Sirindhorn International Institute of Technology
Thammasat University
School of Information, Computer and Communication Technology

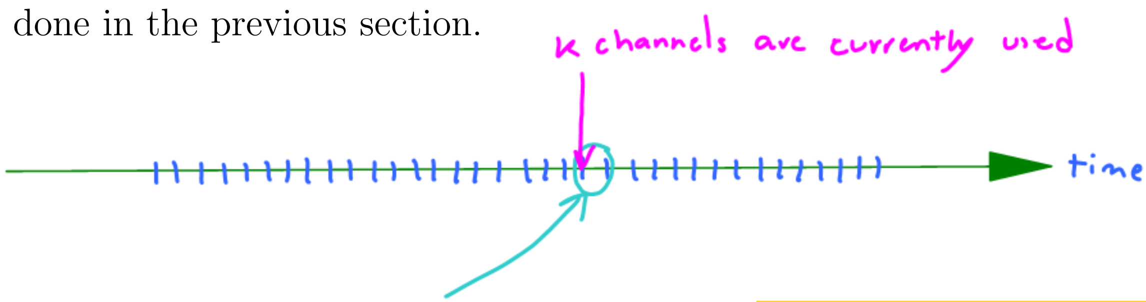
ECS455 2011/2, Chapter 3 Part 2, Dr.Prapun

In the previous section, we discussed Poisson process. In this section, we combine what we know about Poisson process with the assumption on the call length/duration. To do this, we will use (again) the small-slot (discrete-time) approximation.

2.1. For this section, we will assume that there are m channels available in the trunking pool. Therefore, the probability P_b that a call requested by a user will be blocked is given by the probability that none of the m channels are free when a call is requested.

We will consider the long-term behavior of this system, i.e. the system is assumed to have been operating for a long time. In which case, at the instant that somebody is trying to make a call, we don't know how many of the channels are currently free.

2.2. Let's first divide the time into small slots (of the same length δ) as we have done in the previous section.



Then, consider any particular slot. Suppose that at the beginning of this time slot, there are K channels that are currently used.[10] We want to find out how this number K changes as we move forward one slot time. This

[10]The value of K can be any integer from 0 to m.

random variable K will be call the ***state*** of the system[11]. The system moves from one state to another one as we advance one time slot.

Example 2.3. Suppose there are 5 person using the channels at the beginning of a slot. Then, $K = 5$.

(a) Suppose that, by the end of this slot, none of these 5 persons finish their calls.

(b) Suppose also that there is one new person who wants to make a call at some moment of time during this slot.

Then, at the end of this time slot, the number of channels that are used becomes

$$5 - 0 + 1 = 6.$$

So, the state K of the system changes from 5 to 6 when we reach the end of the slot, which can now be regarded as the beginning of the next slot.

2.4. Our current intermediate goal is to study how the state K changes from one slot to the next slot. Note that it might be helpful to label the state K as K_1 (or $K[1]$) for the first slot, K_2 (or $K[2]$) for the second slot, and so on.

As shown in Example 2.3, to determine how the K_i's progress from K_1 to K_2 to K_3 and so on, we need to know two pieces of information:

Q1 How many calls (that are being made at the beginning of the slot under consideration) ***end*** during the slot that we are considering?

Q2 How many ***new*** call requests are made during the slot under consideration?

Note that Q1 depends on the characteristics of the call duration and Q2 depends on the characteristics of the call request/arrival process. After we know the answers to these two question, then we can find K_i via

$$K_{i+1} = K_i - \underbrace{(\text{\# old call ends})}_{\text{Q1}} + \underbrace{(\text{\# new call requests})}_{\text{Q2}}$$

[11]This is the same "state" concept that you have studied in digital circuits class.

2.5. Q2 is easy to answer.

A2: If the interval are small enough (δ is small), then there can be at most one new arrival (new call request) which occurs with probability

$$p_1 = \lambda\delta.$$

2.6. For Q1, we need to consider the ***call duration model***. The M/M/m/m assumption states that the call duration[12] is exponentially distributed. Let's consider the call duration D of a particular call. $\sim \mathcal{E}(\mu)$

Recall that the probability density function (pdf) of an exponential random variable X with parameter μ is given by

$$f_X(x) = \begin{cases} \mu e^{-\mu x}, & x > 0, \\ 0, & x \le 0. \end{cases}$$

and the average (or expected value) is given by

$$\mathbb{E}[X] = \int_0^\infty x f_X(x) dx = \frac{1}{\mu}.$$

You may remember that in the Erlang B formula, we assume that the average call duration is $\mathbb{E}[D] = H = \frac{1}{\mu}$.

An important property of an exponential random variable X is its memoryless property[13]:

$$P[X > x + \delta | X > x] = P[X > \delta].$$

For example,

$$P[X > 7 | X > 5] = P[X > 2].$$

[12] In queueing theory, this is sometimes called the ***service time***.

[13] To see this, first recall the definition of conditional probability:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Therefore,

$$P[X > x + \delta | X > x] = \frac{P[X > x + \delta \text{ and } X > x]}{P[X > x]} = \frac{P[X > x + \delta]}{P[X > x]}.$$

Now,

$$P[X > x] = \int_x^\infty \mu e^{-\mu x} dx = e^{-\mu x}.$$

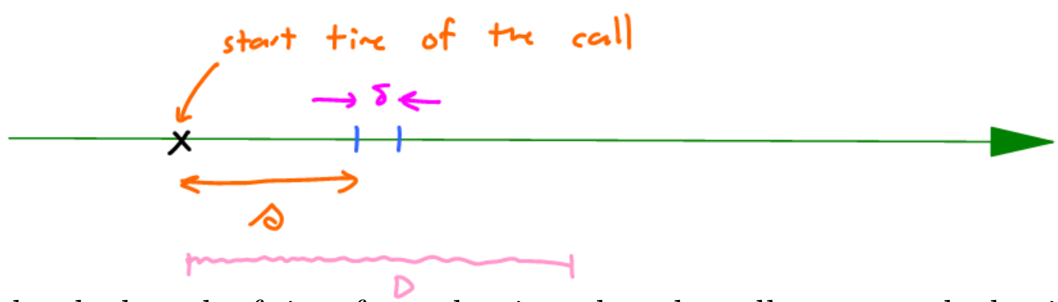
Hence,

$$P[X > x + \delta | X > x] = \frac{e^{-\mu(x+\delta)}}{e^{-\mu x}} = e^{-\mu\delta} = P[X > \delta].$$

What does this memoryless property mean? Suppose you have a lightbulb and you have used it for 5 years already. Let it's lifetime be X. Then, of course, X is a random variable. You know that $X > 5$ because it is still working now. The conditional probability $P[X > 7|X > 5]$ is the probability that it will still work after two more years of use (given the fact that currently it has been working for five years). Now, if X is an exponential random variable, then the memoryless property says that $P[X > 7|X > 5] = P[X > 2]$. In other words, the probability that you can use it for at least two more years is exactly the same as the probability that you can use a new lightbulb for more than two years. So, your old lightbulb essentially forgets that it has been used for 5 years, It always act as a ne lightbulb (statistically). This is we we mean by the memoryless property of an exponential random variable.

2.7. Note that we still haven't answered Q1. We will now return to our small slot approximation. Again, consider one particular slot. At the beginning of our slot, there are $K = k$ ongoing calls. The probability that a particular call, which is still ongoing at the beginning of this slot, will be unfinished by the end of this slot is $e^{-\mu\delta}$.

To see this, consider a particular call. Suppose the duration of this call is D. By assumption, we know that D is exponential with parameter μ.



Let s be the length of time from the time that the call starts to the beginning of our slot. Note that the call is still ongoing. Therefore $D > s$. Now, to say that this call will be unfinished by the end of our slot is equivalent to requiring that $D > s + \delta$. By the memoryless property, we have

$$P[D > s + \delta\,|D > s] = P[D > \delta] = e^{-\mu\delta}. \quad \approx 1 - \mu\delta$$

Recall that we have $K = k$ ongoing calls at the beginning of our slot. So, by the end of our slot, the probability that none of them finishes is

$$\left(e^{-\mu\delta}\right)^k = e^{-k\mu\delta}. \quad \approx 1 - k\mu\delta$$

$e^{x} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$

The probability that exactly one of them finishes is

$$\underbrace{k}_{\substack{\text{pick one}\\ \text{of the } k\\ \text{calls to}\\ \text{end in}\\ \text{our slot}}} \times \underbrace{\left(1-\overbrace{e^{-\mu\delta}}^{\approx 1-\mu\delta}\right)}_{\substack{\text{one call}\\ \text{ends}}} \times \underbrace{\left(e^{-\mu\delta}\right)^{k-1}}_{\substack{k-1 \text{ calls}\\ \text{are still}\\ \text{ongoing}}} \approx k\times(\mu\delta)\times(1-(k-1)\mu\delta)$$

by the end of our slot.

$$\approx k\mu\delta$$

Now, note that $e^x \approx 1 + x$ for small x. Therefore, A1:

(a) the probability that none of the $K = k$ calls ends during our slot is

$$\approx 1 - k\mu\delta$$

(b) the probability that exactly one of them ends during our slot is

$$\approx k\mu\delta$$

Magically, these two probabilities sum to one. So, we don't have to consider other events/cases.

2.8. Summary:

call generation/request/initiation process

$$P[0 \text{ new call}] \approx 1-\lambda\delta$$
$$P[1 \text{ new call}] \approx \lambda\delta$$

call duration process
$$P[0 \text{ old call ends}] \approx 1-k\mu\delta$$
$$P[1 \text{ old call ends}] \approx k\mu\delta$$

So, after one (small) slot, there can be four events:

(i) 0 new call & 0 old-call ends $\to K$ stays the same. $(1-\lambda\delta)(1-k\mu\delta)$

(ii) 0 " & 1 " " $\to K$ decreases by 1 $(1-\lambda\delta)\,k\mu\delta$

(iii) 1 new call & 0 " " $\to K$ increases by 1 $\lambda\delta\,(1-k\mu\delta)$

(iv) 1 " " & 1 " " $\to K$ stays the same. $\lambda\delta\, k\mu\delta$

The corresponding probability for each case is

(i) $(1-\lambda\delta)(1-k\mu\delta) \approx 1-\lambda\delta-k\mu\delta$

(ii) $(1-\lambda\delta)k\mu\delta \approx k\mu\delta$

(iii) $\lambda\delta(1-k\mu\delta) \approx \lambda\delta$

(iv) $\lambda\delta\, k\mu\delta \approx 0$

Therefore, if we have $K = k$ at the beginning of our time slot, the at the end of our slot, K may

(a) remain unchanged with probability $1 - \lambda\delta - k\mu\delta$, or

(b) decrease by 1 with probability $k\mu\delta$, or

(c) increase by 1 with probability $\lambda\delta$.

This can be summarized into the following diagram:

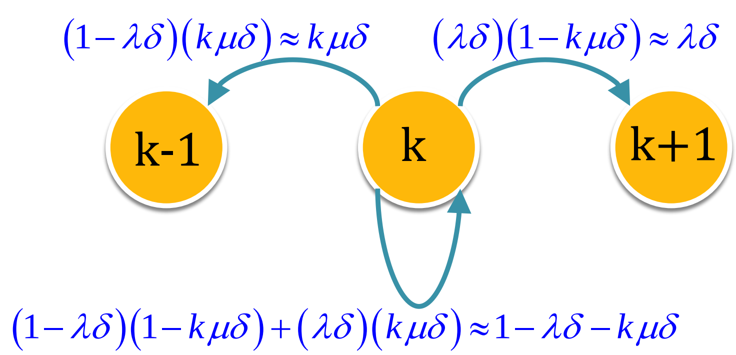
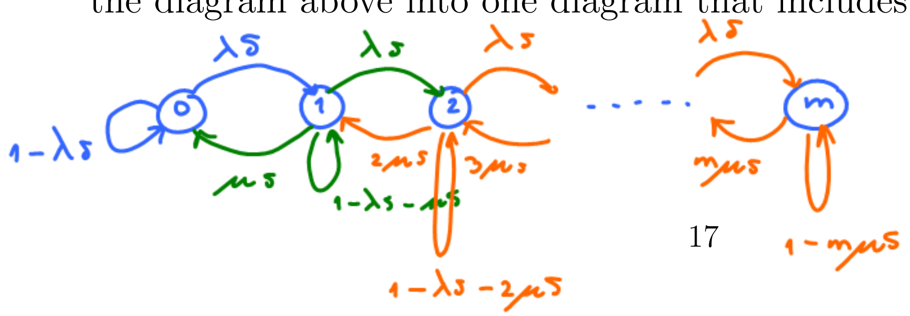


Figure 5: State transition diagram for state k.

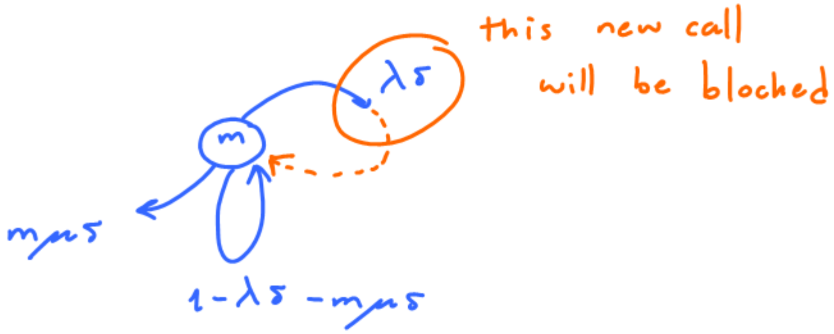
Note that the labels on the arrows indicate transition probabilities which are conditional probabilities of going from some value of K to another value.)

2.9. Given m, the possible values of K are $0, 1, 2, \ldots, m$. We can combine the diagram above into one diagram that includes all possible values of K:

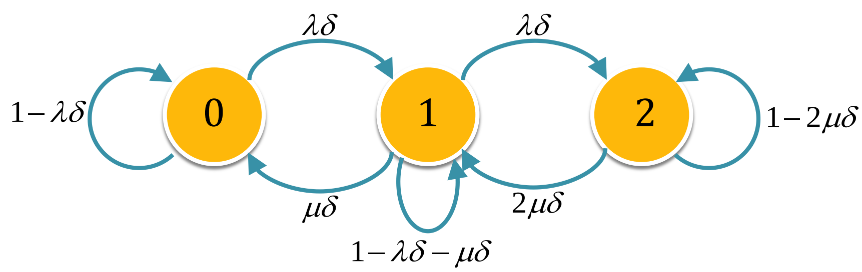


This diagram is called the ***Markov chain state diagram*** for Erlang B.

Note that the arrow $\lambda\delta$ which should go out of state m will return to state m itself because it means blocked calls which do not increase the value of K.



Example 2.10. $m = 2$:



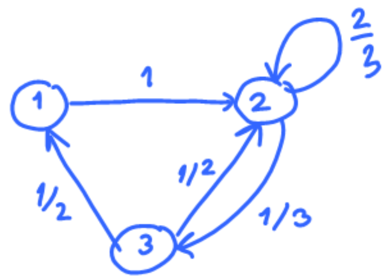
3 Markov Chains

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 Markv chains the most important examples of random processes. In deed, 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 diagrams.

Example 3.4.

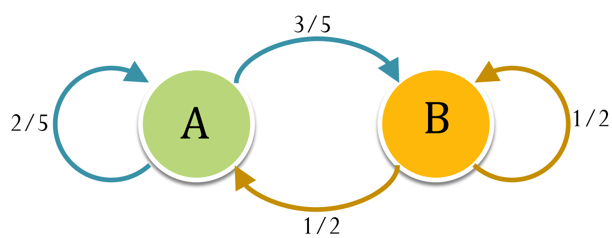


You move from state 1 to state 2 with probability 1. From state 3 you move either to 1 or to 2 with equal probability 1/2, and from 2 you jump to 3 with probability 1/3, otherwise stay at 2.

3.5. The Markov chains that we have seen in the Example 3.4 and in the previous section are all ***discrete-time*** Markov chains. The Poisson process that we have seen earlier is an example of a continuous-time Markov chain. However, with our small-slot approximation (discrete-time approximation), we may analyze the Poisson process as a discrete-time Markov chain.

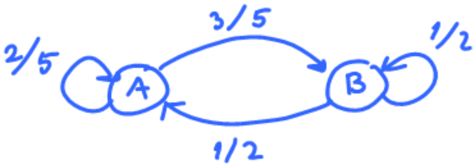
3.6. We will now introduce the concept of stationary distribution, steady-state distribution, equilibrium distribution, and limiting distribution. For the purpose of this class, we will not distinguish these terms. We shall see in the next example that for the Markov chains that we are considering, in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

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



Let's try a thought experiment – imagine that you start with 100,000 trials of these Markov chain, all of which start in state B. So, during slot 1, 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.



100,000 trials

slot	1	2	3	4	5	6	7	8	9	10
	B	A	A	B	A	A	B	B	A	B
	B									
	B									
	B									
	B									
	B									
	⋮									
	B									

10^5 trials

	A	B
slot 1	0	100,000
slot 2	50,000	50,000
slot 3	20,000 + 25,000 = 45,000	30,000 + 25,000 = 55,000
slot 4	18,000 + 27500 = 45,500	27,000 + 27500 = 54,500
slot 5	45,450	54,550
slot 6	45455	54545
slot 7	45454.5	54545.5
⋮		
slot 1,000	n_A	n_B
1,001	n_A	n_B

1/2, 1/2

40%, 60%

$n_B \frac{1}{2}$ $\quad n_A \frac{3}{5}$

$$\left.\begin{aligned} n_A + n_B &= n \\ \frac{1}{2}\frac{n_B}{n} &= \frac{3}{5}\frac{n_A}{n} \end{aligned}\right\}$$

$$\left.\begin{aligned} \frac{1}{2} P_B &= \frac{3}{5} P_A \\ P_A + P_B &= 1 \end{aligned}\right\}$$

$$P_B = \frac{6}{5} P_A$$

$$P_A + \frac{6}{5} P_A = 1 \Rightarrow P_A = \frac{5}{11}$$

These numbers stay roughly the same as you proceed to slot 7, 8, 9, and so on. Note also that it does not matter how you start your 100,000 trials. You may start with 10,000 in state A and 90,000 in state B. Eventually, the same numbers, 45.5% and 5~~5~~4.5%, will emerge.

In conclusion,

(a) If you look at the long run behavior of the Markov chain at a particular slot, then the probability that you will see it in state A is 0.455 and the probability that you will see it in state B is 0.5~~5~~45.

(b) In addition, one can also show that if you look at behavior of a 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 5~~5~~4.5%.

The distribution $(0.455, 0.5\cancel{5}45)$ is what we called stationary distribution, steady-state distribution, equilibrium distribution, or limiting distribution above.

3.8. In [3], the steady-state distribution for the M/M/m/m assumption is derived via the use of the ***global balance equation*** instead of finding the limit of the distribution as we have done in Example 3.7. The basic idea that leads to the global balance equation is simple. When we look back at the numbers that we got in Example 3.7, if we assume that they 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,

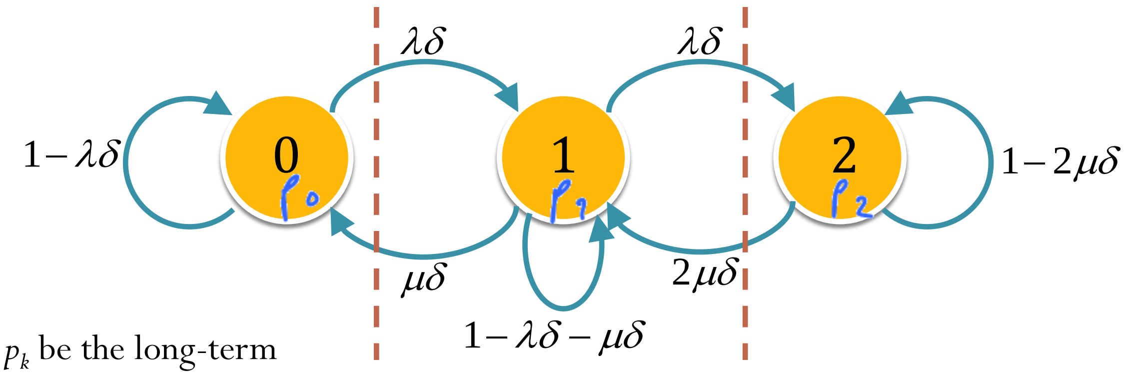
$$\left.\begin{array}{l} n_A + n_B = n \\ \frac{1}{2}\frac{n_B}{n} = \frac{3}{5}\frac{n_A}{n} \end{array}\right\}$$

$$\left.\begin{array}{l} \frac{1}{2}p_B = \frac{3}{5}p_A \\ p_A + p_B = 1 \end{array}\right\}$$

$$p_B = \frac{6}{5}p_A$$

$$p_A + \frac{6}{5}p_A = 1 \Rightarrow p_A = \frac{5}{11}$$

Finally, we can now use what we learned to derive the Erlang B formula.



Let p_k be the long-term probability that $K = k$.

Global Balance equations

$\lambda\cancel{\delta} p_0 = \mu\cancel{\delta} p_1$ $\qquad$ $\lambda\cancel{\delta} p_1 = 2\mu\cancel{\delta} p_2$

$p_1 = \frac{\lambda}{\mu} p_0 = A p_0$ $\qquad$ $p_2 = \frac{\lambda}{2\mu} p_1 = \frac{1}{2} A p_1 = \frac{1}{2} A^2 p_0$

$p_0 + p_1 + p_2 = 1$

$p_0 + Ap_0 + \frac{1}{2}A^2 p_0 = 1$

$$\boxed{p_0 = \frac{1}{1+A+\frac{A^2}{2}},\ p_1 = Ap_0,\ \boxed{p_2 = \frac{1}{2}A^2 p_0}}$$

$$\boxed{p_b = p_m}$$

$$p_2 = \frac{\frac{1}{2}A^2}{1+A+\frac{A^2}{2}}$$

Example 3.9. Let's reconsider Example 2.10 where $m = 2$.

The same process can be used to derive the the Erlang B formula.

3.10. 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, which is the long-run proportion of call requests that get blocked.