# CSCE 619-600
# Networks and Distributed Processing
# Spring 2017

## Markov Chains III

Dmitri Loguinov
Texas A&M University

March 2, 2017

# Agenda

- Renewal theory connection
  - Stationary transition probabilities
- Random 1D walk
  - Null chain example
- Discrete chains continued
  - Boundary transition rates
  - Random walks on graphs

# Renewal Theory Connection

- In this lecture, we examine Markov chains for which stationary distribution $\pi$ exists
  - We will then show that it is the solution to the same equation $\pi = \pi P$
- Let $T_{ij}$ be the first passage time from state $i$ to state $j$:

$$T_{ij} = \min(n \geq 1 : X_n = j | X_0 = i)$$

- For $i = j$, the visit is called a return
  - $T_{ii}$ is the random variable giving the number of steps (delay) before the chain returns to state $i$ after it started in state $i$
  - Notice that $T_{ii}$ does not depend on *how* the chain arrived into state $i$ (Markov property)

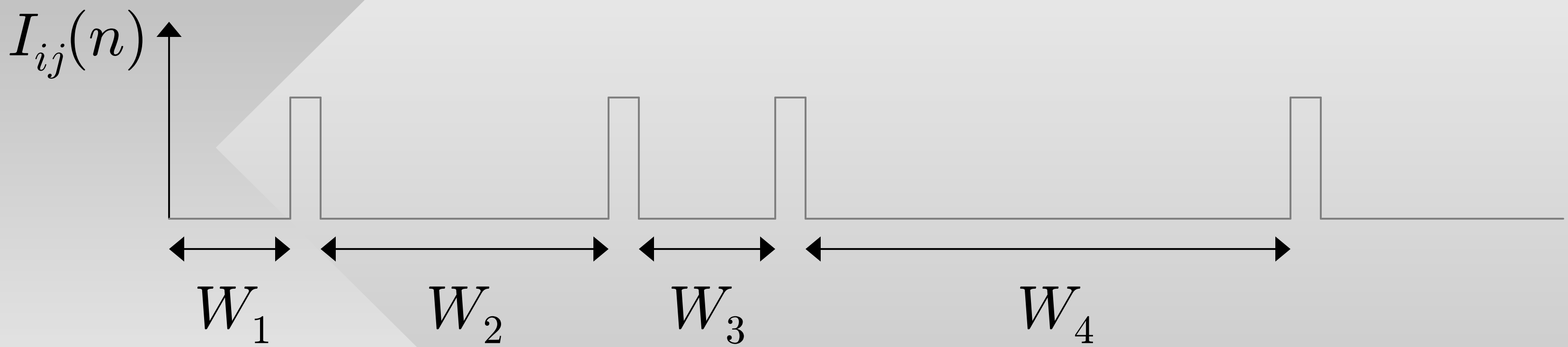
# Renewal Theory Connection 2

- Suppose $f_{ij}$ is the probability that the chain ever visits state $j$ starting in state $i$
- Definition: state $j$ is *recurrent* if $f_{jj} = 1$ and *transient* otherwise (i.e., $f_{jj} < 1$)
- Definition: assume $j$ is recurrent; if $E[T_{jj}] < \infty$ , then $j$ is called *positive*; otherwise, it is called *null*
- Define a renewal process $M_{ij}(n)$ to count number of visits to state $j$ by time $n$ assuming that $X_0 = i$

$$M_{ij}(n) = \sum_{k=1}^{n} I_{ij}(k) \qquad I_{ij}(n) = \begin{cases} 1 & X_n = j \\ 0 & otherwise \end{cases}$$

# Renewal Theory Connection 3

- Graphical illustration:




- Note that $W_1$ has the same distribution as $T_{ij}$
  - All of the remaining $W_k$, $k \geq 2$, are iid random variables $T_{jj}$, which yields from the Elementary Renewal Theorem

$$\lim_{n\to\infty} \frac{E[M_{ij}(n)]}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)} = \frac{1}{E[T_{jj}]}$$

# Renewal Theory Connection 4

- From this point on, assume a positive chain
- Recall that $$a^{(n)} = aP^n$$
  - and $$\pi = \lim_{n\to\infty} \frac{1}{n} \sum_{s=1}^{n} a^{(s)} = \lim_{n\to\infty} \frac{a}{n} \sum_{s=1}^{n} P^s$$
- From the previous slide, this limit exists and thus: $$\pi_i = \frac{1}{E[T_{ii}]}$$
- Additionally, since $$\pi P = \lim_{n\to\infty} \frac{a}{n} \sum_{s=1}^{n} P^{s+1} = \pi$$
  - we get $$\pi = \pi P$$

# Random 1D Walk

- Example

$$P = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & \dots \\ q & 0 & p & \dots & \dots & \dots \\ 0 & q & 0 & p & \dots & \dots \\ 0 & 0 & q & 0 & p & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

- The chain represents a random walk with a reflective (non-absorbing) boundary at 0
  - At each step, it goes left wp $q$ and right wp $p$
  - Except one special case when it bounces off zero with probability $1$

# Random 1D Walk 2

- We assume that $p + q = 1$ and $pq > 0$
- Next, we analyze the stationary distribution of this infinite chain
- Write:

$$\begin{cases} \pi_0 &= q\pi_1 \\ \pi_1 &= \pi_0 + q\pi_2 \\ \pi_j &= p\pi_{j-1} + q\pi_{j+1} \end{cases}$$

- Solving this recursively and using induction:

$$\begin{cases} \pi_1 &= \pi_0/q \\ \pi_2 &= (p/q)\pi_0/q \\ \cdots & \cdots \quad \cdots \\ \pi_j &= (p/q)^{j-1}\pi_0/q, j \geq 1 \end{cases}$$

# Random 1D Walk 3

- Note, however, that we still do not know $\pi_0$
  - This is accomplished using normalization since all $\pi_j$ must sum up to $1$:

$$1 = \sum_{j=0}^{\infty} \pi_j = \pi_0 \left( 1 + \frac{1}{q} \sum_{j=1}^{\infty} \left( \frac{p}{q} \right)^{j-1} \right)$$

  - When $p < q$, the sum is finite and all states are positive (the walk is “pulled” towards zero and $E[T_{jj}]$ is finite for all states $j$):

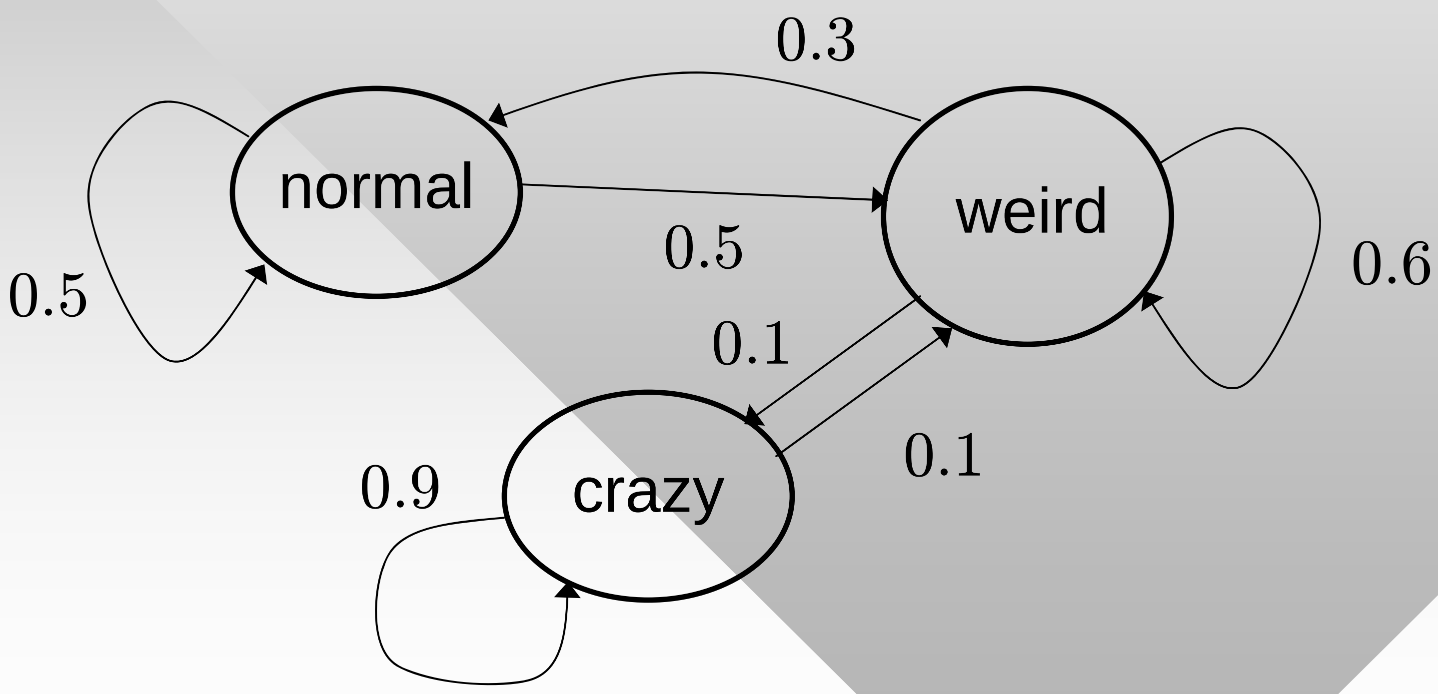
$$\pi_0 = \frac{q-p}{2q}$$

  - If $p = q = 1/2$, then all states are null (the expected duration before return to each state is infinity)
  - Finally, if $p > q$, then all states are transient, i.e., the chain keeps drifting towards infinity

# Example

Note: matlab eig(A) produces *right* eigenvectors; to get the left ones, transpose A first, i.e., eig(A')

- A PhD student goes through 3 states
  - Find the probability that on a given day the student is weird



# Transition Rates

- In what follows, we establish a useful rule that allows a simpler computation of $\pi$
- First, notice that the number of transitions *into* and *out of* a given state $j$ are almost the same
  - Define $A_j(n)$ to be the number of arrivals into $j$ by time $n$ and $D_j(n)$ the number of departures (including self-loops)
  - Clearly, the difference between these two metrics is no more than $1$ at any time $n$
- Thus, arrival and departure rates are asymptotically the same:

$$|r_A - r_D| = \left|\frac{A_j(n)}{n} - \frac{D_j(n)}{n}\right| = \frac{|A_j(n) - D_j(n)|}{n} \to 0$$

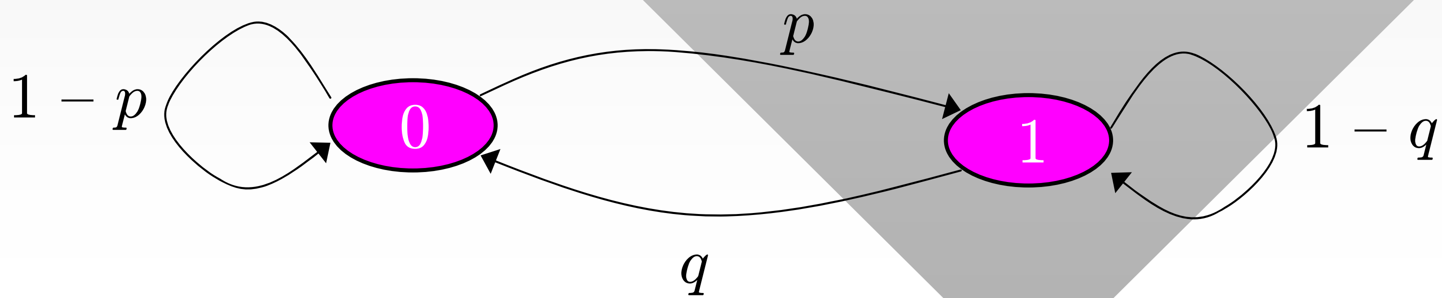
# Transition Rates 2

- Similarly observe the following
  - The probability to find a recurrent chain in state $j$ is equal to the rate of transition from all states (including $j$) into $j$

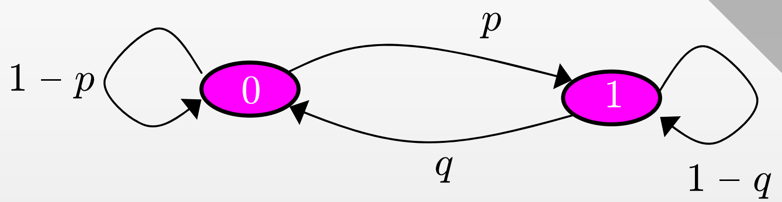
$$\pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij}$$

  - To prove this, notice that this is an expansion of $\pi_j$ from equation $\pi = \pi P$

- Consider the packet-loss chain (note: variables are different from last time to simplify formulas):



# Transition Rates 3




- For this example, we can write:

$$\pi_0 = \pi_0(1-p) + \pi_1 q$$

  - Or in other words: $$\frac{\pi_0}{\pi_1} = \frac{q}{p}$$
  - Since $\pi_0 + \pi_1 = 1$, we have

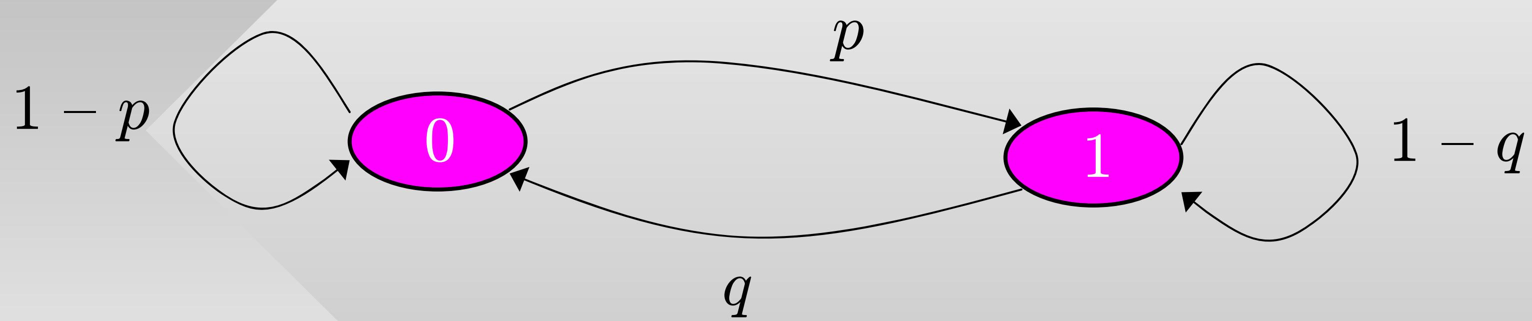
$$\pi_0 = \frac{q}{p+q}$$

- Another way to look at rates is to compute the total transition rate out of and into each state:

$$\pi_i \sum_{j \neq i} p_{ij} = \sum_{j \neq i} \pi_j p_{ji}$$

# Transition Rates 4

- For the same example:



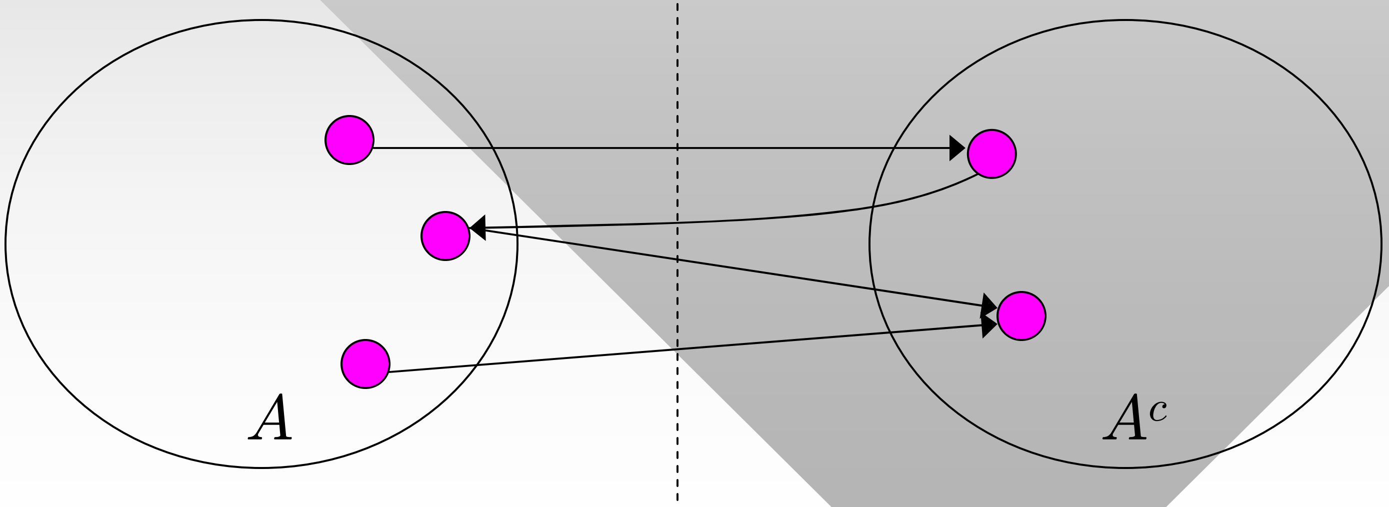

- The transition rate out of state 0 is $\pi_0 p$
  - The rate into the state is $\pi_1 q$
  - Equating the two, we again have:

$$\frac{\pi_0}{\pi_1} = \frac{q}{p}$$

# Transition Rates 5

- In general, transition rates across any boundary must be the same
  - For any set of states $A$ in a recurrent chain, we have:

$$\sum_{i \in A} \sum_{j \in A^c} \pi_i p_{ij} = \sum_{i \in A^c} \sum_{j \in A} \pi_i p_{ij}$$



# Transition Rates 6

- Example:
  - Assume a connected, undirected graph
  - The only thing known about the graph is that the degree of node $i$ is $d_i$
  - A random walk starts at some initial vertex and moves between the nodes uniformly choosing among the neighbors of each current node
- Is this a Markov chain? What is its matrix $P$?
  - Let $N(i)$ be set of all neighbors of node $i$

$$p_{ij} = \begin{cases} 1/d_i & j \in N(i) \\ 0 & otherwise \end{cases}$$

# Transition Rates 7

- Direct solution in Matlab to $\pi = \pi P$ is not possible since it requires the knowledge of $N(i)$ for each $i$
  - Instead, we use the observation that the probability to find the random walk in state $i$ is the combined rate of transitions from all states into $i$

$$\pi_i = \sum_{j=0}^{\infty} \pi_j p_{ji}$$

- Since these terms are non-zero only for neighbors of $i$, we have:

$$\pi_i = \sum_{j \in N(i)} \pi_j p_{ji} = \sum_{j \in N(i)} \frac{\pi_j}{d_j}$$

# Transition Rates 8

- Due to normalization by $d_j$, we can guess the shape of the stationary distribution:

$$\pi_i = \frac{d_i}{C}$$

  - where $C$ is some constant that we determine below (proving that $\pi$ is unique is beyond our scope)

- We next check if this guess is correct:

$$\pi_i = \sum_{j \in N(i)} \frac{\pi_j}{d_j} = \sum_{j \in N(i)} \frac{1}{C} = \frac{d_i}{C}$$

  - and then find out $C$:

$$\sum_k \pi_k = \frac{1}{C} \sum_k d_k = 1 \Rightarrow C = \sum_k d_k$$

# Wrap-up

- For computing $E[T]$, use these hints:
  - Pareto:

$$\int(1-(1-z^{1-\alpha})^k)dz = z\left(1 - {}_2F_1\left(\frac{1}{1-\alpha}, -k, \frac{2-\alpha}{1-\alpha}, z^{1-\alpha}\right)\right)$$

  - Exponential: $\frac{1-z^k}{1-z} = \sum_{i=0}^{k-1} z^i$

- Midterm next Thursday
  - Covers everything since the first lecture