What we need from our model is to find out the behavior of the system in the long run. The complete discussion of the Markov chain itself deserves a whole textbook. (See e.g. [2].) In the next section, we only introduce the concept of Markov chain and some necessary properties.

## 3 Markov Chains

**3.1.** Markov chains are simple mathematical models for random phenomena evolving in time. Their simple structure makes it possible to say a great deal about their behavior. At the same time, the class of Markov chains is rich enough to serve in many applications. This makes Markov chains the most important examples of random processes. Indeed, the whole of the mathematical study of random processes can be regarded as a generalization in one way or another of the theory of Markov chains. [2]

**3.2.** The characteristic property of Markov chain is that it retains no memory of where it has been in the past. This means that only the current state of the process can influence where it goes next.

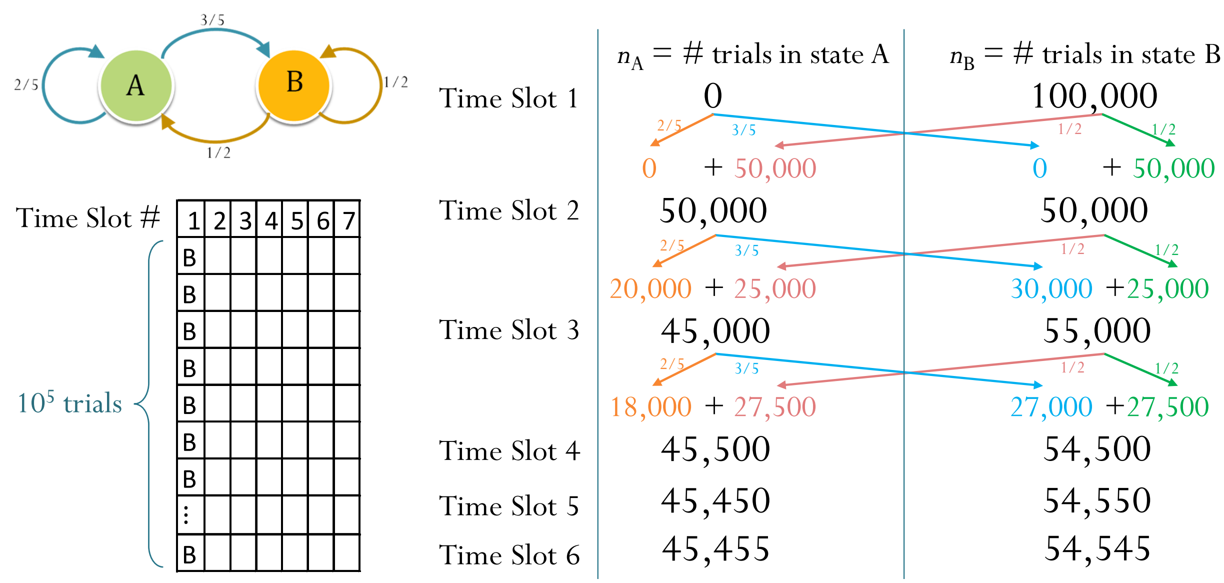
**3.3.** Markov chains are often best described by their (state) diagrams. You have seen a Markov chain in Example 2.10.

**Example 3.4.** Draw the (state) diagram of a Markov chain which has three states: $A$,$B$ and $C$. It moves from state $A$ to state $B$ with probability 1. From state $C$, it moves either to $A$ or to $B$ with equal probability 1/2, and from $B$ it jumps to $C$ with probability 1/3, otherwise it stays at $B$.

**3.5.** The Markov chains that we have just seen in the Example 3.4 and in the previous section (Example 2.10) are all ***discrete-time*** Markov chains. However, the Poisson process and the state $K(t)$ which we have studied earlier are examples of a continuous-time Markov chain. Nonetheless, equipped with our small-slot approximation (discrete-time approximation) technique, we may analyze the Poisson process or any continuous-time Markov chain as a limit of a discrete-time Markov chain as well.

**3.6.** We will now introduce the concept of stationary distribution, steady-state distribution, equilibrium distribution, and limiting distribution. For the purpose of this class, we will not distinguish these terms. We shall see in the next example that for the Markov chains that we are considering, in the long run, it will reach a steady state.

**Example 3.7.** Consider the Markov chain characterized by the state transition diagram below:




Let's try a thought experiment – imagine that you start with $n = 100,000$ trials of these Markov chain, all of which start in state B. So, during slot 1 (the first time slot), all trials will be in state B. For slot 2, about 50% of these will move to state A; but the other 50% of the trials will stay at $B$.

By the time that you reach slot 6, you can observe that out of the 100,000 trials, about 45.5% will be in state A and about 55.5% will be in state B.

Mathematically, this evolution can be calculated simply by

(a) Start with a row vector $\underline{\mathbf{n}}^{(1)} = \left[n_A^{(1)}, n_B^{(1)}\right]$ containing the initial number of trials in each state.

(b) For the $k$th time slot, the new row vector $\underline{\mathbf{n}}^{(k)}$ is the old row vector $\underline{\mathbf{n}}^{(k-1)}$ multiplied by the **transition matrix**

$$\mathbf{P} = \begin{bmatrix} 2/5 & 3/5 \\ 1/2 & 1/2 \end{bmatrix}.$$

```
>> P = sym([2/5 3/5; 1/2 1/2]);
```

```
>> n1 = sym([0 1e5])
n1 =
[ 0, 100000]
>> n2 = n1*P
n2 =
[ 50000, 50000]
>> n3 = n2*P
n3 =
[ 45000, 55000]
>> n4 = n3*P
n4 =
[ 45500, 54500]
>> n5 = n4*P
n5 =
[ 45450, 54550]
>> n6 = n5*P
n6 =
[ 45455, 54545]
```

```
>> n1 = sym([1e4 9e4])
n1 =
[ 10000, 90000]
>> n2 = n1*P
n2 =
[ 49000, 51000]
>> n3 = n2*P
n3 =
[ 45100, 54900]
>> n4 = n3*P
n4 =
[ 45490, 54510]
>> n5 = n4*P
n5 =
[ 45451, 54549]
>> n6 = n5*P
n6 =
[ 45454.9, 54545.1]
```

$\underline{\mathbf{n}}^{(1)} =$ … $\xrightarrow{\times 1/n} = \underline{\mathbf{p}}^{(1)}$

$\underline{\mathbf{n}}^{(2)} =$ … $\xrightarrow{\times 1/n} = \underline{\mathbf{p}}^{(2)}$

$\underline{\mathbf{n}}^{(3)} =$ … $\xrightarrow{\times 1/n} = \underline{\mathbf{p}}^{(3)}$

$\underline{\mathbf{n}}^{(4)} =$ … $\xrightarrow{\times 1/n} = \underline{\mathbf{p}}^{(4)}$

$\underline{\mathbf{n}}^{(5)} =$ … $\xrightarrow{\times 1/n} = \underline{\mathbf{p}}^{(5)}$

$\underline{\mathbf{n}}^{(6)} =$ … $\xrightarrow{\times 1/n} = \underline{\mathbf{p}}^{(6)}$

The relative frequencies (of the two states) can be found from $\frac{1}{n}\underline{\mathbf{n}}^{(k)}$. Recall that when $n$ is large, relative frequencies converge to probabilities.

Turn out that the relative frequencies $[45.5\%, 55.5\%]$ stay roughly the same as you proceed to slot 7, 8, 9, and so on. Note also that it does not matter how we start our 100,000 trials. We may start with 10,000 in state A and 90,000 in state B. Eventually, $[45.5\%, 55.5\%]$ will emerge.

In conclusion,

(a) If we observe the long-run behavior of this Markov chain at a particular slot, then the probability that we will see it in state A is 0.455 and the probability that you will see it in state B is 0.545.

(b) In addition, one can also show that if we observe the behavior of this Markov chain for a long time, then the proportion of time that it stays in state A is 45.5% and the proportion of time that it stays in state B is 54.5%.

The distribution $[0.455, 0.545]$ is what we referred to as stationary distribution, steady-state distribution, equilibrium distribution, or limiting distribution above.

**3.8.** Another way to look at the convergence in Example 3.7 is to first look at $\mathbf{P}^k$. This is because $\underline{\mathbf{n}}^{(k+1)} = \underline{\mathbf{n}}^{(1)}\mathbf{P}^k$.

```
P =
    0.4000    0.6000
    0.5000    0.5000
```

```
>> P^2
ans =
    0.4600    0.5400
    0.4500    0.5500
```

```
>> P^3
ans =
    0.4540    0.5460
    0.4550    0.5450
```

```
>> P^4
ans =
    0.4546    0.5454
    0.4545    0.5455
```

```
>> P^5
ans =
    0.4545    0.5455
    0.4546    0.5454
```

```
>> P^7
ans =
    0.4545    0.5455
    0.4545    0.5455
```

```
>> P^6
ans =
    0.4545    0.5455
    0.4545    0.5455
```

Analytically, the convergence in Example 3.7 can be shown easily by realizing that the transition $\mathbf{P}$ can be decomposed into $\mathbf{P} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$ where $\mathbf{D}$ is a diagonal matrix of eigenvalues and $\mathbf{V}$ is a matrix whose columns contain the corresponding eigenvectors. In `MATLAB`, the matrices $\mathbf{V}$ and $\mathbf{D}$ can be found easily from the command `[V,D] = eig(P)`. For Example 3.7,

$$\mathbf{P} = \begin{bmatrix} 2/5 & 3/5 \\ 1/2 & 1/2 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} -1/10 & 0 \\ 0 & 1 \end{bmatrix}, \text{ and } \mathbf{V} = \begin{bmatrix} -6/5 & 1 \\ 1 & 1 \end{bmatrix}.$$

Therefore,

$$\lim_{k\to\infty} \mathbf{P}^k = \lim_{k\to\infty} \left(\mathbf{V}\mathbf{D}\mathbf{V}^{-1}\right)^k = \lim_{k\to\infty} \mathbf{V}\mathbf{D}^k\mathbf{V}^{-1} = \mathbf{V}\left(\lim_{k\to\infty} \mathbf{D}^k\right)\mathbf{V}^{-1}.$$

From the matrix $\mathbf{D}$ above, it is easy to see that

$$\lim_{k\to\infty} \mathbf{D}^k = \lim_{k\to\infty} \begin{bmatrix} (-1/10)^k & 0 \\ 0 & (1)^k \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

This gives

$$\lim_{k\to\infty} \mathbf{P}^k = \begin{bmatrix} 5/11 & 6/11 \\ 5/11 & 6/11 \end{bmatrix}.$$

Therefore, it does not matter what we have for $\underline{\mathbf{n}}^{(1)}$. As $k \to \infty$,

$$\underline{\mathbf{n}}^{(k+1)} = \underline{\mathbf{n}}^{(1)}\mathbf{P}^k \to \underline{\mathbf{n}}^{(1)} \begin{bmatrix} 5/11 & 6/11 \\ 5/11 & 6/11 \end{bmatrix} = \begin{bmatrix} \frac{5}{11}n & \frac{6}{11}n \end{bmatrix}$$

where $n$ is the total number of trials and

$$\underline{\mathbf{p}}^{(k+1)} = \frac{1}{n}\underline{\mathbf{n}}^{(k+1)} \to \begin{bmatrix} \frac{5}{11} & \frac{6}{11} \end{bmatrix}.$$

**3.9.** From Example 3.7, we can see that Discrete-time Markov chain can be analyzed via its **state transition diagram** or its **probability transition matrix P**.

**Example 3.10.** Consider an evolution of a Markov chain which has two states (1 and 2):

$$2, 2, 1, 1, 1, 2, 2, 1, 2, 1$$

(a) Estimate its transition matrix $\mathbf{P}$.

(b) Let $p_i$ denote the proportion of time that the system spends in state $i$. Estimate $p_i$.

**3.11.** Long-term behavior of a discrete-time Markov chain can be studied in terms of its steady-state (or limiting or equilibrium) probabilities. To analytically find these probabilities, in 3.8, we use a technique called eigendecomposition (spectral decomposition). However, in this class, we will focus on another (easier) technique.

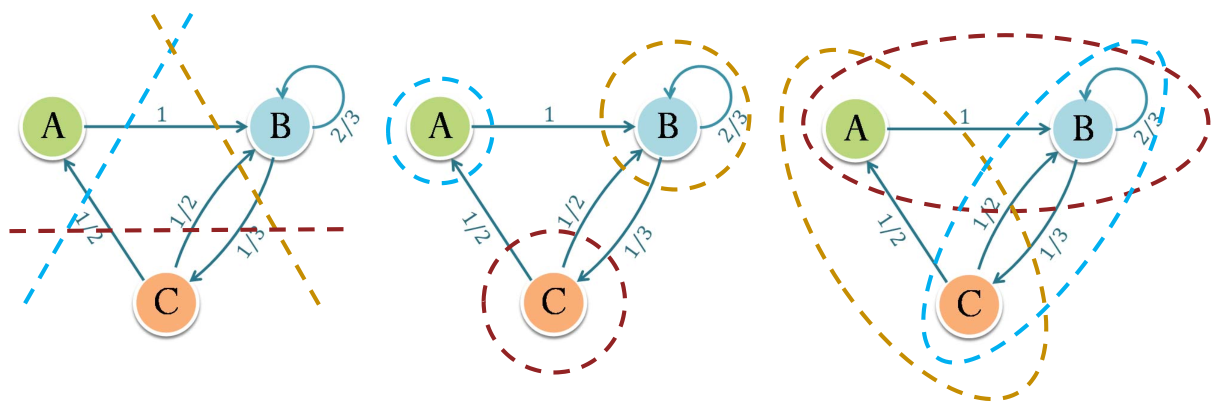
**Example 3.12.** In Example 3.7, instead of finding the limit of the distribution, if we assume that the system will reach some steady-state values, then at the steady-state, we must have roughly the same number of transitions from state A to B and transitions from state B to state A.

Therefore,

**Definition 3.13.** A **balance equation** is an equation that describes the probability flux associated with a Markov chain in and out of states or set of states.

To write down a balance equation, first define a boundary, then consider the transfer of probabilities "in" and "out" of the boundary. To be at equilibrium, there should not be any net transfer.

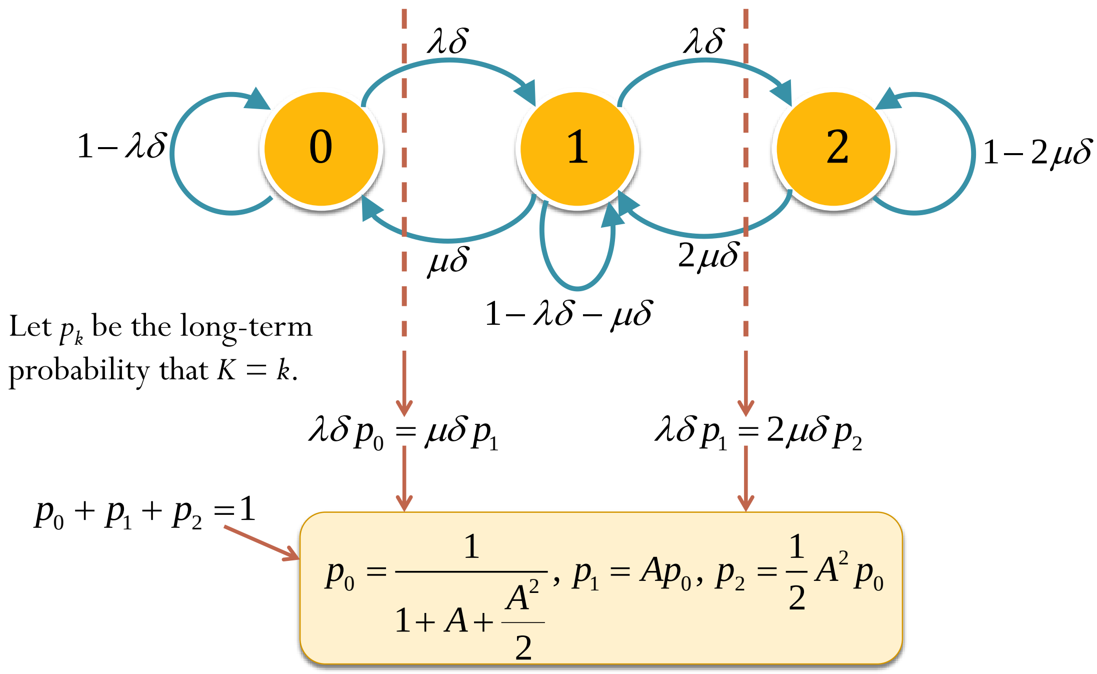
**Example 3.14.** Let's reconsider the state diagram derived in Example 3.4. Write the balance equation for each of the boundaries below.




**3.15.** The steady-state probabilities can be found from the **balance equations** together with the fact that the sum of them must be 1.

**Example 3.16.** For the the state diagram derived in Example 3.4. We can use the balance equations expressed in Example 3.14 to find its steady-state probabilities:

**Example 3.17.** Let's reconsider Example 2.10 where $m = 2$.




**3.18.** Reminder: Two Interpretations of steady-state probabilities: When we let a system governed by a Markov chain evolve for a long time,

(a) at a particular slot, the probability that we will find the system in a particular state can be approximated by its corresponding steady-state probability,

(b) considering the whole evolution up to a particular time, the **proportion of time that the system is in a particular state** can be approximated by its corresponding steady-state probability.

**3.19.** Finally, we can now use what we learned to derive the Erlang B formula. In general, if we have $m$ channels, then

$$p_m = \frac{\frac{A^m}{m!}}{\sum_{k=0}^{m} \frac{A^k}{k!}}.$$

Note that $p_m$ is the (long-run) probability that the system is in state $m$. When the system is in state $m$, all channels are used and therefore any new call request will be blocked and lost.

Here, $p_m$ is the same as call blocking probability $P_b$, which is the long-run proportion of call requests that get blocked.

**3.20.** Convention: In this class, for any question that requires you to get your answers (call blocking probability or steady-state probabilities) "from the Markov chain" or "via the Markov chain", make sure that you

(a) draw the Markov chain (with all the states and transition probabilities),

(b) set up the boundaries,

(c) write down the corresponding balance equations,

(d) use the equations to solve for the interested quantities.

**3.21.** Remark: Beyond this class, mathematicians do have more direct ways to analyze continuous-time Markov chain without the discrete-time approximation that we have been using here. However, the analysis requires more background knowledge and hence we did not try to use it.

Here is a glimpse of what's out there:

(a) Our version:

- Note that the value of $\delta$ itself is not important as long as it is "small enough". When we calculate the steady-state probabilities, $\delta$ disappears anyway.
- The returning-to-the-same-state arrows are not used at all because it will not cross (or cross and then return) any boundary.

(b) The-rest-of-the-world version:

- The label on each arrow indicates probability transition rate instead of transition probability.
- No arrow for returning to the same state