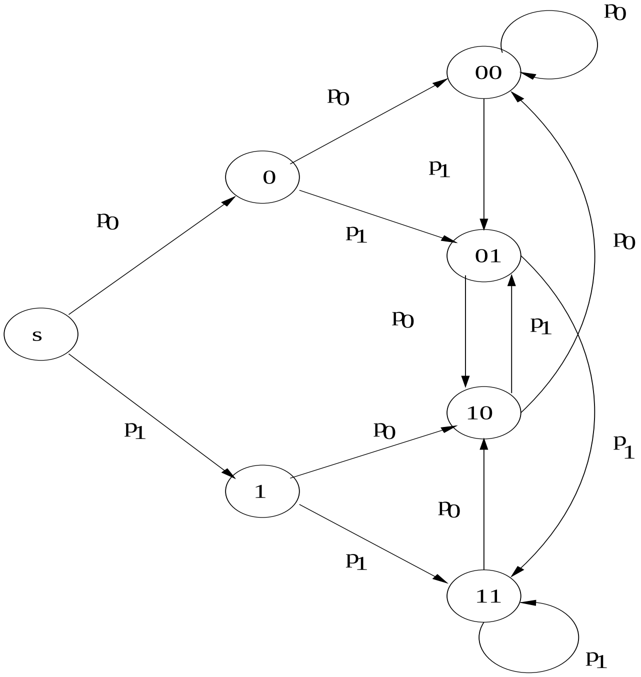
6.262 Discrete Stochastic Processes, Spring 2009
Problem Set 7 — Solutions
due: Wednesday, April 1, 2009

Problem M

a The corresponding chain is the following

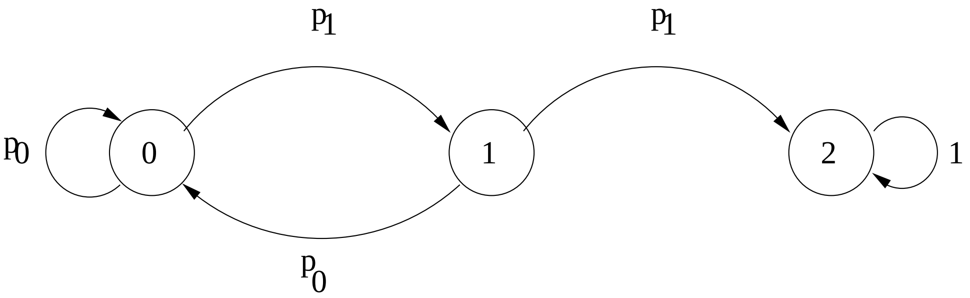


b The meaning of the states 0,1,2 is the following:

0: no symbol yet received (start state) or most recent symbol was 0

1: most recent symbol was 1

2: the pair (1,1) has occurred.



c Consider first the $(0,1)$ case. The simplified chain yields the following system of equations

$$\begin{aligned} v_0 &= 1 + p_1 v_0 + p_0 v_1 \\ v_1 &= 1 + p_0 v_1 + p_1 v_2 \\ v_2 &= 0, \end{aligned}$$

where v_0, v_1, v_2 are the expected first-occurrence times from the respective states $0, 1, 2$. Solving the above equations results in $v_0 = 1/(p_0 p_1)$, as previously obtained using renewal theory. In particular, letting $p_0 = p_1 = 0.5$, the expected time to first occurrence of (0,1) is 4 symbols.

Similarly, for (1,1) we have

$$\begin{aligned} v_0 &= 1 + p_0 v_0 + p_1 v_1 \\ v_1 &= 1 + p_0 v_0 + p_1 v_2 \\ v_2 &= 0, \end{aligned}$$

which yields $v_0 = (1 + p_1)/p_1^2$. When $p_0 = p_1 = 0.5$ the expected time to first occurrence of (1,1) is 6.

The two answers are different when $p_0 = p_1 = 0.5$ because the chains differ at state 1. In particular, although in both chains a '1' leads to a transition into the final state, a '0' pushes the second chain back into the starting state (state 0), while the first chain remains in the intermediary state (state 1). It follows that it takes on average longer to observe a (1,1) than it does to observe a (0,1). (Does this argument hold when $p_0 \neq p_1$?)

Exercise 4.1

a For $r = 1, 2, 3, \ldots$ and any state i, notice that

$$P_{ii}^r \geq P(X_1 = X_2 = \ldots = X_r = i | X_0 = i) = (P_{ii})^r.$$

In other words, the probability of returning to i in r steps is at least the probability of having made r self-transitions. If $P_{ii} > 0$, it follows that $P_{ii}^r > 0$ for all $r \geq 1$.

Recalling that the period of any state j, $d(j)$, is the greatest common divisor of the set $\{n \mid P_{jj}^n > 0\}$, it follows that the period of state i is the greatest common divisor of the set of positive integers. Thus, $d(i) = 1$. By Theorem 4.2, all states in a given class have the same period, which yields that the class containing state i is aperiodic. (Did we need to assume here that the class containing i is recurrent?)

The result of this exercise gives one an easy check for aperiodicity: if any state in a class has a self-loop, then the class is aperiodic.

Exercise 4.5

$[P]$ is doubly stochastic if both $[P]$ and $[P]^T$ are stochastic. This is just a different way of saying that all the rows and all the columns of $[P]$ must sum to one (and all entries are nonnegative). To find the steady state distribution we are looking for the solution of the equation $\pi P = \pi$. Since $[P]^T$ is stochastic, $[P]^T e = e$, where $e = (11\ldots1)^{\mathrm{T}}$. Taking transposes of both sides we get $e^T[P] = e^T$, so e^T is a left eigenvector of $[P]$ associated with $\lambda = 1$. Corollary 4.3 guarantees that for a finite state ergodic chain, this left eigenvector is unique within a scale factor, so scaling e^T to become a probability vector, we obtain $\pi = [1/J, 1/J, ..., 1/J]$.

Exercise 4.6

a First, when $p = 1$, $\pi_0 = \pi_1 = \ldots, \pi_{k-2} = 0$ and $\pi_{k-1} = 1$ since $k-1$ is the only recurrent state. Similarly, for $p = 0$, $\pi_1 = \ldots, \pi_{k-1} = 0$ and $\pi_0 = 1$. For $p \in (0,1)$, the corresponding steady-state equations become

$$\begin{array}{rcl} \pi_0 & = & (1-p)\pi_0 + (1-p)\pi_1 \\ \pi_1 & = & p\pi_0 + (1-p)\pi_2 \\ \pi_2 & = & p\pi_1 + (1-p)\pi_3 \\ \vdots & & \\ \pi_{k-1} & = & p\pi_{k-2} + p\pi_{k-1}. \end{array}$$

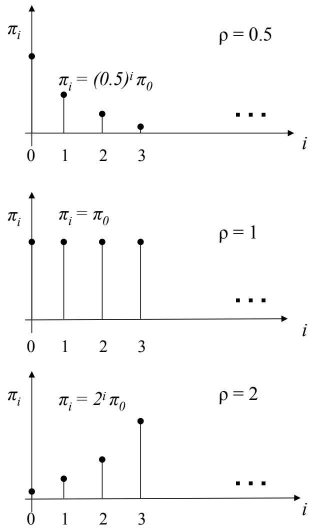
It follows immediately that $\pi_1 = p/(1-p)\pi_0 = \rho\pi_0$. Substituting, we obtain that

$$\pi_i = \rho^i \pi_0, \quad i = 1, \ldots, k-1.$$

Finally, since $\pi_0 + \pi_1 + \ldots + \pi_{k-1} = 1$, it follows that

$$\pi_0 = \frac{1}{1 + \rho + \ldots + \rho^{k-1}} = \frac{1-\rho}{1-\rho^k}.$$

b Sketching the steady-state probabilities for different ρ, we obtain



c If $\rho < 1$, $\lim_{k\to\infty} \pi_i = 0$ for all $i = 1, 2, \ldots$ and $\lim_{k\to\infty} \pi_0 = 1$. If $\rho = 0$, $\lim_{k\to\infty} \pi_i = 0$ for all i, but the distribution remains uniform for all finite k. If $\rho > 1$, $\lim_{k\to\infty} \pi_i = 0$ for all i, but the limiting state escapes out to infinity. (In other words, for large k, the probability mass is concentrated at state $k - 1$.)

Exercise 4.7

a The first Markov chain is symmetrical and hence the state labels can be rotated 4 distinct ways without changing the graph. Thus, all steady state probabilities in the first chain are equal and $\pi_i = 1/4$ for $1 \le i \le 4$. Alternatively, the P matrix, given by

$$P = \begin{pmatrix} 0 & p & 0 & 1-p \\ 1-p & 0 & p & 0 \\ 0 & 1-p & 0 & p \\ p & 0 & 1-p & 0 \end{pmatrix},$$

is doubly stochastic (from Ex. 4.5), so $\pi_i = 1/4$ for $1 \le i \le 4$ is a left eigenvector with unity eigenvalue.

For the second chain, writing out the steady equations, we obtain that $\pi_5 = \pi_4/2$ and $\pi_1 = \pi_4/2$. Furthermore, $\pi_6 = \pi_5, \pi_7 = \pi_6$, etc., so that $\pi_i = \pi_4/2$ for all $i \neq 4, 1 \le i \le 9$. Since $\sum_i \pi_i = 1$, we obtain $\pi_4 = 0.2$ and $\pi_i = 0.1$ for $i \neq 4, 1 \le i \le 9$.

b For the first chain, $P_{11}^2 = 2p(1-p)$; this corresponds to the walks (1,2,1) and (1,4,1) each of which have probability (conditional on starting in 1) of $p(1-p)$. Similarly $P_{ii}^2 = 2p(1-p)$ for $1 \le i \le 4$. Now starting from $X_0 = 1$, after 2 steps X_2 must be either 1 or 3, and $P_{13}^2 = p^2 + (1-p)^2$. Similarly, $P_{31}^2 = P_{24}^2 = P_{42}^2 = p^2 + (1-p)^2$. All other two step transitions, specifically all two step transitions from even to odd or odd to even states, have zero probability.

The first chain, with two step transitions, has two recurrent classes, (1,3) and (2,4). The steady state solution for the first recurrent class is $\pi_1 = \pi_3 = 1/2$ and the steady state solution for the second recurrent class is $\pi_2 = \pi_4 = 1/2$. Letting a be the probability of starting in class (1,3) and $1-a$ in class (2,4), the entire class of steady state solutions is $\pi_1 = \pi_3 = a/2$ and $\pi_2 = \pi_4 = (1-a)1/2$ for any $a, 0 \leq a \leq 1$ (or equivalently letting $b = a/2$ we have $\pi_1 = \pi_3 = b$ and $\pi_2 = \pi_4 = 1/2 - b$ for any $b, 0 \leq b \leq 1/2$). Note that these solutions can also be obtained directly by solving the steady state equations.

For the second chain, $P^2_{46} = P^2_{42} = 1/2$ and $P^2_{68} = P^2_{84} = P^2_{24} = P^2_{57} = P^2_{79} = P^2_{13} = 1$. Finally, $P^2_{95} = P^2_{91} = P^2_{31} = P^2_{35} = 1/2$ and all other two step transitions (including all those from even to odd or odd to even states) have zero probability.

Similarly the second chain has two recurrent classes, the even states and the odd states. The steady state solution for the odd class is $\pi_i = 0.2$ for odd $i, 1 \leq i \leq 9$, and for the even class is $\pi_4 = 0.4$ and $\pi_2 = \pi_6 = \pi_8 = 0.2$. Now, let b be the probability that the chain starts in the even class and $1-b$ the probability that it starts in the odd class. Then, the entire class of steady state solutions is $\pi_4 = (0.4)b, \pi_2 = \pi_6 = \pi_8 = (0.2)b$ and $\pi_1 = \pi_3 = \pi_5 = \pi_7 = \pi_9 = (0.2)(1-b)$, for any $0 \leq b \leq 1$.

The steady state probabilities are not unique because the above chains have multiple recurrent classes and hence depending on where the chain starts these probabilities would differ.

c For each two-step chain, the odd and even class are each recurrent and aperiodic. For the first chain, $\lim_{n\to\infty} P^{2n}_{ij} = 1/2$ for i, j both odd or both even. For chain 2, $\lim_{n\to\infty} P^{2n}_{ij} = 0.2$ for i, j both odd or both even with $j \neq 4$. For $j = 4$ and i even, $\lim_{n\to\infty} P^{2n}_{ij} = 0.4$.

The point here is that for a Markov chain with several recurrent classes, solving the steady-state equations yields averaged steady-state probabilities, weighted by the probability of ending up in each recurrent class. In contrast, the entries of $\lim_{n\to\infty} P^{n}_{ij}$ corresponding to each recurrent class yield correct steady-state probabilities, conditioned on having started in that particular class.

Prof. Wyatt's Detailed Solutions for Problems 6 and 7

Problem 6 (Exercise 5.1)

We are given a Markov chain with states $0 \le i \le J \le \infty$ and the equations

$$F_{ij}(1) = P_{ij}$$

$$F_{ij}(n) = P_{ij} + \sum_{\substack{k=0 \\ k \ne j}}^{J} P_{ik} F_{kj}(n-1), \quad n \ge 2. \tag{5.4}$$

which the quantities $F_{ij}(n)$ must satisfy. We will establish by induction (though the text does this part briefly) that for each $i, j \in [0, \cdots, J]$, the limits $\lim_{n \to \infty} F_{ij}(n) = F_{ij}(\infty)$ exist and satisfy

$$F_{ij}(\infty) = P_{ij} + \sum_{\substack{k=0 \\ k \ne j}}^{J} P_{ik} F_{kj}(\infty). \tag{5.5}$$

While the solutions to eq. (5.4) are unique given the P_{ij}'s, the difficulty with equations (5.5) is that, for any Markov chain and for any choice of $j^* \ge 0$, one solution to (5.5) is $F_{ij^*} = 1, \ \forall \ i \in [0, \cdots, J]$ (even if the all the states are transient, and this solution is therefore definitely wrong). So (5.5) has both the correct values of $F_{ij}(\infty)$ and possible false values as solutions.

We are asked to show that for any choice of $j^* \in [0, \cdots, J]$, the actual values of $F_{ij^*}(\infty)$ constitute the ***smallest*** nonnegative solution vector for (5.5) in the sense that, if $\{x_i, \ \forall i \in [0, \cdots, J]\}$ is also a nonnegative solution to (5.5), i.e., if for some $j^* \ge 0$,

$$x_i = P_{ij^*} + \sum_{\substack{k=0 \\ k \ne j^*}}^{J} P_{ik} x_k, \text{ for all } i \ge 0, \tag{5.1-1}$$

then $F_{ij^*}(\infty) \le x_i, \forall i \in [0, \cdots, J]$. We proceed by induction on n.

n=1

To start the induction, note that for n = 1, since the P_{ij}'s and x's are nonnegative,

$$F_{ij^*}(1) = \mathrm{P}_{\mathrm{ij}^*} \ \le\ \mathrm{P}_{\mathrm{ij}^*} + \sum_{\substack{\mathrm{k}=0\\ \mathrm{k}\neq \mathrm{j}^*}}^{J} P_{ik}x_k = \mathrm{x}_\mathrm{i},\ \forall i \in [0, \cdots, \mathrm{J}]. \tag{5.1-2}$$

n = 2

Similarly, for n = 2, since eq. (5.1-2) shows that $F_{kj^*}(1) \le \mathrm{x}_\mathrm{k}, \forall k \in [0, \cdots, \mathrm{J}]$, we have (using (5.4)) that

$$F_{ij^*}(1) = \mathrm{P}_{\mathrm{ij}^*} \ \le\ F_{ij^*}(2) \ =\ \mathrm{P}_{\mathrm{ij}^*} + \sum_{\substack{\mathrm{k}=0\\ \mathrm{k}\neq \mathrm{j}^*}}^{J} P_{ik}F_{kj^*}(1) \ \le\ \mathrm{P}_{\mathrm{ij}^*} + \sum_{\substack{\mathrm{k}=0\\ \mathrm{k}\neq \mathrm{j}^*}}^{J} P_{ik}x_k \ =\ \mathrm{x}_\mathrm{i},\ \forall i \in [0, \cdots, \mathrm{J}] \tag{5.1-3}$$

This establishes that $F_{ij^*}(2) \le \mathrm{x}_{\mathrm{i},}\ \forall i, j^* \in [0, \cdots, \mathrm{J}]$

Induction step on n

We now assume, for any choice of $n \ge 2$ and $\mathrm{j}^* \in [0, \cdots, \mathrm{J}]$, that $F_{ij^*}(n) \le \mathrm{x}_{\mathrm{i},}\ \forall i \in [0, \cdots, \mathrm{J}]$, as shown above for n = 2. Since $F_{ij^*}(n) = F_{ij^*}(n+1) - f_{ij^*}(n+1) \le F_{ij^*}(n+1)$, we have

$$F_{ij^*}(n) \ \le\ \mathrm{F}_{\mathrm{ij}^*}(\mathrm{n}{+}1) \ =\ \mathrm{P}_{\mathrm{ij}^*} + \sum_{\substack{\mathrm{k}=0\\ \mathrm{k}\neq \mathrm{j}^*}}^{J} P_{ik}\mathrm{F}_{\mathrm{ik}^*}(\mathrm{n}) \le \mathrm{P}_{\mathrm{ij}^*} + \sum_{\substack{\mathrm{k}=0\\ \mathrm{k}\neq \mathrm{j}^*}}^{J} P_{ik}x_k \ =\ \mathrm{x}_\mathrm{i},\ \forall i \in [0, \cdots, \mathrm{J}]. \tag{5.1-4}$$

Since one set of values for $\{x_i,\ \forall i \in [0, \cdots, \mathrm{J}]\}$ that satisfies eq. (1) is $\{x_i = 1, \forall i \in [0, \cdots, \mathrm{J}]\}$, we know that for any choice of j^* and any value of i, the sequence $F_{ij^*}(n)$, n = 1, 2, $\cdots$ is monotone nondecreasing, and bounded above by 1, so $\lim_{\mathrm{n}\to\infty} F_{ij}(n)$ exists. Let $\mathrm{F}_{\mathrm{ij}^*}(\infty)$ denote $\lim_{\mathrm{n}\to\infty} F_{ij^*}(n)$. Since $F_{ij^*}(n) \le \mathrm{x}_{\mathrm{i},}\ \forall \mathrm{n} > 0, \forall i \in [0, \cdots, \mathrm{J}]$, we conclude that $\mathrm{F}_{\mathrm{ij}^*}(\infty) \le \mathrm{x}_\mathrm{i}, \forall i, j^* \ge 0$.

Ñ

Problem 7 **(Exercise 5.2)**

a) We want to show that the probabilities of ever passing to (or, for i = 0, returning to) state 0 from any initial state i are

$$\begin{aligned} &F_{00}(\infty) = 2(1-p), \\ &F_{i0}(\infty) = [{(1-p)}/{p}]^i,\ \mathrm{i} \ \ge\ 1 \end{aligned} \tag{5.2-1}$$

for any $p \geq 1/2$. (This guarantees that state 0 is transient for p > ½ and recurrent for p = ½. By Theorem 5.1, it further guarantees that the entire chain is transient for ½ < p <1 and that the entire chain is recurrent for p = ½. The theorem tells us nothing for p = 1 because the chain does not form a single class in that case, but it is obviously transient when p = 1 as well.)

From the results of Exercise 5.1, we know that to verify the proposed solutions (5.2-1) are correct we have to verify that (i) these values satisfy eq. (5.5), and (ii) they are the smallest solution in the sense that every nonnegative set of numbers $\{x_{i,}\ 0 \leq i\}$ that satisfies (5.5) has $\mathrm{F}_{i0}(\infty) \leq x_i$ for any choice of $i \geq 0$. Since this is a birth-death Markov chain, eqs. (5.5) take the form:

$$\begin{aligned}
&F_{00}(\infty) = (1-p) + pF_{10}(\infty)\\
&F_{10}(\infty) = (1-p) + pF_{20}(\infty)\\
&\vdots\\
&F_{i0}(\infty) = (1-p)F_{(i-1)0}(\infty) + pF_{(i+1)0}(\infty),\ \mathrm{i} \geq 2. && (5.2\text{-}2)
\end{aligned}$$

We check by substitution that the solutions (5.2-1) satisfy equations (5.2-2):

$$\begin{aligned}
&\text{for i = 0, } 2(1-p) = \text{(1-p) + p } [\text{(1-p)}/\text{p}]\\
&\text{for i = 1, } [\text{(1-p)}/\text{p}] = \text{(1-p) + p } [\text{(1-p)}/\text{p}]^2\\
&\vdots\\
&\text{for i} \geq 2,\ [(1-p)/p]^i = \text{(1-p)}[(1-p)/p]^{i-1} + \text{p}[(1-p)/p]^{i+1}. && (5.2\text{-}3)
\end{aligned}$$

To check that (5.2-1) gives the smallest nonnegative solution set for (5.2-2), note first that if p = 1, the solutions (5.2-1) are all zero and are therefore certainly the smallest. Next we will show that for $1/2 \leq \mathrm{p} < 1$, every solution $\{x_i,\ \mathrm{i} \geq 0\}$ of (5.2-2) (with x_i substituted for $F_{i0}(\infty)$) can be characterized as a strictly monotone increasing function of the value chosen for x_0. We solve (5.2-2) for x_1 and x_2:

$$x_1 = \frac{x_0 - (1-p)}{p}$$

$$x_2 = \frac{x_1 - (1-p)}{p} = \frac{\dfrac{x_0 - (1-p)}{p} - (1-p)}{p} = \frac{x_0 - (1-p^2)}{p^2}.$$

If we perturb x_0 from the value 2(1-p) in (5.2-1) to another value $x_0 + \varepsilon$ = 2(1-p) + ε, the resulting changes in x_1 and x_2 are

$$\delta \mathrm{x}_1 = \frac{\varepsilon}{p}$$

$$\delta \mathrm{x}_2 = \frac{\varepsilon}{\mathrm{p}^2} \;, \tag{5.2-4}$$

We note that for $\varepsilon < 0$ and $1/2 \le \mathrm{p} < 1$, $\delta \mathrm{x}_2 < \delta \mathrm{x}_1 < 0$. To establish by induction that for $\varepsilon < 0$ and $1/2 \le \mathrm{p} < 1$, $\delta \mathrm{x}_{(\mathrm{i}+1)} < \delta \mathrm{x}_\mathrm{i} < 0,\ \mathrm{i} \ge 0,$ we solve (5.2-2) (with x_i substituted for $F_{i0}(\infty)$):

$$\mathrm{x}_{(\mathrm{i}+1)} = \frac{x_i - (1-p)x_{(i-1)}}{p}, \ \mathrm{i} \ \ge \ 2, \tag{5.2-5}$$

and assume $\delta x_i < \delta x_{(i-1)} < 0$. Then for p < 1,

$$\delta \mathrm{x}_{(\mathrm{i}+1)} = \frac{\delta x_i - (1-p)\delta x_{(i-1)}}{p} < \frac{\delta x_i - (1-p)\delta x_i}{p} = \delta x_i < 0. \tag{5.2-6}$$

Since for p > ½, $\lim_{i\to\infty} F_{i0}(\infty) = 0$, and since δx_i grows more negative with i for $\varepsilon < 0$, we see that ***decreasing*** x_0 below 2(1-p) by setting $\varepsilon < 0$ causes x_i to eventually become negative, so the solution (5.2-1) for which x_0 = 2(1-p) is the ***smallest nonnegative solution*** for $1/2 < \mathrm{p} \ \le \ 1$. For p = 1/2, $F_{i0}(\infty) = 1, \forall \mathrm{i} \ \ge \ 0$, so a stronger growth condition on the negative $\delta x_i\text{'}s$ is needed. For p = ½, eqs. (5.2-4) and (5.2-5) take the specific form

$$\delta \mathrm{x}_1 = 2\varepsilon$$

$$\delta \mathrm{x}_2 = 4\varepsilon = 2\delta \mathrm{x}_1$$

$$\delta \mathrm{x}_{(\mathrm{i}+1)} = \frac{\delta x_i - (1/2)\delta x_{(i-1)}}{1/2} = 2\delta \mathrm{x}_\mathrm{i} - \delta \mathrm{x}_{(\mathrm{i}\text{-}1)}, \ \mathrm{i} \ \ge \ 2, \tag{5.2-7}$$

These equations have the solution

$$\delta x_i = i\delta x_1, \ \forall \mathrm{i} \ \ge \ 1.$$

Therefore, for p = ½, a perturbation of x_0 from the value 2(1-p) in (5.2-10 to a smaller value 2(1-p) - ε causes a perturbation in x_i from 1 to 1 - $2i\varepsilon$, which is negative for all $i \ \ge \ (1/2\varepsilon)$. Therefore the solution (5.2-1) with x_0 = 2(1-p) is the ***smallest nonnegative solution*** for p = 1/2 as well.

Using the results from Problem 5.1, this guarantees that (5.2-1) gives the correct solution for $F_{i0}(\infty)$, $\forall i \geq 0$, for $1/2 \leq p \leq 1$. Ñ

b) Starting in any given state $j > 0$ and considering some state $i < j$, truncating the chain at state j in order to obtain an equivalent finite-state chain yields a single recurrent class. It follows that there is an eventual return to state j and $F_{ij}(\infty) = 1$.
If on the other hand $i > j$, we can consider a truncated chain with states $0, 1, \ldots, j-1$ removed. Noticing that the probabilistic description of the chain does not change by making j into a starting state conditioned on having started in $i > j$, we obtain $F_{j+i,j}(\infty) = \left(\frac{1-p}{p}\right)^i$.
Finally, from state j, the first transition goes to $j-1$ with probability $1-p$ and to $j+1$ with probability p. If the first transition from j is to $j-1$, considering again a finite-state chain truncated at state j shows that there is an eventual return to state j. If, on the other hand, the first transition from j is to $j+1$, looking at the first return to j, we can consider a truncated chain with states $0, 1, \ldots, j-1$ removed. Noticing that the probabilistic description of the chain does not change by making j into a starting state conditioned on having transitioned to $j+1$, we obtain $F_{j+1,j}(\infty) = \frac{1-p}{p}$. Putting the two cases together, $F_{jj}(\infty) = (1-p) + p\frac{1-p}{p} = 2\,(1-p) = F_{00}(\infty)$.
Ñ

Since the equivalent finite-state chain with K states has a single recurrent class, we have that $F_{ij}^{(K)}(\infty) = 1$ for all i, j. Note that $\lim_{K\to\infty} F_{ij}^{(K)} = 1$ for all i, j, which does not equal $F_{ij}(\infty)$ for the countable birth-death chain.
Point being: beware of describing the behavior of a countable chain by a limiting behavior of a truncated chain, as the limit need not be continuous.

For the case $p = 1/2$, we have $F_{i0}(\infty) = 1$ for all i. In other words, no matter how far from 0 we start, we return to it eventually with probability 1, even if it takes a very long time. Letting state i be associated with having lost a total of $i-1$dollars at some point in the game and state 0 be associated with being 1 dollar ahead, we see that eventually (i.e. with probability 1), we will be 1 dollar ahead no matter where we started.