{1,1} + 0.5\pi_{2,0} + 0.5\pi_{0,3} + 0.5\pi_{1,2} + 0.5\pi_{2,1}$
$\pi_{2,0} + \pi_{2,1} = 0.5\pi_{0,3} + 0.5\pi_{1,2} + 0.5\pi_{2,1}$
$\pi_{0,2} + \pi_{0,3} + \pi_{1,1} + \pi_{1,2} + \pi_{2,0} + \pi_{2,1} = 1$
$\pi_{0,2},\ \pi_{0,3},\ \pi_{1,1},\ \pi_{1,2},\ \pi_{2,0},\ \pi_{2,1} \geq 0$

There exists an optimal solution with $\forall\, i \in S$ at most one $\pi_{i,d} > 0$. The optimal strategy has $\delta(i) = d$. If $\pi_{i,d} = 0\ \forall\, d \in D(i)$ for some state i, then an arbitrary decision can be chosen in that state.