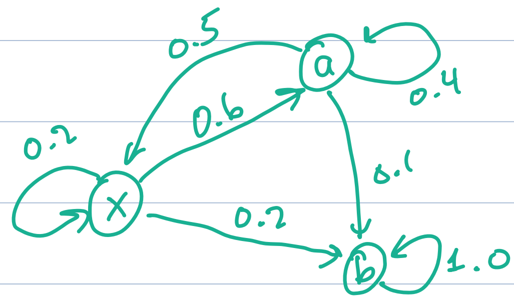
Markov chains and $\boxed{MCMC}$

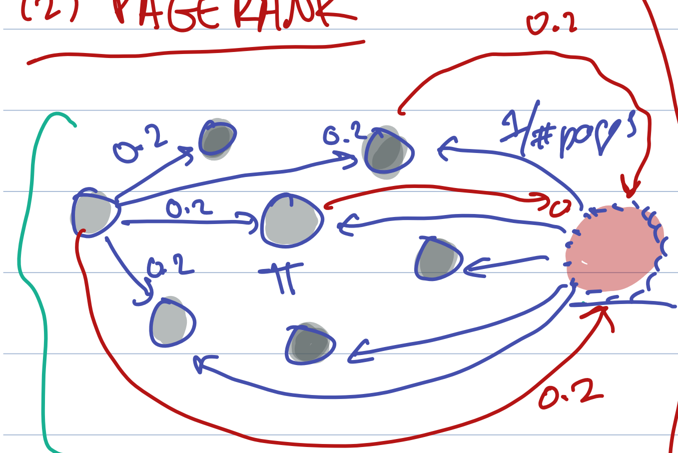


$\Omega = \{x, a, b\}$

$D = (V, A, P) \quad V = \Omega$

$A = \{(x,y) : P(x,y) > 0\}$

(2) PAGERANK



1	2	3	...	50
0	0	0		0
7	~~13~~ 2	1	...	~~2~~ 13

Markov chain:

Ω - state space (finite)

$P : \Omega \times \Omega \to [0,1]$

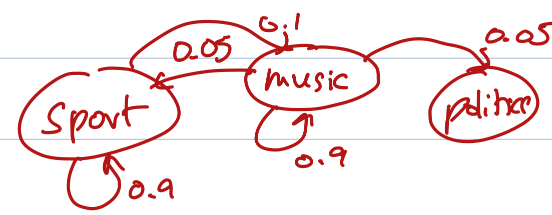
$\forall x \in \Omega : \sum_{y \in \Omega} P(x,y) = 1$

Random walk: $\{X_0, X_1, \dots\} \subseteq \Omega$

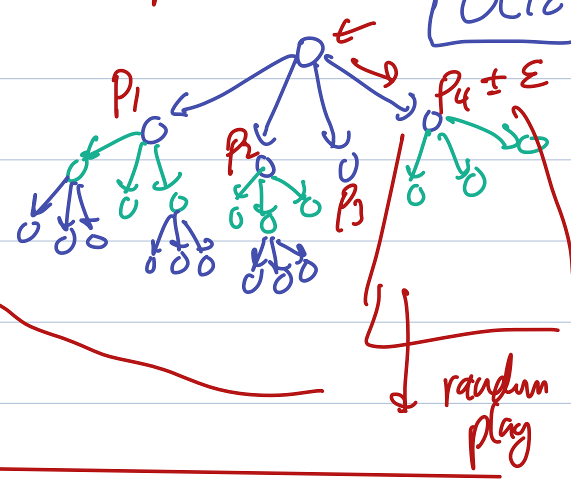
$\mathbb{P}[X_{t+1} = y \mid X_t = x] = P(x,y)$

Applications:

(1) Hidden Markov model

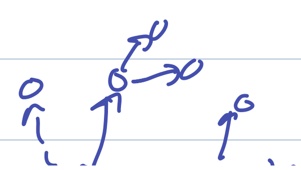


(3) AlphaGO $\qquad \boxed{O(1/\varepsilon^2)}$



(4) Nuclear bombs

(Ulam)

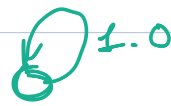


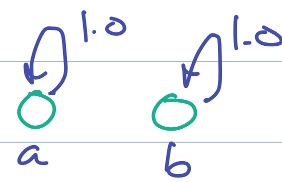
Def: A prob. distribution π on Ω is called the <u>stationary distribution</u> if

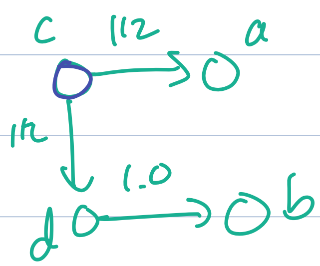
$$P_t(x,y) \longrightarrow \pi(y) \quad \text{as } t \to \infty \quad \forall x,y \in \Omega$$

$$\parallel$$

$$\mathbb{P}[X_t = y \mid X_0 = x]$$





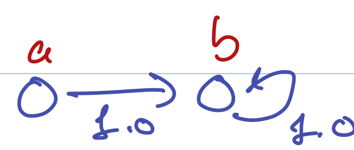


a : 1 : a
b : 1 : b
c : ½ : a , ½ : b
d : 1 : b

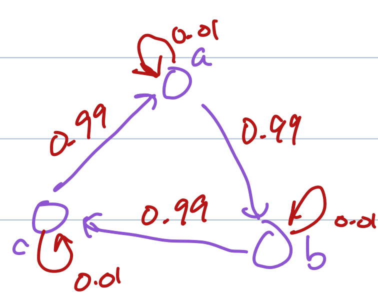
Def: A Markov chain is <u>reducible</u> if its digraph is <u>not</u> strongly connected.

(Irreducible opposite of reducible)

Def: Periodic:



$\pi(b) = 1$
$\pi(a) = 0$



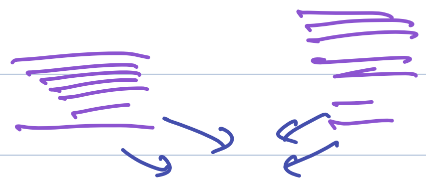
$$\gcd(\{t : P_t(a,a) > 0\}) = 3$$

A Markov chain is <u>aperiodic</u> if

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

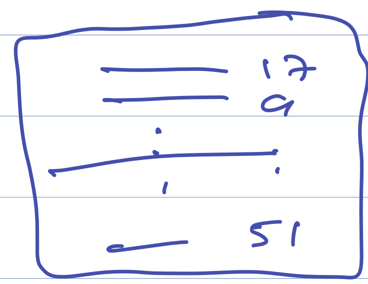
Fundamental thm of Markov chains: If P is aperiodic and irreducible, then P has a unique stationarity distribution.

- What is π? $\quad (P_t(x,\cdot) \simeq \pi)$
- How long does it take for $P_t(x,y) \simeq \pi(y)$ $\forall y \in \Omega$?



$52!$

$\boxed{7 \text{ riffle shuffles}}$



Def: If μ and ν are two prob. distr, then

$$d_{TV}(\mu,\nu) = \frac{1}{2} \sum_{x\in\Omega} |\mu(x) - \nu(x)|$$

$$\left(= \max_{E \subseteq \Omega} |\mu(E) - \nu(E)| \right)$$

$$\tau_{mix} := \min \{ t : d_{TV}(\pi, P_t(x,\cdot)) \leq 1/4 \ \forall x \in \Omega \}$$

$P_t(x,\cdot) \to \pi$ as $t \to \infty$

$d_{TV}(P_t(x,\cdot), \pi) \to 0 \quad \forall x \in \Omega$

(for any $c \geq 1$)

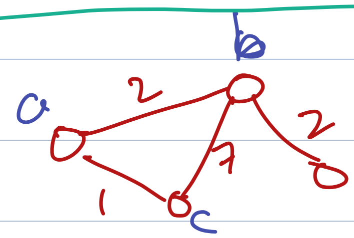
Fact: After $t = c \cdot T_{mix}$ steps,

$$d_{TV}(\pi, P_t(x, \cdot)) \leq e^{-c} \quad \forall x \in \Omega$$

Symmetric chains:

("Reversible" chains)

$P(x,y) = P(y,x)$



$P(a,b) = \frac{2}{2+1} = \frac{2}{3}$

$P(a,c) = 1$

$P(b,a) = \frac{2}{2+7+2} = \frac{2}{11}$

$\Omega = \{1, 2, \ldots, n\}$

$P_{ij} := P(i,j)$

P is real, sym. $n \times n$ matrix

$v = (v_1, v_2, \ldots, v_n)$

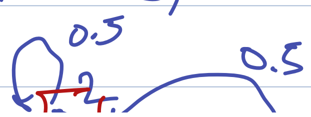
$v_i =$ prob mass on i

$v \in \mathbb{R}^n \quad v_i \geq 0 \ \forall i$

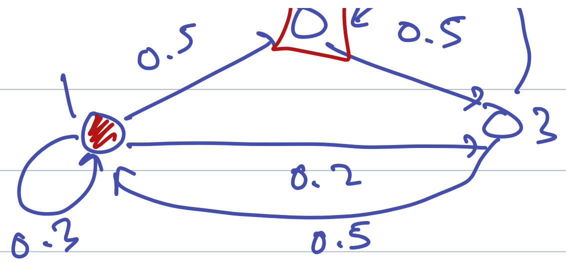
$v_1 + \cdots + v_n = 1$

v prob distr. Then $v^T P$ is the distr after one step.

$v = (1, 0, 0, \ldots, 0)$



$v = (1, 0, 0)$



one step of RW:

$(0.3, 0.5, 0.2)$

two steps of RW:

$0.3\,(0.3, 0.5, 0.2) +$
$0.5\,(0, 0.5, 0.5) +$
$0.2\,(0.5, 0.5, 0)$

$= (0.19, 0.5, 0.31)$

$$(v^T P)_i = \sum_{j=1}^{n} v_j \cdot P_{ji}$$

$$\boxed{v^T P^k} \qquad (v^T P^k)_i = \mathbb{P}[X_k = i \mid X_0 \sim v]$$

$\pi^T P = \pi^T$ left ev.

$\pi^T P^k = \pi^T$ w/ eigenvalue 1.

Eigenvalues of P lie in $[-1, 1]$

$$\underbrace{\dots \le \lambda_2}\ \boxed{\overbrace{\lambda_2 < 1}^{\text{spectral gap}} = \lambda_1}$$

$\underbrace{1 = \lambda_1}_{\text{st. distr. } \pi}$

$v^T P^k$

Reversible Markov chains

mix. time $\Longleftrightarrow$ eigenvalues

$$V = \alpha_1 u^{(1)} + \cdots + \alpha_n u^{(n)}$$

$$\underbrace{\alpha_1 u^{(1)}}_{\Pi} + \alpha_2 \lambda_2^k u^{(2)} + \cdots + \alpha_n \lambda_n^k u^{(n)}$$

$\to \Pi$