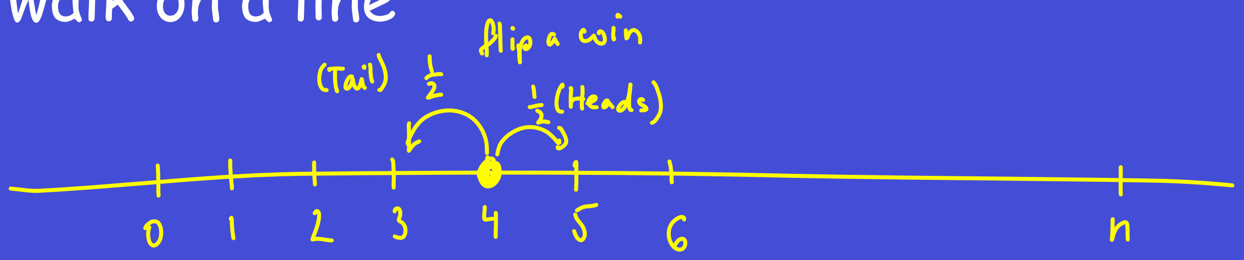
Introduction to Markov chains

Examples of Markov chains:

- Random walk on a line

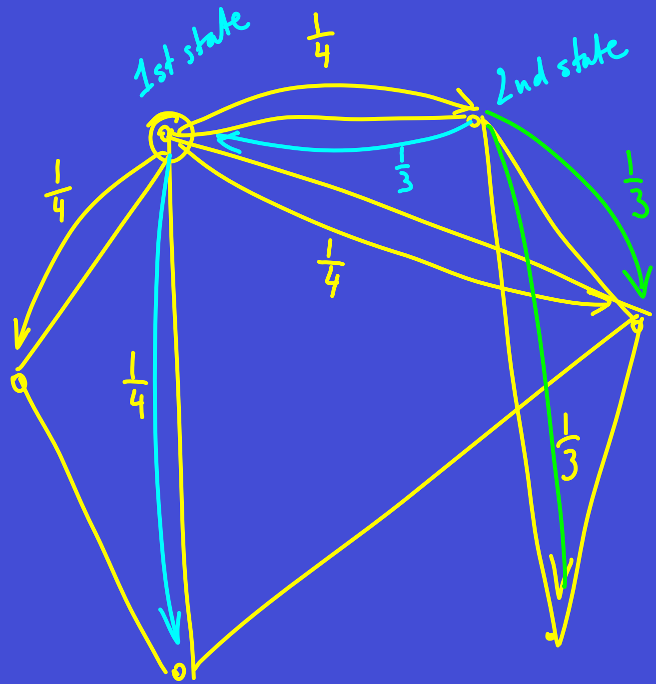


H T T H T H T T T

Introduction to Markov chains

Examples of Markov chains:

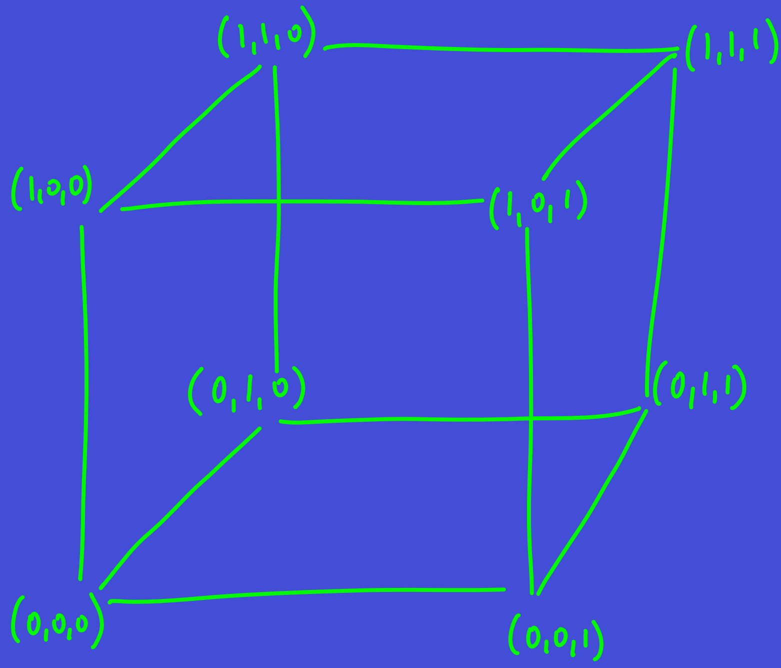
- Random walk on a graph



Introduction to Markov chains

Examples of Markov chains:

- Random walk on a hypercube



states: all bin. vectors of length n

move:

- choose a random position from 1 to n → let be i
- flip i-th bit

Symmetric: prob. of moving from state u to state v = prob. of moving from state v to state u

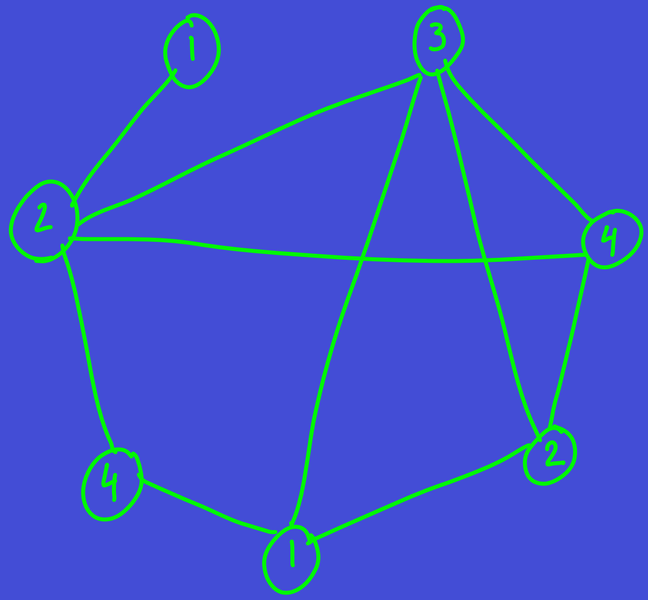
Introduction to Markov chains

Examples of Markov chains:

- Markov chain on graph colorings

states:

valid colorings



colors is: 4

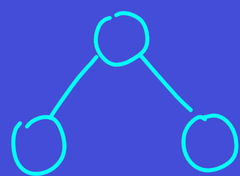
MC step:

- randomly choose a vertex v
- randomly choose a color c
- recolor the vertex v with color c (if valid)

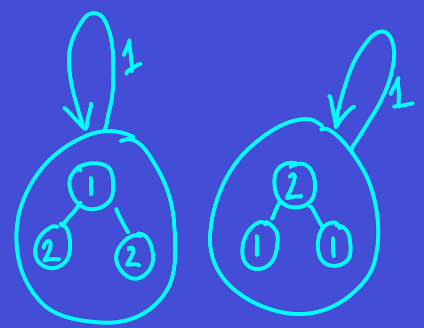
Introduction to Markov chains

Examples of Markov chains:

- Markov chain on graph colorings



# colors	# states
1	0
2	2
3	

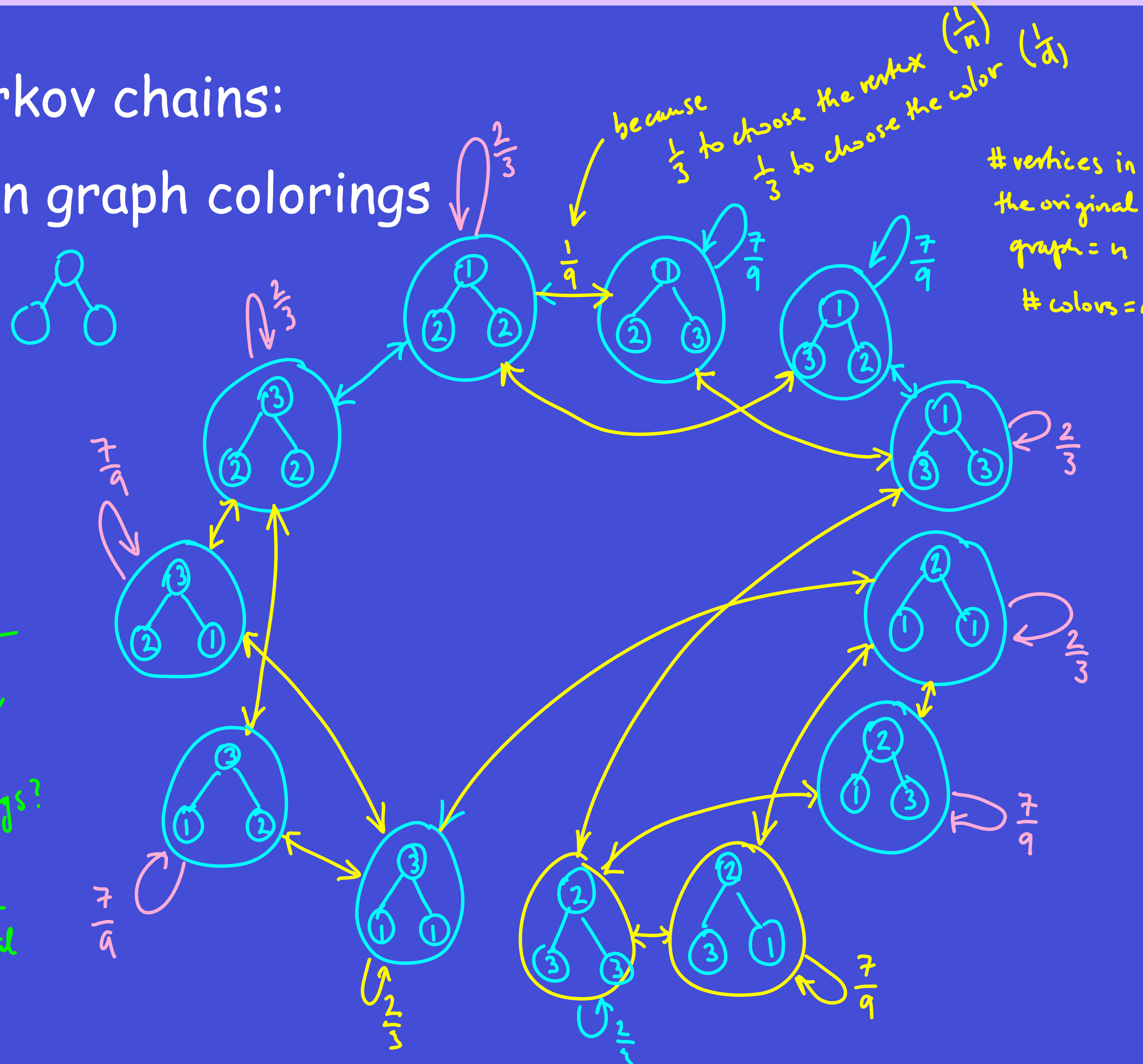


Introduction to Markov chains

Examples of Markov chains:

- Markov chain on graph colorings

#colors: 3



all yellow arcs probability $\frac{1}{9}$

if input graph has n vertices, and we have d colors, how many possible colorings?

up to d^n colorings

exponential in n

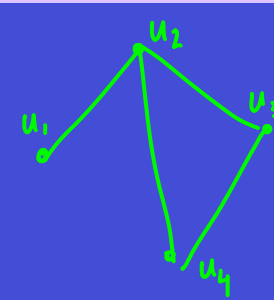
Introduction to Markov chains

Examples of Markov chains:

- Markov chain on matchings of a graph

Definition of Markov chains

- State space Ω ↗ the set of states



random walk on this graph

- Transition matrix P, of size $|\Omega| \mathrm{x} |\Omega|$
  - P(x,y) is the probability of getting from state x to state y
  - $\sum_{y2\Omega} P(x,y) =$ 1 a prob. distribution

    (P is a stochastic matrix)

$\Omega = \{u_1, u_2, u_3, u_4\}$

	u_1	u_2	u_3	u_4
u_1	0	1	0	0
u_2	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{1}{3}$
u_3	0	$\frac{1}{2}$	0	$\frac{1}{2}$
u_4	0	$\frac{1}{2}$	$\frac{1}{2}$	0

if start in state u_1, then prob. of being in state u_1 is 1,
u_j is 0, $j \in \{2,3,4\}$

initial distribution in this case is $(1,0,0,0)$

after 1 step of the MC $(0,1,0,0)$

2nd $(\frac{1}{3}, 0, \frac{1}{3}, \frac{1}{3})$

3rd $(0, \frac{1}{3}+\frac{1}{6}+\frac{1}{6}, \frac{1}{6}, \frac{1}{6})$

$= (0, \frac{2}{3}, \frac{1}{6}, \frac{1}{6})$

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

starting in distribution σ on Ω (given by a vector)
the next distrib. is $\sigma \cdot P$

This course: state space is finite

Definition of Markov chains

- State space Ω
- Transition matrix P, of size $|\Omega| \times |\Omega|$
  - P(x,y) is the probability of getting from state x to state y
  - $\sum_{y2\Omega} P(x,y) = 1$ a prob. distribution

    (P is a stochastic matrix)

random walk on this graph

$\Omega = \{u_1, u_2, u_3, u_4\}$

	u_1	u_2	u_3	u_4
u_1	0	1	0	0
u_2	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{1}{3}$
u_3	0	$\frac{1}{2}$	0	$\frac{1}{2}$
u_4	0	$\frac{1}{2}$	$\frac{1}{2}$	0

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

starting in distribution σ on Ω (given by a vector)
the next distrib. is $\sigma \cdot P$

1st step: $\sigma \cdot P = \sigma \cdot P$

2nd step: $(\sigma \cdot P) \cdot P = \sigma \cdot P^2$

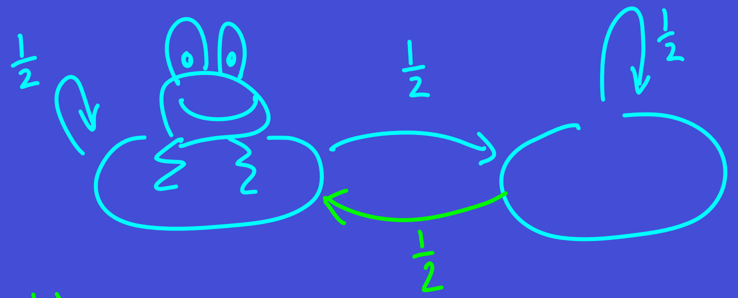
3rd step: $((\sigma \cdot P) \cdot P) \cdot P = \sigma \cdot P^3$

t-th step: $\sigma \cdot P^t$

This course: state space is finite

Definition of Markov chains

Example: frog on lily pads



$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$\sigma = (1,0)$ starting on the 1st lily pad

2nd $\sigma = \left(\frac{1}{2}, \frac{1}{2}\right)$

Definition of Markov chains

Stationary distribution ¼ (on states in Ω):

$$\pi \cdot P = \pi$$

$$\sum_{x \in \Omega} \pi_x = 1 \qquad \text{(a distribution)}$$

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

$$\pi_2 = \frac{p}{p+q}$$

normalizing constant Z (the partition function)

will often write

$\pi_1 \propto q$

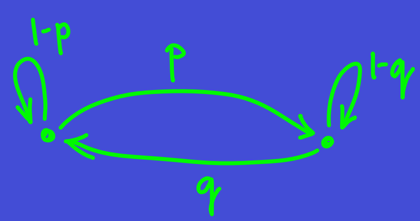
$\pi_2 \propto p$

proportional to

Example: frog with unfair coins

$p, q \in (0,1)$

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



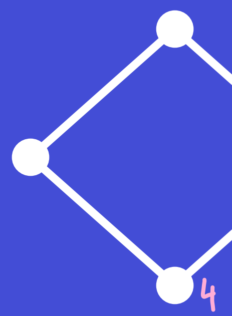
we are looking for $\pi = (\pi_1, \pi_2)$ s.t. $\pi \cdot P = \pi$

$$(\pi_1, \pi_2) \cdot P = (\underbrace{(1-p)\pi_1 + q\pi_2}_{= \pi_1}, \underbrace{p \cdot \pi_1 + (1-q) \cdot \pi_2}_{= \pi_2})$$

Definition of Markov chains

Other examples:

trouble: periodic (will not converge to stationary)



$$P = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

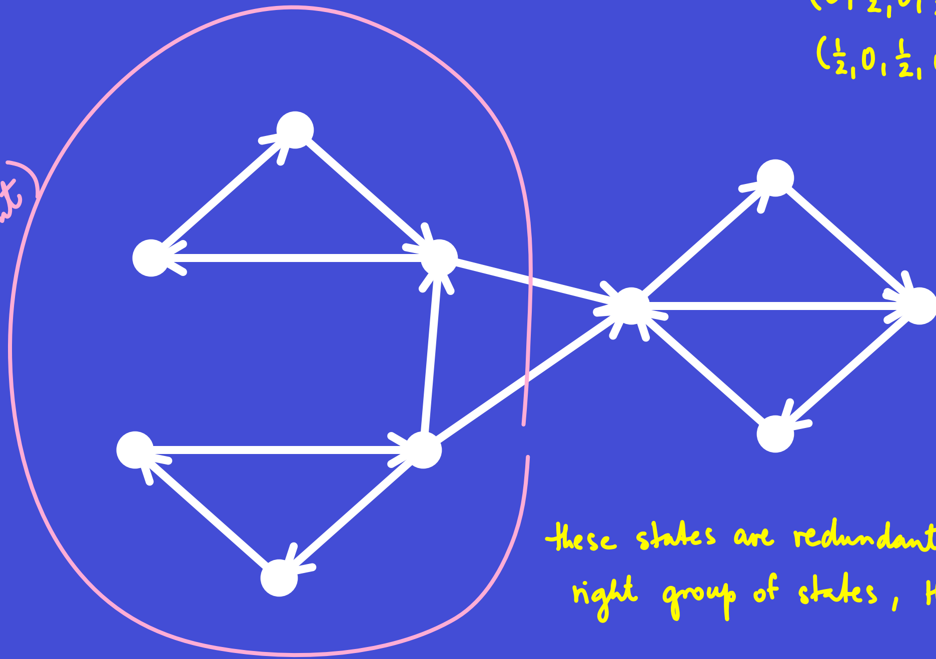
is $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ stationary? YES

if start in $(1,0,0,0)$, then

$(0, \frac{1}{2}, 0, \frac{1}{2})$

$(\frac{1}{2}, 0, \frac{1}{2}, 0)$

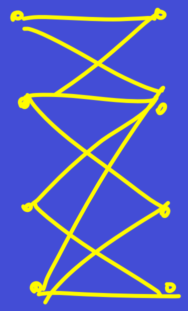
trouble: reducible (some states are redundant)



st. distr. is $(0,0,0,0,0,0,?,?,?,?)$ — left states

these states are redundant; once we get to the right group of states, then never come back

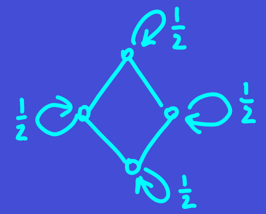
Definition of Markov chains



- aperiodic

$\forall x,y \in \Omega : \quad \gcd\{t \mid P^t(x,y) > 0\} = 1$

quick fix:



all original prob. scaled by ½

LAZY MC (LAZY variant of the original MC)

- irreducible

there exists a sequence of transitions from every state x to every state y

$\forall x,y \in \Omega : \quad \exists t \geq 0 : \quad P^t(x,y) > 0$

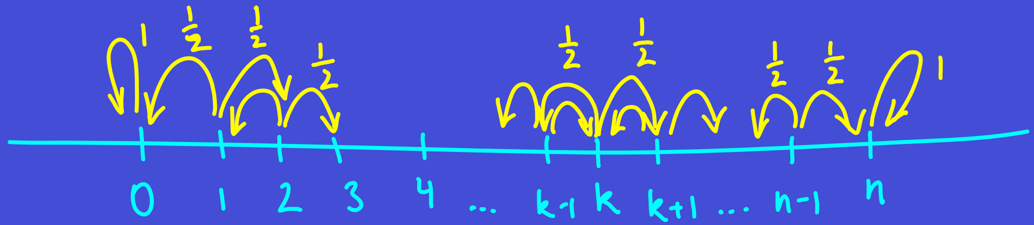
- ergodic: a MC that is both aperiodic and irreducible

Thm (3.6 in Jerrum): *An ergodic MC has a unique stationary distribution. Moreover, it converges to it as time goes to 1.*

Gambler's Ruin

A gambler has k dollars initially and bets on (fair) coin flips.
If she guesses the coin flip: gets 1 dollar.
If not, loses 1 dollar.
Stops either when 0 or n dollars.

Is it a Markov chain ?



How many bets until she stops playing ?
What are the probabilities of 0 vs n dollars ?

Gambler's Ruin

Stationary distribution ?

stat. distr: $\frac{1}{2}$ 0 0 0 . . . 0 $\frac{1}{2}$

0 1 2 3 ... n

not unique!

$(p, 0, 0, 0, \dots, 0, 1-p)$ is stationary for any $p \in [0,1]$

Gambler's Ruin

Starting with k dollars, what is the probability of reaching n ? $\frac{k}{n}$

p_ℓ - the probability of reaching n (not 0) when starting with ℓ dollars

$p_n = 1$

$p_0 = 0$

$p_\ell = \frac{1}{2} \cdot p_{\ell+1} + \frac{1}{2} \cdot p_{\ell-1}$ for $1 \le \ell \le n-1$

(first term: guessed right; second term: not right)

solve the system of lin. equations

$p_\ell = \frac{\ell}{n} \quad \forall \ell \in \{0, \dots, n\}$

verify it:

$p_0 = \frac{0}{n} = 0$ ✓

$p_n = \frac{n}{n} = 1$ ✓

$p_\ell = \frac{\ell}{n} = \frac{1}{2} \cdot \frac{\ell+1}{n} + \frac{1}{2} \frac{\ell-1}{n}$

$= \frac{1}{2} \cdot p_{\ell+1} + \frac{1}{2} \cdot p_{\ell-1}$

What is the probability of reaching 0 ? $\frac{n-k}{n}$

Gambler's Ruin

How many bets until 0 or n ?

f_k – expected # steps to get to 0 or n , starting from k

$f_0 = 0$

$f_n = 0$

$f_\ell = 1 + \frac{1}{2} \cdot f_{\ell+1} + \frac{1}{2} \cdot f_{\ell-1}$ for $\ell \in \{1, 2, \dots, n-1\}$

$f_\ell = \ell\,(n-\ell)$

Coupon Collecting

- n coupons $1, 2, \ldots, n$
- each time get a coupon at random, chosen from 1,…,n (i.e., likely will get some coupons multiple times)
- how long to get all n coupons ?

notice: stat. dist.: $(0, 0, 0, \ldots, 0, 1)$

Is this a Markov chain ?

Option 1:

$\Omega = \mathcal{P}(\{1, 2, \ldots, n\})$ powerset

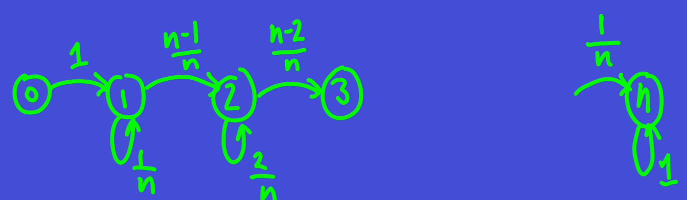
$$P(S_1, S_2) = \begin{cases} \frac{1}{n} & \text{if } S_2 = S_1 \cup \{a\},\ a \notin S_1 \\ \frac{|S_1|}{n} & \text{if } S_1 = S_2 \\ 0 & \text{o/w} \end{cases}$$

Option 2:

$\Omega = \{0, 1, \ldots, n\}$

(how many coupons have so far)

$$P(i, j) = \begin{cases} \frac{n-i}{n} & \text{if } j = i+1 \\ \frac{i}{n} & \text{if } i = j \quad \text{self-loop} \\ 0 & \text{o/w} \end{cases}$$



Coupon Collecting

How long to get all n coupons ?

τ = # steps until collect all n coupons (random variable for) WANT: $E[\tau]$ = $E[\tau_n]$

τ_k = random variable for # steps (of the MC) until collect k distinct coupons

$\tau_n = \tau_1 - \tau_0 + \tau_2 - \tau_1 + \tau_3 - \tau_2 + \ldots + \tau_n - \tau_{n-1}$ notice $\tau_0 = 0$ (no steps to get 0 coupons)

$\tau_{k+1} - \tau_k$ = # steps needed after having k coupons until collecting k+1 coupons (random var.)

$E[\tau_n] = E[\tau_1 - \tau_0] + E[\tau_2 - \tau_1] + \ldots + E[\tau_n - \tau_{n-1}]$ (linearity of expectation)

$E[\tau_{k+1} - \tau_k] = ?$ $\quad \frac{1}{\frac{n-k}{n}} = \frac{n}{n-k}$ (intuitively, expect $\frac{1}{p}$ attempts to make to see an event w. prob. p to happen)

$= \frac{n}{n-k}$ see next slide

$E[\tau_n] = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \ldots + \frac{n}{3} + \frac{n}{2} + \frac{n}{1} = n \sum_{\ell=1}^{n} \frac{1}{\ell} \sim n \cdot \log n$ (log: natural log (ln))

Coupon Collecting

How long to get all n coupons ?

let $p = \frac{n-k}{n}$

the prob. of getting the (k+1)-st coupon, having already collected k, after trying ℓ times and not succeeding $\ell-1$ times: $(1-p)^{\ell-1} \cdot p$

$$E[\tau_{k+1} - \tau_k] = \sum_{\ell=1}^{\infty} (1-p)^{\ell-1} \cdot p \cdot \ell = p \cdot \sum_{\ell=0}^{\infty} q^{\ell} \cdot (\ell+1)$$

where $q = 1-p$

$$= p \cdot \frac{1}{p^2} = \frac{1}{p}$$

$$\sum_{\ell=0}^{\infty} q^{\ell} \cdot (\ell+1) = q^0 + q^1 + q^2 + q^3 + \dots$$
$$+ q^1 + q^2 + q^3 + \dots$$
$$+ q^2 + q^3 + \dots$$
$$+ q^3 + \dots$$
$$+ \ddots$$

$$= \frac{1}{1-q} = \frac{1}{p}$$
$$= q \cdot \frac{1}{1-q} = \frac{q}{p}$$
$$= q^2 \cdot \frac{1}{1-q} = \frac{q^2}{p}$$
$$= q^3 \cdot \frac{1}{1-q} = \frac{q^3}{p}$$

$$= \frac{1}{p}(1 + q + q^2 + q^3 + \dots)$$
$$= \frac{1}{p} \cdot \frac{1}{1-q} = \frac{1}{p^2}$$

Coupon Collecting

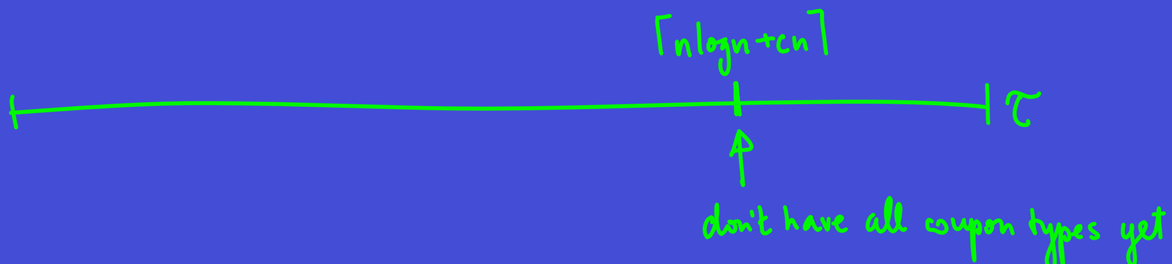
Proposition:
Let ¿ be the random variable for the time needed to collect all n coupons. Then,

Pr(¿ > d n log n + cn e) · e^{-c}.

let A_i be the event that we have not collected the i-th coupon type within the first $\lceil n\log n + cn \rceil$ steps.

$$\Pr(\tau > \lceil n\log n + cn \rceil) = \Pr(\exists i \text{ s.t. } A_i \text{ happened}) \le \sum_{i=1}^{n} \Pr(A_i)$$

(union bound)



$$\Pr(A_i) = \left(1-\frac{1}{n}\right)^{\lceil n\log n + cn \rceil}$$

$$= n\left(1-\frac{1}{n}\right)^{\lceil n\log n + cn \rceil}$$

$$\le e^{-c}$$

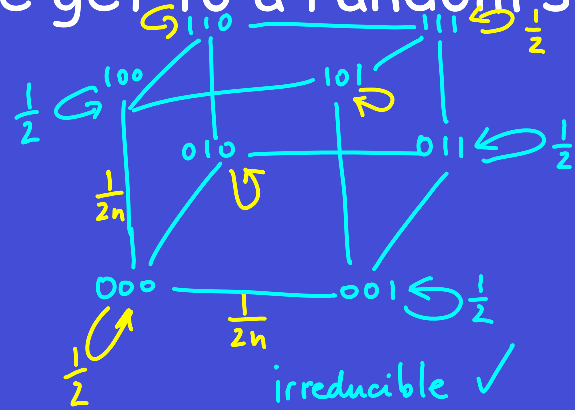
(recall log n natural log)

Random Walk on a Hypercube

How many expected steps until we get to a random state ?
(I.e., what is the mixing time ?)

Ω = all bin. strings of length n

move: choose a random position $k \in \{1, \ldots, n\}$
flip k-th bit

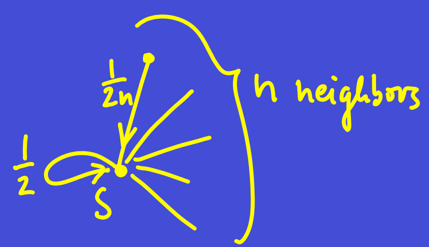


irreducible ✓
not aperiodic !
now OK
(ergodic
⇒ unique stationary distr.)

lazy: move : w. prob. $\frac{1}{2}$ do nothing
else

stat. distr. :

$$\pi(s) = \frac{1}{2^n}$$



verify : $\pi \cdot P = \pi$

chance of getting to s, assuming starting in every state w. prob. $\frac{1}{2^n}$:

$$n \cdot \frac{1}{2^n} \cdot \frac{1}{2n} + \frac{1}{2^n} \cdot \frac{1}{2} = \frac{1}{2^n}$$

Random Walk on a Hypercube

How many expected steps until we get to a random state ?
(I.e., what is the mixing time ?)

rephrased MC step:

randomly choose a position $k \in \{1, \dots, n\}$
w.prob. $\frac{1}{2}$ flip the k-th bit
else do nothing

Meaning: after selecting k, the k-th bit will be random
after selecting all positions, all bits will be random
(i.e. coupon collector!)

Hence: need $\sim n \log n$ steps