Chapter 6

MARKOV PROCESSES WITH COUNTABLE STATE SPACES

6.1 Introduction

Recall that a Markov chain is a discrete-time process $\{X_n;\, n \geq 0\}$ for which the state at each time $n \geq 1$ is an integer-valued random variable (rv) that is statistically dependent on $X_0, \dots X_{n-1}$ only through X_{n-1}. A *countable-state Markov process*[1] (Markov process for short) is a generalization of a Markov chain in the sense that, along with the Markov chain $\{X_n;\, n \geq 1\}$, there is a randomly-varying holding interval in each state which is exponentially distributed with a parameter determined by the current state.

To be more specific, let $X_0{=}i,\ X_1{=}j,\ X_2 = k, \dots$, denote a sample path of the sequence of states in the Markov chain (henceforth called the *embedded Markov chain*). Then the *holding interval* U_n between the time that state $X_{n-1} = \ell$ is entered and X_n is entered is a nonnegative exponential rv with parameter ν_ℓ, *i.e.*, for all $u \geq 0$,

$$\Pr\{U_n \leq u \mid X_{n-1} = \ell\} = 1 - \exp(-\nu_\ell u). \tag{6.1}$$

Furthermore, U_n, conditional on X_{n-1}, is jointly independent of X_m for all $m \neq n-1$ and of U_m for all $m \neq n$.

If we visualize starting this process at time 0 in state $X_0 = i$, then the first transition of the embedded Markov chain enters state $X_1 = j$ with the transition probability P_{ij} of the embedded chain. This transition occurs at time U_1, where U_1 is independent of X_1 and exponential with rate ν_i. Next, conditional on $X_1 = j$, the next transition enters state $X_2 = k$ with the transition probability P_{jk}. This transition occurs after an interval U_2, *i.e.*, at time $U_1 + U_2$, where U_2 is independent of X_2 and exponential with rate ν_j. Subsequent transitions occur similarly, with the new state, say $X_n = i$, determined from the old state, say $X_{n-1} = \ell$, via $P_{\ell i}$, and the new holding interval U_n determined via the exponential rate ν_ℓ. Figure 6.1 illustrates the statistical dependencies between the rv's $\{X_n;\, n \geq 0\}$ and $\{U_n;\, n \geq 1\}$.

[1]These processes are often called *continuous-time* Markov chains.

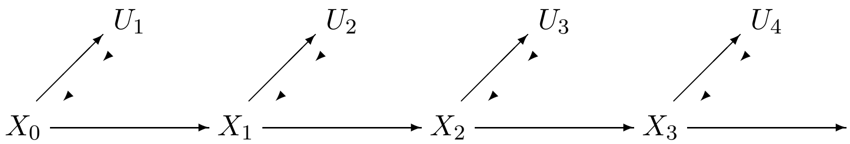


Figure 6.1: The statistical dependencies between the rv's of a Markov process. Each holding interval U_i, conditional on the current state X_{i-1}, is independent of all other states and holding intervals.

The epochs at which successive transitions occur are denoted $S_1, S_2, \dots$, so we have $S_1 = U_1$, $S_2 = U_1 + U_2$, and in general $S_n = \sum_{m=1}^{n} U_m$ for $n \geq 1$ wug $S_0 = 0$. The state of a Markov process at any time $t > 0$ is denoted by $X(t)$ and is given by

$$X(t) = X_n \quad \text{for } S_n \leq t < S_{n+1} \qquad \text{for each } n \geq 0.$$

This defines a stochastic process $\{X(t);\ t \geq 0\}$ in the sense that each sample point $\omega \in \Omega$ maps to a sequence of sample values of $\{X_n;\ n \geq 0\}$ and $\{S_n;\ n \geq 1\}$, and thus into a sample function of $\{X(t);\ t \geq 0\}$. This stochastic process is what is usually referred to as a Markov process, but it is often simpler to view $\{X_n;\ n \geq 0\}, \{S_n;\ n \geq 1\}$ as a characterization of the process. Figure 6.2 illustrates the relationship between all these quantities.

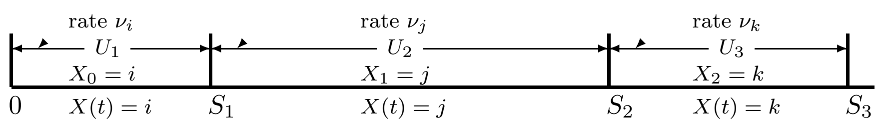


Figure 6.2: The relationship of the holding intervals $\{U_n; n \geq 1\}$ and the epochs $\{S_n; n \geq 1\}$ at which state changes occur. The state $X(t)$ of the Markov process and the corresponding state of the embedded Markov chain are also illustrated. Note that if $X_n = i$, then $X(t) = i$ for $S_n \leq t < S_{n+1}$

This can be summarized in the following definition.

Definition 6.1.1. *A countable-state Markov process $\{X(t);\ t \geq 0\}$ is a stochastic process mapping each nonnegative real number t to the nonnegative integer-valued rv $X(t)$ in such a way that for each $t \geq 0$,*

$$X(t) = X_n \quad \textit{for } S_n \leq t < S_{n+1}; \quad S_0 = 0;\ S_n = \sum_{m=1}^{n} \} U_m \quad \textit{for } n \geq 1, \tag{6.2}$$

where $\{X_n;\ n \geq 0\}$ is a Markov chain with a countably infinite or finite state space and each U_n, given $X_{n-1} = i$, is exponential with rate $\nu_i > 0$ and is conditionally independent of all other U_m and X_m.

The tacit assumptions that the state space is the set of nonnegative integers and that the process starts at $t = 0$ are taken only for notational simplicity but will serve our needs here.

We assume throughout this chapter (except in a few places where specified otherwise) that the embedded Markov chain has no self transitions, *i.e.*, $P_{ii} = 0$ for all states i. One reason for this is that such transitions are invisible in $\{X(t);\, t \geq 0\}$. Another is that with this assumption, the sample functions of $\{X(t);\, t \geq 0\}$ and the joint sample functions of $\{X_n;\, n \geq 0\}$ and $\{U_n;\, n \geq 1\}$ uniquely specify each other.

We are not interested for the moment in exploring the probability distribution of $X(t)$ for given values of t, but one important feature of this distribution is that for any times $t > \tau > 0$ and any states i, j,

$$\Pr\{X(t){=}j \mid X(\tau){=}i, \{X(s) = x(s);\, s < \tau\}\} = \Pr\{X(t{-}\tau){=}j \mid X(0){=}i\}\,. \tag{6.3}$$

This property arises because of the memoryless property of the exponential distribution. That is, if $X(\tau) = i$, it makes no difference how long the process has been in state i before τ; the time to the next transition is still exponential with rate ν_i and subsequent states and holding intervals are determined as if the process starts in state i at time 0. This will be seen more clearly in the following exposition. This property is the reason why these processes are called Markov, and is often taken as the defining property of Markov processes.

Example 6.1.1. The M/M/1 queue: An M/M/1 queue has Poisson arrivals at a rate denoted by λ and has a single server with an exponential service distribution of rate $\mu > \lambda$ (see Figure 6.3). Successive service times are independent, both of each other and of arrivals. The state $X(t)$ of the queue is the total number of customers either in the queue or in service. When $X(t) = 0$, the time to the next transition is the time until the next arrival, *i.e.*, $\nu_0 = \lambda$. When $X(t) = i$, $i \geq 1$, the server is busy and the time to the next transition is the time until either a new arrival occurs or a departure occurs. Thus $\nu_i = \lambda + \mu$. For the embedded Markov chain, $P_{01} = 1$ since only arrivals are possible in state 0, and they increase the state to 1. In the other states, $P_{i,i-1} = \mu/(\lambda{+}\mu)$ and $P_{i,i+1} = \lambda/(\lambda{+}\mu)$.

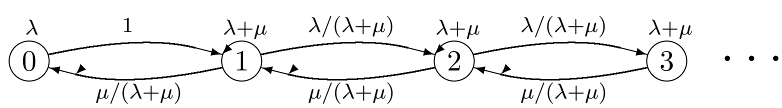


Figure 6.3: The embedded Markov chain for an M/M/1 queue. Each node i is labeled with the corresponding rate ν_i of the exponentially distributed holding interval to the next transition. Each transition, say i to j, is labeled with the corresponding transition probability P_{ij} in the embedded Markov chain.

The embedded Markov chain is a Birth-death chain, and its steady state probabilities can be calculated easily using (5.25). The result is

$$\pi_0 = \frac{1-\rho}{2} \qquad \text{where } \rho = \frac{\lambda}{\mu}$$
$$\pi_n = \frac{1-\rho^2}{2}\rho^{n-1} \qquad \text{for } n \geq 1. \tag{6.4}$$

Note that if $\lambda << \mu$, then π_0 and π_1 are each close to 1/2 (*i.e.*, the embedded chain mostly alternates between states 0 and 1, and higher ordered states are rarely entered), whereas because of the large holding interval in state 0, the process spends most of its time in state 0 waiting for arrivals. The steady-state probability π_i of state i in the embedded chain is the long-term fraction of the total transitions that go to state i. We will shortly learn how to find the long term *fraction of time* spent in state i as opposed to this fraction of transitions, but for now we return to the general study of Markov processes.

The evolution of a Markov process can be visualized in several ways. We have already looked at the first, in which for each state $X_{n-1} = i$ in the embedded chain, the next state X_n is determined by the probabilities $\{P_{ij};\ j \geq 0\}$ of the embedded Markov chain, and the holding interval U_n is independently determined by the exponential distribution with rate ν_i.

For a second viewpoint, suppose an independent Poisson process of rate $\nu_i > 0$ is associated with each state i. When the Markov process enters a given state i, the next transition occurs at the next arrival epoch in the Poisson process for state i. At that epoch, a new state is chosen according to the transition probabilities P_{ij}. Since the choice of next state, given state i, is independent of the interval in state i, this view describes the same process as the first view.

For a third visualization, suppose, for each pair of states i and j, that an independent Poisson process of rate $\nu_i P_{ij}$ is associated with a possible transition to j conditional on being in i. When the Markov process enters a given state i, both the time of the next transition and the choice of the next state are determined by the set of i to j Poisson processes over all possible next states j. The transition occurs at the epoch of the first arrival, for the given i, to any of the i to j processes, and the next state is the j for which that first arrival occurred. Since such a collection of independent Poisson processes is equivalent to a single process of rate ν_i followed by an independent selection according to the transition probabilities P_{ij}, this view again describes the same process as the other views.

It is convenient in this third visualization to define the rate from any state i to any other state j as

$$q_{ij} = \nu_i P_{ij}.$$

If we sum over j, we see that ν_i and P_{ij} are also uniquely determined by $\{q_{ij};\ i,j \geq 0\}$ as

$$\nu_i = \sum_j q_{ij}; \qquad P_{ij} = q_{ij}/\nu_i. \tag{6.5}$$

This means that the fundamental characterization of the Markov process in terms of the P_{ij} and the ν_i can be replaced by a characterization in terms of the set of transition rates q_{ij}. In many cases, this is a more natural approach. For the M/M/1 queue, for example, $q_{i,i+1}$ is simply the arrival rate λ. Similarly, $q_{i,i-1}$ is the departure rate μ when there are customers to be served, *i.e.*, when $i > 0$. Figure 6.4 shows Figure 6.3 incorporating this notational simplification.

Note that the interarrival density for the Poisson process, from any given state i to other state j, is given by $q_{ij} \exp(-q_{ij}x)$. On the other hand, given that the process is in state i, the probability density for the interval until the next transition, whether conditioned on the next state or not, is $\nu_i \exp(-\nu_i x)$ where $\nu_i = \sum_j q_{ij}$. One might argue, incorrectly, that, conditional on the next transition being to state j, the time to that transition has density $q_{ij} \exp(-q_{ij}x)$. Exercise 6.1 uses an M/M/1 queue to provide a guided explanation of why this argument is incorrect.

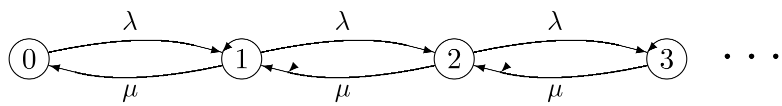


Figure 6.4: The Markov process for an M/M/1 queue. Each transition (i, j) is labelled with the corresponding transition rate q_{ij}.

6.1.1 The sampled-time approximation to a Markov process

As yet another way to visualize a Markov process, consider approximating the process by viewing it only at times separated by a given increment size δ. The Poisson processes above are then approximated by Bernoulli processes where the transition probability from i to j in the sampled-time chain is defined to be $q_{ij}\delta$ for all $j \neq i$.

The Markov process is then approximated by a Markov chain. Since each δq_{ij} decreases with decreasing δ, there is an increasing probability of no transition out of any given state in the time increment δ. These must be modeled with self-transition probabilities, say $P_{ii}(\delta)$ which must satisfy

$$P_{ii}(\delta) = 1 - \sum_j q_{ij}\delta = 1 - \nu_i\delta \qquad \text{for each } i \geq 0.$$

This is illustrated in Figure 6.5 for the M/M/1 queue. Recall that this sampled-time M/M/1 Markov chain was analyzed in Section 5.4 and the steady-state probabilities were shown to be

$$p_i(\delta) = (1-\rho)\rho^i \qquad \text{for all } i \geq 0 \quad \text{where } \rho = \lambda/\mu. \tag{6.6}$$

We have denoted the steady-state probabilities here by $p_i(\delta)$ to avoid confusion with the steady-state probabilities for the embedded chain. As discussed later, the steady-state probabilities in (6.6) do not depend on δ, so long as δ is small enough that the self-transition probabilities are nonnegative.

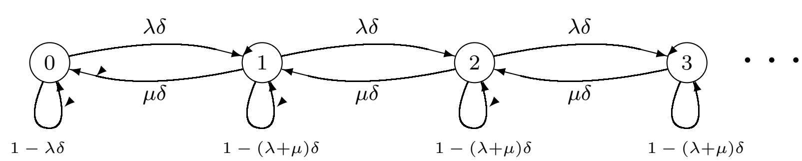


Figure 6.5: Approximating an M/M/1 queue by a sampled-time Markov chain.

This sampled-time approximation is an approximation in two ways. First, transitions occur only at integer multiples of the increment δ, and second, $q_{ij}\delta$ is an approximation to $\Pr\{X(\delta)=j \mid X(0)=i\}$. From (6.3), $\Pr\{X(\delta)=j \mid X(0) = i\} = q_{ij}\delta + o(\delta)$, so this second approximation is increasingly good as $\delta \to 0$.

As observed above, the steady-state probabilities for the sampled-time approximation to an M/M/1 queue do not depend on δ. As seen later, whenever the embedded chain is positive recurrent and a sampled-time approximation exists for a Markov process, then the steady-state probability of each state i is independent of δ and represents the limiting fraction of time spent in state i with probability 1.

Figure 6.6 illustrates the sampled-time approximation of a generic Markov process. Note that $P_{ii}(\delta)$ is equal to $1 - \delta\nu_i$ for each i in any such approximation, and thus it is necessary for δ to be small enough to satisfy $\nu_i\delta \leq 1$ for all i. For a finite state space, this is satisfied for any $\delta \leq [\max_i \nu_i]^{-1}$. For a countably infinite state space, however, the sampled-time approximation requires the existence of some finite B such that $\nu_i \leq B$ for all i. The consequences of having no such bound are explored in the next section.

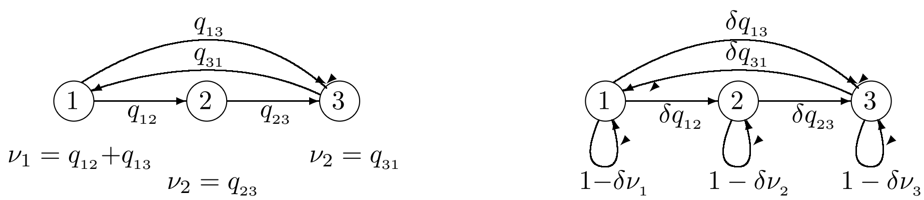


Figure 6.6: Approximating a generic Markov process by its sampled-time Markov chain.

6.2 Steady-state behavior of irreducible Markov processes

Definition 6.2.1 (Irreducible Markov processes). *An irreducible Markov process is a Markov process for which the embedded Markov chain is irreducible (*i.e.*, all states are in the same class).*

The analysis in this chapter is restricted almost entirely to irreducible Markov processes. The reason for this restriction is not that Markov processes with multiple classes of states are unimportant, but rather that they can usually be best understood by first looking at the embedded Markov chains for the various classes making up that overall chain.

We will ask the same types of steady-state questions for Markov processes as we asked about Markov chains. In particular, under what conditions is there a set of steady-state probabilities, $p_0, p_1, \dots$ with the property that for any given starting state i, the limiting fraction of of time spent in any given state j is p_j with probability 1? Do these probabilities also have the property that $p_j = \lim_{t\to\infty} \Pr\{X(t) = j \mid X_0 = i\}$?

We will find that simply having a positive-recurrent embedded Markov chain is not quite enough to ensure that such a set of probabilities exists. It is also necessary for the embedded-chain steady-state probabilities $\{\pi_i;\, i \geq 0\}$ and the holding-interval parameters $\{\nu_i; i \geq 0\}$ to satisfy $\sum_i \pi_i/\nu_i < \infty$. We will interpret this latter condition as asserting that the limiting long-term rate at which transitions occur must be strictly positive. Finally we will show that when these conditions are satisfied, the steady-state probabilities for the process are related to those of the embedded chain by

$$p_j = \frac{\pi_j/\nu_j}{\sum_k \pi_k/\nu_k}. \tag{6.7}$$

Definition 6.2.2. *The steady-state process probabilities, $p_0, p_1, \dots$ for a Markov process are a set of numbers satisfying (6.7), where $\{\pi_i;\ i \geq 0\}$ and $\{\nu_i;\ i \geq 0\}$ are the steady-state probabilities for the embedded chain and the holding-interval rates respectively.*

As one might guess, the appropriate approach to answering these questions comes from applying renewal theory to various renewal processes associated with the Markov process. Many of the needed results for this have already been developed in looking at the steady-state behavior of countable-state Markov chains.

We start with a very technical lemma that will perhaps appear obvious, and the reader is welcome to ignore the proof until perhaps questioning the issue later. The lemma is not restricted to irreducible processes, although we only use it in that case.

Lemma 6.2.1. *Consider a Markov process for which the embedded chain starts in some given state i. Then the holding time intervals, $U_1, U_2, \dots$ are all rv's. Let $M_i(t)$ be the number of transitions made by the process up to and including time t. Then with probability 1 (WP1),*

$$\lim_{t\to\infty} M_i(t) = \infty. \tag{6.8}$$

Proof: The first holding interval U_1 is exponential with rate $\nu_i > 0$, so it is clearly a rv (*i.e.*, not defective). In general the state after the $(n-1)$th transition has the PMF P_{ij}^{n-1}, so the complementary distribution function of U_n is

$$\begin{aligned}
\Pr\{U_n > u\} &= \lim_{k\to\infty} \sum_{j=1}^{k} P_{ij}^{n-1} \exp(-\nu_j u). \\
&\leq \sum_{j=1}^{k} P_{ij}^{n-1} \exp(-\nu_j u) + \sum_{j=k+1}^{\infty} P_{ij}^{n+1} \qquad \text{for every } k.
\end{aligned}$$

For each k, the first sum above approaches 0 with increasing u and the second sum approaches 0 with increasing k so the limit as $u \to \infty$ must be 0 and U_n is a rv.

It follows that each $S_n = U_1 + \cdots + U_n$ is also a rv. Now $\{S_n;\ n \geq 1\}$ is the sequence of arrival epochs in an arrival process, so we have the set equality $\{S_n \leq t\} = \{M_i(t) \geq n\}$ for each choice of n. Since S_n is a rv, we have $\lim_{t\to\infty} \Pr\{S_n \leq t\} = 1$ for each n. Thus $\lim_{t\to\infty} \Pr\{M_i(t) \geq n\} = 1$ for all n. This means that the set of sample points ω for which $\lim_{t\to\infty} M_i(t,\omega) < n$ has probability 0 for all n, and thus $\lim_{t\to\infty} M_i(t,\omega) = \infty$ WP1. □

6.2.1 Renewals on successive entries to a given state

For an irreducible Markov process with $X_0 = i$, let $M_{ij}(t)$ be the number of transitions into state j over the interval $(0, t]$. We want to find when this is a delayed renewal counting process. It is clear that the sequence of epochs at which state j is entered form renewal points, since they form renewal points in the embedded Markov chain and the holding intervals between transitions depend only on the current state. The questions are whether the first entry to state j must occur within some finite time, and whether recurrences to j must occur within finite time. The following lemma answers these questions for the case where the embedded chain is recurrent (either positive recurrent or null recurrent).

Lemma 6.2.2. *Consider a Markov process with an irreducible recurrent embedded chain* $\{X_n;\, n \geq 0\}$. *Given* $X_0 = i$, *let* $\{M_{ij}(t);\, t \geq 0\}$ *be the number of transitions into a given state* j *in the interval* $(0, t]$. *Then* $\{M_{ij}(t);\, t \geq 0\}$ *is a delayed renewal counting process (or an ordinary renewal counting process if* $j = i$*).*

Proof: Given $X_0 = i$, let $N_{ij}(n)$ be the number of transitions into state j that occur in the embedded Markov chain by the n^{th} transition of the embedded chain. From Lemma 5.1.4, $\{N_{ij}(n);\, n \geq 0\}$ is a delayed renewal process, so from Lemma 4.8.2, $\lim_{n\to\infty} N_{ij}(n) = \infty$ with probability 1. Note that $M_{ij}(t) = N_{ij}(M_i(t))$, where $M_i(t)$ is the total number of state transitions (between all states) in the interval $(0, t]$. Thus, with probability 1,

$$\lim_{t\to\infty} M_{ij}(t) = \lim_{t\to\infty} N_{ij}(M_i(t)) = \lim_{n\to\infty} N_{ij}(n) = \infty.$$

where we have used Lemma 6.2.1, which asserts that $\lim_{t\to\infty} M_i(t) = \infty$ with probability 1.

It follows that the interval W_1 until the first transition to state j, and the subsequent interval W_2 until the next transition to state j, are both finite with probability 1. Subsequent intervals have the same distribution as W_2, and all intervals are independent, so $\{M_{ij}(t);\, t \geq 0\}$ is a delayed renewal process with inter-renewal intervals $\{W_k;\, k \geq 1\}$. If $i = j$, then all W_k are identically distributed and we have an ordinary renewal process, completing the proof. □

The inter-renewal intervals $W_2, W_3, \dots$ for $\{M_{ij}(t);\, t \geq 0\}$ above are well-defined nonnegative IID rv's whose distribution depends on j but not i. They either have an expectation as a finite number or have an infinite expectation. In either case, this expectation is denoted as $\mathsf{E}[W(j)] = \overline{W}(j)$. This is the mean time between successive entries to state j, and we will see later that in some cases this mean time can be infinite.

6.2.2 The limiting fraction of time in each state

In order to study the fraction of time spent in state j, we define a delayed renewal-reward process, based on $\{M_{ij}(t);\, t \geq 0\}$, for which unit reward is accumulated whenever the process is in state j. That is (given $X_0 = i$), $R_{ij}(t) = 1$ when $X(t) = j$ and $R_{ij}(t) = 0$ otherwise. (see Figure 6.7). Given that transition $n-1$ of the embedded chain enters state j, the interval U_n is exponential with rate ν_j, so $\mathsf{E}[U_n \mid X_{n-1}{=}j] = 1/\nu_j$.

Let $p_j(i)$ be the limiting time-average fraction of time spent in state j. We will see later that such a limit exists WP1, that the limit does not depend on i, and that it is equal to

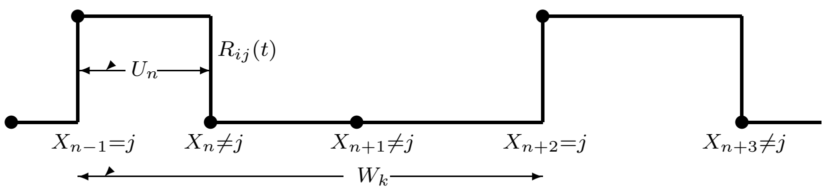


Figure 6.7: The delayed renewal-reward process $\{R_{ij}(t);\, t \geq 0\}$ for time in state j. The reward is one whenever the process is in state j, *i.e.*, $R_{ij}(t) = \mathbb{I}_{\{X(t)=j\}}$. A renewal occurs on each entry to state j, so the reward starts at each such entry and continues until a state transition, assumed to enter a state other than j. The reward then ceases until the next renewal, *i.e.*, the next entry to state j. The figure illustrates the kth inter-renewal interval, of duration W_k, which is assumed to start on the $n-1$st state transition. The expected interval over which a reward is accumulated is ν_j and the expected duration of the inter-renewal interval is $\overline{W}(j)$.

the steady-state probability p_j in (6.7). Since $\overline{U}(j) = 1/\nu_j$, Theorems 4.4.1 and 4.8.4, for ordinary and delayed renewal-reward processes respectively, state that[2]

$$p_j(i) = \lim_{t\to\infty} \frac{\int_0^t R_{ij}(\tau)d\tau}{t} \qquad \text{WP1} \tag{6.9}$$

$$= \frac{\overline{U}(j)}{\overline{W}(j)} = \frac{1}{\nu_j \overline{W}(j)}. \tag{6.10}$$

This shows that the limiting time average, $p_j(i)$, exists with probability 1 and is independent of the starting state i. We show later that it is the steady-state process probability given by (6.7). We can also investigate the limit, as $t \to \infty$, of the probability that $X(t) = j$. This is equal to $\lim_{t\to\infty} \mathsf{E}[R(t)]$ for the renewal-reward process above. Because of the exponential holding intervals, the inter-renewal times are non-arithmetic, and from Blackwell's theorem, in the form of (4.105),

$$\lim_{t\to\infty} \Pr\{X(t) = j\} = \frac{1}{\nu_j \overline{W}(j)} = p_j(i). \tag{6.11}$$

We summarize these results in the following lemma.

Lemma 6.2.3. *Consider an irreducible Markov process with a recurrent embedded Markov chain starting in $X_0 = i$. Then with probability 1, the limiting time average in state j is given by $p_j(i) = \frac{1}{\nu_j \overline{W}(j)}$. This is also the limit, as $t \to \infty$, of $\Pr\{X(t) = j\}$.*

6.2.3 Finding $\{p_j(i);\, j \geq 0\}$ in terms of $\{\pi_j;\, j \geq 0\}$

Next we must express the mean inter-renewal time, $\overline{W}(j)$, in terms of more accessible quantities that allow us to show that $p_j(i) = p_j$ where p_j is the steady-state process probability

[2] Theorems 4.4.1 and 4.8.4 do not cover the case where $\overline{W}(j) = \infty$, but, since the expected reward per renewal interval is finite, it is not hard to verify (6.9) in that special case.

of Definition 6.2.2. We now assume that the embedded chain is not only recurrent but also positive recurrent with steady-state probabilities $\{\pi_j;\ j \geq 0\}$. We continue to assume a given starting state $X_0 = i$. Applying the strong law for delayed renewal processes (Theorem 4.8.1) to the Markov process,

$$\lim_{t\to\infty} M_{ij}(t)/t = 1/\overline{W}(j) \qquad \text{WP1.} \tag{6.12}$$

As before, $M_{ij}(t) = N_{ij}(M_i(t))$. Since $\lim_{t\to\infty} M_i(t) = \infty$ with probability 1,

$$\lim_{t\to\infty} \frac{M_{ij}(t)}{M_i(t)} = \lim_{t\to\infty} \frac{N_{ij}(M_i(t))}{M_i(t)} = \lim_{n\to\infty} \frac{N_{ij}(n)}{n} = \pi_j \qquad \text{WP1.} \tag{6.13}$$

In the last step, we applied the same strong law to the embedded chain. Combining (6.12) and (6.13), the following equalities hold with probability 1.

$$\begin{aligned} \frac{1}{\overline{W}(j)} &= \lim_{t\to\infty} \frac{M_{ij}(t)}{t} \\ &= \lim_{t\to\infty} \frac{M_{ij}(t)}{M_i(t)} \frac{M_i(t)}{t} \\ &= \pi_j \lim_{t\to\infty} \frac{M_i(t)}{t}. \end{aligned} \tag{6.14}$$

This tells us that $\overline{W}(j)\pi_j$ is the same for all j. Also, since $\pi_j > 0$ for a positive recurrent chain, it tell us that if $\overline{W}(j) < \infty$ for one state j, it is finite for all states. Also these expected recurrence times are finite if and only if $\lim_{t\to\infty} M_i(t)/t > 0$. Finally, it says implicitly that $\lim_{t\to\infty} M_i(t)/t$ exists WP1 and has the same value for all starting states i.

There is relatively little left to do, and the following theorem does most of it.

Theorem 6.2.1. *Consider an irreducible Markov process with a positive recurrent embedded Markov chain. Let $\{\pi_j;\ j \geq 0\}$ be the steady-state probabilities of the embedded chain and let $X_0 = i$ be the starting state. Then, with probability 1, the limiting time-average fraction of time spent in any arbitrary state j is the steady-state process probability in (6.7),* i.e.,

$$p_j(i) = p_j = \frac{\pi_j/\nu_j}{\sum_k \pi_k/\nu_k}. \tag{6.15}$$

The expected time between returns to state j is

$$\overline{W}(j) = \frac{\sum_k \pi_k/\nu_k}{\pi_j}, \tag{6.16}$$

and the limiting rate at which transitions take place is independent of the starting state and given by

$$\lim_{t\to\infty} \frac{M_i(t)}{t} = \frac{1}{\sum_k \pi_k/\nu_k} \qquad \text{WP1.} \tag{6.17}$$

Discussion: Recall that $p_j(i)$ was defined as a time average WP1, and we saw earlier that this time average existed with a value independent of i. The theorem states that this time average (and the limiting ensemble average) is given by the steady-state process probabilities in (6.7). Thus, after the proof, we can stop distinguishing these quantities.

At a superficial level, the theorem is almost obvious from what we have done. In particular, substituting (6.14) into (6.9), we see that

$$p_j(i) = \frac{\pi_j}{\nu_j} \lim_{t\to\infty} \frac{M_i(t)}{t} \qquad \text{WP1.} \tag{6.18}$$

Since $p_j(i) = \lim_{t\to\infty} \Pr\{X(t) = j\}$, and since $X(t)$ is in some state at all times, we would conjecture (and even insist if we didn't read on) that $\sum_j p_j(i) = 1$. Adding that condition to normalize (6.18), we get (6.15), and (6.16) and (6.17) follow immediately. The trouble is that if $\sum_j \pi_j/\nu_j = \infty$, then (6.15) says that $p_j = 0$ for all j, and (6.17) says that $\lim M_i(t)/t = 0$, *i.e.*, the process 'gets tired' with increasing t and the rate of transitions decreases toward 0. The rather technical proof to follow deals with these limits more carefully.

Proof: We have seen in (6.14) that $\lim_{t\to\infty} M_i(t)/t$ is equal to a constant, say α, with probability 1 and that this constant is the same for all starting states i. We first consider the case where $\alpha > 0$. In this case, from (6.14), $\overline{W}(j) < \infty$ for all j. Choosing any given j and any positive integer ℓ, consider a renewal-reward process with renewals on transitions to j and a reward $R_{ij}^{\ell}(t) = 1$ when $X(t) \leq \ell$. This reward is independent of j and equal to $\sum_{k=1}^{\ell} R_{ik}(t)$. Thus, from (6.9), we have

$$\lim_{t\to\infty} \frac{\int_0^t R_{ij}^{\ell}(\tau)d\tau}{t} = \sum_{k=1}^{\ell} p_k(i). \tag{6.19}$$

If we let $\mathsf{E}\left[\mathrm{R}_j^{\ell}\right]\}$ be the expected reward over a renewal interval, then, from Theorem 4.8.4,

$$\lim_{t\to\infty} \frac{\int_0^t R_{ij}^{\ell}(\tau)d\tau}{t} = \frac{\mathsf{E}\left[\mathrm{R}_j^{\ell}\right]\}}{\overline{U}_j}. \tag{6.20}$$

Note that $\mathsf{E}\left[\mathrm{R}_j^{\ell}\right]\}$ above is non-decreasing in ℓ and goes to the limit $\overline{W}(j)$ as $\ell \to \infty$. Thus, combining (6.19) and (6.20), we see that

$$\lim_{\ell\to\infty} \sum_{k=1}^{\ell} p_k(i) = 1.$$

With this added relation, (6.15), (6.16), and (6.17) follow as in the discussion. This completes the proof for the case where $\alpha > 0$.

For the remaining case, where $\lim_{t\to\infty} M_i(t)/t = \alpha = 0$, (6.14) shows that $\overline{W}(j) = \infty$ for all j and (6.18) then shows that $p_j(i) = 0$ for all j. We give a guided proof in Exercise 6.6 that, for $\alpha = 0$, we must have $\sum_i \pi_i/\nu_i = \infty$. It follows that (6.15), (6.16), and (6.17) are all satisfied. □

This has been quite a difficult proof for something that might seem almost obvious for simple examples. However, the fact that these time-averages are valid over all sample points with probability 1 is not obvious and the fact that $\pi_j \overline{W}(j)$ is independent of j is certainly not obvious.

The most subtle thing here, however, is that if $\sum_i \pi_i/\nu_i = \infty$, then $p_j = 0$ for all states j. This is strange because the time-average state probabilities do not add to 1, and also strange because the embedded Markov chain continues to make transitions, and these transitions, in steady state for the Markov chain, occur with the probabilities π_i. Example 6.2.1 and Exercise 6.3 give some insight into this. Some added insight can be gained by looking at the embedded Markov chain starting in steady state, *i.e.*, with probabilities $\{\pi_i;\, i \geq 0\}$. Given $X_0 = i$, the expected time to a transition is $1/\nu_i$, so the unconditional expected time to a transition is $\sum_i \pi_i/\nu_i$, which is infinite for the case under consideration. This is not a phenomenon that can be easily understood intuitively, but Example 6.2.1 and Exercise 6.3 will help.

6.2.4 Solving for the steady-state process probabilities directly

Let us return to the case where $\sum_k \pi_k/\nu_k < \infty$, which is the case of virtually all applications. We have seen that a Markov process can be specified in terms of the time-transitions $q_{ij} = \nu_i P_{ij}$, and it is useful to express the steady-state equations for p_j directly in terms of q_{ij} rather than indirectly in terms of the embedded chain. As a useful prelude to this, we first express the π_j in terms of the p_j. Denote $\sum_k \pi_k/\nu_k$ as $\beta < \infty$. Then, from (6.15), $p_j = \pi_j/\nu_j\beta$, so $\pi_j = p_j\nu_j\beta$. Expressing this along with the normalization $\sum_k \pi_k = 1$, we obtain

$$\pi_i = \frac{p_i \nu_i}{\sum_k \; p_k \nu_k}. \tag{6.21}$$

Thus, $\beta = 1/\sum_k p_k\nu_k$, so

$$\sum_k \pi_k/\nu_k = \frac{1}{\sum_k p_k \nu_k}. \tag{6.22}$$

We can now substitute π_i as given by (6.21) into the steady-state equations for the embedded Markov chain, *i.e.*, $\pi_j = \sum_i \pi_i P_{ij}$ for all j, obtaining

$$p_j \nu_j = \sum_i p_i \nu_i P_{ij}$$

for each state j. Since $\nu_i P_{ij} = q_{ij}$,

$$p_j \nu_j = \sum_i p_i q_{ij}; \qquad \sum_i p_i = 1. \tag{6.23}$$

This set of equations is known as the steady-state equations for the Markov process. The normalization condition $\sum_i p_i = 1$ is a consequence of (6.22) and also of (6.15). Equation (6.23) has a nice interpretation in that the term on the left is the steady-state rate at which transitions occur out of state j and the term on the right is the rate at which transitions occur into state j. Since the total number of entries to j must differ by at most 1 from the exits from j for each sample path, this equation is not surprising.

The embedded chain is positive recurrent, so its steady-state equations have a unique solution with all $\pi_i > 0$. Thus (6.23) also has a unique solution with all $p_i > 0$ under the the added condition that $\sum_i \pi_i/\nu_i < \infty$. However, we would like to solve (6.23) directly without worrying about the embedded chain.

If we find a solution to (6.23), however, and if $\sum_i p_i\nu_i < \infty$ in that solution, then the corresponding set of π_i from (6.21) must be the unique steady-state solution for the embedded chain. Thus the solution for p_i must be the corresponding steady-state solution for the Markov process. This is summarized in the following theorem.

Theorem 6.2.2. *Assume an irreducible Markov process and let $\{p_i;\, i \geq 0\}$ be a solution to (6.23). If $\sum_i\ p_i\nu_i < \infty$, then, first, that solution is unique, second, each p_i is positive, and third, the embedded Markov chain is positive recurrent with steady-state probabilities satisfying (6.21). Also, if the embedded chain is positive recurrent, and $\sum_i \pi_i/\nu_i < \infty$ then the set of p_i satisfying (6.15) is the unique solution to (6.23).*

6.2.5 The sampled-time approximation again

For an alternative view of the probabilities $\{p_i;\, i \geq 0\}$, consider the special case (but the typical case) where the transition rates $\{\nu_i;\, i \geq 0\}$ are bounded. Consider the sampled-time approximation to the process for a given increment size $\delta \leq [\max_i \nu_i]^{-1}$ (see Figure 6.6). Let $\{p_i(\delta);\, i \geq 0\}$ be the set of steady-state probabilities for the sampled-time chain, assuming that they exist. These steady-state probabilities satisfy

$$p_j(\delta) = \sum_{i\neq j} p_i(\delta)q_{ij}\delta + p_j(\delta)(1-\nu_j\delta); \quad p_j(\delta) \geq 0; \quad \sum_j p_j(\delta) = 1. \tag{6.24}$$

The first equation simplifies to $p_j(\delta)\nu_j = \sum_{i\neq j} p_i(\delta)q_{ij}$, which is the same as (6.23). It follows that the steady-state probabilities $\{p_i;\, i \geq 0\}$ for the process are the same as the steady-state probabilities $\{p_i(\delta);\, i \geq 0\}$ for the sampled-time approximation. Note that this is not an approximation; $p_i(\delta)$ is exactly equal to p_i for all values of $\delta \leq 1/\sup_i \nu_i$. We shall see later that the dynamics of a Markov process are not quite so well modeled by the sampled time approximation except in the limit $\delta \to 0$.

6.2.6 Pathological cases

Example 6.2.1 (Zero transition rate). Consider the Markov process with a positive-recurrent embedded chain in Figure 6.8. This models a variation of an M/M/1 queue in which the server becomes increasingly rattled and slow as the queue builds up, and the customers become almost equally discouraged about entering. The downward drift in the transitions is more than overcome by the slow-down in large numbered states. Transitions continue to occur, but the number of transitions per unit time goes to 0 with increasing time. Although the embedded chain has a steady-state solution, the process can not be viewed as having any sort of steady state. Exercise 6.3 gives some added insight into this type of situation.

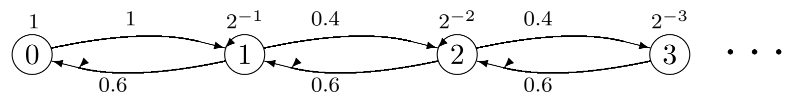


Figure 6.8: The Markov process for a variation on M/M/1 where arrivals and services get slower with increasing state. Each node i has a rate $\nu_i = 2^{-i}$. The embedded chain transition probabilities are $P_{i,i+1} = 0.4$ for $i \geq 1$ and $P_{i,i-1} = 0.6$ for $i \geq 1$, thus ensuring that the embedded Markov chain is positive recurrent. Note that $q_{i,i+1} > q_{i+1,i}$, thus ensuring that the Markov process drifts to the right.

It is also possible for (6.23) to have a solution for $\{p_i;\, i \geq 0\}$ with $\sum_i \; p_i = 1$, but $\sum_i \; p_i\nu_i = \infty$. This is not possible for a positive recurrent embedded chain, but is possible both if the embedded Markov chain is transient and if it is null recurrent. A transient chain means that there is a positive probability that the embedded chain will *never* return to a state after leaving it, and thus there can be no sensible kind of steady-state behavior for the process. These processes are characterized by arbitrarily large transition rates from the various states, and these allow the process to transit through an infinite number of states in a finite time.

Processes for which there is a non-zero probability of passing through an infinite number of states in a finite time are called *irregular*. Exercises 6.8 and 6.7 give some insight into irregular processes. Exercise 6.9 gives an example of a process that is not irregular, but for which (6.23) has a solution with $\sum_i p_i = 1$ and the embedded Markov chain is null recurrent. We restrict our attention in what follows to irreducible Markov chains for which (6.23) has a solution, $\sum \} p_i = 1$, and $\sum \} p_i\nu_i < \infty$. This is slightly more restrictive than necessary, but processes for which $\sum_i \; p_i\nu_i = \infty$ (see Exercise 6.9) are not very robust.

6.3 The Kolmogorov differential equations

Let $P_{ij}(t)$ be the probability that a Markov process $\{X(t);\, t \geq 0\}$ is in state j at time t given that $X(0) = i$,

$$P_{ij}(t) = P\{X(t){=}j \mid X(0){=}i\}. \tag{6.25}$$

$P_{ij}(t)$ is analogous to the n^{th} order transition probabilities P_{ij}^n for Markov chains. We have already seen that $\lim_{t\to\infty} P_{ij}(t) = p_j$ for the case where the embedded chain is positive recurrent and $\sum_i \; \pi_i/\nu_i < \infty$. Here we want to find the transient behavior, and we start by deriving the Chapman-Kolmogorov equations for Markov processes. Let s and t be arbitrary times, $0 < s < t$. By including the state at time s, we can rewrite (6.25) as

$$\begin{aligned} P_{ij}(t) &= \sum_k \Pr\{X(t){=}j,\ X(s){=}k \mid X(0){=}i\} \\ &= \sum_k \}\Pr\{X(s){=}k \mid X(0){=}i\}\ \Pr\{X(t){=}j \mid X(s){=}k\}\,; \qquad \text{all } i,j, \end{aligned} \tag{6.26}$$

where we have used the Markov condition, (6.3). Given that $X(s) = k$, the residual time until the next transition after s is exponential with rate ν_k, and thus the process starting at time s in state k is statistically identical to that starting at time 0 in state k. Thus, for any s, $0 \le s \le t$, we have

$$P\{X(t){=}j \mid X(s){=}k\} = P_{kj}(t-s).$$

Substituting this into (6.26), we have the *Chapman-Kolmogorov equations* for a Markov process,

$$P_{ij}(t) = \sum_k P_{ik}(s)P_{kj}(t-s). \tag{6.27}$$

These equations correspond to (3.8) for Markov chains. We now use these equations to derive two types of sets of differential equations for $P_{ij}(t)$. The first are called the *Kolmogorov forward differential equations*, and the second the *Kolmogorov backward differential equations*. The forward equations are obtained by letting s approach t from below, and the backward equations are obtained by letting s approach 0 from above. First we derive the forward equations.

For $t-s$ small and positive, $P_{kj}(t-s)$ in (6.27) can be expressed as $(t{-}s)q_{kj} + \mathrm{o}(t{-}s)$ for $k \ne j$. Similarly, $P_{jj}(t-s)$ can be expressed as $1 - (t{-}s)\nu_j + \mathrm{o}(s)$. Thus (6.27) becomes

$$P_{ij}(t) = \sum_{k \ne j} [P_{ik}(s)(t-s)q_{kj}] + P_{ij}(s)[1-(t-s)\nu_j] + \mathrm{o}(t-s) \tag{6.28}$$

We want to express this, in the limit $s \to t$, as a differential equation. To do this, subtract $P_{ij}(s)$ from both sides and divide by $t-s$.

$$\frac{P_{ij}(t) - P_{ij}(s)}{t-s} = \sum_{k \ne j} (P_{ik}(s)q_{kj}) \ - P_{ij}(s)\nu_j + \frac{\mathrm{o}(s)}{s}$$

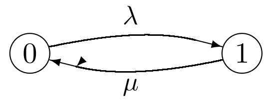
Taking the limit as $s \to t$ from below,[3] we get the Kolmogorov forward equations,

$$\frac{dP_{ij}(t)}{dt} = \sum_{k \ne j} (P_{ik}(t)q_{kj}) - P_{ij}(t)\nu_j, \tag{6.29}$$

The first term on the right side of (6.29) is the rate at which transitions occur into state j at time t and the second term is the rate at which transitions occur out of state j. Thus the difference of these terms is the net rate at which transitions occur into j, which is the rate at which $P_{ij}(t)$ is increasing at time t.

[3] We have assumed that the sum and the limit in (6.3) can be interchanged. This is certainly valid if the state space is finite, which is the only case we analyze in what follows.

Example 6.3.1 (A queueless M/M/1 queue). Consider the following 2-state Markov process where $q_{01} = \lambda$ and $q_{10} = \mu$.



This can be viewed as a model for an M/M/1 queue with no storage for waiting customers. When the system is empty (state 0), memoryless customers arrive at rate λ, and when the server is busy, an exponential server operates at rate μ, with the system returning to state 0 when service is completed.

To find $P_{01}(t)$, the probability of state 1 at time t conditional on state 0 at time 0, we use the Kolmogorov forward equations for $P_{01}(t)$, getting

$$\frac{dP_{01}(t)}{dt} \;=\; P_{00}(t)q_{01} - P_{01}(t)\nu_1 \;=\; P_{00}(t)\lambda - P_{01}(t)\mu.$$

Using the fact that $P_{00}(t) = 1 - P_{01}(t)$, this becomes

$$\frac{dP_{01}(t)}{dt} = \lambda - P_{01}(t)(\lambda + \mu).$$

Using the boundary condition $P_{01}(0) = 0$, the solution is

$$P_{01}(t) = \frac{\lambda}{\lambda+\mu}\left[1 - e^{-(\lambda+\mu)t}\right]\} \tag{6.30}$$

Thus $P_{01}(t)$ is 0 at $t = 0$ and increases as $t \to \infty$ to its steady-state value in state 1, which is $\lambda/(\lambda+\mu)$.

In general, for any given starting state i in a Markov process with M states, (6.29) provides a set of M simultaneous linear differential equations, one for each j, $1 \le j \le \mathsf{M}$. As we saw in the example, one of these is redundant because $\sum_{j=1}^{\mathsf{M}} P_{ij}(t) = 1$. This leaves $\mathsf{M} - 1$ simultaneous linear differential equations to be solved.

For more than 2 or 3 states, it is more convenient to express (6.29) in matrix form. Let $[P(t)]$ (for each $t > 0$) be an M by M matrix whose i,j element is $P_{ij}(t)$. Let $[Q]$ be an M by M matrix whose i,j element is q_{ij} for each $i \ne j$ and $-\nu_j$ for $i = j$. Then (6.29) becomes

$$\frac{d[P(t)]}{dt} = [P(t)][Q]. \tag{6.31}$$

For Example 6.3.1, $P_{ij}(t)$ can be calculated for each i,j as in (6.30), resulting in

$$[P(t)] = \begin{bmatrix} \}\frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t} & \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t} \\ \}\frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t} & \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t} \end{bmatrix}\} \qquad [Q] = \begin{bmatrix} \}-\lambda & \lambda \\ \}\ \mu & -\mu \end{bmatrix}\}$$

In order to provide some insight into the general solution of (6.31), we go back to the sampled time approximation of a Markov process. With an increment of size δ between samples, the probability of a transition from i to j, $i \ne j$, is $q_{ij}\delta + o(\delta)$, and the probability of remaining in state i is $1-\nu_i\delta+o(\delta)$. Thus, in terms of the matrix $[Q]$ of transition rates, the transition probability matrix in the sampled time model is $[I]+\delta[Q]$, where $[I]$ is the identity matrix. We denote this matrix by $[W_\delta]=[I]+\delta[Q]$. Note that λ is an eigenvalue of $[Q]$ if and only if $1+\lambda\delta$ is an eigenvalue of $[W_\delta]$. Also the eigenvectors of these corresponding eigenvalues are the same. That is, if $\boldsymbol{\nu}$ is a right eigenvector of $[Q]$ with eigenvalue λ, then $\boldsymbol{\nu}$ is a right eigenvector of $[W_\delta]$ with eigenvalue $1+\lambda\delta$, and conversely. Similarly, if $\boldsymbol{p}$ is a left eigenvector of $[Q]$ with eigenvalue λ, then $\boldsymbol{p}$ is a left eigenvectorof $[W_\delta]$ with eigenvalue $1+\lambda\delta$, and conversely.

We saw in Section 6.2.5 that the steady-state probability vector $\boldsymbol{p}$ of a Markov process is the same as that of any sampled-time approximation. We have now seen that, in addition, *all* the eigenvectors are the same and the eigenvalues are simply related. Thus study of these differential equations can be largely replaced by studying the sampled-time approximation.

The following theorem uses our knowledge of the eigenvalues and eigenvectors of transition matrices such as $[W_\delta]$ in Section 3.4, to be more specific about the properties of $[Q]$.

Theorem 6.3.1. *Consider an irreducible finite-state Markov process with* M *states. Then the matrix* $[Q]$ *for that process has an eigenvalue* λ *equal to* 0. *That eigenvalue has a right eigenvector* $\boldsymbol{e}=(1,1,\dots,1)^{\mathsf{T}}$ *which is unique within a scale factor. It has a left eigenvector* $\boldsymbol{p}=(p_1,\dots,p_{\mathsf{M}})$ *that is positive, sums to* 1, *satisfies (6.23), and is unique within a scale factor. All the other eigenvalues of* $[Q]$ *have strictly negative real parts.*

Proof: Since all M states communicate, the sampled time chain is recurrent. From Theorem 3.4.1, $[W_\delta]$ has a unique eigenvalue $\lambda=1$. The corresponding right eigenvector is $\boldsymbol{e}$ and the left eigenvector is the steady-state probability vector $\boldsymbol{p}$ as given in (3.9). Since $[W_\delta]$ is recurrent, the components of $\boldsymbol{p}$ are strictly positive. From the equivalence of (6.23) and (6.24), $\boldsymbol{p}$, as given by (6.23), is the steady-state probability vector of the process. Each eigenvalue λ_δ of $[W_\delta]$ corresponds to an eigenvalue λ of $[Q]$ with the correspondence $\lambda_\delta=1+\lambda\delta$, *i.e.*, $\lambda=(\lambda_\delta-1)/\delta$. Thus the eigenvalue 1 of $[W_\delta]$ corresponds to the eigenvalue 0 of $[Q]$. Since $|\lambda_\delta|\le 1$ and $\lambda_\delta\ne 1$ for all other eigenvalues, the other eigenvalues of $[Q]$ all have strictly negative real parts, completing the proof. □

We complete this section by deriving the Komogorov backward equations. For s small and positive, the Chapman-Kolmogorov equations in (6.27) become

$$\begin{aligned} P_{ij}(t) &= \sum_k P_{ik}(s)P_{kj}(t-s) \\ &= \sum_{k\ne i} sq_{ik}P_{kj}(t-s)+(1-s\nu_i)P_{ij}(t-s)+o(s) \end{aligned}$$

Subtracting $P_{ij}(t-s)$ from both sides and dividing by s,

$$\begin{aligned} \frac{P_{ij}(t)-P_{ij}(t-s)}{s} &= \sum_{k\ne i} q_{ik}P_{kj}(t-s)-\nu_iP_{ij}(t-s)+\frac{o(s)}{s} \\ \frac{dP_{ij}(t)}{dt} &= \sum_{k\ne i} q_{ik}P_{kj}(t)-\nu_iP_{ij}(t) \end{aligned} \tag{6.32}$$

In matrix form, this is expressed as

$$\frac{d[P(t)]}{dt} = [Q][P(t)] \tag{6.33}$$

By comparing (6.33 and (6.31), we see that $[Q][P(t)] = [P(t)][Q]$, *i.e.*, that the matrices $[Q]$ and $[P(t)]$ commute. Simultaneous linear differential equations appear in so many applications that we leave the further exploration of these forward and backward equations as simple differential equation topics rather than topics that have special properties for Markov processes.

6.4 Uniformization

Up until now, we have discussed Markov processes under the assumption that $q_{ii} = 0$ (*i.e.*, no transitions from a state into itself are allowed). We now consider what happens if this restriction is removed. Suppose we start with some Markov process defined by a set of transition rates q_{ij} with $q_{ii} = 0$, and we modify this process by some arbitrary choice of $q_{ii} \geq 0$ for each state i. This modification changes the embedded Markov chain, since ν_i is increased from $\sum_{k \neq i} q_{ik}$ to $\sum_{k \neq i} q_{ik} + q_{ii}$. From (6.5), P_{ij} is changed to q_{ij}/ν_i for the new value of ν_i for each i, j. Thus the steady-state probabilities π_i for the embedded chain are changed. The Markov process $\{X(t);\, t \geq 0\}$ is not changed, since a transition from i into itself does not change $X(t)$ and does not change the distribution of the time until the next transition to a different state. The steady-state probabilities for the process still satisfy

$$p_j \nu_j = \sum_k \}p_k q_{kj} \quad ; \quad \sum_i \}p_i = 1. \tag{6.34}$$

The addition of the new term q_{jj} increases ν_j by q_{jj}, thus increasing the left hand side by $p_j\, q_{jj}$. The right hand side is similarly increased by $p_j\, q_{jj}$, so that the solution is unchanged (as we already determined it must be).

A particularly convenient way to add self-transitions is to add them in such a way as to make the transition rate ν_j the same for all states. Assuming that the transition rates $\{\nu_i;\, i \geq 0\}$ are bounded, we define ν^* as $\sup_j\ \nu_j$ for the original transition rates. Then we set $q_{jj} = \nu^* - \sum_{k \neq j} q_{jk}$ for each j. With this addition of self-transitions, all transition rates become ν^*. From (6.21), we see that the new steady state probabilities, π_i^*, in the embedded Markov chain become equal to the steady-state process probabilities, p_i. Naturally, we could also choose any ν greater than ν^* and increase each q_{jj} to make all transition rates equal to that value of ν. When the transition rates are changed in this way, the resulting embedded chain is called a *uniformized chain* and the Markov process is called the *uniformized process*. The uniformized process is the same as the original process, except that quantities like the number of transitions over some interval are different because of the self transitions.

Assuming that all transition rates are made equal to ν^*, the new transition probabilities in the embedded chain become $P_{ij}^* = q_{ij}/\nu^*$. Let $N(t)$ be the total number of transitions that occur from 0 to t in the uniformized process. Since the rate of transitions is the same from all states and the inter-transition intervals are independent and identically exponentially distributed, $N(t)$ is a Poisson counting process of rate ν^*. Also, $N(t)$ is independent of the sequence of transitions in the embedded uniformized Markov chain. Thus, given that $N(t) = n$, the probability that $X(t) = j$ given that $X(0) = i$ is just the probability that the embedded chain goes from i to j in ν steps, *i.e.*, P_{ij}^{*n}. This gives us another formula for calculating $P_{ij}(t)$, (*i.e.*, the probability that $X(t) = j$ given that $X(0) = i$).

$$P_{ij}(t) = \sum_{n=0}^{\infty} \} P_{ij}^{*n} \, \frac{e^{-\nu^* t}(\nu^* t)^n}{n!}. \tag{6.35}$$

Another situation where the uniformized process is useful is in extending Markov decision theory to Markov processes, but we do not pursue this.

6.5 Birth-death processes

Birth-death processes are very similar to the birth-death Markov chains that we studied earlier. Here transitions occur only between neighboring states, so it is convenient to define λ_i as $q_{i,i+1}$ and μ_i as $q_{i,i-1}$ (see Figure 6.9). Since the number of transitions from i to $i+1$ is within 1 of the number of transitions from $i+1$ to i for every sample path, we conclude that

$$p_i \lambda_i = p_{i+1}\mu_{i+1}. \tag{6.36}$$

This can also be obtained inductively from (6.23) using the same argument that we used earlier for birth-death Markov chains.

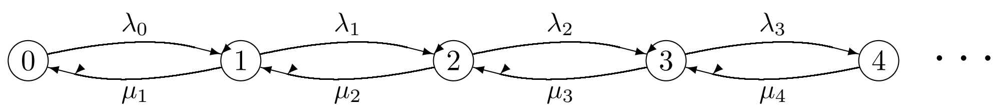


Figure 6.9: Birth-death process.

Define ρ_i as λ_i/μ_{i+1}. Then applying (6.36) iteratively, we obtain the steady-state equations

$$p_i = p_0 \prod_{j=0}^{i-1} \} \rho_j \; ; \quad i \geq 1. \tag{6.37}$$

We can solve for p_0 by substituting (6.37) into $\sum_i p_i$, yielding

$$p_0 = \frac{1}{1 + \sum_{i=1}^{\infty} \prod_{j=0}^{i-1} \rho_j}. \tag{6.38}$$

For the M/M/1 queue, the state of the Markov process is the number of customers in the system (*i.e.*, customers either in queue or in service). The transitions from i to $i+1$ correspond to arrivals, and since the arrival process is Poisson of rate λ, we have $\lambda_i = \lambda$ for all $i \geq 0$. The transitions from i to $i-1$ correspond to departures, and since the service time distribution is exponential with parameter μ, say, we have $\mu_i = \mu$ for all $i \geq 1$. Thus, (6.38) simplifies to $p_0 = 1 - \rho$, where $\rho = \lambda/\mu$ and thus

$$p_i = (1-\rho)\rho^i; \qquad i \geq 0. \tag{6.39}$$

We assume that $\rho < 1$, which is required for positive recurrence. The probability that there are i or more customers in the system in steady state is then given by $P\{X(t) \geq i\} = \rho^i$ and the expected number of customers in the system is given by

$$\mathsf{E}[X(t)] = \sum_{i=1}^{\infty}\}P\{X(t) \geq i\} = \frac{\rho}{1-\rho}. \tag{6.40}$$

The expected time that a customer spends in the system in steady state can now be determined by Little's formula (Theorem 4.5.3).

$$\mathsf{E}[\text{System time}] = \frac{\mathsf{E}[X(t)]}{\lambda} = \frac{\rho}{\lambda(1-\rho)} = \frac{1}{\mu-\lambda}. \tag{6.41}$$

The expected time that a customer spends in the queue (*i.e.*, before entering service) is just the expected system time less the expected service time, so

$$\mathsf{E}[\text{Queueing time}] = \frac{1}{\mu-\lambda} - \frac{1}{\mu} = \frac{\rho}{\mu-\lambda}. \tag{6.42}$$

Finally, the expected number of customers in the queue can be found by applying Little's formula to (6.42),

$$\mathsf{E}[\text{Number in queue}] = \frac{\lambda\rho}{\mu-\lambda}. \tag{6.43}$$

Note that the expected number of customers in the system and in the queue depend only on ρ, so that if the arrival rate and service rate were both speeded up by the same factor, these expected values would remain the same. The expected system time and queueing time, however would decrease by the factor of the rate increases. Note also that as ρ approaches 1, all these quantities approach infinity as $1/(1-\rho)$. At the value $\rho = 1$, the embedded Markov chain becomes null-recurrent and the steady-state probabilities (both $\{\pi_i; i \geq 0\}$ and $\{p_i; i \geq 0\}$) can be viewed as being all 0 or as failing to exist.

There are many types of queueing systems that can be modeled as birth-death processes. For example the arrival rate could vary with the number in the system and the service rate could vary with the number in the system. All of these systems can be analyzed in steady state in the same way, but (6.37) and (6.38) can become quite messy in these more complex systems. As an example, we analyze the M/M/m system. Here there are m servers, each with exponentially distributed service times with parameter μ. When i customers are in the system, there are i servers working for $i < m$ and all m servers are working for $i \geq m$. With i servers working, the probability of a departure in an incremental time δ is $i\mu\delta$, so that μ_i is $i\mu$ for $i < m$ and $m\mu$ for $i \geq m$ (see Figure 6.10).

Define $\rho = \lambda/(m\mu)$. Then in terms of our general birth-death process notation, $\rho_i = m\rho/(i+1)$ for $i < m$ and $\rho_i = \rho$ for $i \geq m$. From (6.37), we have

$$p_i = p_0 \frac{m\rho}{1}\frac{m\rho}{2}\cdots\frac{m\rho}{i} = \frac{p_0(m\rho)^i}{i!}; \qquad i \leq m \tag{6.44}$$

$$p_i = \frac{p_0\rho^i m^m}{m!}; \qquad i \geq m. \tag{6.45}$$

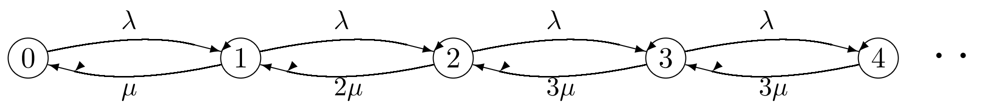


Figure 6.10: M/M/m queue for $m = 3$..

We can find p_0 by summing p_i and setting the result equal to 1; a solution exists if $\rho < 1$. Nothing simplifies much in this sum, except that $\sum_{i\geq m} p_i = p_0(\rho m)^m/[m!(1-\rho)]$, and the solution is

$$p_0 = \left[\} \frac{(m\rho)^m}{m!(1-\rho)} + \sum_{i=0}^{m-1}\} \frac{(m\rho)^i}{i!}\right]^{-1}. \tag{6.46}$$

6.6 Reversibility for Markov processes

In Section 5.3 on reversibility for Markov chains, (5.37) showed that the backward transition probabilities P^*_{ij} in steady state satisfy

$$\pi_i P^*_{ij} = \pi_j P_{ji}. \tag{6.47}$$

These equations are then valid for the embedded chain of a Markov process. Next, consider backward transitions in the process itself. Given that the process is in state i, the probability of a transition in an increment δ of time is $\nu_i\delta + o(\delta)$, and transitions in successive increments are independent. Thus, if we view the process running backward in time, the probability of a transition in each increment δ of time is also $\nu_i\delta + o(\delta)$ with independence between increments. Thus, going to the limit $\delta \to 0$, the distribution of the time backward to a transition is exponential with parameter ν_i. This means that the process running backwards is again a Markov process with transition probabilities P^*_{ij} and transition rates ν_i. Figure 6.11 helps to illustrate this.

Since the steady-state probabilities $\{p_i;\, i \geq 0\}$ for the Markov process are determined by

$$p_i = \frac{\pi_i/\nu_i}{\sum_k \pi_k/\nu_k}, \tag{6.48}$$

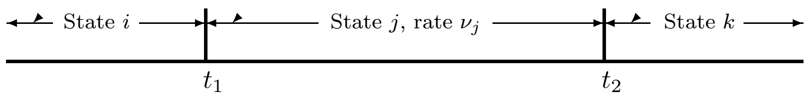


Figure 6.11: The forward process enters state j at time t_1 and departs at t_2. The backward process enters state j at time t_2 and departs at t_1. In any sample function, as illustrated, the interval in a given state is the same in the forward and backward process. Given $X(t) = j$, the time forward to the next transition and the time backward to the previous transition are each exponential with rate ν_j.

and since $\{\pi_i;\ i \geq 0\}$ and $\{\nu_i;\ i \geq 0\}$ are the same for the forward and backward processes, we see that the steady-state probabilities in the backward Markov process are the same as the steady-state probabilities in the forward process. This result can also be seen by the correspondence between sample functions in the forward and backward processes.

The *transition rates* in the backward process are defined by $q^*_{ij} = \nu_i P^*_{ij}$. Using (6.47), we have

$$q^*_{ij} = \nu_j P^*_{ij} = \frac{\nu_i \pi_j P_{ji}}{\pi_i} = \frac{\nu_i \pi_j q_{ji}}{\pi_i \nu_j}. \qquad (6.49)$$

From (6.48), we note that $p_j = \alpha\pi_j/\nu_j$ and $p_i = \alpha\pi_i/\nu_i$ for the same value of α. Thus the ratio of π_j/ν_j to π_i/ν_i is p_j/p_i. This simplifies (6.49) to $q^*_{ij} = p_j q_{ji}/p_i$, and

$$p_i q^*_{ij} = p_j q_{ji}. \qquad (6.50)$$

This equation can be used as an alternate definition of the backward transition rates. To interpret this, let δ be a vanishingly small increment of time and assume the process is in steady state at time t. Then $\delta p_j q_{ji} \approx \Pr\{X(t) = j\} \Pr\{X(t+\delta) = i \mid X(t) = j\}$ whereas $\delta p_i q^*_{ij} \approx \Pr\{X(t+\delta) = i\} \Pr\{X(t) = j \mid X(t+\delta) = i\}$.

A Markov process is defined to be *reversible* if $q^*_{ij} = q_{ij}$ for all i, j. If the embedded Markov chain is reversible, (*i.e.*, $P^*_{ij} = P_{ij}$ for all i, j), then one can repeat the above steps using P_{ij} and q_{ij} in place of P^*_{ij} and q^*_{ij} to see that $p_i q_{ij} = p_j q_{ji}$ for all i, j. Thus, if the embedded chain is reversible, the process is also. Similarly, if the Markov process is reversible, the above argument can be reversed to see that the embedded chain is reversible. Thus, we have the following useful lemma.

Lemma 6.6.1. *Assume that steady-state probabilities $\{p_i;\ i \geq 0\}$ exist in an irreducible Markov process (*i.e.*, (6.23) has a solution and $\sum p_i \nu_i < \infty$). Then the Markov process is reversible if and only if the embedded chain is reversible.*

One can find the steady-state probabilities of a reversible Markov process and simultaneously show that it is reversible by the following useful theorem (which is directly analogous to Theorem 5.3.2 of chapter 5).

Theorem 6.6.1. *For an irreducible Markov process, assume that* $\{p_i; i \geq 0\}$ *is a set of nonnegative numbers summing to* 1*, satisfying* $\sum_i p_i \nu_i \leq \infty$*, and satisfying*

$$p_i q_{ij} = p_j q_{ji} \qquad \text{for all } i, j. \tag{6.51}$$

Then $\{p_i; i \geq 0\}$ *is the set of steady-state probabilities for the process,* $p_i > 0$ *for all* i*, the process is reversible, and the embedded chain is positive recurrent.*

Proof: Summing (6.51) over i, we obtain

$$\sum_i p_i q_{ij} = p_j \nu_j \qquad \text{for all } j.$$

These, along with $\sum_i p_i = 1$ are the steady-state equations for the process. These equations have a solution, and by Theorem 6.2.2, $p_i > 0$ for all i, the embedded chain is positive recurrent, and $p_i = \lim_{t\to\infty} \Pr\{X(t) = i\}$. Comparing (6.51) with (6.50), we see that $q_{ij} = q^*_{ij}$, so the process is reversible. □

There are many irreducible Markov processes that are not reversible but for which the backward process has interesting properties that can be deduced, at least intuitively, from the forward process. Jackson networks (to be studied shortly) and many more complex networks of queues fall into this category. The following simple theorem allows us to use whatever combination of intuitive reasoning and wishful thinking we desire to guess both the transition rates q^*_{ij} in the backward process and the steady-state probabilities, and to then verify rigorously that the guess is correct. One might think that guessing is somehow unscientific, but in fact, the art of educated guessing and intuitive reasoning leads to much of the best research.

Theorem 6.6.2. *For an irreducible Markov process, assume that a set of positive numbers* $\{p_i; i \geq 0\}$ *satisfy* $\sum_i p_i = 1$ *and* $\sum_i p_i \nu_i < \infty$*. Also assume that a set of nonnegative numbers* $\{q^*_{ij}\}$ *satisfy the two sets of equations*

$$\sum_j q_{ij} = \sum_j q^*_{ij} \quad \text{for all } i \tag{6.52}$$

$$p_i q_{ij} = p_j q^*_{ji} \quad \text{for all } i, j. \tag{6.53}$$

Then $\{p_i\}$ *is the set of steady-state probabilities for the process,* $p_i > 0$ *for all* i*, the embedded chain is positive recurrent, and* $\{q^*_{ij}\}$ *is the set of transition rates in the backward process.*

Proof: Sum (6.53) over i. Using the fact that $\sum_j q_{ij} = \nu_i$ and using (6.52), we obtain

$$\sum_i p_i q_{ij} = p_j \nu_j \qquad \text{for all} j. \tag{6.54}$$

These, along with $\sum_i p_i = 1$, are the steady-state equations for the process. These equations thus have a solution, and by Theorem 6.2.2, $p_i > 0$ for all i, the embedded chain is positive recurrent, and $p_i = \lim_{t\to\infty} \Pr\{X(t) = i\}$. Finally, q^*_{ij} as given by (6.53) is the backward transition rate as given by (6.50) for all i, j. □

We see that Theorem 6.6.1 is just a special case of Theorem 6.6.2 in which the guess about q^*_{ij} is that $q^*_{ij} = q_{ij}$.

Birth-death processes are all reversible if the steady-state probabilities exist. To see this, note that Equation (6.36) (the equation to find the steady-state probabilities) is just (6.51) applied to the special case of birth-death processes. Due to the importance of this, we state it as a theorem.

Theorem 6.6.3. *For a birth-death process, if there is a solution $\{p_i;\, i \geq 0\}$ to (6.36) with $\sum_i p_i = 1$ and $\sum_i p_i\nu_i < \infty$, then the process is reversible, and the embedded chain is positive recurrent and reversible.*

Since the M/M/1 queueing process is a birth-death process, it is also reversible. Burke's theorem, which was given as Theorem 5.4.1 for sampled-time M/M/1 queues, can now be established for continuous-time M/M/1 queues. Note that the theorem here contains an extra part, part c).

Theorem 6.6.4 (Burke's theorem). *Given an M/M/1 queueing system in steady state with $\lambda < \mu$,*

a) the departure process is Poisson with rate λ,

b) the state $X(t)$ at any time t is independent of departures prior to t, and

c) for FCFS service, given that a customer departs at time t, the arrival time of that customer is independent of the departures prior to t.

Proof: The proofs of parts a) and b) are the same as the proof of Burke's theorem for sampled-time, Theorem 5.4.1, and thus will not be repeated. For part c), note that with FCFS service, the m^{th} customer to arrive at the system is also the m^{th} customer to depart. Figure 6.12 illustrates that the association between arrivals and departures is the same in the backward system as in the forward system (even though the customer ordering is reversed in the backward system). In the forward, right moving system, let τ be the epoch of some given arrival. The customers arriving after τ wait behind the given arrival in the queue, and have no effect on the given customer's service. Thus the interval from τ to the given customer's service completion is independent of arrivals after τ.

Since the backward, left moving, system is also an M/M/1 queue, the interval from a given backward arrival, say at epoch t, moving left until the corresponding departure, is independent of arrivals to the left of t. From the correspondence between sample functions in the right moving and left moving systems, given a departure at epoch t in the right moving system, the departures before time t are independent of the arrival epoch of the given customer departing at t; this completes the proof. □

Part c) of Burke's theorem does not apply to sampled-time M/M/1 queues because the sampled time model does not allow for both an arrival and departure in the same increment of time.

Note that the proof of Burke's theorem (including parts a and b from Section 5.4) does not make use of the fact that the transition rate $q_{i,i-1} = \mu$ for $i \geq 1$ in the M/M/1 queue. Thus

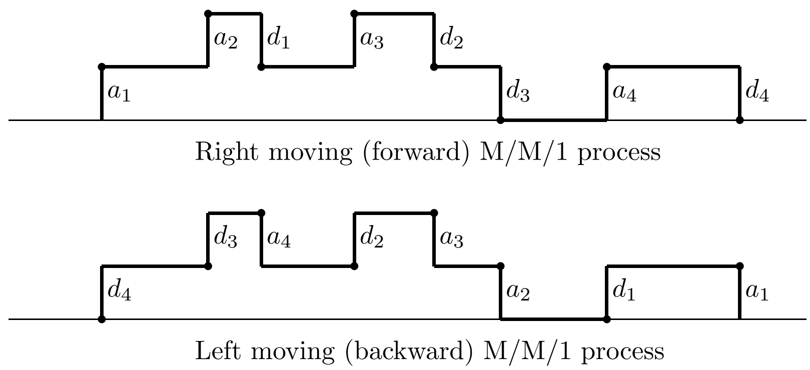


Figure 6.12: FCFS arrivals and departures in right and left moving M/M/1 processes.

Burke's theorem remains true for any birth-death Markov process in steady state for which $q_{i,i+1} = \lambda$ for all $i \geq 0$. For example, parts a and b are valid for M/M/m queues; part c is also valid (see [22]), but the argument here is not adequate since the first customer to enter the system might not be the first to depart.

We next show how Burke's theorem can be used to analyze a tandem set of queues. As shown in Figure 6.13, we have an M/M/1 queueing system with Poisson arrivals at rate λ and service at rate μ_1. The departures from this queueing system are the arrivals to a second queueing system, and we assume that a departure from queue 1 at time t instantaneously enters queueing system 2 at the same time t. The second queueing system has a single server and the service times are IID and exponentially distributed with rate μ_2. The successive service times at system 2 are also independent of the arrivals to systems 1 and 2, and independent of the service times in system 1. Since we have already seen that the departures from the first system are Poisson with rate λ, the arrivals to the second queue are Poisson with rate λ. Thus the second system is also M/M/1.

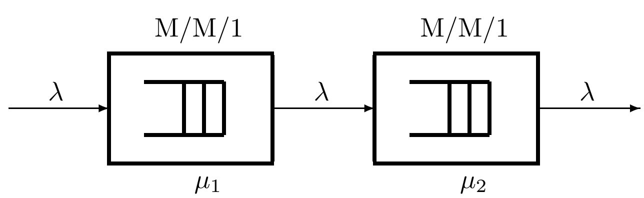


Figure 6.13: A tandem queueing system. Assuming that $\lambda > \mu_1$ and $\lambda > \mu_2$, the departures from each queue are Poisson of rate λ.

Let X(t) be the state of queueing system 1 and $Y(t)$ be the state of queueing system 2. Since $X(t)$ at time t is independent of the departures from system 1 prior to t, $X(t)$ is independent of the arrivals to system 2 prior to time t. Since $Y(t)$ depends only on the arrivals to system 2 prior to t and on the service times that have been completed prior to t, we see that $X(t)$ is independent of $Y(t)$. This leaves a slight nit-picking question about what happens at the instant of a departure from system 1. We have considered the state $X(t)$ at the instant of a departure to be the number of customers remaining in system 1 not counting the departing customer. Also the state $Y(t)$ is the state in system 2 including the new arrival at instant t. The state $X(t)$ then is independent of the departures up to and including t, so that $X(t)$ and $Y(t)$ are still independent.

Next assume that both systems use FCFS service. Consider a customer that leaves system 1 at time t. The time at which that customer arrived at system 1, and thus the waiting time in system 1 for that customer, is independent of the departures prior to t. This means that the state of system 2 immediately before the given customer arrives at time t is independent of the time the customer spent in system 1. It therefore follows that the time that the customer spends in system 2 is independent of the time spent in system 1. Thus the total system time that a customer spends in both system 1 and system 2 is the sum of two independent random variables.

This same argument can be applied to more than 2 queueing systems in tandem. It can also be applied to more general networks of queues, each with single servers with exponentially distributed service times. The restriction here is that there can not be any cycle of queueing systems where departures from each queue in the cycle can enter the next queue in the cycle. The problem posed by such cycles can be seen easily in the following example of a single queueing system with feedback (see Figure 6.14).

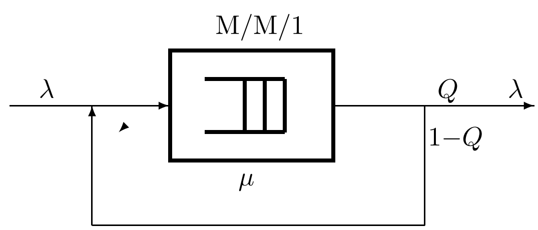


Figure 6.14: A queue with feedback. Assuming that $\mu > \lambda/Q$, the exogenous output is Poisson of rate λ.

We assume that the queueing system in Figure 6.14 has a single server with IID exponentially distributed service times that are independent of arrival times. The exogenous arrivals from outside the system are Poisson with rate λ. With probability Q, the departures from the queue leave the entire system, and, alternatively, with probability $1 - Q$, they return instantaneously to the input of the queue. Successive choices between leaving the system and returning to the input are IID and independent of exogenous arrivals and of service times. Figure 6.15 shows a sample function of the arrivals and departures in the case in which the service rate μ is very much greater than the exogenous arrival rate λ. Each exogenous arrival spawns a geometrically distributed set of departures and simultaneous re-entries. Thus the overall arrival process to the queue, counting both exogenous arrivals and feedback from the output, is not Poisson. Note, however, that if we look at the Markov process description, the departures that are fed back to the input correspond to self loops from one state to itself. Thus the Markov process is the same as one without the self loops with a service rate equal to μQ. Thus, from Burke's theorem, the exogenous departures are Poisson with rate λ. Also the steady-state distribution of $X(t)$ is $P\{X(t) = i\} = (1 - \rho)\rho^i$ where $\rho = \lambda/(\mu Q)$ (assuming, of course, that $\rho < 1$).

Figure 6.15: Sample path of arrivals and departures for queue with feedback.

The tandem queueing system of Figure 6.13 can also be regarded as a combined Markov process in which the state at time t is the pair $(X(t),\ Y(t))$. The transitions in this process correspond to, first, exogenous arrivals in which $X(t)$ increases, second, exogenous departures in which $Y(t)$ decreases, and third, transfers from system 1 to system 2 in which $X(t)$ decreases and $Y(t)$ simultaneously increases. The combined process is not reversible since there is no transition in which $X(t)$ increases and $Y(t)$ simultaneously decreases. In the next section, we show how to analyze these combined Markov processes for more general networks of queues.

6.7 Jackson networks

In many queueing situations, a customer has to wait in a number of different queues before completing the desired transaction and leaving the system. For example, when we go to the registry of motor vehicles to get a driver's license, we must wait in one queue to have the application processed, in another queue to pay for the license, and in yet a third queue to obtain a photograph for the license. In a multiprocessor computer facility, a job can be queued waiting for service at one processor, then go to wait for another processor, and so forth; frequently the same processor is visited several times before the job is completed. In a data network, packets traverse multiple intermediate nodes; at each node they enter a queue waiting for transmission to other nodes.

Such systems are modeled by a network of queues, and Jackson networks are perhaps the simplest models of such networks. In such a model, we have a network of k interconnected queueing systems which we call nodes. Each of the k nodes receives customers (*i.e.*, tasks or jobs) both from outside the network (exogenous inputs) and from other nodes within the network (endogenous inputs). It is assumed that the exogenous inputs to each node i form a Poisson process of rate r_i and that these Poisson processes are independent of each other. For analytical convenience, we regard this as a single Poisson input process of rate λ_0, with each input independently going to each node i with probability $Q_{0i} = r_i/\lambda_0$.

Each node i contains a single server, and the successive service times at node i are IID random variables with an exponentially distributed service time of rate μ_i. The service times at each node are also independent of the service times at all other nodes and independent of the exogenous arrival times at all nodes. When a customer completes service at a given

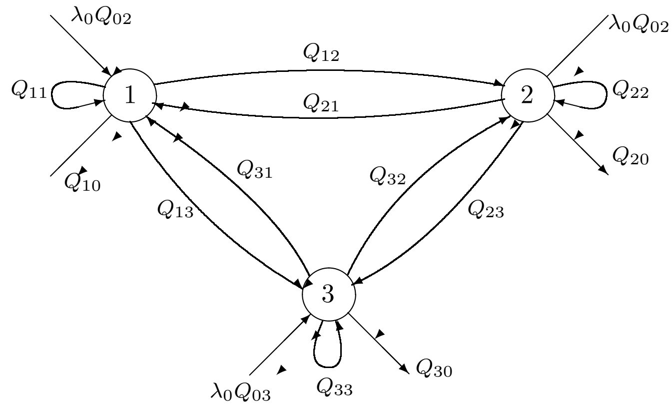


Figure 6.16: A Jackson network with 3 nodes. Given a departure from node i, the probability that departure goes to node j (or, for $j = 0$, departs the system) is Q_{ij}. Note that a departure from node i can re-enter node i with probability Q_{ii}. The overall exogenous arrival rate is λ_0, and, conditional on an arrival, the probability the arrival enters node i is Q_{0i}.

node i, that customer is routed to node j with probability Q_{ij} (see Figure 6.16). It is also possible for the customer to depart from the network entirely (called an exogenous departure), and this occurs with probability $Q_{i0} = 1 - \sum_{j \geq 1} Q_{ij}$. For a customer departing from node i, the next node j is a random variable with PMF $\{Q_{ij}, 0 \leq j \leq k\}$.

Successive choices of the next node for customers at node i are IID, independent of the customer routing at other nodes, independent of all service times, and independent of the exogenous inputs. Notationally, we are regarding the outside world as a fictitious node 0 from which customers appear and to which they disappear.

When a customer is routed from node i to node j, it is assumed that the routing is instantaneous; thus at the epoch of a departure from node i, there is a simultaneous endogenous arrival at node j. Thus a node j receives Poisson exogenous arrivals from outside the system at rate $\lambda_0 Q_{0j}$ and receives endogenous arrivals from other nodes according to the probabilistic rules just described. We can visualize these combined exogenous and endogenous arrivals as being served in FCFS fashion, but it really makes no difference in which order they are served, since the customers are statistically identical and simply give rise to service at node j at rate μ_j whenever there are customers to be served.

The Jackson queueing network, as just defined, is fully described by the exogenous input rate λ_0, the service rates $\{\mu_i\}$, and the routing probabilities $\{Q_{ij};\, 0 \leq i, j \leq k\}$. The network as a whole is a Markov process in which the state is a vector $\boldsymbol{m} = (m_1, m_2, \ldots, m_k)$, where $m_i, 1 \leq i \leq k$, is the number of customers at node i. State changes occur upon exogenous arrivals to the various nodes, exogenous departures from the various nodes, and departures from one node that enter another node. In a vanishingly small interval δ of time, given that the state at the beginning of that interval is $\boldsymbol{m}$, an exogenous arrival at node j occurs in the interval with probability $\lambda_0 Q_{0j}\delta$ and changes the state to $\boldsymbol{m}' = \boldsymbol{m} + \boldsymbol{e}_j$ where $\boldsymbol{e}_j$ is a unit vector with a one in position j. If $m_i > 0$, an exogenous departure from node i occurs in the interval with probability $\mu_i Q_{i0}\delta$ and changes the state to $\boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i$. Finally, if $m_i > 0$, a departure from node i entering node j occurs in the interval with probability $\mu_i Q_{ij}\delta$ and changes the state to $\boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i + \boldsymbol{e}_j$. Thus, the transition rates are given by

$$
\begin{aligned}
q_{\boldsymbol{m},\boldsymbol{m}'} &= \lambda_0 Q_{0j} && \text{for } \boldsymbol{m}' = \boldsymbol{m} + \boldsymbol{e}_j, \quad 1 \leq i \leq k && (6.55)\\
&= \mu_i Q_{i0} && \text{for } \boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{m}_i, \quad m_i > 0, \quad 1 \leq i \leq k && (6.56)\\
&= \mu_i Q_{ij} && \text{for } \boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i + \boldsymbol{e}_j, \quad m_i > 0, \quad 1 \leq i, j \leq k && (6.57)\\
&= 0 && \text{for all other choices of } \boldsymbol{m}'.
\end{aligned}
$$

Note that a departure from node i that re-enters node i causes a transition from state $\boldsymbol{m}$ back into state $\boldsymbol{m}$; we disallowed such transitions in sections 6.1 and 6.2, but showed that they caused no problems in our discussion of uniformization. It is convenient to allow these self-transitions here, partly for the added generality and partly to illustrate that the single node network with feedback of Figure 6.14 is an example of a Jackson network.

Our objective is to find the steady-state probabilities $p(\boldsymbol{m})$ for this type of process, and our plan of attack is in accordance with Theorem 6.6.2; that is, we shall guess a set of transition rates for the backward Markov process, use these to guess $p(\boldsymbol{m})$, and then verify that the guesses are correct. Before making these guesses, however, we must find out a little more about how the system works, so as to guide the guesswork. Let us define λ_i for each i, $1 \leq i \leq k$, as the time-average overall rate of arrivals to node i, including both exogenous and endogenous arrivals. Since λ_0 is the rate of exogenous inputs, we can interpret λ_i/λ_0 as the expected number of visits to node i per exogenous input. The endogenous arrivals to node i are not necessarily Poisson, as the example of a single queue with feedback shows, and we are not even sure at this point that such a time-average rate exists in any reasonable sense. However, let us assume for the time being that such rates exist and that the time-average rate of departures from each node equals the time-average rate of arrivals (*i.e.*, the queue sizes do not grow linearly with time). Then these rates must satisfy the equation

$$
\lambda_j = \sum_{j=0}^{k} \} \lambda_i Q_{ij}; \qquad 1 \leq j \leq k. \tag{6.58}
$$

To see this, note that $\lambda_0 Q_{0j}$ is the rate of exogenous arrivals to j. Also λ_i is the time-average rate at which customers depart from queue i, and $\lambda_i Q_{ij}$ is the rate at which customers go from node i to node j. Thus, the right hand side of (6.58) is the sum of the exogenous and endogenous arrival rates to node j. Note the distinction between the time-average rate of customers going from i to j in (6.58) and the rate $q_{\boldsymbol{m},\boldsymbol{m}'} = \mu_i Q_{ij}$ for $\boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i + \boldsymbol{e}_j, m_i > 0$ in (6.57). The rate in (6.57) is conditioned on a state $\boldsymbol{m}$ with $m_i > 0$, whereas that in (6.58) is the overall time-average rate, averaged over all states.

Note that $\{Q_{ij}; 0 \leq i, j \leq k\}$ forms a stochastic matrix and (6.58) is formally equivalent to the equations for steady-state probabilities (except that steady-state probabilities sum to 1). The usual equations for steady-state probabilities include an equation for $j = 0$, but that equation is redundant. Thus we know that, if there is a path between each pair of nodes (including the fictitious node 0), then (6.58) has a solution for $\{\lambda_i; 0 \leq i \leq k$, and that solution is unique within a scale factor. The known value of λ_0 determines this scale factor and makes the solution unique. Note that we don't have to be careful at this point about whether these rates are time-averages in any nice sense, since this will be verified later; we do have to make sure that (6.58) has a solution, however, since it will appear in our solution for $p(\boldsymbol{m})$. Thus we assume in what follows that a path exists between each pair of nodes, and thus that (6.58) has a unique solution as a function of λ_0.

We now make the final necessary assumption about the network, which is that $\mu_i > \lambda_i$ for each node i. This will turn out to be required in order to make the process positive recurrent. We also define ρ_i as λ_i/μ_i. We shall find that, even though the inputs to an individual node i are not Poisson in general, there is a steady-state distribution for the number of customers at i, and that distribution is the same as that of an M/M/1 queue with the parameter ρ_i.

Now consider the backward time process. We have seen that only three kinds of transitions are possible in the forward process. First, there are transitions from $\boldsymbol{m}$ to $\boldsymbol{m}' = \boldsymbol{m} + \boldsymbol{e}_j$ for any j, $1 \leq j \leq k$. Second, there are transitions from $\boldsymbol{m}$ to $\boldsymbol{m} - \boldsymbol{e}_i$ for any i, $1 \geq i \leq k$, such that $m_i > 0$. Third, there are transitions from $\boldsymbol{m}$ to $\boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i + \boldsymbol{e}_j$ for $1 \leq i, j \leq k$ with $m_i > 0$. Thus in the backward process, transitions from $\boldsymbol{m}'$ to $\boldsymbol{m}$ are possible only for the $\boldsymbol{m}$, $\boldsymbol{m}'$ pairs above. Corresponding to each arrival in the forward process, there is a departure in the backward process; for each forward departure, there is a backward arrival; and for each forward passage from i to j, there is a backward passage from j to i.

We now make the conjecture that the backward process is itself a Jackson network with Poisson exogenous arrivals at rates $\{\lambda_0 Q^*_{0j}\}$, service times that are exponential with rates $\{\mu_i\}$, and routing probabilities $\{Q^*_{ij}\}$. The backward routing probabilities $\{Q^*_{ij}\}$ must be chosen to be consistent with the transition rates in the forward process. Since each transition from i to j in the forward process must correspond to a transition from j to i in the backward process, we should have

$$\lambda_i Q_{ij} = \lambda_j Q^*_{ji} \quad ; \quad 0 \leq i, j \leq k. \tag{6.59}$$

Note that $\lambda_i Q_{ij}$ represents the rate at which forward transitions go from i to j, and λ_i represents the rate at which forward transitions *leave* node i. Equation (6.59) takes advantage of the fact that λ_i is also the rate at which forward transitions *enter* node i, and thus the rate at which backward transitions *leave* node i. Using the conjecture that the backward time system is a Jackson network with routing probabilities $\{Q^*_{ij};\, 0 \leq i, j \leq k\}$, we can write down the backward transition rates in the same way as (6.55-6.57),

$$\begin{aligned}
q^*_{\boldsymbol{m},\boldsymbol{m}'} &= \lambda_0 Q^*_{0j} && \text{for } \boldsymbol{m}' = \boldsymbol{m} + \boldsymbol{e}_j && (6.60)\\
&= \mu_i Q^*_{i0} && \text{for } \boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i, m_i > 0,\ 1 \leq i \leq k && (6.61)\\
&= \mu_i Q^*_{ij} && \text{for } \boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i + \boldsymbol{e}_j,\ \ m - i > 0,\ \ 1 \leq i,\ \ j \leq k. && (6.62)
\end{aligned}$$

If we substitute (6.59) into (6.60)-(6.62), we obtain

$$\begin{aligned}
q^*_{\boldsymbol{m},\boldsymbol{m}'} &= \lambda_j Q_{j0} && \text{for } \boldsymbol{m}' = \boldsymbol{m} + \boldsymbol{e}_j, \quad 1 \leq j \leq k && (6.63)\\
&= (\mu_i/\lambda_i)\lambda_0 Q_{0i} && \text{for } \boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i, m_i > 0, \quad 1 \leq i \leq k && (6.64)\\
&= (\mu_i/\lambda_i)\lambda_j Q_{ji} && \text{for } \boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i + \boldsymbol{e}_j, \qquad m_i > 0,\ 1 \leq i, j \leq k. && (6.65)
\end{aligned}$$

This gives us our hypothesized backward transition rates in terms of the parameters of the original Jackson network. To use theorem 6.6.2, we must verify that there is a set of positive numbers, $p(\boldsymbol{m})$, satisfying $\sum_{\boldsymbol{m}} p(\boldsymbol{m}) = 1$ and $\sum_{\boldsymbol{m}} \; \nu_{\boldsymbol{m}} p_{\boldsymbol{m}} < \infty$, and a set of nonnegative numbers $q^*_{\boldsymbol{m}',\boldsymbol{m}}$ satisfying the following two sets of equations:

$$
\begin{aligned}
p(\boldsymbol{m})q_{\boldsymbol{m},\boldsymbol{m}'} &= p(\boldsymbol{m}')q^*_{\boldsymbol{m}',\boldsymbol{m}} \qquad \text{for all } \boldsymbol{m}, \boldsymbol{m}' && (6.66)\\
\sum_{\boldsymbol{m}} q_{\boldsymbol{m},\boldsymbol{m}'} &= \sum_{\boldsymbol{m}'} q^*_{\boldsymbol{m},\boldsymbol{m}'} \qquad \text{for all } \boldsymbol{m}. && (6.67)
\end{aligned}
$$

We verify (6.66) by substituting (6.55)-(6.57) on the left side of (6.66) and (6.63)-(6.65) on the right side. Recalling that ρ_i is defined as λ_i/μ_i, and cancelling out common terms on each side, we have

$$
\begin{aligned}
p(\boldsymbol{m}) &= p(\boldsymbol{m}')/\rho_j \qquad \text{for } \boldsymbol{m}' = \boldsymbol{m} + \boldsymbol{e}_j && (6.68)\\
p(\boldsymbol{m}) &= p(\boldsymbol{m}')\rho_i \qquad \text{for } \boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i, m_i > 0 && (6.69)\\
p(\boldsymbol{m}) &= p(\boldsymbol{m}')\rho_i/\rho_j \qquad \text{for } \boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i + \boldsymbol{e}_j, \qquad m_i > 0. && (6.70)
\end{aligned}
$$

Looking at the case $\boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i$, and using this equation repeatedly to get from state $(0, 0, \dots, 0)$ up to an arbitrary m, we obtain

$$
p(\boldsymbol{m}) = p(0, 0, \dots, 0) \prod_{i=1}^{k} \} \rho_i^{m_i}. \qquad (6.71)
$$

It is easy to verify that (6.71) satisfies (6.68)-(6.70) for all possible transitions. Summing over all $\boldsymbol{m}$ to solve for $p(0, 0, \dots, 0)$, we get

$$
\begin{aligned}
1 = \sum_{m_1, m_2, \dots, m_k} p(\boldsymbol{m}) &= p(0, 0, \dots, 0) \sum_{m_1} \rho_1^{m_1} \sum_{m_2} \rho_2^{m_2} \dots \sum_{m_k} \rho_k^{m_k}\\
&= p(0, 0 \dots, 0)(1-\rho_1)^{-1}(1-\rho_2)^{-1} \dots (1-\rho_k)^{-1}.
\end{aligned}
$$

Thus, $p(0, 0, \dots, 0) = (1-\rho_1)(1-\rho_2)\dots(1-\rho_k)$, and substituting this in (6.71), we get

$$
p(\boldsymbol{m}) = \prod_{i=1}^{k} \} p_i(m_i) = \prod_{i=1}^{k} \Big[(1-\rho_i)\rho_i^{m_i}\Big]. \qquad (6.72)
$$

where $p_i(m) = (1-\rho_i)\rho_i^m$ is the steady-state distribution of a single M/M/1 queue. Now that we have found the steady-state distribution implied by our assumption about the backward process being a Jackson network, our remaining task is to verify (6.67)

To verify (6.67), *i.e.*, $\sum_{\boldsymbol{m}'} q_{\boldsymbol{m},\boldsymbol{m}'} = \sum_{\boldsymbol{m}'} q^*_{\boldsymbol{m},\boldsymbol{m}'}$, first consider the right side. Using (6.60) to sum over all $\boldsymbol{m}' = \boldsymbol{m} + \boldsymbol{e}_j$, then (6.61) to sum over $\boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i$ (for i such that $m_i > 0$), and finally (6.62) to sum over $\boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i + \boldsymbol{e}_j$, (again for i such that $m_i > 0$), we get

$$
\sum_{\boldsymbol{m}'} q^*_{\boldsymbol{m},\boldsymbol{m}'} = \sum_{j=1}^{k} \} \lambda_0 Q^*_{0j} + \sum_{i:m_j>0} \mu_i Q^*_{i0} + \sum_{i:m_i>0} \mu_i \sum_{j=1}^{k} Q^*_{ij}. \qquad (6.73)
$$

Using the fact Q^* is a stochastic matrix, then,

$$
\sum_{\boldsymbol{m}'} q^*_{\boldsymbol{m},\boldsymbol{m}'} = \lambda_0 + \sum_{i:m_i>0} \mu_i. \qquad (6.74)
$$

The left hand side of (6.67) can be summed in the same way to get the result on the right side of (6.74), but we can see that this must be the result by simply observing that λ_0 is the rate of exogenous arrivals and $\sum_{i:m_i>0} \mu_i$ is the overall rate of service completions in state $\boldsymbol{m}$. Note that this also verifies that $\nu_{\boldsymbol{m}} = \sum_{\boldsymbol{m}'} q_{\boldsymbol{m},\boldsymbol{m}'} \geq \lambda_0 + \sum_i \mu_i$, and since $\nu_{\boldsymbol{m}}$ is bounded, $\sum_{\boldsymbol{m}} \nu_{\boldsymbol{m}} p(\boldsymbol{m}) < \infty$. Since all the conditions of Theorem 6.6.2 are satisfied, $p(\boldsymbol{m})$, as given in (6.72), gives the steady-state probabilities for the Jackson network. This also verifies that the backward process is a Jackson network, and hence the exogenous departures are Poisson and independent.

Although the exogenous arrivals and departures in a Jackson network are Poisson, the endogenous processes of customers travelling from one node to another are typically not Poisson if there are feedback paths in the network. Also, although (6.72) shows that the numbers of customers at the different nodes are independent random variables at any given time in steady-state, it is not generally true that the number of customers at one node at one time is independent of the number of customers at another node at another time.

There are many generalizations of the reversibility arguments used above, and many network situations in which the nodes have independent states at a common time. We discuss just two of them here and refer to Kelly, [13], for a complete treatment.

For the first generalization, assume that the service time at each node depends on the number of customers at that node, *i.e.*, μ_i is replaced by μ_{i,m_i}. Note that this includes the M/M/m type of situation in which each node has several independent exponential servers. With this modification, the transition rates in (6.56) and (6.57) are modified by replacing μ_i with μ_{i,m_i}. The hypothesized backward transition rates are modified in the same way, and the only effect of these changes is to replace ρ_i and ρ_j for each i and j in (6.68)-(6.70) with $\rho_{i,m_i} = \lambda_i/\mu_{i,m_i}$ and $\rho_{j,m_j} = \lambda_j/\mu_{j,m_j}$. With this change, (6.71) becomes

$$p(\boldsymbol{m}) = \prod_{i=1}^{k} \} p_i(m_i) = \prod_{i=1}^{k} \} p_i(0) \prod_{j=0}^{m_i} \} \rho_{i,j} \qquad (6.75)$$

$$p_i(0) = \left[1 + \sum_{m=1}^{\infty} \prod_{j=0}^{m-1} \rho_{i,j} \right]^{-1}. \qquad (6.76)$$

Thus, $p(\boldsymbol{m})$ is given by the product distribution of k individual birth-death systems.

6.7.1 Closed Jackson networks

The second generalization is to a network of queues with a fixed number M of customers in the system and with no exogenous inputs or outputs. Such networks are called *closed Jackson* networks, whereas the networks analyzed above are often called *open Jackson* networks. Suppose a k node closed network has routing probabilities Q_{ij}, $1 \leq i,j \leq k$, where $\sum_j Q_{ij} = 1$, and has exponential service times of rate μ_i (this can be generalized to μ_{i,m_i} as above). We make the same assumptions as before about independence of service variables and routing variables, and assume that there is a path between each pair of nodes. Since $\{Q_{ij};\ 1 \leq i,j \leq k\}$ forms an irreducible stochastic matrix, there is a one dimensional set of solutions to the steady-state equations

$$\lambda_j = \sum_i \lambda_i Q_{ij}; \qquad 1 \le j \le k. \tag{6.77}$$

We interpret λ_i as the time-average rate of transitions that go into node i. Since this set of equations can only be solved within an unknown multiplicative constant, and since this constant can only be determined at the end of the argument, we define $\{\pi_i;\, 1 \le i \le k\}$ as the particular solution of (6.77) satisfying

$$\pi_j = \sum_i \pi_i Q_{ij};\ 1 \le j \le k; \quad \sum_i \pi_i = 1. \tag{6.78}$$

Thus, for all i, $\lambda_i = \alpha\pi_i$, where α is some unknown constant. The state of the Markov process is again taken as $\boldsymbol{m} = (m_1, m_2, \dots, m_k)$ with the condition $\sum_i\ m_i = M$. The transition rates of the Markov process are the same as for open networks, except that there are no exogenous arrivals or departures; thus (6.55)-(6.57) are replaced by

$$q_{\boldsymbol{m},\boldsymbol{m}'} = \mu_i Q_{ij} \quad \text{for } \boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i + \boldsymbol{e}_j, \qquad m_i > 0,\ \ 1 \le i,j \le k. \tag{6.79}$$

We hypothesize that the backward time process is also a closed Jackson network, and as before, we conclude that if the hypothesis is true, the backward transition rates should be

$$q^*_{\boldsymbol{m},\boldsymbol{m}'} = \mu_i Q^*_{ij} \quad \text{for } \boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i + \boldsymbol{e}_j,\ m_i > 0,\ 1 \le i,j \le k \tag{6.80}$$

$$\text{where } \lambda_i Q_{ij} = \lambda_j Q^*_{ji} \qquad \text{for } 1 \le i,j \le k. \tag{6.81}$$

In order to use Theorem 6.6.2 again, we must verify that a PMF $p(\boldsymbol{m})$ exists satisfying $p(\boldsymbol{m})q_{\boldsymbol{m},\boldsymbol{m}'} = p(\boldsymbol{m}')q^*_{\boldsymbol{m}',\boldsymbol{m}}$ for all possible states and transitions, and we must also verify that $\sum_{\boldsymbol{m}'} q_{\boldsymbol{m},\boldsymbol{m}'} = \sum_{\boldsymbol{m}'} q^*_{\boldsymbol{m},\boldsymbol{m}'}$ for all possible $\boldsymbol{m}$. This latter verification is virtually the same as before and is left as an exercise. The former verification, with the use of (72), (73), and (74), becomes

$$p(\boldsymbol{m})(\mu_i/\lambda_i) = p(\boldsymbol{m}')(\mu_j/\lambda_j) \quad \text{for } \boldsymbol{m}' = \boldsymbol{m} - \boldsymbol{e}_i + \boldsymbol{e}_j, \qquad m_i > 0. \tag{6.82}$$

Using the open network solution to guide our intuition, we see that the following choice of $p(\boldsymbol{m})$ satisfies (6.82) for all possible $\boldsymbol{m}$ (*i.e.*, all $\boldsymbol{m}$ such that $\sum_i\ m_i = M$)

$$p(\boldsymbol{m}) = A \prod_{i=1}^{k} \}(\lambda_i/\mu_i)^{m_i}; \quad \text{for } \boldsymbol{m} \text{ such that } \sum_i \} m_i = M. \tag{6.83}$$

The constant A is a normalizing constant, chosen to make $p(\boldsymbol{m})$ sum to unity. The problem with (6.83) is that we do not know λ_i (except within a multiplicative constant independent of i). Fortunately, however, if we substitute π_i/α for λ_i, we see that α is raised to the power $-M$, independent of the state $\boldsymbol{m}$. Thus, letting $A' = A\alpha - M$, our solution becomes

$$p(\boldsymbol{m}) = A' \prod_{i=1}^{K} \}(\pi_i/\mu_i)^{m_i}; \ \text{for } \boldsymbol{m} \text{ such that } \sum_i \} m_i = M. \tag{6.84}$$

$$\frac{1}{A'} = \sum_{\boldsymbol{m}:\sum_i m_i = M} \ \prod_{i=1}^{k} \} \ \frac{\pi_i}{\mu_i} \Big)^{m_i}. \tag{6.85}$$

Note that the steady-state distribution of the closed Jackson network has been found without solving for the time-average transition rates. Note also that the steady-state distribution looks very similar to that for an open network; that is, it is a product distribution over the nodes with a geometric type distribution within each node. This is somewhat misleading, however, since the constant A' can be quite difficult to calculate. It is surprising at first that the parameter of the geometric distribution can be changed by a constant multiplier in (6.84) and (6.85) (*i.e.*, π_i could be replaced with λ_i) and the solution does not change; the important quantity is the relative values of π_i/μ_i from one value of i to another rather than the absolute value.

In order to find λ_i (and this is important, since it says how quickly the system is doing its work), note that $\lambda_i = \mu_i \Pr\{m_i > 0\}$). Solving for $\Pr\{m_i > 0\}$ requires finding the constant A' in (6.79). In fact, the major difference between open and closed networks is that the relevant constants for closed networks are tedious to calculate (even by computer) for large networks and large M.

6.8 Semi-Markov processes

Semi-Markov processes are generalizations of Markov processes in which the time intervals between transitions have an arbitrary distribution rather than an exponential distribution. To be specific, there is an embedded Markov chain, $\{X_n;\ n \geq 0\}$ with a finite or countably infinite state space, and a sequence $\{U_n;\ n \geq 1\}$ of holding intervals between state transitions. The epochs at which state transitions occur are then given, for $n \geq 1$, as $S_n = \sum_{m=1}^{n} U_n$. The process starts at time 0 with S_0 defined to be 0. The semi-Markov process is then the continuous-time process $\{X(t);\ t \geq 0\}$ where, for each $n \geq 0$, $X(t) = X_n$ for t in the interval $S_n \leq X_n < S_{n+1}$. Initially, $X_0 = i$ where i is any given element of the state space.

The holding intervals $\{U_n;\ n \geq 1\}$ are nonnegative rv's that depend only on the current state X_{n-1} and the next state X_n. More precisely, given $X_{n-1} = j$ and $X_n = k$,say, the interval U_n is independent of $\{U_m;\ m < n\}$ and independent of $\{X_m;\ m < n-1\}$ The conditional distribution function for such an interval U_n is denoted by $G_{ij}(u)$, *i.e.*,

$$\Pr\{U_n \leq u \mid X_{n-1} = j, X_n = k\} = G_{jk}(u). \tag{6.86}$$

The dependencies between the rv's $\{X_n;\ n \geq 0\}$ and $\{U_n;\ n \geq 1$ is illustrated in Figure 6.17.

The conditional mean of U_n, conditional on $X_{n-1} = j$, $X_n = k$, is denoted $\overline{U}(j,k)$, i.e.,

$$\overline{U}(i,j) = \mathsf{E}\left[U_n \mid X_{n-1} = i, X_n = j\right] = \int_{u \geq 0} [1 - G_{ij}(u)] du. \tag{6.87}$$

A semi-Markov process evolves in essentially the same way as a Markov process. Given an initial state, $X_0 = i$ at time 0, a new state $X_1 = j$ is selected according to the embedded chain with probability P_{ij}. Then $U_1 = S_1$ is selected using the distribution $G_{ij}(u)$. Next a new state $X_2 = k$ is chosen according to the probability P_{jk}; then, given $X_1 = j$ and

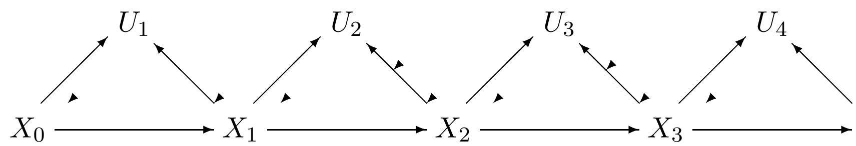


Figure 6.17: The statistical dependencies between the rv's of a semi-Markov process. Each holding interval U_n, conditional on the current state X_{n-1} and next state X_n, is independent of all other states and holding intervals. Note that, conditional on X_n, the holding intervals $U_n, U_{n-1}, \dots$, are statistically independent of $U_{n+1}, X_{n+2}, \dots$.

$X_2 = k$, the interval U_2 is selected with distribution function $G_{jk}(u)$. Successive state transitions and transition times are chosen in the same way.

The steady-state behavior of semi-Markov processes can be analyzed in virtually the same way as Markov processes. We outline this in what follows, and often omit proofs where they are the same as the corresponding proof for Markov processes. First, since the holding intervals, U_n, are rv's, the transition epochs, $S_n = \sum_{m=1}^{n} U_n$, are also rv's. The following lemma then follows in the same way as Lemma 6.2.1 for Markov processes.

Lemma 6.8.1. *Let $M_i(t)$ be the number of transitions in a semi-Markov process in the interval $(0, t]$ for some given initial state $X_0 = i$. Then* $\lim_{t\to\infty} M_i(t) = \infty$ WP1.

In what follows, we assume that the embedded Markov chain is irreducible and positive-recurrent. We want to find the limiting fraction of time that the process spends in any given state, say j. We will find that this limit exists WP1, and will find that it depends only on the steady-state probabilities of the embedded Markov chain and on the expected holding interval in each state. This is given by

$$\overline{U}(j) = \mathsf{E}\left[U_n \mid X_{n-1} = j\right] = \sum_k P_{jk} \mathsf{E}\left[U_n \mid X_{n-1} = j, X_n = k\right] = \sum_k P_{jk} \overline{U}(j,k), \qquad (6.88)$$

where $\overline{U}(j,k)$ is given in 6.87. The steady-state probabilities $\{\pi_i;\, i \geq 0\}$ for the embedded chain tell us the fraction of transitions that enter any given state i. Since $\overline{U}(i)$ is the expected holding interval in i per transition into i, we would guess that the fraction of time spent in state i should be proportional to $\pi_i \overline{U}(i)$. Normalizing, we would guess that the time-average probability of being in state i should be

$$p_j = \frac{\pi_j \overline{U}(j)}{\sum_k \pi_k \overline{U}(k)}. \qquad (6.89)$$

Identifying the mean holding interval, $\overline{U}_j$ with $1/\nu_j$, this is the same result that we established for the Markov process case. Using the same arguments, we find this is valid for the semi-Markov case. It is valid both in the conventional case where each p_j is positive and $\sum_j p_j = 1$, and also in the case where $\sum_k \pi_k \overline{U}(k) = \infty$, where each $p_j = 0$. The analysis is based on the fact that successive transitions to some given state, say j, given $X_0 = i$, form a delayed renewal process.

Lemma 6.8.2. *Consider a semi-Markov process with an irreducible recurrent embedded chain* $\{X_n; n \geq 0\}$*. Given* $X_0 = i$*, let* $\{M_{ij}(t); t \geq 0\}$ *be the number of transitions into a given state* j *in the interval* $(0, t]$*. Then* $\{M_{ij}(t); t \geq 0\}$ *is a delayed renewal process (or, if* $j = i$*, is an ordinary renewal process).*

This is the same as Lemma 6.2.2, but it is not quite so obvious that successive intervals between visits to state j are statistically independent. This can be seen, however, from Figure 6.17, which makes it clear that, given $X_n = j$, the future holding intervals, $U_n, U_{n+1}, \dots$, are independent of the past intervals $U_{n-1}, U_{n-2}, \dots$.

Next, using the same renewal reward process as in Lemma 6.2.3, assigning reward 1 whenever $X(t) = j$, we define W_n as the interval between the $n-1$st and the nth entry to state j and get the following lemma:

Lemma 6.8.3. *Consider a semi-Markov process with an irreducible, recurrent, embedded Markov chain starting in* $X_0 = i$*. Then with probability 1, the limiting time-average in state* j *is given by* $p_j(i) = \frac{\overline{U}_j}{\overline{W}(j)}$.

This lemma has omitted any assertion about the limiting ensemble probability of state j, *i.e.*, $\lim_{t\to\infty} \Pr\{X(t) = j\}$. This follows easily from Blackwell's theorem, but depends on whether the successive entries to state j, *i.e.*, $\{W_n;\ n \geq 1\}$, are arithmetic or non-arithmetic. This is explored in Exercise 6.33. The lemma shows (as expected) that the limiting time-average in each state is independent of the starting state, so we henceforth replace $p_j(i)$ with p_j.

Next, let $M_i(t)$ be the total number of transitions in the semi-Markov process up to and including time n, given $X_0 = i$. This is not a renewal counting process, but, as with Markov processes, it provides a way to combine the time-average results for all states j. The following theorem is the same as that for Markov processes, except for the omission of ensemble average results.

Theorem 6.8.1. *Consider a semi Markov process with an irreducible, positive-recurrent, embedded Markov chain. Let* $\{\pi_j;\ j \geq 0\}$ *be the steady-state probabilities of the embedded chain and let* $X_0 = i$ *be the starting state. Then, with probability 1, the limiting time-average fraction of time spent in any arbitrary state* j *is given by*

$$p_j = \frac{\pi_j \overline{U}(j)}{\sum_k \pi_k \overline{U}(j)}. \tag{6.90}$$

The expected time between returns to state j *is*

$$\overline{W}(j) = \frac{\sum_k \pi_k \overline{U}(k)}{\pi_j}, \tag{6.91}$$

and the rate at which transitions take place is independent of X_0 *and given by*

$$\lim_{t\to\infty} \frac{M_i(t)}{t} = \frac{1}{\sum_k \pi_k \overline{U}(k)} \qquad \text{WP1.} \tag{6.92}$$

For a semi-Markov process, knowing the steady state probability of $X(t) = j$ for large t does not completely specify the steady-state behavior. Another important steady state question is to determine the fraction of time involved in i to j transitions. To make this notion precise, define $Y(t)$ as the residual time until the next transition after time t (i e., $t + Y(t)$ is the epoch of the next transition after time t). We want to determine the fraction of time t over which $X(t) = i$ and $X(t + Y(t)) = j$. Equivalently, for a non-arithmetic process, we want to determine $\Pr\{X(t) = i, X(t + Y(t) = j)\}$ in the limit as $t \to \infty$. Call this limit $Q(i, j)$.

Consider a renewal process, starting in state i and with renewals on transitions to state i. Define a reward $R(t) = 1$ for $X(t) = i$, $X(t + Y(t)) = j$ and $R(t) = 0$ otherwise (see Figure 6.18). That is, for each n such that $X(S_n) = i$ and $X(S_{n+1}) = j$, $R(t) = 1$ for $S_n \le t < S_{n+1}$. The expected reward in an inter-renewal interval is then $P_{ij}\overline{U}(i, j)$. It follows that $Q(i, j)$ is given by

$$Q(i, j) = \lim_{t\to\infty} \frac{\int_0^t \ R(\tau)(d\tau}{t} = \frac{P_{ij}\overline{U}(i, j)}{\overline{W}(i)} = \frac{p_i P_{ij}\overline{U}(i, j)}{\overline{U}(i)}. \tag{6.93}$$

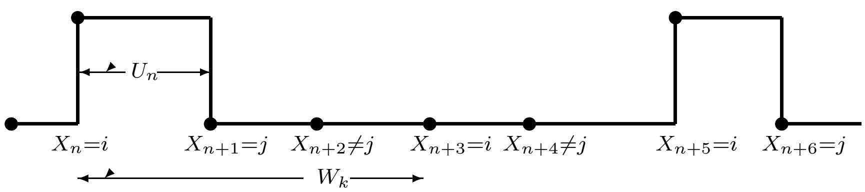


Figure 6.18: The renewal-reward process for i to j transitions. The expected value of U_n if $X_n = i$ and $X_{n+1} = j$ is $\overline{U}(i, j)$ and the expected interval between entries to i is $\overline{W}(i)$.

6.8.1 Example — the M/G/1 queue

As one example of a semi-Markov chain, consider an M/G/1 queue. Rather than the usual interpretation in which the state of the system is the number of customers in the system, we view the state of the system as changing only at departure times; the new state at a departure time is the number of customers left behind by the departure. This state then remains fixed until the next departure. New customers still enter the system according to the Poisson arrival process, but these new customers are not considered as part of the state until the next departure time. The number of customers in the system at arrival epochs does not in general constitute a "state" for the system, since the age of the current service is also necessary as part of the statistical characterization of the process.

One purpose of this example is to illustrate that it is often more convenient to visualize the transition interval $U_n = S_n - S_{n-1}$ as being chosen first and the new state X_n as being chosen second rather than choosing the state first and the transition time second. For the M/G/1 queue, first suppose that the state is some $i > 0$. In this case, service begins on the next customer immediately after the old customer departs. Thus, U_n, conditional on $X_n = i$ for $i > 0$, has the distribution of the service time, say $G(u)$. The mean interval until a state transition occurs is

$$\overline{U}(i) = \int_0^\infty [1 - G(u)]du; \quad i > 0. \tag{6.94}$$

Given the interval u for a transition from state $i > 0$, the number of arrivals in that period is a Poisson random variable with mean λu, where λ is the Poisson arrival rate. Since the next state j is the old state i, plus the number of new arrivals, minus the single departure,

$$\Pr\{X_{n+1} = j \mid X_n = i, U_n = u\} = \frac{(\lambda u)^{j+i+1} \exp(-\lambda u)}{(j - i + 1)!}. \tag{6.95}$$

for $j \geq i - 1$. For $j < i - 1$, the probability above is 0. The unconditional probability P_{ij} of a transition from i to j can then be found by multiplying the right side of (6.95) by the probability density $g(u)$ of the service time and integrating over u.

$$P_{ij} = \int_0^\infty \frac{G(u)(\lambda u)^{j-i+1} \exp(-\lambda u)}{(j - i + 1)} du; \quad j \geq i - 1, i > 0. \tag{6.96}$$

For the case $i = 0$, the server must wait until the next arrival before starting service. Thus the expected time from entering the empty state until a service completion is

$$\overline{U}(0) = (1/\lambda) + \int_0^\infty [1 - G(u)]du. \tag{6.97}$$

We can evaluate P_{0j} by observing that the departure of that first arrival leaves j customers in this system if and only if j customers arrive during the service time of that first customer; *i.e.*, the new state doesn't depend on how long the server waits for a new customer to serve, but only on the arrivals while that customer is being served. Letting $g(u)$ be the density of the service time,

$$P_{0j} = \int_0^\infty \frac{g(u)(\lambda u)^j \exp(-\lambda u)}{j!} du; \quad j \geq 0. \tag{6.98}$$

6.9 Summary

We have seen that Markov processes with countable state spaces are remarkably similar to Markov chains with countable state spaces, and throughout the chapter, we frequently made use of both the embedded chain corresponding to the process and to the sampled time approximation to the process.

For irreducible processes, the steady-state equations, (6.23) and $\sum_i p_i \nu_i = 1$, were found to specify the steady-state probabilities, p_i, which have significance both as time-averages and as limiting probabilities. If the transition rates ν_i are bounded, then the sampled-time approximation exists and has the same steady-state probabilities as the Markov process itself. If the transition rates ν_i are unbounded but $\sum_i p_i \nu_i < \infty$, then the embedded chain is positive recurrent and has steady-state probabilities, but the sampled-time approximation does not exist. We assumed throughout the remainder of the chapter that $\sum_i p_i\nu_i < \infty$. This ruled out irregular processes in which there is no meaningful steady state, and also some peculiar processes such as that in Exercise 6.9 where the embedded chain is null recurrent.

Section 6.3 developed the Kolmogoroff backward and forward differential equations for the transient probabilities $P_{ij}(t)$ of being in state j at time t given state i at time 0. We showed that for finite-state processes, these equations can be solved by finding the eigenvalues and eigenvectors of the transition rate matrix Q. There are close analogies between this analysis and the algebraic treatment of finite-state Markov chains in chapter 3, and exercise 6.7 showed how the transients of the process are related to the transients of the sampled time approximation.

For irreducible processes with bounded transition rates, uniformization was introduced as a way to simplify the structure of the process. The addition of self transitions does not change the process itself, but can be used to adjust the transition rates ν_i to be the same for all states. This changes the embedded Markov chain, and the steady-state probabilities for the embedded chain become the same as those for the process. The epochs at which transitions occur then form a Poisson process which is independent of the set of states entered. This yields a separation between the transition epochs and the sequence of states.

The next two sections analyzed birth-death processes and reversibility. The results about birth-death Markov chains and reversibility for Markov chains carried over almost without change to Markov processes. These results are central in queueing theory, and Burke's theorem allowed us to look at simple queueing networks with no feedback and to understand how feedback complicates the problem.

Jackson networks were next discussed. These are important in their own right and also provide a good example of how one can solve complex queueing problems by studying the reverse time process and making educated guesses about the steady-state behavior. The somewhat startling result here is that in steady state, and at a fixed time, the number of customers at each node is independent of the number at each other node and satisfies the same distribution as for an M/M/1 queue. Also the exogenous departures from the network are Poisson and independent from node to node. We emphasized that the number of customers at one node at one time is often dependent on the number at other nodes at other times. The independence holds only when all nodes are viewed at the same time.

Finally, semi-Markov processes were introduced. Renewal theory again provided the key to analyzing these systems. Theorem 6.8.1 showed how to find the steady-state probabilities of these processes, and it was shown that these probabilities could be interpreted both as time-averages and, in the case of non-arithmetic transition times, as limiting probabilities in time.

For further reading on Markov processes, see [13], [16], [22], and [8].

6.10 Exercises

Exercise 6.1. Consider an M/M/1 queue as represented in Figure 6.4. Assume throughout that $X_0 = i$ where $i > 0$. The purpose of this exercise is to understand the relationship between the holding interval until the next state transition and the interval until the next arrival to the M/M/1 queue. Your explanations in the following parts can and should be very brief.

a) Explain why the expected holding interval $\mathsf{E}[U_1|X_0 = i]$ until the next state transition is $1/(\lambda + \mu)$.

b) Explain why the expected holding interval U_1, conditional on $X_0 = i$ and $X_1 = i + 1$, is

$$\mathsf{E}[U_1|X_0 = i,\, X_1 = i + 1] = 1/(\lambda + \mu).$$

Show that $\mathsf{E}[U_1|X_0 = i, X_1 = i - 1]$ is the same.

c) Let V be the time of the first arrival after time 0 (this may occur either before or after the time W of the first departure.) Show that

$$\begin{aligned}\mathsf{E}[V|X_0 = i,\, X_1 = i + 1] &= \frac{1}{\lambda + \mu} \\ \mathsf{E}[V|X_0 = i,\, X_1 = i - 1] &= \frac{1}{\lambda + \mu} + \frac{1}{\lambda}.\end{aligned}$$

Hint: In the second equation, use the memorylessness of the exponential rv and the fact that V under this condition is the time to the first departure plus the remaining time to an arrival.

d) Use your solution to part c) plus the probability of moving up or down in the Markov chain to show that $\mathsf{E}[V] = 1/\lambda$. (Note: you already know that $\mathsf{E}[V] = 1/\lambda$. The purpose here is to show that your solution to part c) is consistent with that fact.)

Exercise 6.2. Consider a Markov process for which the embedded Markov chain is a birth-death chain with transition probabilities $P_{i,i+1} = 2/5$ for all $i \geq 0$, $P_{i,i-1} = 3/5$ for all $i \geq 1$, $P_{01} = 1$, and $P_{ij} = 0$ otherwise.

a) Find the steady-state probabilities $\{\pi_i;\, i \geq 0\}$ for the embedded chain.

b) Assume that the transition rate ν_i out of state i, for $i \geq 0$, is given by $\nu_i = 2^i$. Find the transition rates $\{q_{ij}\}$ between states and find the steady-state probabilities $\{p_i\}$ for the Markov process. Explain heuristically why $\pi_i \neq p_i$.

c) Explain why there is no sampled-time approximation for this process. Then truncate the embedded chain to states 0 to m and find the steady-state probabilities for the sampled-time approximation to the truncated process.

d) Show that as $m \to \infty$, the steady-state probabilities for the sequence of sampled-time approximations approaches the probabilities p_i in part b).

Exercise 6.3. Consider a Markov process for which the embedded Markov chain is a birth-death chain with transition probabilities $P_{i,i+1} = 2/5$ for all $i \geq 1$, $P_{i,i-1} = 3/5$ for all $i \geq 1$, $P_{01} = 1$, and $P_{ij} = 0$ otherwise.

a) Find the steady-state probabilities $\{\pi_i;\ i \geq 0\}$ for the embedded chain.

b) Assume that the transition rate out of state i, for $i \geq 0$, is given by $\nu_i = 2^{-i}$. Find the transition rates $\{q_{ij}\}$ between states and show that there is no probability vector solution $\{p_i;\ i \geq 0\}$ to (6.23).

e) Argue that the expected time between visits to any given state i is infinite. Find the expected number of transitions between visits to any given state i. Argue that, starting from any state i, an eventual return to state i occurs with probability 1.

f) Consider the sampled-time approximation of this process with $\delta = 1$. Draw the graph of the resulting Markov chain and argue why it must be null-recurrent.

Exercise 6.4. Consider the Markov process for the M/M/1 queue, as given in Figure 6.4.

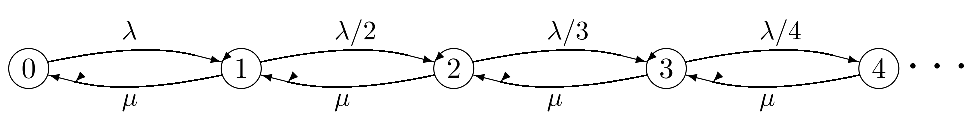
a) Find the steady state process probabilities (as a function of $\rho = \lambda/\mu$) from (6.15) and also as the solution to (6.23). Verify that the two solutions are the same.

b) For the remaining parts of the exercise, assume that $\rho = 0.01$, thus ensuring (for aiding intuition) that states 0 and 1 are much more probable than the other states. Assume that the process has been running for a very long time and is in steady state. Explain in your own words the difference between π_1 (the steady-state probability of state 1 in the embedded chain) and p_1 (the steady-state probability that the process is in state 1. More explicitly, what experiments could you perform (repeatedly) on the process to measure π_1 and p_1.

c) Now suppose you want to start the process in steady state. Show that it is impossible to choose initial probabilities so that both the process and the embedded chain start in steady state. Which version of steady state is closest to your intuitive view? (There is no correct answer here, but it is important to realiize that the notion of steady state is not quite as simple as you might imagine).

d) Let $M(t)$ be the number of transitions (counting both arrivals and departures) that take place by time t in this Markov process and assume that the embedded Markov chain starts in steady state at time 0. Let $U_1, U_2, \ldots$, be the sequence of holding intervals between transitions (with U_1 being the time to the first transition). Show that these rv's are identically distributed. Show by example that they are not independent (*i.e.*, $M(t)$ is not a renewal process).

Exercise 6.5. Consider the Markov process illustrated below. The transitions are labelled by the rate q_{ij} at which those transitions occur. The process can be viewed as a single server queue where arrivals become increasingly discouraged as the queue lengthens. The word *time-average* below refers to the limiting time-average over each sample-path of the process, except for a set of sample paths of probability 0.



a) Find the time-average fraction of time p_i spent in each state $i > 0$ in terms of p_0 and then solve for p_0. Hint: First find an equation relating p_i to p_{i+1} for each i. It also may help to recall the power series expansion of e^x.

b) Find a closed form solution to $\sum_i p_i \nu_i$ where ν_i is the departure rate from state i. Show that the process is positive recurrent for all choices of $\lambda > 0$ and $\mu > 0$ and explain intuitively why this must be so.

c) For the embedded Markov chain corresponding to this process, find the steady-state probabilities π_i for each $i \geq 0$ and the transition probabilities P_{ij} for each i, j.

d) For each i, find both the time-average interval and the time-average number of overall state transitions between successive visits to i.

Exercise 6.6. (Detail from proof of Theorem 6.2.1) **a)** Let U_n be the nth holding interval for a Markov process starting in steady state, with $\Pr\{X_0 = i\} = \pi_i$. Show that $\mathsf{E}[U_n] = \sum_k \pi_k/\nu_k$ for each integer n.

b) Let S_n be the epoch of the nth transition. Show that $\Pr\{S_n \geq n\beta\} \leq \frac{\sum_k \pi_k/\nu_k}{\beta}$ for all $\beta > 0$.

c) Let $M(t)$ be the number of transitions of the Markov process up to time t, given that X_0 is in steady state. Show that $\Pr\{M(n\beta) < n\} > 1 - \frac{\sum_k \pi_k/\nu_k}{\beta}$.

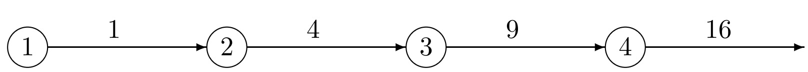
d) Show that if $\sum_k \pi_k/\nu_k$ is finite, then $\lim_{t\to\infty} M(t)/t = 0$ WP1 is impossible. (Note that this is equivalent to showing that $\lim_{t\to\infty} M(t)/t = 0$ WP1 implies $\sum_k \pi_k/\nu_k = \infty$).

e) Let $M_i(t)$ be the number of transitions by time t starting in state i. Show that if $\sum_k \pi_k/\nu_k$ is finite, then $\lim_{t\to\infty} M_i(t)/t = 0$ WP1 is impossible.

Exercise 6.7. **a)** Consider the process in the figure below. The process starts at $X(0) = 1$, and for all $i \geq 1$, $P_{i,i+1} = 1$ and $\nu_i = i^2$ for all i. Let S_n be the epoch when the n^{th} transition occurs. Show that

$$\mathsf{E}[S_n] = \sum_{i=1}^{n} \} i^{-2} < 2 \text{ for all } n.$$

Hint: Upper bound the sum from $i = 2$ by integrating x^{-2} from $x = 1$.



b) Use the Markov inequality to show that $\Pr\{S_n > 4\} \leq 1/2$ for all n. Show that the probability of an infinite number of transitions by time 4 is at least 1/2.

Exercise 6.8. **a)** Consider a Markov process with the set of states $\{0, 1, \dots\}$ in which the transition rates $\{q_{ij}\}$ between states are given by $q_{i,i+1} = (3/5)2^i$ for $i \geq 0$, $q_{i,i-1} = (2/5)2^i$

for $i \geq 1$, and $q_{ij} = 0$ otherwise. Find the transition rate ν_i out of state i for each $i \geq 0$ and find the transition probabilities $\{P_{ij}\}$ for the embedded Markov chain.

b) Find a solution $\{p_i;\, i \geq 0\}$ with $\sum_i p_i = 1$ to (6.23).

c) Show that all states of the embedded Markov chain are transient.

d) Explain in your own words why your solution to part b) is not in any sense a set of steady-state probabilities

Exercise 6.9. Let $q_{i,i+1} = 2^{i-1}$ for all $i \geq 0$ and let $q_{i,i-1} = 2^{i-1}$ for all $i \geq 1$. All other transition rates are 0.

a) Solve the steady-state equations and show that $p_i = 2^{-i-1}$ for all $i \geq 0$.

b) Find the transition probabilities for the embedded Markov chain and show that the chain is null recurrent.

c) For any state i, consider the renewal process for which the Markov process starts in state i and renewals occur on each transition to state i. Show that, for each $i \geq 1$, the expected inter-renewal interval is equal to 2. Hint: Use renewal-reward theory.

d) Show that the expected number of transitions between each entry into state i is infinite. Explain why this does *not* mean that an infinite number of transitions can occur in a finite time.

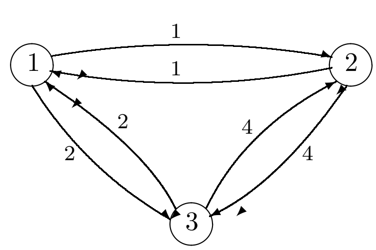
Exercise 6.10. **a)** Consider the two state Markov process of Example 6.3.1 with $q_{01} = \lambda$ and $q_{10} = \mu$. Find the eigenvalues and eigenvectors of the transition rate matrix $[Q]$.

b) If $[Q]$ has M distinct eigenvalues, the differential equation $d[P(t)]/dt = [Q][P(t)]$ can be solved by the equation

$$[P(t)] = \sum_{i=1}^{\mathsf{M}} \}\boldsymbol{\nu}_i e^{t\lambda_i} \boldsymbol{p}^{\mathsf{T}}{}_i,$$

where $\boldsymbol{p}_i$ and $\boldsymbol{\nu}_i$ are the left and right eigenvectors of eigenvalue λ_i. Show that this equation gives the same solution as that given for Example 6.3.1.

Exercise 6.11. Consider the three state Markov process below; the number given on edge (i, j) is q_{ij}, the transition rate from i to j. Assume that the process is in steady state.



a) Is this process reversible?

b) Find p_i, the time-average fraction of time spent in state i for each i.

c) Given that the process is in state i at time t, find the mean delay from t until the process leaves state i.

d) Find π_i, the time-average fraction of all transitions that go into state i for each i.

e) Suppose the process is in steady state at time t. Find the steady state probability that the next state to be entered is state 1.

f) Given that the process is in state 1 at time t, find the mean delay until the process first returns to state 1.

g) Consider an arbitrary irreducible finite-state Markov process in which $q_{ij} = q_{ji}$ for all i, j. Either show that such a process is reversible or find a counter example.

Exercise 6.12. a) Consider an M/M/1 queueing system with arrival rate λ, service rate μ, $\mu > \lambda$. Assume that the queue is in steady state. Given that an arrival occurs at time t, find the probability that the system is in state i immediately *after* time t.

b) Assuming FCFS service, and conditional on i customers in the system immediately after the above arrival, characterize the time until the above customer departs as a sum of random variables.

c) Find the unconditional probability density of the time until the above customer departs. **Hint:** You know (from splitting a Poisson process) that the sum of a geometrically distributed number of IID exponentially distributed random variables is exponentially distributed. Use the same idea here.

Exercise 6.13. a) Consider an M/M/1 queue in steady state. Assume $\rho = \lambda/\mu < 1$. Find the probability $Q(i, j)$ for $i \geq j > 0$ that the system is in state i at time t and that $i - j$ departures occur before the next arrival.

b) Find the PMF of the state immediately before the first arrival after time t.

c) There is a well-known queueing principle called PASTA, standing for "Poisson arrivals see time-averages". Given your results above, give a more precise statement of what that principle means in the case of the M/M/1 queue.

Exercise 6.14. A small bookie shop has room for at most two customers. Potential customers arrive at a Poisson rate of 10 customers per hour; they enter if there is room and are turned away, never to return, otherwise. The bookie serves the admitted customers in order, requiring an exponentially distributed time of mean 4 minutes per customer.

a) Find the steady-state distribution of number of customers in the shop.

b) Find the rate at which potential customers are turned away.

c) Suppose the bookie hires an assistant; the bookie and assistant, working together, now serve each customer in an exponentially distributed time of mean 2 minutes, but there is only room for one customer (*i.e.*, the customer being served) in the shop. Find the new rate at which customers are turned away.

Exercise 6.15. This exercise explores a continuous time version of a simple branching process.

Consider a population of primitive organisms which do nothing but procreate and die. In particular, the population starts at time 0 with one organism. This organism has an exponentially distributed lifespan T_0 with rate μ (*i.e.*, $\Pr\{T_0 \geq \tau\} = e^{-\mu\tau}$). While this organism is alive, it gives birth to new organisms according to a Poisson process of rate λ. Each of these new organisms, while alive, gives birth to yet other organisms. The lifespan and birthrate for each of these new organisms are independently and identically distributed to those of the first organism. All these and subsequent organisms give birth and die in the same way, again indepedently of all other organisms.

a) Let $X(t)$ be the number of (live) organisms in the population at time t. Show that $\{X(t);\ t \geq 0\}$ is a Markov process and specify the transition rates between the states.

b) Find the embedded Markov chain $\{X_n;\ n \geq 0\}$ corresponding to the Markov process in part a). Find the transition probabilities for this Markov chain.

c) Explain why the Markov process and Markov chain above are not irreducible. Note: The majority of results you have seen for Markov processes assume the process is irreducible, so be careful not to use those results in this exercise.

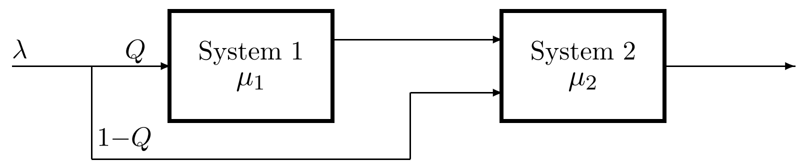
d) For purposes of analysis, add an additional transition of rate λ from state 0 to state 1. Show that the Markov process and the embedded chain are irreducible. Find the values of λ and μ for which the modified chain is positive recurrent, null-recurrent, and transient.

e) Assume that $\lambda < \mu$. Find the steady state process probabilities for the modified Markov process.

f) Find the mean recurrence time between visits to state 0 for the modified Markov process.

g) Find the mean time $\overline{T}$ for the population in the original Markov process to die out. Note: We have seen before that removing transitions from a Markov chain or process to create a trapping state can make it easier to find mean recurrence times. This is an example of the opposite, where adding an exit from a trapping state makes it easy to find the recurrence time.

Exercise 6.16. Consider the job sharing computer system illustrated below. Incoming jobs arrive from the left in a Poisson stream. Each job, independently of other jobs, requires pre-processing in system 1 with probability Q. Jobs in system 1 are served FCFS and the service times for successive jobs entering system 1 are IID with an exponential distribution of mean $1/\mu_1$. The jobs entering system 2 are also served FCFS and successive service times are IID with an exponential distribution of mean $1/\mu_2$. The service times in the two systems are independent of each other and of the arrival times. Assume that $\mu_1 > \lambda Q$ and that $\mu_2 > \lambda$. Assume that the combined system is in steady state.



a Is the input to system 1 Poisson? Explain.

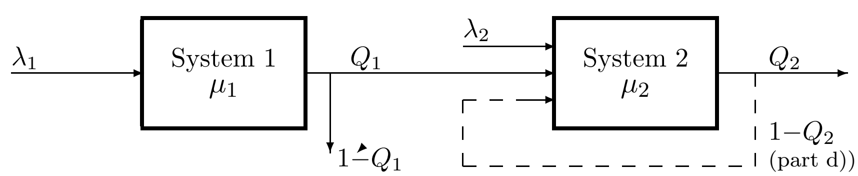
b) Are each of the two input processes coming into system 2 Poisson? Explain.

d) Give the joint steady-state PMF of the number of jobs in the two systems. Explain briefly.

e) What is the probability that the first job to leave system 1 after time t is the same as the first job that entered the entire system after time t?

f) What is the probability that the first job to leave system 2 after time t both passed through system 1 and arrived at system 1 after time t?

Exercise 6.17. Consider the following combined queueing system. The first queue system is M/M/1 with service rate μ_1. The second queue system has IID exponentially distributed service times with rate μ_2. Each departure from system 1 independently $1 - Q_1$. System 2 has an additional Poisson input of rate λ_2, independent of inputs and outputs from the first system. Each departure from the second system independently leaves the combined system with probability Q_2 and re-enters system 2 with probability $1 - Q_2$. For parts a), b), c) assume that $Q_2 = 1$ (*i.e.*, there is no feedback).



a) Characterize the process of departures from system 1 that enter system 2 and characterize the overall process of arrivals to system 2.

b) Assuming FCFS service in each system, find the steady state distribution of time that a customer spends in each system.

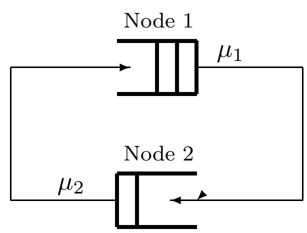
c) For a customer that goes through both systems, show why the time in each system is independent of that in the other; find the distribution of the combined system time for such a customer.

d) Now assume that $Q_2 < 1$. Is the departure process from the combined system Poisson? Which of the three input processes to system 2 are Poisson? Which of the input processes are independent? Explain your reasoning, but do not attempt formal proofs.

Exercise 6.18. Suppose a Markov chain with transition probabilities $\{P_{ij}\}$ is reversible. Suppose we change the transition probabilities out of state 0 from $\{P_{0j}; j \geq 0\}$ to $\{P'_{0j}; j \geq$

0}. Assuming that all P_{ij} for $i \neq 0$ are unchanged, what is the most general way in which we can choose $\{P'_{0j};\ j \geq 0\}$ so as to maintain reversibility? Your answer should be explicit about how the steady-state probabilities $\{\pi_i;\ i \geq 0\}$ are changed. Your answer should also indicate what this problem has to do with uniformization of reversible Markov processes, if anything. Hint: Given P_{ij} for all i, j, a single additional parameter will suffice to specify P'_{0j} for all j.

Exercise 6.19. Consider the closed queueing network in the figure below. There are three customers who are doomed forever to cycle between node 1 and node 2. Both nodes use FCFS service and have exponentially distributed IID service times. The service times at one node are also independent of those at the other node and are independent of the customer being served. The server at node i has mean service time $1/\mu_i$, $i = 1, 2$. Assume to be specific that $\mu_2 < \mu_1$.



a) The system can be represented by a four state Markov process. Draw its graphical representation and label it with the individual states and the transition rates between them.

b) Find the steady-state probability of each state.

c) Find the time-average rate at which customers leave node 1.

d) Find the time-average rate at which a given customer cycles through the system.

e) Is the Markov process reversible? Suppose that the backward Markov process is interpreted as a closed queueing network. What does a departure from node 1 in the forward process correspond to in the backward process? Can the transitions of a single customer in the forward process be associated with transitions of a single customer in the backward process?

Exercise 6.20. Consider an M/G/1 queueing system with last come first serve (LCFS) pre-emptive resume service. That is, customers arrive according to a Poisson process of rate λ. A newly arriving customer interrupts the customer in service and enters service itself. When a customer is finished, it leaves the system and the customer that had been interrupted by the departing customer resumes service from where it had left off. For example, if customer 1 arrives at time 0 and requires 2 units of service, and customer 2 arrives at time 1 and requires 1 unit of service, then customer 1 is served from time 0 to 1; customer 2 is served from time 1 to 2 and leaves the system, and then customer 1 completes service from time 2 to 3. Let X_i be the service time required by the i^{th} customer; the X_i are IID random variables with expected value $\mathsf{E}[X]$; they are independent of customer arrival times. Assume $\lambda\,\mathsf{E}[X] < 1$.

a) Find the mean time between busy periods (*i.e.*, the time until a new arrival occurs after the system becomes empty).

b) Find the time-average fraction of time that the system is busy.

c) Find the mean duration, $\mathsf{E}[B]$, of a busy period. Hint: use a) and b).

d) Explain briefly why the customer that starts a busy period remains in the system for the entire busy period; use this to find the expected system time of a customer given that that customer arrives when the system is empty.

e) Is there any statistical dependence between the system time of a given customer (*i.e.*, the time from the customer's arrival until departure) and the number of customers in the system when the given customer arrives?

f) Show that a customer's expected system time is equal to $\mathsf{E}[B]$. Hint: Look carefully at your answers to **d)** and **e)**.

g) Let C be the expected system time of a customer conditional on the service time X of that customer being 1. Find (in terms of C) the expected system time of a customer conditional on $X = 2$; (Hint: compare a customer with $X = 2$ to two customers with $X = 1$ each); repeat for arbitrary $X = x$.

h) Find the constant C. Hint: use **f)** and **g)**; don't do any tedious calculations.

Exercise 6.21. Consider a queueing system with two classes of customers, type A and type B. Type A customers arrive according to a Poisson process of rate λ_A and customers of type B arrive according to an independent Poisson process of rate λ_B. The queue has a FCFS server with exponentially distributed IID service times of rate $\mu > \lambda_A + \lambda_B$. Characterize the departure process of class A customers; explain carefully. Hint: Consider the combined arrival process and be judicious about how to select between A and B types of customers.

Exercise 6.22. Consider a pre-emptive resume last come first serve (LCFS) queueing system with two classes of customers. Type A customer arrivals are Poisson with rate λ_A and Type B customer arrivals are Poisson with rate λ_B. The service time for type A customers is exponential with rate μ_A and that for type B is exponential with rate μ_B. Each service time is independent of all other service times and of all arrival epochs.

Define the "state" of the system at time t by the string of customer types in the system at t, in order of arrival. Thus state AB means that the system contains two customers, one of type A and the other of type B; the type B customer arrived later, so is in service. The set of possible states arising from transitions out of AB is as follows:

ABA if another type A arrives.

ABB if another type B arrives.

A if the customer in service (B) departs.

Note that whenever a customer completes service, the next most recently arrived resumes service, so the state changes by eliminating the final element in the string.

a) Draw a graph for the states of the process, showing all states with 2 or fewer customers and a couple of states with 3 customers (label the empty state as E). Draw an arrow, labelled by the rate, for each state transition. Explain why these are states in a Markov process.

b) Is this process reversible. Explain. Assume positive recurrence. Hint: If there is a transition from one state S to another state S', how is the number of transitions from S to S' related to the number from S' to S?

c) Characterize the process of type A departures from the system (*i.e.*, are they Poisson?; do they form a renewal process?; at what rate do they depart?; etc.)

d) Express the steady-state probability $\Pr\{A\}$ of state A in terms of the probability of the empty state $\Pr\{E\}$. Find the probability $\Pr\{AB\}$ and the probability $\Pr\{ABBA\}$ in terms of $\Pr\{E\}$. Use the notation $\rho_A = \lambda_A/\mu_A$ and $\rho_B = \lambda_B/\mu_B$.

e) Let Q_n be the probability of n customers in the system, as a function of $Q_0 = \Pr\{E\}$. Show that $Q_n = (1-\rho)\rho^n$ where $\rho = \rho_A + \rho_B$.

Exercise 6.23. a) Generalize Exercise 6.22 to the case in which there are m types of customers, each with independent Poisson arrivals and each with independent exponential service times. Let λ_i and μ_i be the arrival rate and service rate respectively of the i^{th} user. Let $\rho_i = \lambda_i/\mu_i$ and assume that $\rho = \rho_1 + \rho_2 + \cdots + \rho_m < 1$. In particular, show, as before that the probability of n customers in the system is $Q_n = (1-\rho)\rho^n$ for $0 \leq n < \infty$.

b) View the customers in part a) as a single type of customer with Poisson arrivals of rate $\lambda = \sum_i \lambda_i$ and with a service density $\sum_i (\lambda_i/\lambda)\mu_i \exp(-\mu_i x)$. Show that the expected service time is ρ/λ. Note that you have shown that, if a service distribution can be represented as a weighted sum of exponentials, then the distribution of customers in the system for LCFS service is the same as for the M/M/1 queue with equal mean service time.

Exercise 6.24. Consider a k node Jackson type network with the modification that each node i has s servers rather than one server. Each server at i has an exponentially distributed service time with rate μ_i. The exogenous input rate to node i is $\rho_i = \lambda_0 Q_{0i}$ and each output from i is switched to j with probability Q_{ij} and switched out of the system with probability Q_{i0} (as in the text). Let $\lambda_i, 1 \leq i \leq k$, be the solution, for given λ_0, to

$$\lambda_j = \sum_{i=0}^{k} \}\lambda_i Q_{ij};$$

$1 \leq j \leq k$ and assume that $\lambda_i < s\mu_i$; $1 \leq i \leq k$. Show that the steady-state probability of state $\boldsymbol{m}$ is

$$\Pr\{\boldsymbol{m}\} = \prod_{i=1}^{k} \} p_i(m_i),$$

where $p_i(m_i)$ is the probability of state m_i in an (M, M, s) queue. Hint: simply extend the argument in the text to the multiple server case.

Exercise 6.25. Suppose a Markov process with the set of states A is reversible and has steady-state probabilities p_i; $i \in A$. Let B be a subset of A and assume that the process is changed by setting $q_{ij} = 0$ for all $i \in B, j \notin B$. Assuming that the new process (starting in B and remining in B) is irreducible, show that the new process is reversible and find its steady-state probabilities.

Exercise 6.26. Consider a queueing system with two classes of customers. Type A customer arrivals are Poisson with rate λ_A and Type B customer arrivals are Poisson with rate λ_B. The service time for type A customers is exponential with rate μ_A and that for type B is exponential with rate μ_B. Each service time is independent of all other service times and of all arrival epochs.

a) First assume there are infinitely many identical servers, and each new arrival immediately enters an idle server and begins service. Let the state of the system be (i, j) where i and j are the numbers of type A and B customers respectively in service. Draw a graph of the state transitions for $i \leq 2, j \leq 2$. Find the steady-state PMF, $\{p(i, j); i, j \geq 0\}$, for the Markov process. Hint: Note that the type A and type B customers do not interact.

b) Assume for the rest of the exercise that there is some finite number m of servers. Customers who arrive when all servers are occupied are turned away. Find the steady-state PMF, $\{p(i, j); i, j \geq 0, i + j \leq m\}$, in terms of $p(0, 0)$ for this Markov process. Hint: Combine part a) with the result of Exercise 6.25.

c) Let Q_n be the probability that there are n customers in service at some given time in steady state. Show that $Q_n = p(0, 0)\rho_n/n!$ for $0 \leq n \leq m$ where $\rho = \rho_A + \rho_B$, $\rho_A = \lambda_A/\mu_A$, and $\rho_B = \lambda_B/\mu_B$. Solve for $p(0, 0)$.

Exercise 6.27. a) Generalize Exercise 6.26 to the case in which there are K types of customers, each with independent Poisson arrivals and each with independent exponential service times. Let λ_k and μ_k be the arrival rate and service rate respectively for the k^h user type, $1 \leq k \leq K$. Let $\rho_k = \lambda_k/\mu_k$ and $\rho = \rho_1 + \rho_2 + \cdots + \rho_K$. In particular, show, as before, that the probability of n customers in the system is $Q_n = p(0, \dots, 0)\rho_n/n!$ for $0 \leq n \leq m$.

b) View the customers in part a) as a single type of customer with Poisson arrivals of rate $\lambda = \sum_k \lambda_k$ and with a service density $\sum_k (\lambda_k/\lambda)\mu_k \exp(-\mu_k x)$. Show that the expected service time is ρ/λ. Note that what you have shown is that, if a service distribution can be represented as a weighted sum of exponentials, then the distribution of customers in the system is the same as for the M/M/m,m queue with equal mean service time.

Exercise 6.28. Consider a sampled-time M/D/m/m queueing system. The arrival process is Bernoulli with probability $\lambda << 1$ of arrival in each time unit. There are m servers; each arrival enters a server if a server is not busy and otherwise the arrival is discarded. If an arrival enters a server, it keeps the server busy for d units of time and then departs; d is some integer constant and is the same for each server.

Let n, $0 \leq n \leq m$ be the number of customers in service at a given time and let x_i be the number of time units that the i^{th} of those n customers (in order of arrival) has been in service. Thus the state of the system can be taken as $(n, \boldsymbol{x}) = (n, x_1, x_2, \dots, x_n)$ where $0 \le n \le m$ and $1 \le x_1 < x_2 < \dots < x_n \le d$.

Let $A(n, \boldsymbol{x})$ denote the next state if the present state is $(n, \boldsymbol{x})$ and a new arrival enters service. That is,

$$A(n, \boldsymbol{x}) = (n+1, 1, x_1+1, x_2+1, \dots, x_n+1) \qquad \text{for } n < m \text{ and } x_n < d \quad (6.99)$$
$$A(n, \boldsymbol{x}) = (n, 1, x_1+1, x_2+1, \dots, x_{n-1}+1) \qquad \text{for } n \le m \text{ and } x_n = d. \quad (6.100)$$

That is, the new customer receives one unit of service by the next state time, and all the old customers receive one additional unit of service. If the oldest customer has received d units of service, then it leaves the system by the next state time. Note that it is possible for a customer with d units of service at the present time to leave the system and be replaced by an arrival at the present time (*i.e.*, (6.100) with $n = m$, $x_n = d$). Let $B(n, \boldsymbol{x})$ denote the next state if either no arrival occurs or if a new arrival is discarded.

$$B(n, \boldsymbol{x}) = (n, x_1+1, x_2+1, \dots, x_n+1) \qquad \text{for } x_n < d \quad (6.101)$$
$$B(n, \boldsymbol{x}) = (n-1, x_1+1, x_2+1, \dots, x_{n-1}+1) \quad \text{for } x_n = d. \quad (6.102)$$

a) Hypothesize that the *backward* chain for this system is also a sampled-time M/D/m/m queueing system, but that the state $(n, x_1, \dots, x_n)(0 \le n \le m, 1 \le x_1 < x_2 < \dots < x_n \le d)$ has a different interpretation: n is the number of customers as before, but x_i is now the remaining service required by customer i. Explain how this hypothesis leads to the following steady-state equations:

$$\lambda \pi_{n,\boldsymbol{x}} = (1-\lambda)\pi_{A(n,\boldsymbol{x})} \qquad ; n < m, x_n < d \quad (6.103)$$
$$\lambda \pi_{n,\boldsymbol{x}} = \lambda \pi_{A(n,\boldsymbol{x})} \qquad ; n \le m, x_n = d \quad (6.104)$$
$$(1-\lambda) \pi_{n,\boldsymbol{x}} = \lambda \pi_{B(n,\boldsymbol{x})} \qquad ; n \le m, x_n = d \quad (6.105)$$
$$(1-\lambda) \pi_{n,\boldsymbol{x}} = (1-\lambda)\pi_{B(n,\boldsymbol{x})} \qquad ; n \le m, x_n < d. \quad (6.106)$$

b) Using this hypothesis, find $\pi_{n,\boldsymbol{x}}$ in terms of π_0, where π_0 is the probability of an empty system. Hint: Use (6.105) and (6.106); your answer should depend on n, but not $\boldsymbol{x}$.

c) Verify that the above hypothesis is correct.

d) Find an expression for π_0.

e) Find an expression for the steady-state probability that an arriving customer is discarded.

Exercise 6.29. A taxi alternates between three locations. When it reaches location 1 it is equally likely to go next to either 2 or 3. When it reaches 2 it will next go to 1 with probability 1/3 and to 3 with probability 2/3. From 3 it always goes to 1. The mean time between locations i and j are $t_{12} = 20$, $t_{13} = 30$, $t_{23} = 30$. Assume $t_{ij} = t_{ji}$).

Find the (limiting) probability that the taxi's most recent stop was at location i, $i = 1, 2, 3$?

What is the (limiting) probability that the taxi is heading for location 2?

What fraction of time is the taxi traveling from 2 to 3. Note: Upon arrival at a location the taxi immediately departs.

Exercise 6.30. (Semi-Markov continuation of Exercise 6.5) **a)** Assume that the Markov process in Exercise 6.5 is changed in the following way: whenever the process enters state 0, the time spent before leaving state 0 is now a *uniformly distributed* rv, taking values from 0 to $2/\lambda$. All other transitions remain the same. For this new process, determine whether the successive epochs of entry to state 0 form renewal epochs, whether the successive epochs of exit from state 0 form renewal epochs, and whether the successive entries to any other given state i form renewal epochs.

a) For each i, find both the time-average interval and the time-average number of overall state transitions between successive visits to i.

b) Is this modified process a Markov process in the sense that $\Pr\{X(t) = i \mid X(\tau) = j, X(s) = k\} = \Pr\{X(t) = i \mid X(\tau) = j\}$ for all $0 < s < \tau < t$ and all i, j, k? Explain.

Exercise 6.31. Consider an M/G/1 queueing system with Poisson arrivals of rate λ and expected service time $\mathsf{E}[X]$. Let $\rho = \lambda\mathsf{E}[X]$ and assume $\rho < 1$. Consider a semi-Markov process model of the M/G/1 queueing system in which transitions occur on departures from the queueing system and the state is the number of customers immediately following a departure.

a) Suppose a colleague has calculated the steady-state probabilities $\{p_i\}$ of being in state i for each $i \geq 0$. For each $i \geq 0$, find the steady-state probability pi_i of state i in the embedded Markov chain. Give your solution as a function of ρ, π_i, and p_0.

b) Calculate p_0 as a function of ρ.

c) Find π_i as a function of ρ and p_i.

d) Is p_i the same as the steady-state probability that the queueing system contains i customers at a given time? Explain carefully.

Exercise 6.32. Consider an M/G/1 queue in which the arrival rate is λ and the service time distributin is uniform $(0, 2W)$ with $\lambda W < 1$. Define a semi-Markov chain following the framework for the M/G/1 queue in Section 6.8.1.

a) Find $P_{0j}; j \geq 0$.

b) Find P_{ij} for $i > 0$; $j \geq i - 1$.

Exercise 6.33. Consider a semi-Markov process for which the embedded Markov chain is irreducible and positive-recurrent. Assume that the distribution of inter-renewal intervals for one state j is arithmetic with span d. Show that the distribution of inter-renewal intervals for all states is arithmetic with the same span.