Finite Markov Chains & Random walks

Today
- intro Markov chains
- Algorithms
  - 2-SAT
  - bipartite matching in d-regular graph
- random walks on graphs

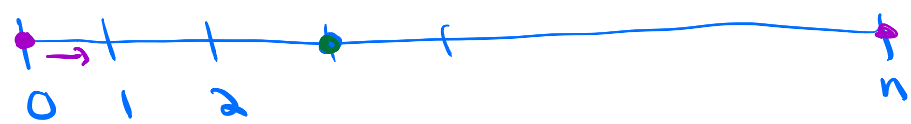
2-SAT Algorithm

- start w/ initial assignment
- Repeat up to $100n^2$ times until all clauses are satisfied
  - choose an arbitrary clause that is not satisfied
  - pick a variable in that clause at random & switch its value
- Report satisfying assignment if found; else return "unsatisfiable"

How to analyze?
Fix satisfying assignment S
Think of alg as random walk on line

n variables.

Fix satisfying assignment.



state # vars on which current assignment agrees with S

$(\ , \ )$

X_t: position after t steps
move to $X_t - 1$ or $X_t + 1$

$$\Pr(X_{t+1} = i+1 \mid X_t = i) \geq \frac{1}{2} \qquad = \frac{1}{2}$$

succeed when $X_t = n$

h_j = expected # steps for r.w to reach n starting from posn j

$$h_j = \underline{\tfrac{1}{2}} h_{j+1} + \underline{\tfrac{1}{2}} h_{j-1} + 1 \Rightarrow h_j - h_{j+1} = h_{j-1} - h_j + 2$$

$$h_0 = 1 + h_1 \Longrightarrow$$

$h_n = 0$

$h_0 - h_1 = 1$

By induction $h_j - h_{j+1} = 2j+1$

$$h_0 = h_0 - h_n = \sum_{i=0}^{n-1} h_i - h_{i+1} = \sum_{j=0}^{n-1} (2j+1)$$

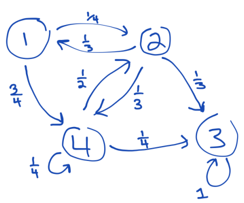
$$= n^2$$

$\Pr(\underline{\text{takes} > 2n^2}) \le \frac{1}{2}$

$100n^2$, 50 trials Prob failure $\le \frac{1}{2^{50}}$

Finite Markov Chains

- random walk on directed graph
- each vertex is a "state" of MC.
- each arc describes corresponding transition probability



$$\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & & & & \\ 2 & & & & \\ 3 & & & & \\ 4 & & & & \end{array} \quad \begin{pmatrix} 0 & \frac{1}{4} & 0 & \frac{3}{4} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

Transition Prob Matrix

Use X_t to denote state at time t

$$Pr(X_{t+1}=j \mid X_t=i) = p_{ij}$$

$P = (p_{ij})$ transition prob matrix

$\vec{p}^t = (p_1^t, p_2^t, \ldots, p_n^t)$ describes prob distn over states at time t

$$p_i^t = Pr(X_t = i)$$

$p_1^t \ldots p_n^t \; \Big(\; p_{1j}, p_{2j}, \ldots, p_{nj} \; \Big)$ (j)

$\vec{p}^0 = (1, 0, \ldots 0)$ means start in state 1

$\vec{p}^0 = (\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n})$ means start in uniformly random state

$$p_j^{t+1} = \sum_{i=1}^{n} p_i^t \, p_{ij}$$

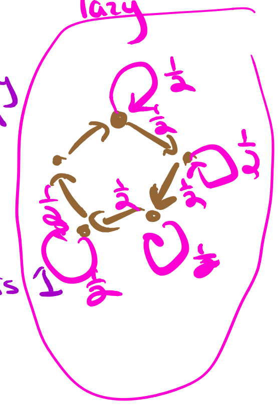
$$p^{t+1} = p^t P$$

$$\vec{p}^{t+m} = p^t P^m$$

Irreducible Markov chain

$\equiv$ corresponding graph strongly connected

"lazy"



period of state i $= \gcd\{n \geq 1 \mid P_{ii}^n > 0\}$

Markov chain is aperiodic if period of every state is 1

All Markov chains we will consider will be finite, irreducible & aperiodic

$\Rightarrow \; \exists N > 0$ s.t. P^n strictly positive $\forall n \geq N$

$$P_{ij}^n = Prob(X_{t+n} = j \mid X_t = i)$$

A $\boxed{\text{stationary}}$ distn $\vec{\pi}$ of a M.C. is a prob distn s.t.

$$\vec{\pi} = \vec{\pi} P \qquad \text{"fixed point"}$$

$$\sum_{i=1}^{n} \pi_i = 1$$

$$\forall j \quad \pi_j = \sum_i \pi_i p_{ij}$$

$$\vec{\pi} = \vec{\pi} P^t$$

$$(\pi_1, \pi_2, \ldots, \pi_n)$$

Fundamental Thm of Markov Chains

For any finite, irreducible, aperiodic MC

1. $\exists$ stationary distn $\vec{\pi}$ (with $\pi_i > 0 \; \forall i$)
2. $\vec{\pi}$ is unique
3. $\pi_i = \frac{1}{h_{i,i}}$ $\qquad h_{i,i} = \frac{1}{\pi_i}$
4. $\forall i,j \quad \lim_{t \to \infty} P_{ij}^t = \pi_j$

$$\Downarrow$$

$$\Pr(X_t = j \mid X_0 = i)$$

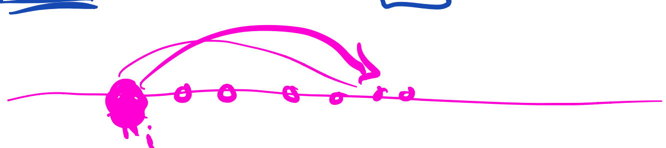
Notation

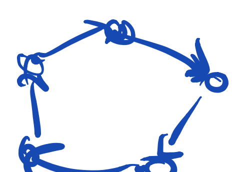
$$T_{ij} = \min_{t \geq 1} \{ X_t = j \mid X_0 = i \}$$

$$h_{ij} = E(T_{ij})$$

h_{ii} = expected first return time.

aperiodic: $\exists N$ s.t. $\forall i,j, \; P_{ij}^t > 0 \quad \forall t \geq N$





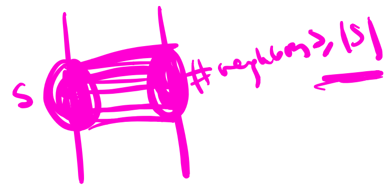
$$\left(\tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}\right)$$

$$\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Maximum Matching in Regular Bipartite Graphs

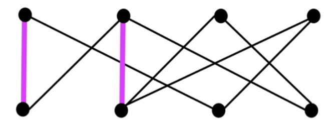
regular means all vertices have same degree d.

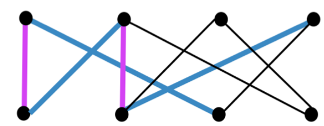
By Hall's Marriage Thm, regular bipartite graphs always have perfect matching



Traditional approach: augmenting path alg
repeatedly find one: can be done in $O(m)$ steps
using BFS

$O(mn)$ alg.



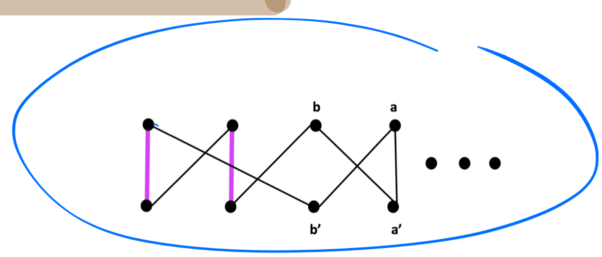


Prior to this best alg $O(m)$

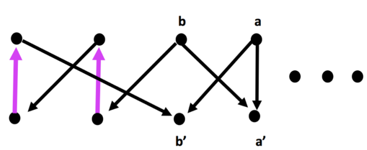
m # edges

Random walk based alg

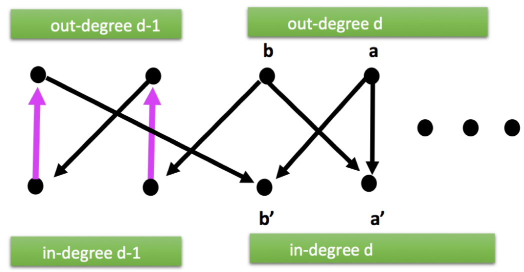
Find AP using random walk



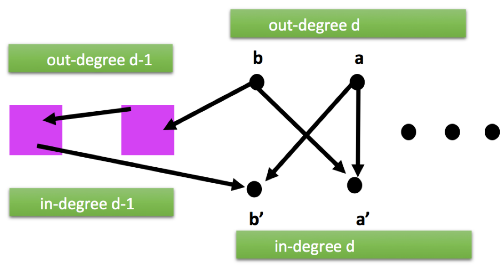
∃ AP iff ∃ directed path from top unmatched vertex to bottom unmatched vertex



$d=2$



contract matched edges

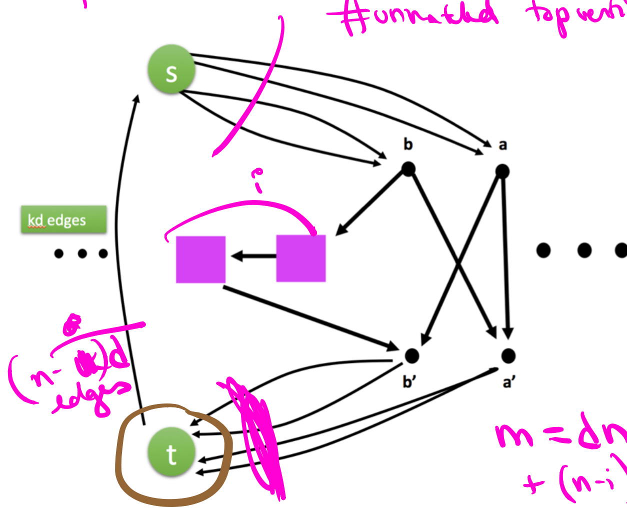


$n-i$ unmatched top vertices

unmatched top vertices $\cdot d$

add d edges from s to each unmatched node on top

add d edges from each unmatched node on bottom to t



$m = dn + (n-i)d \cdot 3$

Claim: AP in original graph iff $\exists$ cycle in graph above from s to itself

Alg: do r.w. starting from s till get back to s.

exp time starting from s to get back to s

$$h_{ss} = \frac{1}{\pi_s}$$

need to find stationary distn.

Claim

If $\boxed{\forall v \quad d_{in}(v) = d_{out}(v)}$ (in directed r.w.).

Claim $\forall v$

$$\pi_v = \frac{d^{out}(v)}{m}$$

m total # of edges in graph

$$\pi = \pi P$$

$$\sum \pi_v = 1$$

$$\pi_v \stackrel{?}{=} \sum_{u|u\to v} \pi_u \cdot P_{uv}$$

$$= \sum_{u|u\to v} \frac{d^{out}(u)}{m} \cdot \frac{1}{d^{out}(u)} = \sum_{u|u\to v} \frac{1}{m} = \frac{d_{in}(v)}{m} = \frac{d_{out}(v)}{m}$$

$$h_{ss} = \frac{1}{\pi_s} = \frac{m}{d^{out}(s)}$$

Suppose in this iteration $\exists$ i matched edges

$$d^{out}(s) = d(n-i)$$

$$m = dn + 3d(n-i)$$

$$\Rightarrow \quad h_{ss} \le \frac{dn + 3d(n-i)}{d(n-i)} = \boxed{\frac{n}{n-i}} + 3$$

exp time to go from i matched vertices to i+1 matched vertices

$$\frac{n}{n-i} + 3$$

total exp running time = $\sum_{i=0}^{n-1} \left(\frac{n}{n-i} + 3 \right)$

$= O(n \log n)$

$O(m)$

could be $o(m)$

d much bigger than $\log n$

---

can add self loops.

$\pi = \pi P$

$\Downarrow$

$\frac{1}{2}(I + P)$

$\pi = \pi I$