Mathematics for Informatics 4a

José Figueroa-O'Farrill

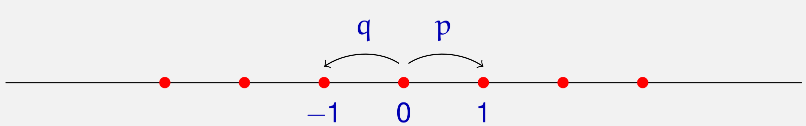
Lecture 18
23 March 2012

The story of the film so far...

- We have been studying **stochastic processes**; i.e., systems whose time evolution has an element of chance
- In particular, **Markov processes**, whose future only depends on the present and not on how we got there
- Particularly tractable examples are the **Markov chains**, which are discrete both in "time" and "space":
  - random walks
  - branching processes
- We can answer basic questions: **steady-state distributions**, **hitting times**, **extinction probabilities**,...
- One example of Markov chains are random walks on graphs: e.g., Google's **PageRank** is the steady-state distribution of a random walk on the world wide web
- In today's lecture we will look at random walks on simpler graphs and will focus on a different sort of questions

Random walk revisited

Let us consider again the random walk on the integers:



The jumps J_i are independent random variables with

$$\mathbb{P}(J_i = 1) = p \qquad \mathbb{P}(J_i = -1) = q = 1 - p$$

Starting at 0, $X_n = \sum_{i=1}^n J_i$ is the position after n steps. Define the **hitting times**

$$T_r = \begin{cases} \text{number of steps until we visit } r \text{ for the first time,} & r \neq 0 \\ \text{number of steps until we revisit } 0, & r = 0. \end{cases}$$

Question: How the T_r are distributed? i.e., $\mathbb{P}(T_r = n) = ?$

Hitting times for random walks I

- Let $r > 0$ and let T_r be the number of steps until we visit r for the first time, starting at 0.
- Let $T_{k,k+1}$ be the number of steps needed to reach $k+1$ having reached k. Then $T_{0,1} = T_1$ and the $T_{k,k+1}$ are i.i.d.
- $T_r = T_{0,1} + T_{1,2} + \cdots + T_{r-1,r}$ and by independence

$$\mathbb{E}(s^{T_r}) = \mathbb{E}(s^{T_1})^r$$

- Conditioning on the first jump,

$$\begin{aligned} \mathbb{E}(s^{T_1}) &= \mathbb{E}(s^{T_1} \mid J_1 = 1)\mathbb{P}(J_1 = 1) + \mathbb{E}(s^{T_1} \mid J_1 = -1)\mathbb{P}(J_1 = -1) \\ &= sp + s\mathbb{E}(s^{T_{-1,0} + T_{0,1}})q \\ &= sp + sq\mathbb{E}(s^{T_1})^2 \end{aligned}$$

which we solve for $\mathbb{E}(s^{T_1})$

Hitting times for random walks II

- Let $\mathbb{E}(s^{T_1}) = x$ and we must solve $x = sp + sqx^2$
- Assuming that $q \neq 0$, there are two solutions:

$$x = \frac{1 \pm \sqrt{1 - 4pqs^2}}{2sq}$$

but only one has a power series expansion around $s = 0$:

$$\mathbb{E}(s^{T_1}) = \frac{1 - \sqrt{1 - 4pqs^2}}{2sq}$$

- Hence for $r > 0$,

$$\mathbb{E}(s^{T_r}) = \left(\frac{1 - \sqrt{1 - 4pqs^2}}{2sq}\right)^r$$

and for $r < 0$ we simply replace $p \leftrightarrow q$

Hitting times for random walks III

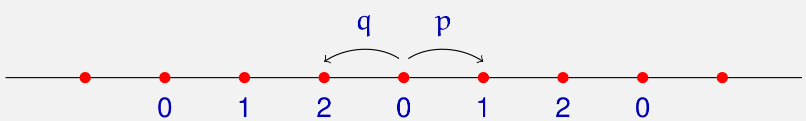
- How about $\mathbb{E}(s^{T_0})$?
- We condition on the first jump:

$$\begin{aligned}
\mathbb{E}(s^{T_0}) &= \mathbb{E}(s^{T_0} \mid J_1 = 1)\mathbb{P}(J_1 = 1) + \mathbb{E}(s^{T_0} \mid J_1 = -1)\mathbb{P}(J_1 = -1) \\
&= s\mathbb{E}(s^{T_{1,0}})p + s\mathbb{E}(s^{T_{-1,0}})q \\
&= sp\mathbb{E}(s^{T_{-1}}) + sq\mathbb{E}(s^{T_1}) \\
&= sp\left(\frac{1 - \sqrt{1 - 4pqs^2}}{2sp}\right) + sq\left(\frac{1 - \sqrt{1 - 4pqs^2}}{2sq}\right) \\
&= 1 - \sqrt{1 - 4pqs^2}
\end{aligned}$$

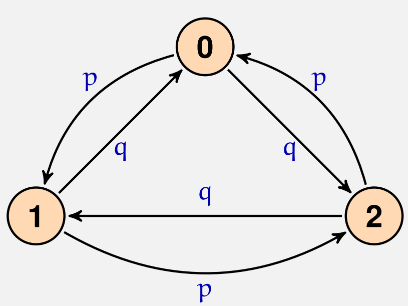
$$\therefore \mathbb{E}(T_0) = \frac{4pq}{\sqrt{1 - 4pq}} \longrightarrow \infty \qquad \text{if } p = q$$

Random walk on a triangle

Let us consider again the random walk on the integers, but this time we only keep track of the position *modulo* 3:



This is equivalent to the following 3-state Markov chain describing a random walk on a triangle:



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

$$\boldsymbol{\pi} = \left(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}\right)$$

Mean return times I

A typical quantity of interest is the **mean return time**: say we start at the vertex 0 of the triangle. Let T_0 denote the number of steps until we first revisit 0. *What is $\mathbb{E}(T_0)$?*
Let T_1 (resp. T_2) be the number of steps until we visit 0 starting from 1 (resp. 2). And let $\tau_i = \mathbb{E}(T_i)$ for $i = 0, 1, 2$.
We fill find τ_i by conditioning on the first move: $\circlearrowright$ with probability p and $\circlearrowleft$ with probability $q = 1 - p$. Therefore

$$\begin{aligned}
\tau_0 &= \mathbb{E}(T_0 \mid \circlearrowright)\mathbb{P}(\circlearrowright) + \mathbb{E}(T_0 \mid \circlearrowleft)\mathbb{P}(\circlearrowleft) \\
&= (1 + \mathbb{E}(T_1))p + (1 + \mathbb{E}(T_2))q \\
&= 1 + p\tau_1 + q\tau_2
\end{aligned}$$

Similarly,

$$\begin{aligned}
\tau_1 &= \mathbb{E}(T_1 \mid \circlearrowright)p + \mathbb{E}(T_1 \mid \circlearrowleft)q \\
&= (1 + \tau_2)p + q \\
&= 1 + p\tau_2
\end{aligned}
\qquad \text{and} \qquad
\begin{aligned}
\tau_2 &= \mathbb{E}(T_2 \mid \circlearrowright)p + \mathbb{E}(T_2 \mid \circlearrowleft)q \\
&= p + (1 + \tau_1)q \\
&= 1 + q\tau_1
\end{aligned}$$

Mean return times II

Solving the system

$$\begin{aligned}\tau_0 &= 1 + p\tau_1 + q\tau_2 \\ \tau_1 &= 1 + p\tau_2 \\ \tau_2 &= 1 + q\tau_1\end{aligned} \qquad \Longrightarrow \qquad (\tau_0, \tau_1, \tau_2) = \left(3, \frac{1+p}{1-pq}, \frac{1+q}{1-pq}\right)$$

Notice that $\tau_0 = 3 = \frac{1}{\pi_0}$, where (π_0, π_1, π_2) is the steady-state distribution of the Markov chain!
This is **not** a coincidence:

Theorem

The steady-state distribution (π_i) of a regular Markov chain is such that $\pi_i = \frac{1}{\tau_i}$, where τ_i is the mean return time to state i.

(We will not prove it; although it is not hard.)

Hitting times distribution I

It is not just the mean return time that can be calculated, but in fact the whole probability distribution of the T_i: namely, $\mathbb{E}(s^{T_i})$.
We do this again by conditioning on the first move:

$$\begin{aligned}\mathbb{E}(s^{T_0}) &= \mathbb{E}(s^{T_0} | \circlearrowleft)p + \mathbb{E}(s^{T_0} | \circlearrowright)q \\ &= \mathbb{E}(s^{1+T_1})p + \mathbb{E}(s^{1+T_2})q \\ &= sp\mathbb{E}(s^{T_1}) + sq\mathbb{E}(s^{T_2})\end{aligned}$$

and similarly

$$\begin{aligned}\mathbb{E}(s^{T_1}) &= \mathbb{E}(s^{1+T_2})p + sq \\ &= sp\mathbb{E}(s^{T_2}) + sq\end{aligned} \qquad \text{and} \qquad \begin{aligned}\mathbb{E}(s^{T_2}) &= p + \mathbb{E}(s^{1+T_1})q \\ &= sp + sq\mathbb{E}(s^{T_1})\end{aligned}$$

In terms of $\gamma_i = \mathbb{E}(s^{T_i})$, we have

$$\gamma_0 = s(p\gamma_1 + q\gamma_2) \qquad \gamma_1 = s(q + p\gamma_2) \qquad \gamma_2 = s(p + q\gamma_1)$$

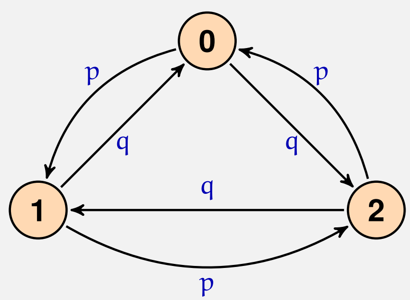
Hitting times distribution II

The unique solution is

$$\mathbb{E}(s^{T_0}) = \frac{s^2(2pq+p^3s+q^3s)}{1-pqs^2} = 2pqs^2 + \left(p^3 + q^3\right)s^3 + 2p^2q^2s^4 + \cdots$$

$$\mathbb{E}(s^{T_1}) = \frac{s(q+p^2s)}{1-pqs^2} = qs + p^2s^2 + pq^2s^3 + \cdots$$

$$\mathbb{E}(s^{T_2}) = \frac{s(p+q^2s)}{1-pqs^2} = ps + q^2s^2 + p^2qs^3 + \cdots$$



Independent random walks I

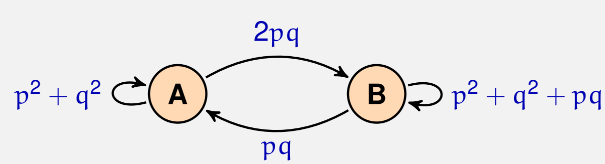
Now suppose that we have two independent random walks on the triangle. Assume that both particles start in the same vertex and let T denote the number of steps until they again share a vertex. *What is* $\mathbb{E}(T)$*?*
We can turn this into a Markov chain with two states:

(A) the two particles share the same vertex
(B) the two particles are in different vertices

with transitions

initial	$\circlearrowleft\circlearrowleft$	$\circlearrowright\circlearrowright$	$\circlearrowleft\circlearrowright$	$\circlearrowright\circlearrowleft$
A	A	B	B	A
B	B	B	A	B

and probabilities



Independent random walks II

In terms of $\theta = pq$, the transition matrix is

$$\mathbf{P} = \begin{pmatrix} 1-2\theta & 2\theta \\ \theta & 1-\theta \end{pmatrix} \implies \boldsymbol{\pi} = \left(\tfrac{1}{3}, \tfrac{2}{3}\right) \implies \mathbb{E}(T) = \frac{1}{\pi_A} = 3$$

We can also derive this directly by conditioning. Let U denote the number of steps until the particles first share a vertex, starting from different vertices. Then

$$\begin{aligned} \mathbb{E}(T) &= \mathbb{E}(T \mid \circlearrowleft\circlearrowleft)p^2 + \mathbb{E}(T \mid \circlearrowleft\circlearrowright)pq + \mathbb{E}(T \mid \circlearrowright\circlearrowleft)pq + \mathbb{E}(T \mid \circlearrowright\circlearrowright)q^2 \\ &= p^2 + 2(1+\mathbb{E}(U))pq + q^2 = 1 + 2pq\mathbb{E}(U) \end{aligned}$$

$$\begin{aligned} \mathbb{E}(U) &= \mathbb{E}(U \mid \circlearrowleft\circlearrowleft)p^2 + \mathbb{E}(U \mid \circlearrowleft\circlearrowright)pq + \mathbb{E}(U \mid \circlearrowright\circlearrowleft)pq + \mathbb{E}(U \mid \circlearrowright\circlearrowright)q^2 \\ &= (1+\mathbb{E}(U))p^2 + (1+\mathbb{E}(U))pq + pq + (1+\mathbb{E}(U))q^2 \\ &= 1 + (p^2+q^2+pq)\mathbb{E}(U) = 1 + (1-pq)\mathbb{E}(U) \end{aligned}$$

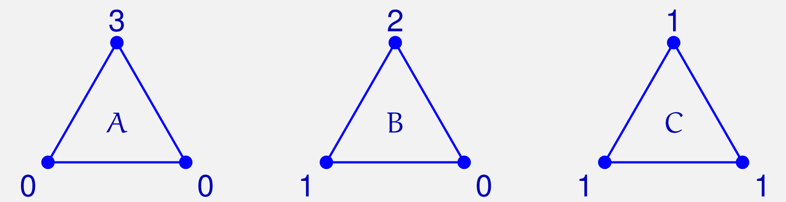
$$\implies \mathbb{E}(U) = \frac{1}{pq} \implies \mathbb{E}(T) = 3$$

Three independent random walks I

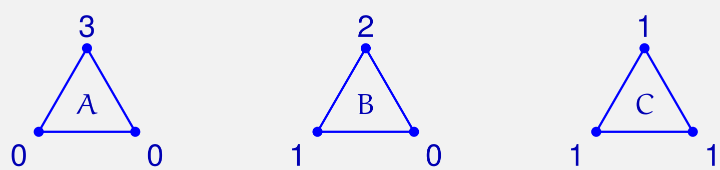
Now consider *three* particles moving in a triangle, but let us assume for simplicity that the random walk is symmetric, so that $p = q = \frac{1}{2}$. Assuming that the particles start at the same vertex, let T denote the number of steps until they once again share a vertex. *What is* $\mathbb{E}(T)$*?*

We can turn this into a Markov chain with 3 states:

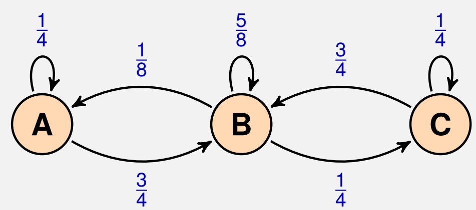
(A) all particles share a vertex
(B) precisely two particles share a vertex
(C) all particles in different vertices



Three independent random walks II



The resulting Markov chain is



$$\mathbf{P} = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} & 0 \\ \frac{1}{8} & \frac{5}{8} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{pmatrix}$$

The steady-state has distribution $\boldsymbol{\pi} = \left(\frac{1}{9}, \frac{2}{3}, \frac{2}{9}\right)$.
Therefore $\mathbb{E}(T) = 9$.

Three independent random walks III

We can also solve this directly by conditioning. Let U (resp. V) denote the number of steps needed to reach a configuration of type A starting from a configuration of type B (resp. C).The conditioning on the first step,

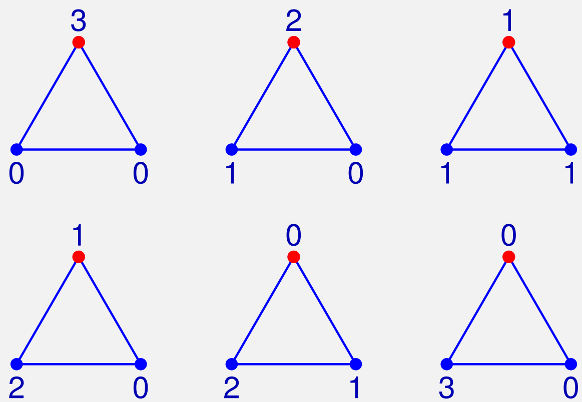
$$\begin{aligned} \mathbb{E}(T) &= \mathbb{E}(T \mid A \to A)\mathbb{P}(A \to A) + \mathbb{E}(T \mid A \to B)\mathbb{P}(A \to B) \\ &\quad + \mathbb{E}(T \mid A \to C)\mathbb{P}(A \to C) \\ &= \tfrac{1}{4} + (1+\mathbb{E}(U))\tfrac{3}{4} = 1 + \tfrac{3}{4}\mathbb{E}(U) \end{aligned}$$

and similarly

$$\begin{aligned} \mathbb{E}(U) &= \tfrac{1}{8} + (1+\mathbb{E}(U))\tfrac{5}{8} + (1+\mathbb{E}(V))\tfrac{1}{4} \\ &= 1 + \tfrac{5}{8}\mathbb{E}(U) + \tfrac{1}{4}\mathbb{E}(V) \\ \mathbb{E}(V) &= (1+\mathbb{E}(U))\tfrac{3}{4} + (1+\mathbb{E}(V))\tfrac{1}{4} \\ &= 1 + \tfrac{3}{4}\mathbb{E}(U) + \tfrac{1}{4}\mathbb{E}(V) \end{aligned} \implies \begin{aligned} \mathbb{E}(T) &= 9 \\ \mathbb{E}(U) &= \tfrac{32}{3} \\ \mathbb{E}(V) &= 12 \end{aligned}$$

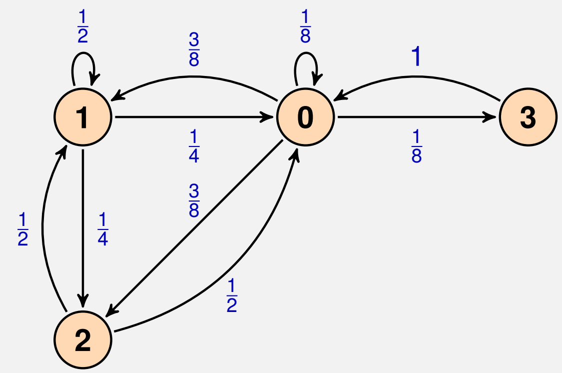
Three independent random walks III

Let Z be the number of steps until the particles meet again at the *starting* vertex. *What is* $\mathbb{E}(Z)$*?*
We now have four types of configurations, labelled by how many of the particles are at the starting vertex: 0, 1, 2 or 3.



Three independent random walks IV

The corresponding 4-state Markov chain is:



$$\mathbf{P} = \begin{pmatrix} \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\pi = \left(\tfrac{8}{27}, \tfrac{4}{9}, \tfrac{2}{9}, \tfrac{1}{27}\right)$$
$$\implies \mathbb{E}(Z) = 27$$

Summary

- We have looked at random walks on the integers and answered questions about **return times** and **hitting times**
- We did the same for finite graphs, e.g., triangle
- We considered independent random walks on a graph and computed a variety of **mean return times**
- The main techniques are:
  1. Map the problem to a finite-state Markov chain, compute the steady-state distribution and the mean return time
  2. Directly by conditioning on the first move to obtain recursion relations and/or linear equations
- Pick your favourite graphs and work out a couple of examples: the last assignment in this course will ask you to do this for K_4