6 Networks—B

Application: Social Networks, Communication Networks
Topics: Continuous-Time Markov Chains, Product-Form Queueing Networks

6.1 Social Networks

We provide the proofs of the theorems in Sect. 5.1.

Theorem 6.1 (Spreading of a Message) *Let Z be the number of nodes that eventually receive the message.*

(a) If $\mu < 1$, then $P(Z < \infty) = 1$ and $E(Z) < \infty$;
(b) If $\mu > 1$, then $P(Z = \infty) > 0$.

Proof For part (a), let X_n be the number of nodes that are n steps from the root. If $X_n = k$, we can write $X_{n+1} = Y_1 + \cdots + Y_k$ where Y_j is the number of children of node j at level n. By assumption, $E(Y_j) = \mu$ for all j. Hence,

$$E[X_{n+1} \mid X_n = k] = E(Y_1 + \cdots + Y_k) = \mu k.$$

Hence, $E[X_{n+1} \mid X_n] = \mu X_n$. Taking expectations shows that $E(X_{n+1}) = \mu E(X_n), n \geq 0$. Consequently,

$$E(X_n) = \mu^n, n \geq 0.$$

J. Walrand, *Probability in Electrical Engineering and Computer Science*,
https://doi.org/10.1007/978-3-030-49995-2_6

Now, the sequence $Z_n = X_0 + \cdots + X_n$ is nonnegative and increases to $Z = \sum_{n=0}^{\infty} Z_n$. By MCT, it follows that $E(Z_n) \to Z$. But

$$E(Z_n) = \mu_0 + \cdots + \mu^n = \frac{1 - \mu^{n+1}}{1 - \mu}.$$

Hence, $E(Z) = 1/(1 - \mu) < \infty$. Consequently, $P(Z < \infty) = 1$.

For part (b), one first observes that the theorem does not state that $P(Z = \infty) = 1$. For instance, assume that each node has three children with probability 0.5 and has no child otherwise. Then $\mu = 1.5 > 1$ and $P(Z = 1) = P(X_1 = 0) = 0.5$, so that $P(Z = \infty) \leq 0.5 < 1$. We define X_n, Y_j, and Z_n as in the proof of part (a).

Let $\alpha_n = P(X_n > 0)$. Consider the X_1 children of the root. Since α_{n+1} is the probability that there is one survivor after $n + 1$ generations, it is the probability that at least one of the X_1 children of the root has a survivor after n generations. Hence,

$$1 - \alpha_{n+1} = E((1 - \alpha_n)^{X_1}), n \geq 0.$$

Indeed, if $X_1 = k$, the probability that none of the k children of the root has a survivor after n generations is $(1 - \alpha_n)^k$. Hence,

$$\alpha_{n+1} = 1 - E((1 - \alpha_n)^{X_1}) =: g(\alpha_n), n \geq 0.$$

Also, $\alpha_0 = 1$. As $n \to \infty$, one has $\alpha_n \to \alpha^* = P(X_n > 0, \text{ for all } n)$. Figure 6.1 shows that $\alpha^* > 0$. The key observations are that

$$\begin{aligned} g(0) &= 0 \\ g(1) &= P(X_1 > 0) < 1 \\ g'(0) &= E(X_1(1 - \alpha)^{X_1 - 1}) \mid_{\alpha=0} = \mu > 1 \\ g'(1) &= E(X_1(1 - \alpha)^{X_1 - 1}) \mid_{\alpha=1} = 0, \end{aligned}$$

so that the figure is as drawn. □

Theorem 6.2 (Cascades) *Assume $p_k = p \in (0, 1]$ for all $k \geq 1$. Then, all nodes turn red with probability at least equal to θ where*

$$\theta = \exp\left\{-\frac{1 - p}{p}\right\}.$$

Proof The probability that node n does not listen to anyone is $a_n = (1 - p)^n$. Let X be the index of the first node that does not listen to anyone. Then

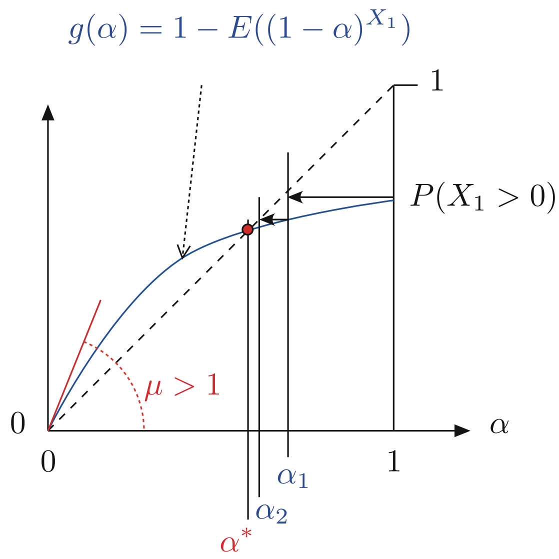


Fig. 6.1 The proof that $\alpha^* > 0$

$$P(X > n) = (1 - a_1)(1 - a_2) \cdots (1 - a_n) \leq \exp\{-a_1 - \cdots - a_n\}$$
$$= \exp\left\{-\frac{1}{p}((1 - p) - (1 - p)^{n+1})\right\}.$$

Now,

$$P(X = \infty) = \lim_n P(X > n) \geq \exp\left\{-\frac{1 - p}{p}\right\} = \theta.$$

Thus, with probability at least θ, every node listens to at least one previous node. When that is the case, all the nodes turn red. To see this, assume that n is the first blue node. That is not possible since it listened to some previous nodes that are all red. □

6.2 Continuous-Time Markov Chains

Our goal is to understand networks where packets travel from node to node until they reach their destination. In particular, we want to study the delay of packets from source to destination and the backlog in the nodes.

It turns out that the analysis of such systems is much easier in continuous time than in discrete time. To carry out such analysis, we have to introduce continuous-time Markov chains. We do this on a few simple examples.

6.2.1 Two-State Markov Chain

Figure 6.2 illustrates a random process $\{X_t, t \geq 0\}$ that takes values in $\{0, 1\}$. A random process is a collection of random variables indexed by $t \geq 0$. Saying that such a random process is defined means that one can calculate the probability that $\{X_{t_1} = x_1, X_{t_2} = x_2, \ldots, X_{t_n} = x_n\}$ for any value of $n \geq 1$, any $0 \leq t_1 \leq \cdots \leq t_n$, and $x_1, \ldots, x_n \in \{0, 1\}$. We explain below how one could calculate such a probability.

We call X_t the *state* of the process at time t. The possible values $\{0, 1\}$ are also called states. The state X_t evolves according to rules characterized by two positive numbers λ and μ. As Fig. 6.2 shows, if $X_0 = 0$, the state remains equal to zero for a random time T_0 that is exponentially distributed with parameter λ, thus with mean $1/\lambda$. The state X_t then jumps to 1 where it stays for a random time T_1 that is exponentially distributed with rate μ, independent of T_0, and so on. The definition is similar if $X_0 = 1$. In that case, X_t keeps the value 1 for an exponentially distributed time with rate μ, then jumps to 0, etc.

Thus, the pdf of T_0 is

$$f_{T_0}(t) = \lambda \exp\{-\lambda t\} 1\{t \geq 0\}.$$

In particular,

$$P(T_0 \leq \epsilon) \approx f_{T_0}(0)\epsilon = \lambda\epsilon, \text{ for } \epsilon \ll 1.$$

Throughout this chapter, the symbol $\approx$ means "up to a quantity negligible compared to ϵ." It is shown in Theorem 15.3 that exponentially distributed random variable is *memoryless*. That is,

$$P[T_0 > t + s \mid T_0 > t] = P(T_0 > s), s, t \geq 0.$$

The memoryless property and the independence of the exponential times T_k imply that $\{X_t, t \geq 0\}$ *starts afresh* from X_s at time s. Figure 6.3 illustrates that property. Mathematically, it says that given $\{X_t, t \leq s\}$ with $X_s = k$, the process $\{X_{s+t}, t \geq 0\}$ has the same properties as $\{X_t, t \geq 0\}$ given that $X_0 = k$, for $k = 0, 1$ and for any $s \geq 0$. Indeed, if $X_s = 0$, then the residual time that X_t remains in 0 is exponentially distributed with rate λ and is independent of what happened before

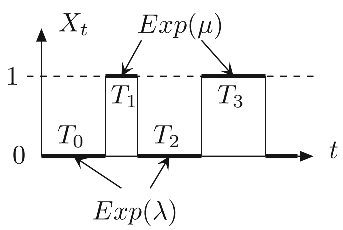


Fig. 6.2 A random process on $\{0, 1\}$

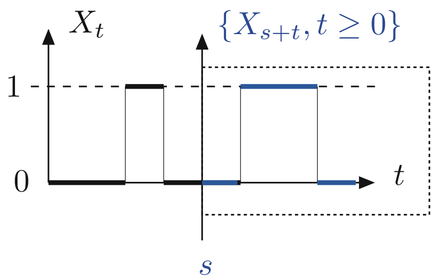


Fig. 6.3 The process X_t starts afresh from X_s at time s

time s, because the time in 0 is memoryless and independent of the previous times in 0 and 1. This property is written as

$$P[\{X_{s+t}, t \geq 0\} \in A \mid X_s = k; X_t, t \leq s] = P[\{X_t, t \geq 0\} \in A \mid X_0 = k],$$

for $k = 0, 1$, for all $s \geq 0$, and for all sets A of possible trajectories. A generic set A of trajectories is

$$A = \{(x_t, t \geq 0) \in C_+ \mid x_{t_1} = i_1, \ldots, x_{t_n} = i_n\}$$

for given $0 < t_1 < \cdots < t_n$ and $i_1, \ldots, i_n \in \{0, 1\}$. Here, C_+ is the set of right-continuous functions of $t \geq 0$ that take values in $\{0, 1\}$.

This property is the continuous-time version of the Markov property for Markov chains. One says that the process X_t satisfies the *Markov property* and one calls $\{X_t, t \geq 0\}$ is a *continuous-time Markov chain* (CTMC).

For instance,

$$P[X_{s+2.5} = 1, X_{s+4} = 0, X_{s+5.1} = 0 \mid X_s = 0; X_t, t \leq s]$$
$$= P[X_{2.5} = 1, X_4 = 0, X_{5.1} = 0 \mid X_0 = 0].$$

The Markov property generalizes to situations where s is replaced by a random τ that is defined by a *causal* rule, i.e., a rule that does not look ahead. For instance, as in Fig. 6.4, τ can be the second time that X_t visits state 0. Or τ could be the first time that it visits state 0 after having spent at least 3 time units in state 1. The property does not extend to non-causal times such as one time unit before X_t visits state 1. Random times τ defined by causal rules are called *stopping times*. This more general property is called the *strong Markov property*. To prove this property, one conditions on the value s of τ and uses the fact that the future evolution does not depend on this value since the event $\{\tau = s\}$ depends only on $\{X_t, t \leq s\}$.

For $0 < \epsilon \ll 1$ one has

$$P[X_{t+\epsilon} = 1 \mid X_t = 0] \approx \lambda\epsilon.$$

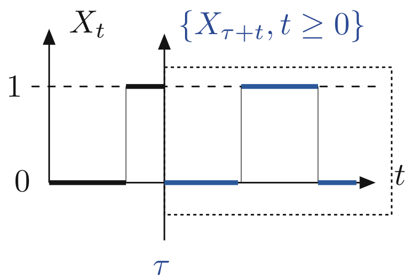


Fig. 6.4 The process X_t starts afresh from X_τ at the stopping time τ

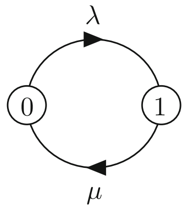


Fig. 6.5 The state transition diagram

Indeed, the process jumps from 0 to 1 in ϵ time units if the exponential time in 0 is less than ϵ, which has probability approximately $\lambda\epsilon$.

Similarly,

$$P[X_{t+\epsilon} = 0 \mid X_t = 1] \approx \mu\epsilon.$$

We say that the *transition rate* from 0 to 1 is equal to λ and that from 1 to 0 is equal to μ to indicate that the probability of a transition from 0 to 1 in ϵ units of time is approximately $\lambda\epsilon$ and that from 1 to 0 is approximately $\mu\epsilon$.

Figure 6.5 illustrates these transition rates. This figure is called the *state transition diagram*.

The previous two identities imply that

$$\begin{aligned} P(X_{t+\epsilon} = 1) &= P(X_t = 0, X_{t+\epsilon} = 1) + P(X_t = 1, X_{t+\epsilon} = 1) \\ &= P(X_t{=}0)P[X_{t+\epsilon}{=}1 \mid X_t{=}0]{+}P(X_t{=}1)P[X_{t+\epsilon}{=}1 \mid X_t{=}1] \\ &\approx P(X_t = 0)\lambda\epsilon + P(X_t = 1)(1 - P[X_{t+\epsilon} = 0 \mid X_t = 1]) \\ &\approx P(X_t = 0)\lambda\epsilon + P(X_t = 1)(1 - \mu\epsilon). \end{aligned}$$

Also, similarly, one finds that

$$P(X_{t+\epsilon} = 0) \approx P(X_t = 0)(1 - \lambda\epsilon) + P(X_t = 1)\mu\epsilon.$$

We can write these identities in a convenient matrix notation as follows. For $t \geq 0$, one defines the row vector π_t as

$$\pi_t = [P(X_t = 0), P(X_t = 1)].$$

One also defines the *transition rate matrix* Q as follows:

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}.$$

With that notation, the previous identities can be written as

$$\pi_{t+\epsilon} \approx \pi_t(\mathbf{I} + Q\epsilon),$$

where $\mathbf{I}$ is the identity matrix. Subtracting π_t from both sides, dividing by ϵ, and letting $\epsilon \to 0$, we find

$$\frac{d}{dt}\pi_t = \pi_t Q. \tag{6.1}$$

By analogy with the scalar equation $dx_t/dt = ax_t$ whose solution is $x_t = x_0 \exp\{at\}$, we conclude that

$$\pi_t = \pi_0 \exp\{Qt\}, \tag{6.2}$$

where

$$\exp\{Qt\} := \mathbf{I} + Qt + \frac{1}{2!}Q^2t^2 + \frac{1}{3!}Q^3t^3 + \cdots.$$

Note that

$$\frac{d}{dt}\exp\{Qt\} = \mathbf{0} + Q + Q^2t + \frac{1}{2!}Q^3t^2 + \cdots = Q\exp\{Qt\}.$$

Observe also that $\pi_t = \pi$ for all $t \geq 0$ if and only if $\pi_0 = \pi$ and

$$\pi Q = 0. \tag{6.3}$$

Indeed, if $\pi_t = \pi$ for all t, then (6.1) implies that $0 = \frac{d}{dt}\pi_t = \pi_t Q = \pi Q$. Conversely, if $\pi_0 = \pi$ with $\pi Q = 0$, then

$$\pi_t = \pi_0 \exp\{Qt\} = \pi \exp\{Qt\} = \pi\left(\mathbf{I} + Qt + \frac{1}{2!}Q^2t^2 + \frac{1}{3!}Q^3t^3 + \cdots\right) = \pi.$$

These equations $\pi Q = 0$ are called the *balance equations*. They are

$$[\pi(0), \pi(1)]\begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix} = 0,$$

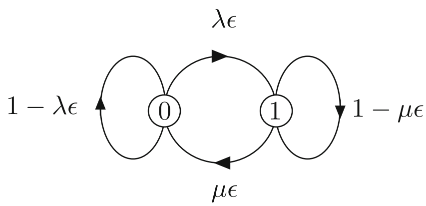


Fig. 6.6 A discrete-time approximation of X_t

i.e.,

$$\pi(0)(-\lambda) + \pi(1)\mu = 0$$
$$\pi(0)\lambda - \pi(1)\mu = 0.$$

These two equations are identical. To determine π, we use the fact that $\pi(0) + \pi(1) = 1$. Combined with the previous identity, we find

$$[\pi(0), \pi(1)] = \left[\frac{\mu}{\lambda+\mu}, \frac{\lambda}{\lambda+\mu}\right].$$

The identity $\pi_{t+\epsilon} \approx \pi_t(\mathbf{I} + Q\epsilon)$ shows that one can view $\{X_{n\epsilon}, n = 0, 1, \ldots\}$ as a discrete-time Markov chain with transition matrix $P = \mathbf{I} + Q\epsilon$. Figure 6.6 shows the transition diagram that corresponds to this transition matrix. The invariant distribution for P is such that $\pi P = \pi$, i.e., $\pi(\mathbf{I} + Q\epsilon) = \pi$, so that $\pi Q = 0$, not surprisingly.

Note that this discrete-time Markov chain is aperiodic because states have self-loops. Thus, we expect that

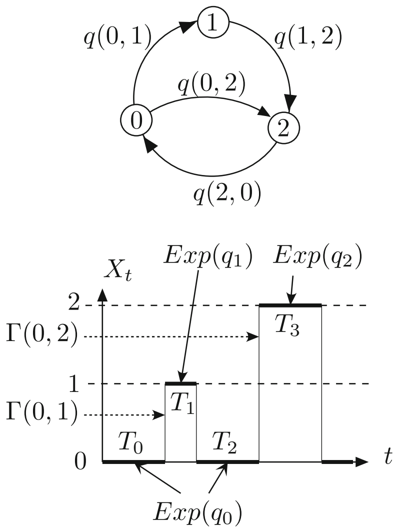
$$\pi_{n\epsilon} \to \pi, \text{ as } n \to \infty.$$

Consequently, we expect that, in continuous time,

$$\pi_t \to \pi, \text{ as } t \to \infty.$$

6.2.2 Three-State Markov Chain

The previous Markov chain alternates between the states 0 and 1. More general Markov chains visit states in a random order. We explain that feature in our next example with 3 states. Fortunately, this example suffices to illustrate the general case. We do not have to look at Markov chains with 4, 5, ... states to describe the general model.



$$q_0 = q(0,1) + q(0,2) \quad \Gamma(0,1) = q(0,1)/q_0$$
$$q_1 = q(1,2) \quad \Gamma(0,2) = q(0,2)/q_0$$
$$q_2 = q(2,0)$$

Fig. 6.7 A three-state Markov chain

In the example shown in Fig. 6.7, the rules of evolution are characterized by positive numbers $q(0, 1), q(0, 2), q(1, 2)$, and $q(2, 0)$. One also defines $q_0, q_1, q_2, \Gamma(0, 1)$, and $\Gamma(0, 2)$ as in the figure.

If $X_0 = 0$, the state X_t remains equal to 0 for some random time T_0 that is exponentially distributed with rate q_0. At time T_0, the state jumps to 1 with probability $\Gamma(0, 1)$ or to state 2 otherwise, with probability $\Gamma(0, 2)$. If X_t jumps to 1, it stays there for an exponentially distributed time T_1 with rate q_1 that is independent of T_0. More generally, when X_t enters state k, it stays there for a random time that is exponentially distributed with rate q_k that is independent of the past evolution. From this definition, it should be clear that the process X_t satisfies the Markov property.

Define $\pi_t = [\pi_t(0), \pi_t(1), \pi_t(2)]$ where $\pi_t(k) = P(X_t = k)$ for $k = 0, 1, 2$. One has, for $0 < \epsilon \ll 1$,

$$P[X_{t+\epsilon} = 1 \mid X_t = 0] \approx q_0 \epsilon \Gamma(0, 1) = q(0, 1)\epsilon.$$

Indeed, the process jumps from 0 to 1 in ϵ time units if the exponential time with rate q_0 is less than ϵ and if the process then jumps to 1 instead of jumping to 2.

Similarly,

$$P[X_{t+\epsilon} = 2 \mid X_t = 0] \approx q_0 \epsilon \Gamma(0, 2) = q(0, 2)\epsilon.$$

Also,

$$P[X_{t+\epsilon} = 1 \mid X_t = 1] \approx 1 - q_1\epsilon,$$

since this is approximately the probability that the exponential time with rate q_1 is larger than ϵ. Moreover,

$$P[X_{t+\epsilon} = 1 \mid X_t = 2] \approx 0,$$

because the probability that both the exponential time with rate q_2 in state 2 and the exponential time with rate q_0 in state 0 are less than ϵ is roughly $(q_2\epsilon) \times (q_1\epsilon)$, and this is negligible compared to ϵ.

These observations imply that

$$\begin{aligned}
\pi_{t+\epsilon}(1) &= P(X_t = 0, X_{t+\epsilon} = 1) + P(X_t = 1, X_{t+\epsilon} = 1) + P(X_t = 2, X_{t+\epsilon} = 1)\\
&= P(X_t{=}0)P[X_{t+\epsilon}{=}1 \mid X_t{=}0] + P(X_t{=}1)P[X_{t+\epsilon}{=}1 \mid X_t{=}1]\\
&\quad + P(X_t = 2)P[X_{t+\epsilon} = 1 \mid X_t = 2]\\
&\approx \pi_t(0)q(0,1)\epsilon + \pi_t(1)(1 - q_1\epsilon).
\end{aligned}$$

Proceeding in a similar way shows that

$$\begin{aligned}
\pi_{t+\epsilon}(0) &\approx \pi_t(0)(1 - q_0\epsilon) + \pi_t(2)q(2,0)\epsilon\\
\pi_{t+\epsilon}(2) &\approx \pi_t(1)q(1,2)\epsilon + \pi_t(2)(1 - q_2\epsilon.
\end{aligned}$$

Similarly to the two-state example, let us define the rate matrix Q as follows:

$$Q = \begin{bmatrix} -q_0 & q(0,1) & q(0,2) \\ 0 & -q_1 & q(0,1) \\ q(2,0) & 0 & -q_2 \end{bmatrix}.$$

The previous identities can then be written as follows:

$$\pi_{t+\epsilon} \approx \pi_t[\mathbf{I} + Q\epsilon].$$

Subtracting π_t from both sides, dividing by ϵ, and letting $\epsilon \to 0$ then shows that

$$\frac{d}{dt}\pi_t = \pi_t Q.$$

As before, the solution of this equation is

$$\pi_t = \pi_0 \exp\{Qt\}, t \geq 0.$$

The distribution π is invariant if and only if

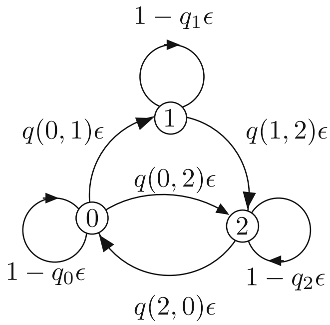


Fig. 6.8 The transition matrix of the discrete-time approximation

$$\pi Q = 0.$$

Once again, we note that $\{X_{n\epsilon}, n = 0, 1, \ldots\}$ is approximately a discrete-time Markov chain with transition matrix $P = I + Q\epsilon$ shown in Fig. 6.8. This Markov chain is aperiodic, and we conclude that

$$P(X_{n\epsilon} = k) \to \pi(k), \text{ as } n \to \infty.$$

Thus, we can expect that

$$\pi_t \to \pi, \text{ as } t \to \infty.$$

Also, since $X_{n\epsilon}$ is irreducible, the long-term fraction of time that it spends in the different states converge to π, and we can then expect the same for X_t.

6.2.3 General Case

Let $\mathscr{X}$ be a countable or finite set. The process $\{X_t, t \geq 0\}$ is defined as follows. One is given a probability distribution π on $\mathscr{X}$ and a *rate matrix* $Q = \{q(i, j), i, j \in \mathscr{X}\}$.

By definition, Q is such that

$$q(i, j) \geq 0, \forall i \neq j \text{ and } \sum_j q(i, j) = 0, \forall x.$$

Definition 6.1 (Continuous-Time Markov Chain) A continuous-time Markov chain with initial distribution π and rate matrix Q is a process $\{X_t, t \geq 0\}$ such that $P(X_0 = i) = \pi(i)$. Also,

$$P[X_{t+\epsilon} = j | X_t = i, X_u, u < t] = 1\{i = j\} + \epsilon q(i, j) + o(\epsilon).$$

◇

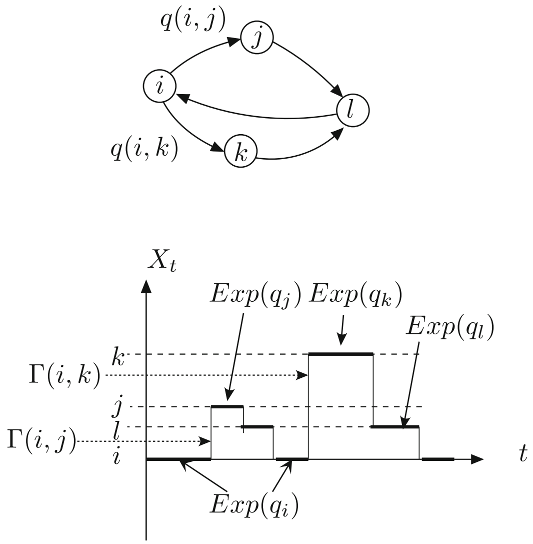


Fig. 6.9 Construction of a continuous-time Markov chain

This definition means that the process jumps from i to $j \neq i$ with probability $q(i, j)\epsilon$ in $\epsilon \ll 1$ time units. Thus, $q(i, j)$ is the probability of jumping from i to j, per unit of time. Note that the sum of these expressions over all j gives 1, as should be.

One construction of this process is as follows. Say that $X_t = i$. One then chooses a random time τ that is exponentially distributed with rate $q_i := -q(i, i)$. At time $t + \tau$, the process jumps and goes to state y with probability $\Gamma(i, j) = q(i, j)/q_i$ for $j \neq i$ (Fig. 6.9).

Thus, if $X_t = i$, the probability that $X_{t+\epsilon} = j$ is the probability that the process jumps in $(t, t+\epsilon)$, which is $q_i\epsilon$, times the probability that it then jumps to j, which is $\Gamma(i, j)$. Hence,

$$P[X_{t+\epsilon} = j | X_t = i] = q_i \epsilon \frac{q(i, j)}{q_i} = q(i, j)\epsilon,$$

up to $o(\epsilon)$. Thus, the construction yields the correct transition probabilities.

As we observed in the examples,

$$\frac{d}{dt}\pi_t = \pi_t Q,$$

so that

$$\pi_t = \pi_0 \exp\{Qt\}.$$

Moreover, a distribution π is invariant if and only if it solves the balance equations

$$0 = \pi Q.$$

These equations, state by state, say that

$$\pi(i)q_i = \sum_{j \neq i} \pi(j)q(j,i), \forall i \in \mathscr{X}.$$

These equations express the equality of the rate of leaving a state and the rate of entering that state.

Define

$$P_t(i,j) = P[X_{s+t} = j \mid X_s = i], \text{ for } i, j \in \mathscr{X} \text{ and } s, t \geq 0.$$

The Markov property implies that

$$P(X_{t_1} = i_1, \ldots, X_{t_n} = i_n) = P(X_{t_1} = i_1)P_{t_2-t_1}(i_1, i_2)P_{t_3-t_2}(i_2, i_3) \cdots P_{t_n-t_{n-1}}(i_{n-1}, i_n),$$

for all $i_1, \ldots, i_n \in \mathscr{X}$ and all $0 < t_1 < \cdots < t_n$.

Moreover, this identity implies the Markov property. Indeed, if it holds, one has

$$\begin{aligned} &P[X_{t_{m+1}} = i_{m+1}, \ldots, X_{t_n} = i_n \mid X_{t_1} = i_1, \ldots, X_{t_m} = i_m] \\ &= \frac{P(X_{t_1} = i_1, \ldots, X_{t_n} = i_n)}{P(X_{t_1} = i_1, \ldots, X_{t_m} = i_m)} \\ &= \frac{P(X_{t_1} = i_1)P_{t_2-t_1}(i_1, i_2)P_{t_3-t_2}(i_2, i_3) \cdots P_{t_n-t_{n-1}}(i_{n-1}, i_n)}{P(X_{t_1} = i_1)P_{t_2-t_1}(i_1, i_2)P_{t_3-t_2}(i_2, i_3) \cdots P_{t_{m-1}-t_{m-2}}(i_{m-2}, i_{m-1})} \\ &= P_{t_m-t_{m-1}}(i_{m-1}, i_m) \cdots P_{t_n-t_{n-1}}(i_{n-1}, i_n). \end{aligned}$$

Hence,

$$\begin{aligned} &P[X_{t_{m+1}} = i_{m+1}, \ldots, X_{t_n} = i_n \mid X_{t_1} = i_1, \ldots, X_{t_m} = i_m] \\ &= \frac{P(X_{t_{m-1}} = i_{m-1})P_{t_m-t_{m-1}}(i_{m-1}, i_m) \cdots P_{t_n-t_{n-1}}(i_{n-1}, i_n)}{P(X_{t_{m-1}} = i_{m-1})P_{t_m-t_{m-1}}(i_{m-1}, i_m)} \\ &= \frac{P(X_{t_{m-1}} = i_{m-1}, \ldots, X_{t_n} = i_n)}{P(X_{t_{m-1}} = i_{m-1})} \\ &= P[X_{t_m} = i_m, \ldots, X_{t_n} = i_n \mid X_{t_{m-1}} = i_{m-1}]. \end{aligned}$$

If X_t has the invariant distribution, one has

$$P(X_{t_1} = i_1, \ldots, X_{t_n} = i_n) = \pi(i_1)P_{t_2-t_1}(i_1, i_2)P_{t_3-t_2}(I_2, i_3) \cdots P_{t_n-t_{n-1}}(i_{n-1}, i_n),$$

for all $i_1, \ldots, i_n \in \mathscr{X}$ and all $0 < t_1 < \cdots < t_n$.

Here is the result that corresponds to Theorem 15.1. We define irreducibility, transience, and null and positive recurrence as in discrete time. There is no notion of periodicity in continuous time.

Theorem 6.1 (Big Theorem for Continuous-Time Markov Chains)
Consider a continuous-time Markov chain.

(a) If the Markov chain is irreducible, the states are either all transient, all positive recurrent, or all null recurrent. We then say that the Markov chain is transient, positive recurrent, or null recurrent, respectively.
(b) If the Markov chain is positive recurrent, it has a unique invariant distribution π and $\pi(i)$ is the long-term fraction of time that X_t is equal to i. Moreover, the probability $\pi_t(i)$ that the Markov chain X_t is in state i converges to $\pi(i)$.
(c) If the Markov chain is not positive recurrent, it does not have an invariant distribution and the fraction of time that it spends in any state goes to zero.

■

6.2.4 Uniformization

We saw earlier that a CTMC can be approximated by a discrete-time Markov chain that has a time step $\epsilon \ll 1$. There are two other DTMCs that have a close relationship with the CTMC: the jump chain and the uniformized chain. We explain these chains for the CTMC X_t in Fig. 6.7.

The *jump chain* is X_t observed when it jumps. As Fig. 6.7 shows, this DTMC has a transition matrix equal to Γ where

$$\Gamma(i, j) = \begin{cases} q(i, j)/q_i, & \text{if } i \neq j \\ 0, & \text{if } i = j. \end{cases}$$

Let ν be the invariant distribution of this jump chain. That is, $\nu = \nu\Gamma$. Since $\nu(i)$ is the long-term fraction of time that the jump chain is in state i, and since the CTMC X_t spends an average time $1/q_i$ in state i whenever it visits that state, the fraction of time that X_t spends in state i should be proportional to $\nu(i)/q_i$. That is, one expects

$$\pi(i) = A\nu(i)/q_i$$

for some constant A. That is, one should have

$$\sum_j [A\nu(i)/q_i]q(i, j) = 0.$$

To verify that equality, we observe that

$$\sum_j [\nu(i)/q_i]q(i,j) = \sum_{j \neq i} \nu(i)\Gamma(i,j) + \nu(i)q(i,i)/q_i = \nu(i) - \nu(i) = 0.$$

We used the fact that $\nu\Gamma = \nu$ and $q(i,i) = -q_i$.

The *uniformized chain* is not the jump chain. It is a discrete-time Markov chain obtained from the CTMC as follows. Let $\lambda \geq q_i$ for all i. The rate at which X_t changes state is q_i when it is in state i. Let us add a dummy jump from i to i with rate $\lambda - q_i$. The rate of jumps, including these dummy jumps, of this new Markov chain Y_t is now constant and equal to λ.

The transition matrix P of Y_t is such that

$$P(i,j) = \begin{cases} (\lambda - q_i)/\lambda, & \text{if } i = j \\ q(i,j)/\lambda, & \text{if } i \neq j. \end{cases}$$

To see this, assume that $Y_t = i$. The next jump will occur with rate λ. With probability $(\lambda - q_i)/\lambda$, it is a dummy jump from i to i. With probability q_i/λ it is an actual jump where Y_t jumps to $j \neq i$ with probability $\Gamma(i,j)$. Hence, Y_t jumps from i to i with probability $(\lambda - q_i)/\lambda$ and from i to $j \neq i$ with probability $(q_i/\lambda)\Gamma(i,j) = q(i,j)/\lambda$.

Note that

$$P = \mathbf{I} + \frac{1}{\lambda}Q,$$

where $\mathbf{I}$ is the identity matrix.

Now, define Z_n to be the jump chain of Y_t, i.e., the Markov chain with transition matrix P. Since the jumps of Y_t occur at rate λ, independently of the value of the state Y_t, we can simulate Y_t as follows. Let N_t be a Poisson process with rate λ. The jump times $\{t_1, t_2, \ldots\}$ of N_t will be the jump times of Y_t. The successive values of Y_t are those of Z_n. Formally,

$$Y_t = Z_{N_t}.$$

That is, if $N_t = n$, then we define $Y_t = Z_n$. Since the CTMC Y_t spends $1/\lambda$ on average between jumps, the invariant distribution of Y_t should be the same as that of X_t, i.e., π. To verify this, we check that $\pi P = \pi$, i.e., that

$$\pi\left(\mathbf{I} + \frac{1}{\lambda}Q\right) = \pi.$$

That identity holds since $\pi Q = 0$. Thus, the DTMC Z_n has the same invariant distribution as X_t. Observe that Z_n is not the same as the jump chain of X_t. Also, it is not a discrete-time approximation of X_t. This DTMC shows that a CTMC can be seen as a DTMC where one replaces the constant time steps by i.i.d. exponentially distributed time steps between the jumps.

6.2.5 Time Reversal

As a preparation for our study of networks of queues, we note the following result.

Theorem 6.2 (Kelly's Lemma) *Let Q be the rate matrix of a Markov chain on $\mathcal{X}$. Let also $\tilde{Q}$ be another rate matrix on $\mathcal{X}$. Assume that π is a distribution on $\mathcal{X}$ and that*

$$q_i = \tilde{q}_i, i \in \mathcal{X} \text{ and}$$
$$\pi(i)q(i, j) = \pi(j)\tilde{q}(j, i), \forall i \neq j.$$

Then $\pi Q = 0$.

■

Proof We have

$$\sum_{j \neq i} \pi(j)q(j, i) = \sum_{j \neq i} p(i)\tilde{q}(i, j) = p(i) \sum_{j \neq i} \tilde{q}(i, j) = p(i)\tilde{q}_i = p(i)q_i,$$

so that $\pi Q = 0$. □

The following result explains the meaning of $\tilde{Q}$ in the previous theorem. We state it without proof.

Theorem 6.3 *Assume that X_t has the invariant distribution π. Then X_t reversed in time is a Markov chain with rate matrix $\tilde{Q}$ given by*

$$\tilde{q}(i, j) = \frac{\pi(j)q(j, i)}{\pi(i)}.$$

■

6.3 Product-Form Networks

Theorem 6.4 (Invariant Distribution of Network) *Assume $\lambda_k < \mu_k$ and let $\rho_k = \lambda_k/\mu_k$, for $k = 1, 2, 3$. Then the Markov chain X_t has a unique invariant distribution π that is given by*

$$\pi(x_1, x_2, x_3) = \pi_1(x_1)\pi_2(x_2)\pi_3(x_3)$$
$$\pi_k(n) = (1 - \rho_k)\rho_k^n, n \geq 0, k = 1, 2$$
$$\pi_3(a_1, a_2, \ldots, a_n) = p(a_1)p(a_2)\cdots p(a_n)(1 - \rho_3)\rho_3^n,$$
$$n \geq 0, a_k \in \{1, 2\}, k = 1, \ldots, n,$$

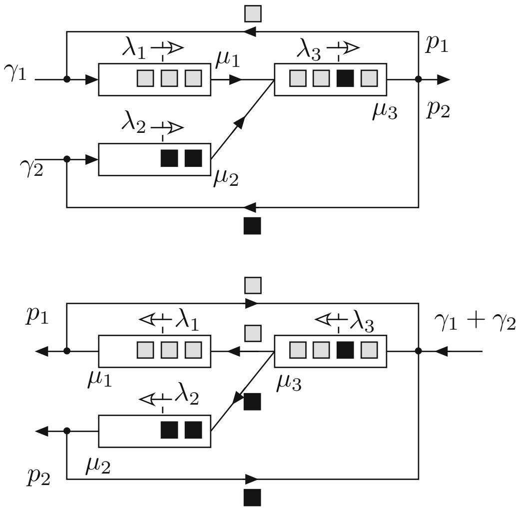


Fig. 6.10 The network (top) and a guess for its time-reversal (bottom). The bottom network is obtained from the top one by reversing the flows of customers. It is a bold guess that the arrivals have exponential inter-arrival times and their rates are independent of the current queue lengths

where $p(1) = \lambda_1/(\lambda_1 + \lambda_2)$ *and* $p(2) = \lambda_2/(\lambda_1 + \lambda_2)$.

Proof Figure 6.10 shows a guess for the time-reversal of the network.

Let Q be the rate matrix of the top network and $\tilde{Q}$ that of the bottom one. Let also π be as stated in the theorem. We show that π, Q, $\tilde{Q}$ satisfy the conditions of Kelly's Lemma.

For instance, we verify that

$$\pi([3, 2, [1, 1, 2, 1]])q([3, 2, [1, 1, 2, 1]], [4, 2, [1, 1, 2]])$$
$$= \pi([4, 2, [1, 1, 2]])\tilde{q}([4, 2, [1, 1, 2]], [3, 2, [1, 1, 2, 1]]).$$

Looking at the figure, we can see that

$$q([3, 2, [1, 1, 2, 1]], [4, 2, [1, 1, 2]]) = \mu_3 p_1$$
$$\tilde{q}([4, 2, [1, 1, 2]], [3, 2, [1, 1, 2, 1]] = \mu_1 p_1.$$

Thus, the previous identity reads

$$\pi([3, 2, [1, 1, 2, 1]])\mu_3 p_1 = \pi([4, 2, [1, 1, 2]])\mu_1 p_1,$$

i.e.,

$$\pi([3, 2, [1, 1, 2, 1]])\mu_3 = \pi([4, 2, [1, 1, 2]])\mu_1.$$

Given the expression for π, this is

$$(1-\rho_1)\rho_1^3 \times (1-\rho_2)\rho_2^2 \times p(1)p(1)p(2)p(1)(1-\rho_3)\rho_3^4\mu_3$$
$$= (1-\rho_1)\rho_1^4 \times (1-\rho_2)\rho_2^2 \times p(1)p(1)p(2)(1-\rho_3)\rho_3^3\mu_1.$$

After simplifications, this identity is seen to be equivalent to

$$p(1)\rho_3\mu_3 = \rho_1\mu_1,$$

i.e.,

$$\frac{\lambda_1}{\lambda_3}\frac{\lambda_3}{\mu_3}\mu_3 = \frac{\lambda_1}{\mu_1}\mu_1$$

and this equation is seen to be satisfied. A similar argument shows that Kelly's lemma is satisfied for all pairs of states. □

6.4 Proof of Theorem 5.7

The first step in using the theorem is to solve the flow conservation equations. Let us call class 1 that of the white jobs and class 2 that of the gray job. Then we see that

$$\lambda_1^1 = \lambda_2^1 = \gamma, \lambda_1^2 = \lambda_2^2 = \alpha$$

solve the flow conservation equations for any $\alpha > 0$. We have to assume $\gamma < \mu$ for the services to be able to keep up with the white jobs. With this assumption, we can choose α small enough so that $\lambda_1 = \lambda_2 = \lambda := \gamma + \alpha < \min\{\mu_1, \mu_2\}$.

The second step is to use the theorem to obtain the invariant distribution. It is

$$\pi(x_1, x_2) = Ah(x_1)h(x_2)$$

with

$$h(x_i) = \left(\frac{\gamma}{\mu}\right)^{n_1(x_i)} \left(\frac{\alpha}{\mu}\right)^{n_2(x_i)} = \rho_1^{n_1(x_i)}\rho_2^{n_2(x_i)},$$

where $\rho_1 = \gamma/\mu$, $\rho_2 = \alpha/\mu$, and $n_c(x)$ is the number of jobs of class c in x_i, for $c = 1, 2$. To calculate A, we note that there are $n + 1$ states x_i with n class 1 jobs and 1 class 2 job, and 1 state x_i with n classes 1 jobs and no class 2 job. Indeed, the class 2 customer can be in $n+1$ positions in the queue with the n customers of class 1.

Also, all the possible pairs (x_1, x_2) must have one class 2 customer either in queue 1 or in queue 2. Thus,

$$1 = \sum_{(x_1,x_2)} \pi(x_1, x_2) = A \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} G(m, n),$$

where

$$G(m, n) = (m + 1)\rho_1^{m+n}\rho_2 + (n + 1)\rho_1^{m+n}\rho_2.$$

In this expression, the first term corresponds to the states with m class 1 customers and one class 2 customer in queue 1 and n customers of class 1 in queue 2; the second term corresponds to the states with m customer of class 1 in queue 1, and n customers of class 1 and one customer of class 2 in queue 2. Thus, $AG(m, n)$ is the probability that there are m customers of class 1 in the first queue and n customers of class 1 in the second queue.

Hence,

$$1 = A \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} [(m + 1)\rho_1^{m+n}\rho_2 + (n + 1)\rho_1^{m+n}\rho_2] = 2A \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} (m + 1)\rho_1^{m+n}\rho_2,$$

by symmetry of the two terms. Thus,

$$1 = 2A\rho_2 \left[\sum_{m=0}^{\infty} (m + 1)\rho_1^m \right] \left[\sum_{n=0}^{\infty} \rho_1^n \right].$$

To compute the sum, we use the following identities:

$$\sum_{n=0}^{\infty} \rho^n = (1 - \rho)^{-1}, \text{ for } 0 < \rho < 1$$

and

$$\sum_{n=0}^{\infty} (n + 1)\rho^n = \frac{\partial}{\partial \rho} \sum_{n=0}^{\infty} \rho^{n+1} = \frac{\partial}{\partial \rho} [(1 - \rho)^{-1} - 1] = (1 - \rho)^{-2}.$$

Thus, one has

$$1 = 2A\rho_2(1 - \rho_1)^{-3},$$

so that

$$A = \frac{(1 - \rho_1)^3}{2\rho_2}.$$

Third, we calculate the expected number L of jobs of class 1 in the two queues. One has

$$
\begin{aligned}
L &= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} A(m+n)G(m,n) \\
&= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} A(m+n)(m+1)\rho_1^{m+n}\rho_2 + \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} A(m+n)(n+1)\rho_1^{m+n}\rho_2 \\
&= 2\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} A(m+n)(m+1)\rho_1^{m+n}\rho_2,
\end{aligned}
$$

where the last identity follows from the symmetry of the two terms. Thus,

$$
\begin{aligned}
L &= 2\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} Am(m+1)\rho_1^{m+n}\rho_2 + 2\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} An(m+1)\rho_1^{m+n}\rho_2 \\
&= 2A\rho_2\left[\sum_{m=0}^{\infty} m(m+1)\rho_1^m\right]\left[\sum_{n=0}^{\infty}\rho_1^n\right] + 2A\rho_2\left[\sum_{m=0}^{\infty}(m+1)\rho_1^m\right]\left[\sum_{n=0}^{\infty} n\rho_1^n\right] \\
&= 2A\rho_2\left[\sum_{m=0}^{\infty} m(m+1)\rho_1^m\right](1-\rho_1)^{-1} + 2A\rho_2(1-\rho)^{-2}\left[\sum_{n=0}^{\infty} n\rho_1^n\right].
\end{aligned}
$$

To calculate the sums, we use the fact that

$$
\begin{aligned}
\sum_{m=0}^{\infty} m(m+1)\rho^m &= \rho\sum_{m=0}^{\infty} m(m+1)\rho^{m-1} \\
&= \rho\frac{\partial^2}{\partial\rho^2}\sum_{m=0}^{\infty}\rho^{m+1} = \rho\frac{\partial^2}{\partial\rho^2}[(1-\rho)^{-1}-1] \\
&= 2\rho(1-\rho)^{-3}.
\end{aligned}
$$

Also,

$$
\sum_{n=0}^{\infty} n\rho_1^n = \rho_1\sum_{n=0}^{\infty} n\rho_1^{n-1} = \rho_1\sum_{n=0}^{\infty}(n+1)\rho_1^n = \rho_1(1-\rho_1)^{-2}.
$$

Hence,

$$
\begin{aligned}
L &= 2A\rho_2 \times 2\rho(1-\rho)^{-3} \times (1-\rho_1)^{-1} + 2A\rho_2(1-\rho)^{-2} \times \rho_1(1-\rho_1)^{-2} \\
&= 6A\rho_2\rho_1(1-\rho_1)^{-4}.
\end{aligned}
$$

Substituting the value for A that we derived above, we find

$$L = 3\frac{\rho_1}{1-\rho_1}.$$

Finally, we get the average time W that jobs of class 1 spend in the network: $W = L/\gamma$.

Without the gray job, the expected delay W' of the white jobs would be the sum of delays in two M/M/1 queues, i.e., $W' = L'/\gamma$ where

$$L' = 2\frac{\rho_1}{1-\rho_1}.$$

Hence, we find that

$$W = 1.5W',$$

so that using a hello message increases the average delay of the class 1 customers by 50%.

6.5 References

The time-reversal arguments are developed in Kelly (1979). That book also explains many other models that can be analyzed using that approach. See also Bremaud (2008), Lyons and Perez (2017), Neely (2010).