CS5314
Randomized Algorithms

Lecture 23: Markov Chains
(Gambler's Ruin, Random Walks)

Objectives

- Discuss Gambler's ruin
  - A study of the game between two gamblers until one is ruined (no money left)
- Introduce stationary distribution
  - and a sufficient condition when a Markov chain has stationary distribution
- Analyze random walks on graph

The Game

- Consider two players, one have L_1 dollars, and the other has L_2 dollars. Player 1 will continue to throw a fair coin, such that

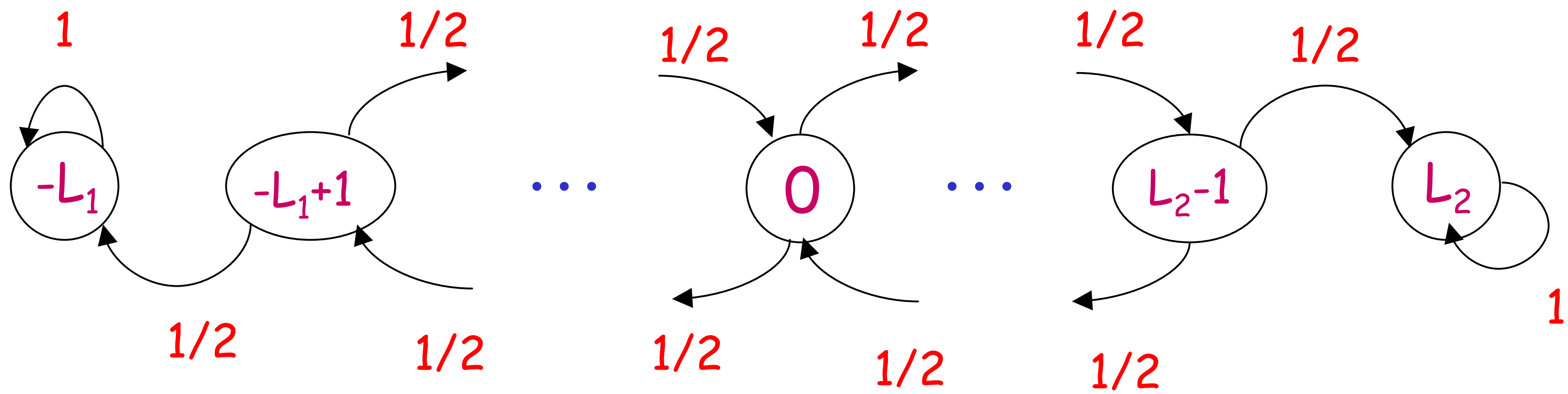
  ** if head comes up, he wins one dollar

  ** if tail comes up, he loses one dollar

- Suppose the game is played until one player goes bankrupt. What is the chances that Player 1 survives?

The Markov Chain Model

- The previous game can be modeled by the following Markov chain:



The Markov Chain Model

Initially, the chain is at state 0

Let $P_j^{(t)}$ denote the probability that after t steps, the chain is at state j

Also, let q be the probability that the game ends with Player 1 winning L_2 dollars

We can see that

$\lim_{t\to\infty} P_j^{(t)} = 0$, for $j \neq -L_1, L_2$

$\lim_{t\to\infty} P_j^{(t)} = 1-q$, for $j = -L_1$

$\lim_{t\to\infty} P_j^{(t)} = q$, for $j = L_2$

The Analysis

- Now, let W_t to be the money Player 1 has won after t steps
- By linearity of expectation,

$$E[W_t] = 0$$

- On the other hand,

$$E[W_t] = \sum j \, P_j^{(t)}$$

The Analysis (2)

- By taking limits, we have

$$0 = \lim_{t\to\infty} E[W_t]$$

$$= \lim_{t\to\infty} \sum j\, P_j^{(t)}$$

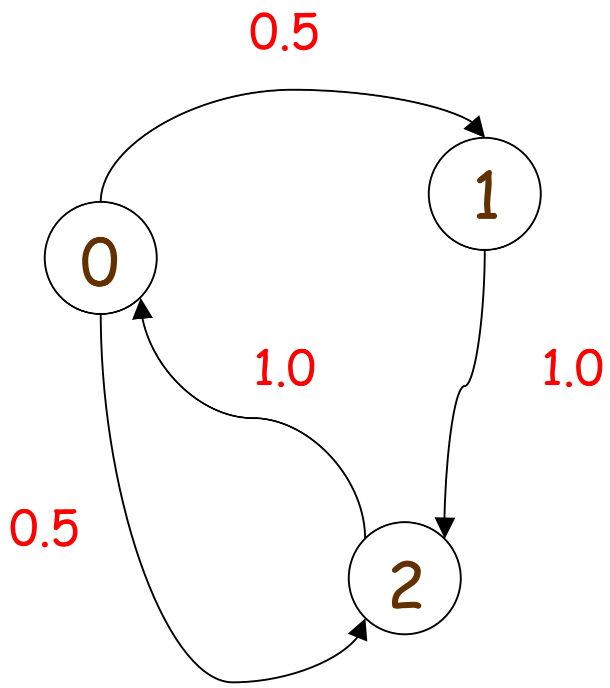
$$= (-L_1)\,(1-q) + 0 + 0 + \ldots + 0 + (L_2)\, q$$

- Re-arranging terms, we obtain

$$q = L_1 / (L_1+L_2)$$

Stationary Distribution

Consider the following Markov chain:



Let $p_j(t)$ = probability that the chain is at state j at time t, and let

$$\langle p(t) \rangle = (p_0(t), p_1(t), p_2(t))$$

Suppose that

$$\langle p(t) \rangle = (0.4, 0.2, 0.4)$$

Question: In this case, what will be $\langle p(t+1) \rangle$?

Stationary Distribution (2)

- After some calculation, we get

$$\langle p(t+1) \rangle = (0.4, 0.2, 0.4)$$

  which is the same as $\langle p(t) \rangle$!!!

- We can see that in the previous example, the Markov chain has entered an "equilibrium" condition at time t, where

$$\langle p(n) \rangle \text{ remains } (0.4, 0.2, 0.4) \text{ for all } n \geq t$$

  → this probability distribution is called a stationary distribution

Stationary Distribution (3)

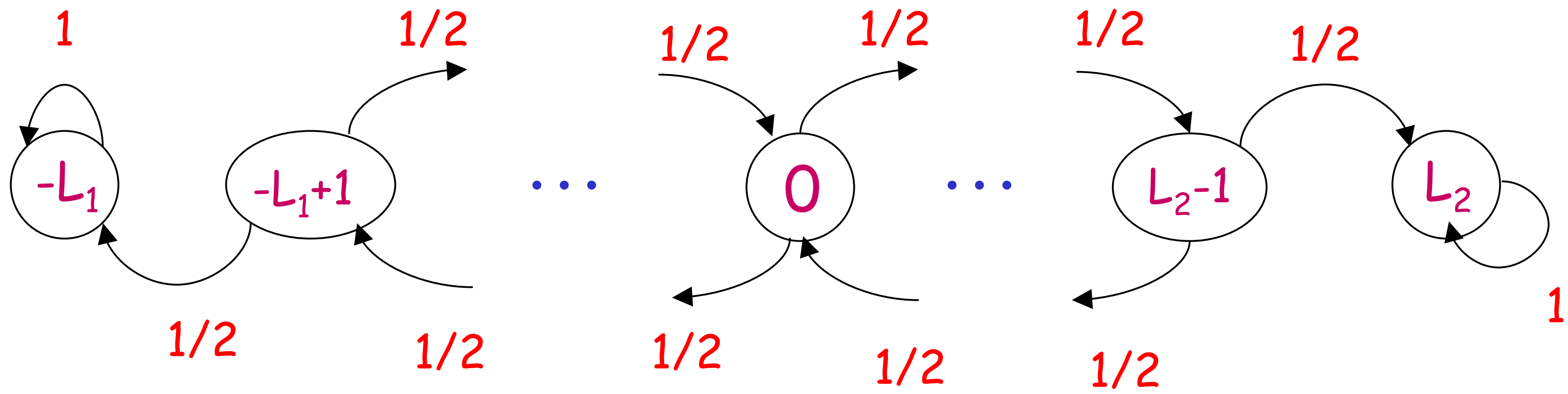
- Precisely, let P be the transition matrix of a Markov chain

Definition: If $\langle p(t+1) \rangle = \langle p(t) \rangle P = \langle p(t) \rangle$, then $\langle p(t) \rangle$ is a **stationary distribution** of the Markov chain

Question: How many stationary distribution can a Markov chain have? Can it be more than one? Can it be none?

Stationary Distribution (4)

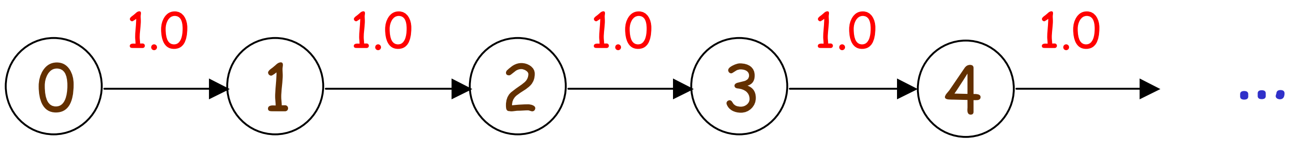
Ans. Can be more than one. For example,



In this case, both (1,0,0,…,0) and (0,0,…,0,1) are stationary distribution

Stationary Distribution (5)

Ans. Can be none. For example,



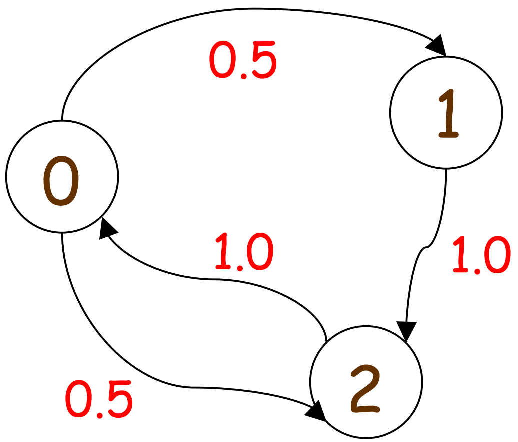
Here, no stationary distribution exists

Question: Are there some conditions that can tell if a Markov chain has a unique stationary distribution?

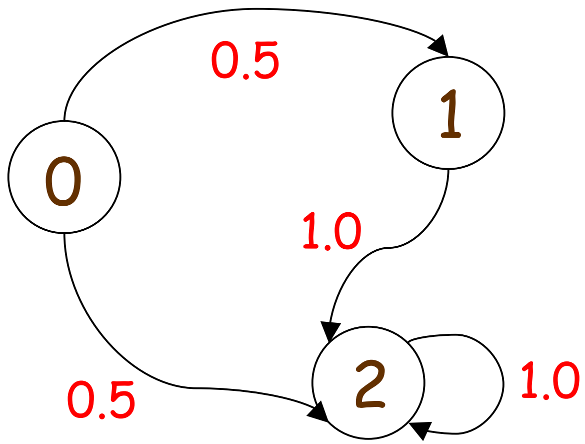
Special Markov Chains

Definition: A Markov chain is irreducible if its directed representation is a strongly connected component. That is, every state j can reach any state k

For example:



irreducible



not irreducible

Special Markov Chains (2)

Definition: A Markov chain is periodic if there exists some state j and some integer $d > 1$ such that:

$$Pr(X_{t+s} = j \mid X_t = j) = 0$$

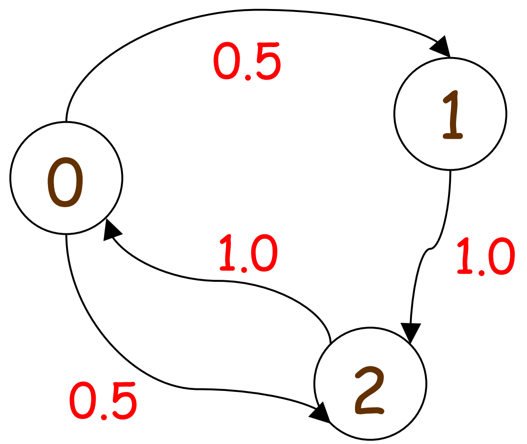
unless s is divisible by d

In other words, once we start at state j, we can only return to j after a multiple of d steps

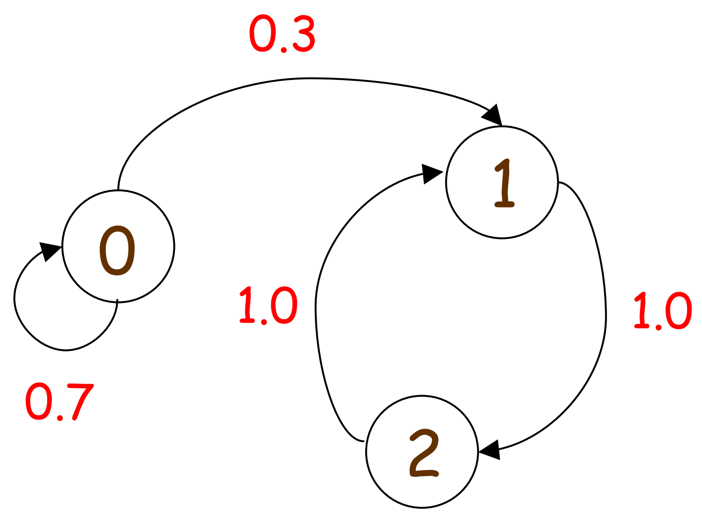
- If a Markov chain is not periodic, then it is called aperiodic

Special Markov Chains (3)

For example,



aperiodic



periodic

Sufficient Conditions

A simplified version of an important result of Ch. 7 is stated as follows:

Theorem: Suppose a Markov chain is finite with states 0,1,…,n. If it is irreducible and aperiodic, then

1. The chain has a unique stationary distribution $\langle \pi \rangle = (\pi_0, \pi_1, \ldots, \pi_n)$;
2. $\pi_k = 1/h_{k,k}$, where

   $h_{k,k}$ = expected # of steps to return to state k, when starting at state k

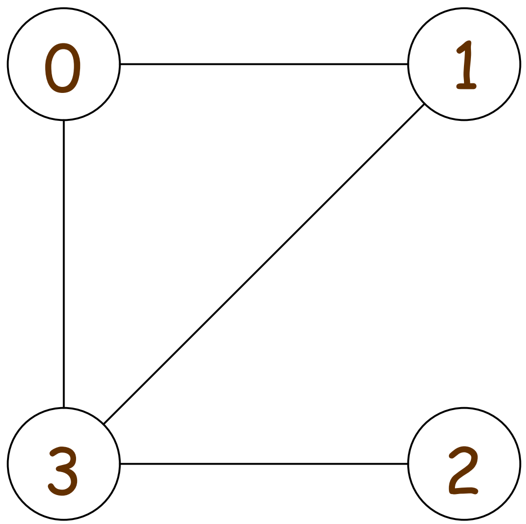
Random Walk

- Let G be a finite, undirected, and connected graph
- Let $D(G)$ be a directed graph formed by replacing each undirected edge $\{u,v\}$ of G by two directed edges (u,v) and (v,u)

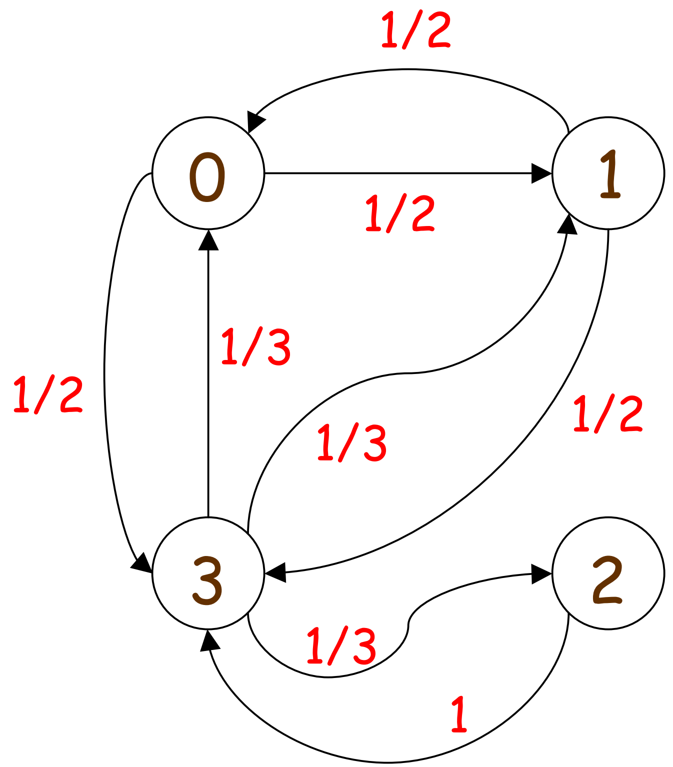
Definition: *A* random walk on G is a Markov chain whose directed representation is $D(G)$, and for each edge (u,v), the transition probability is $1/\deg(u)$

Random Walk (2)

For example,



G



Representation for random walk on *G*

Random Walk (3)

- Since G is connected, it is easy to check that $D(G)$ is strongly connected ... [why?]
  ➔ Random walk on G is irreducible

- The lemma below gives a simple criterion for a random walk on G to be aperiodic

Lemma: A random walk on G is aperiodic if and only if G is not bipartite

How to prove??

(→) If G is bipartite, all cycles have even number of edges. Then, if we start at any vertex in $D(G)$, it can only come back to itself in even steps

So, $d = 2$, and the chain is periodic

(←) If G is not bipartite, there exists an odd-length cycle C. Let w be any vertex in C. Then, for any vertex u in $D(G)$, it can come back to itself in 2 steps (via (u,v) then (v,u) for some v), and also in odd steps (via a path from u to w, then C, then a path from w to u).

So, $d = 1$, and the chain is aperiodic

Random Walk (4)

Then, we have the following theorem:

Lemma: If G =(V,E) is not bipartite, the random walk on G has a unique stationary distribution $\langle\pi\rangle$. Moreover, for vertex v, the corresponding probability in $\langle\pi\rangle$ is:

$$\pi_v = deg(v) / (2|E|)$$

Proof: The first statement follows immediately since the random walk is finite, irreducible, and aperiodic …

Proof (cont)

For the second statement, we first see that

$$\sum_{v \in G} \pi_v = \sum_{v \in G} \deg(v) / (2|E|) = 1$$

so that $\langle \pi \rangle$ is a valid probability distribution

Next, let P be the transition matrix of the random walk. Let $N(v)$ be the set of the neighbors of v (so $|N(v)| = \deg(v)$)

Then, the v^{th} entry of $\langle \pi \rangle P$ is:

Proof (cont)

$\sum_{u \in G} \pi_u P_{u,v}$

$= \sum_{u \in N(v)} \pi_u P_{u,v} + \sum_{u \notin N(v)} \pi_u 0 = \sum_{u \in N(v)} \pi_u P_{u,v}$

$= \sum_{u \in N(v)} (deg(u)/(2|E|)) (1/deg(u))$

$= deg(v) / (2|E|) = \pi_v$

➔ $\langle \pi \rangle = \langle \pi \rangle P$, so $\langle \pi \rangle$ is a (unique) stationary distribution of the random walk

Random Walk (5)

From now on, we assume G is not bipartite.

Recall that $h_{v,u}$ = expected number of steps to reach u, starting from v

Then we have the following corollary:

Corollary: In the random walk on G, for any vertex v,

$$h_{v,v} = 1/\pi_v = 2|E| / \deg(v)$$

Next, we give a lemma for bounding $h_{u,v}$

Random Walk (6)

Lemma: For any edge $(u,v) \in E$,

$$h_{u,v} < 2|E|$$

Proof: Let $N(v)$ be the neighbor-set of v.
Then, $h_{v,v}$ can be expressed by:

$$(1/\deg(v)) \sum_{u \in N(v)} (1 + h_{u,v})$$

Then, by previous corollary, we see that

$$2|E| = \sum_{u \in N(v)} (1 + h_{u,v})$$

➔ Lemma thus follows

Cover Time

- For a graph G, we denote cover(v) to be the expected number of steps to visit all nodes in G by a random walk, starting at v

Definition: The cover time of G is defined as $\max_{v \in G}$ { cover(v) }

- Consider any spanning tree T of G
  - Let ρ = an Eulerian tour that traverses each edge of T once in each direction

Cover Time (2)

- Let v_0, v_1, …, $v_{2|V|-3}$, $v_{2|V|-2}$ be the sequence of vertices in ρ (Note: $v_{2|V|-2} = v_0$)
- Now, based on ρ, consider a tour on G that starts at v_0, then by random walk reaches v_1, then by random walk reaches v_2, and so on, and it finally reaches v_0
- Because in this tour, we cannot stop even if we have visited every vertices of G ➔the expected time for this tour must be at least cover(v_0)

Cover Time (3)

- In fact, we can start from any vertex of ρ and obtain similar arguments
  ➔ The expected time for this tour must be at least the cover time of G
- For the expected time for the tour, it is:

$$\sum_{k=0 \text{ to } 2|V|-3} h_{v_k,v_{k+1}} < (2|V|-2)\,2|E| < 4|V||E|$$

This gives the following theorem:

Theorem: The cover time of $G \le 4|V||E|$

ST Connectivity

- Given an undirected graph $G = (V,E)$ and two vertices s and t, we want to know if s and t are connected

The following succeeds with prob $\geq 1/2$

Step 1: Start a random walk on G from s

Step 2: If the walk reaches t in $4|V|^3$ steps, return true. Otherwise, return false

Proof: If s and t are connected, expected time to reach t from s is at most $2|V|^3$. Then apply Markov Inequality

ST Connectivity (2)

- This algorithm is very space-efficient!
- Apart for the input G (which is read-only here) and assume that we can choose a random neighbor to visit at each step, the algorithm just needs $O(\log |V|)$ bits to store the current position!!!

Remark 1: Recently, Reingold (2006) shows that even without the random bits, ST connectivity in undirected graph can be done in $O(\log |V|)$ bits

Remark 2: If graph is directed, it becomes the hardest problem solvable by an NTM using $O(\log |V|)$ bits