Random Sums and Graphs

Probability

Events, probabilities,
conditional probabilities,
union bound

Random variables,
expected value,
linearity of expectation

Markov's inequality

Algorithms

Approx. 3-SAT,

Big idea: we only want the average

Linearity of Expectation
$E[X+Y] = E[X] + E[Y]$ (ridiculously useful)
"average sum = sum of averages"

let $Y_i = \begin{cases} 1 & \text{if ith clause satisfied} \\ 0 & \text{otherwise} \end{cases}$ $(i = 1, \dots, m)$

$E[Y_i] = 7/8$ (1 clause case)

$$E\left[\begin{matrix}\text{\# clauses} \\ \text{satisfied}\end{matrix}\right] = E\left[\sum_{i=1}^{m} Y_i\right] = \sum_{i=1}^{m} E[Y_i] = 7/8\, m$$

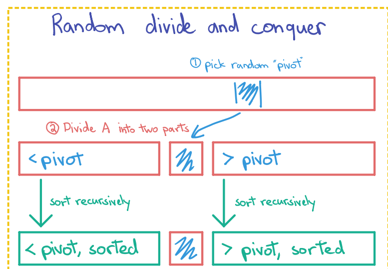
let $Y_i = \begin{cases} 1 & \text{if ith clause satisfied} \\ 0 & \text{otherwise} \end{cases}$

$$Z = \sum_{i=1}^{m} Y_i = \text{\# clauses satisfied}$$

$$E[Z] = \tfrac{1}{2} E[Z \mid x_1 = t] + \tfrac{1}{2} E[Z \mid x_1 = f]$$

$\Rightarrow$ either $E[Z \mid x_1 = t]$ or $E[Z \mid x_1 = f]$ is $\geq E[Z]$!

Quicksort,



Expected running time of QuickSort

For each $i, j \in [n]$ w/ $i < j$,

$X_{ij} = \begin{cases} 1 & \text{if ith smallest element is compared to jth sma\dots} \\ 0 & \text{otherwise} \end{cases}$

$\sum_{i<j} X_{ij}$ = total # comparisons made by the algorithm
$\approx$ total running time

$E\left[\sum_{i<j} X_{ij}\right]$ = average running time

$E\left[\sum_{i<j} X_{ij}\right] = \sum_{i,j} E[X_{ij}]$

Linearity of Expectation
$E[X+Y] = E[X] + E[Y]$ (ridiculously useful)
"average sum = sum of averages"

now analyze $E[X_{ij}]$ in isolation

$E[X_{ij}] = \Pr[X_{ij} = 1]$
$= \Pr\left[\begin{matrix}\text{ith element is} \\ \text{compared to jth}\end{matrix}\right]$

$\boxed{i}\ \boxed{}\ \boxed{}\ \boxed{} \cdots \boxed{}\ \boxed{j}$

ith elem is compared w/ jth
$\Leftrightarrow$
ith or jth elem selected before any element in between

happens w/ probability $\frac{2}{j-i+1}$

$$E[X_{ij}] = \frac{2}{j-i+1}$$

$$E\left[\sum_{i<j} X_{ij}\right] = \sum_{i,j} E[X_{ij}]$$ by linearity of expectation

$$= \sum_{i=1}^{n} \sum_{j=i+1}^{n} \frac{2}{j-i+1} = \sum_{j=1}^{n} \sum_{i=1}^{j} \frac{2}{j-i+1}$$

$$\leq 2n\left(\frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n}\right)$$

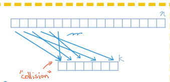
$$\leq O(n \log n)$$ nth harmonic number

Quickselect (HW)

Probability

Hash Functions,

Universal Hash Functions



$h : [n] \to k$ s.t. for all pairs $i, j \in [n]$,

$$\Pr[h(i) = h(j)] \leq \frac{1}{k}$$

(much weaker than ideal hash)

e.g. $h(x) = ((ax + b) \bmod p) \bmod k$

where
- p prime, $> k$
- a in $\{1, \ldots, p\}$ unif. random
- b in $\{0, \ldots, p-1\}$ unif. random

Markov's Inequality

Markov's Inequality (special case)

Let $X \geq 0$ be nonnegative RV

$$\Pr[X \geq 2E[X]] \leq \frac{1}{2}$$

Algorithms

Frequency Estimation, Heavy Hitters (in streams)

A_1, h_1

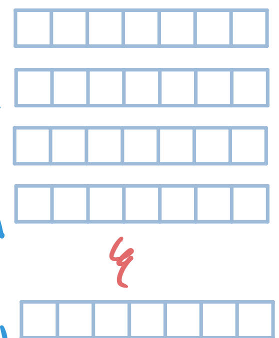
A_2, h_2

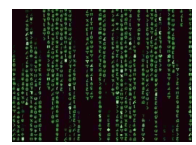
A_3, h_3

A_4, h_4

$\vdots$

A_d, h_d





estimate f_i w/ $\min_{j=1,\ldots,d} A_j[h_j(i)]$

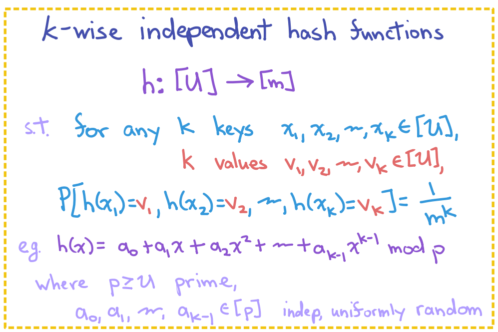
Probability

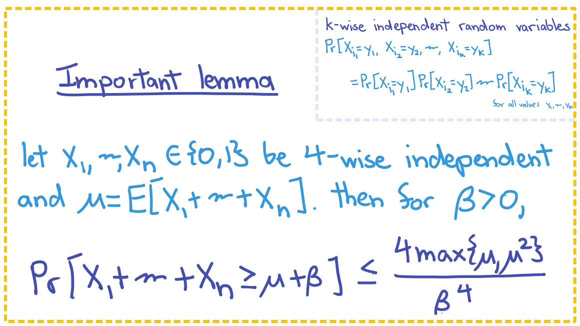
k-wise independence, higher order moments

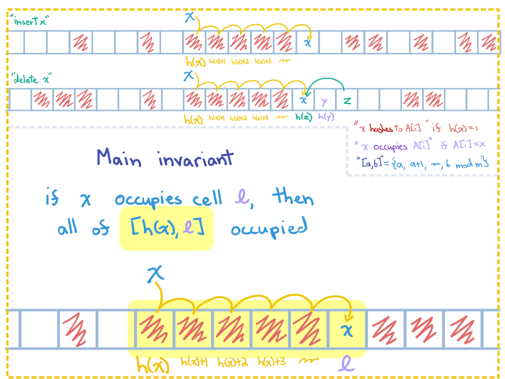
(generalizating universal hash functions and Markov's inequality)

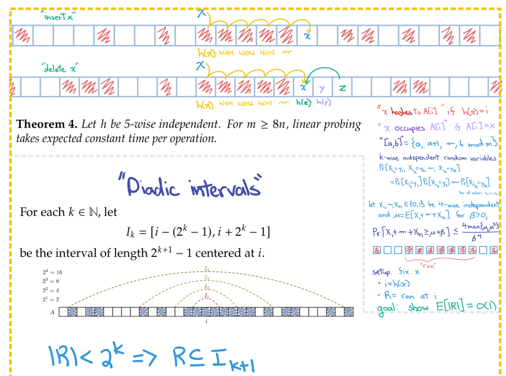
Algorithms

Hash tables w/ chaining + linear probing





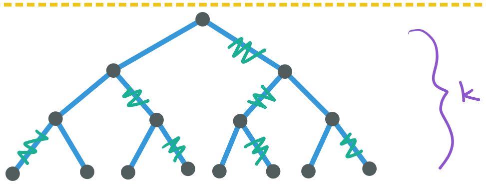




Probability

Algorithms

Galton-Watson processes

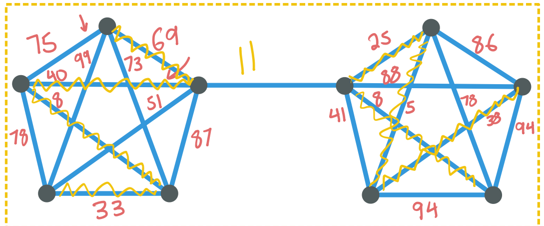


let T = complete binary tree of height k,
w/ every edge deleted w/ prob $\le 1/2$

Fact: $P[\text{some leaf connected to root}] \ge \frac{1}{k}$

"Galton-Watson process"

Random Contractions for min cut



Randomized contractions

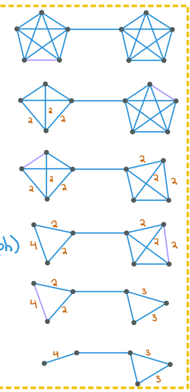
Input: $G=(V,E)$, weights $c \in \mathbb{R}^E_{\ge 0}$ (capacities)

① while $|E|>1$

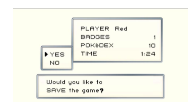
Ⓐ sample $e \in E$ in proportion to c

Ⓑ "contract" e

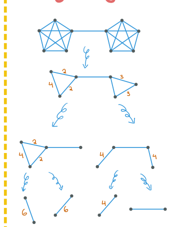
② return set of edges (in original graph) contracted to remaining edge



Amplification by branching



Instead of restarting at the beginning, restart partway through



High-level algorithm

① randomly contract until prob. of ruining the min-cut is $\sim 1/2$

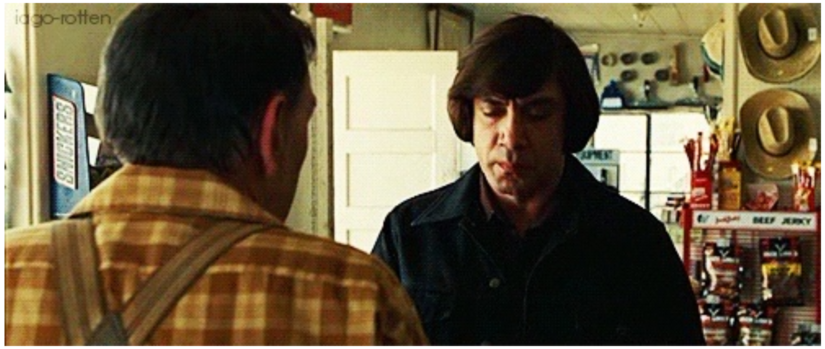
② *Branch:*

make two copies of current graph

recurse on both copies

Probability	Algorithms (/ results)
Multiplicative Chernoff Bound	Random graphs & the giant component

Probability	Algorithms (/ results)
Multiplicative Chernoff Bound	Random graphs & the giant component

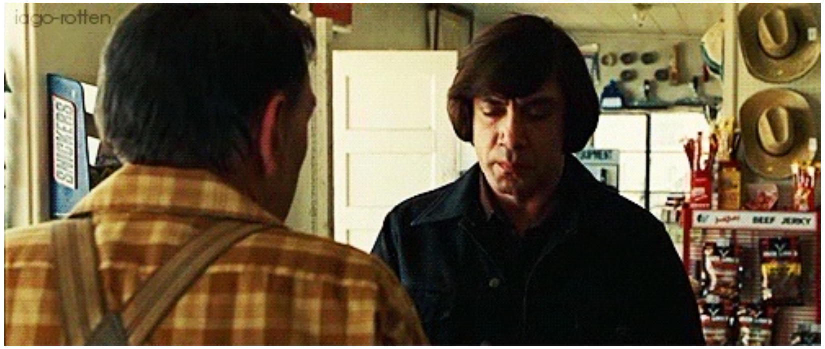


Flip 100 coins.

50 come up heads.

Should you be surprised?

(Maybe a little: ≈ 8% chance)

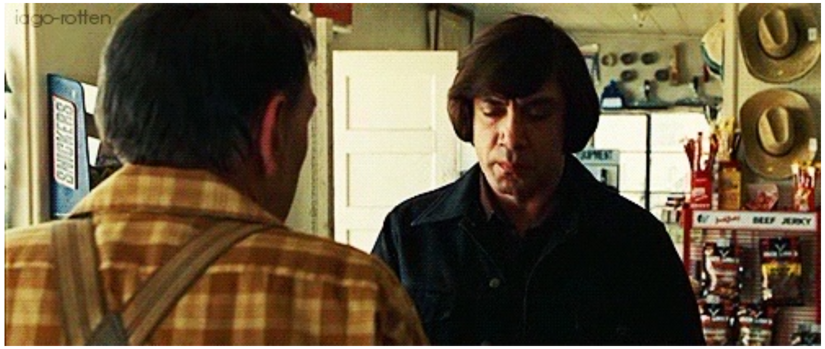


Flip 100 coins.

0 come up heads.

Should you be surprised?

Absolutely: $\frac{1}{2^{100}}$ chance



$.0038$

10^{-20}

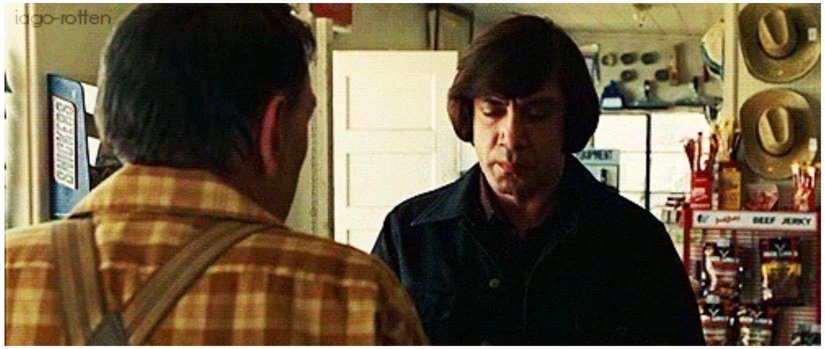
10^{-5}

Flip 100 coins.

≤ 25 come up heads.

Should you be surprised?

Yes: $\approx 2.818 \times 10^{-7}$

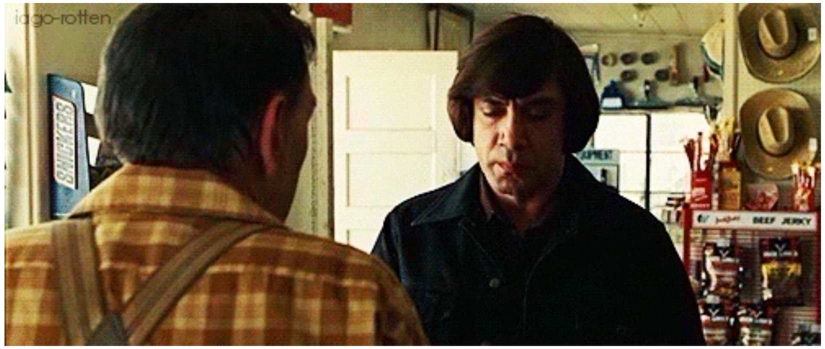


Flip 10 coins.

≤ 4 come up heads.

Should you be surprised?

No: $\approx 37.7\%$



10^{-20}

10^{-9}

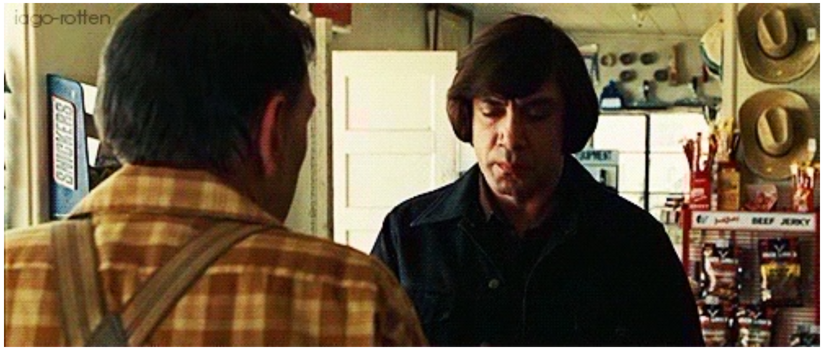
$.00377$

Flip 1000 coins.

$\leq 40\%$ come up heads.

Should you be surprised?

Yes: $\approx 1.364 \times 10^{-10}$ probability



10 coins.

≤ 4 heads.

$\approx 37.7\%$

1000 coins.

$\leq 40\%$ heads.

1.364×10^{-10} prob.

scale matters

Concentration phenomena

let $X_1, \cdots, X_n \in [0,1]$ be independent

$$\mu = E[X_1 + \cdots + X_n]$$

Q: probability that

$X_1 + \cdots + X_n$ is close to μ?

(multiplicatively)

A: error probability exponentially small in μ

Multiplicative Chernoff bound

let $X_1, \dots, X_n \in [0,1]$ be independent

$\mu = E[X_1 + \dots + X_n]$

$\varepsilon \in (0,1)$

- $P[X_1 + \dots + X_n \geq (1+\varepsilon)\mu] \leq e^{-\varepsilon^2 \mu / 3}$
- $P[X_1 + \dots + X_n \leq (1-\varepsilon)\mu] \leq e^{-\varepsilon^2 \mu / 2}$

(proof at end if time permits)

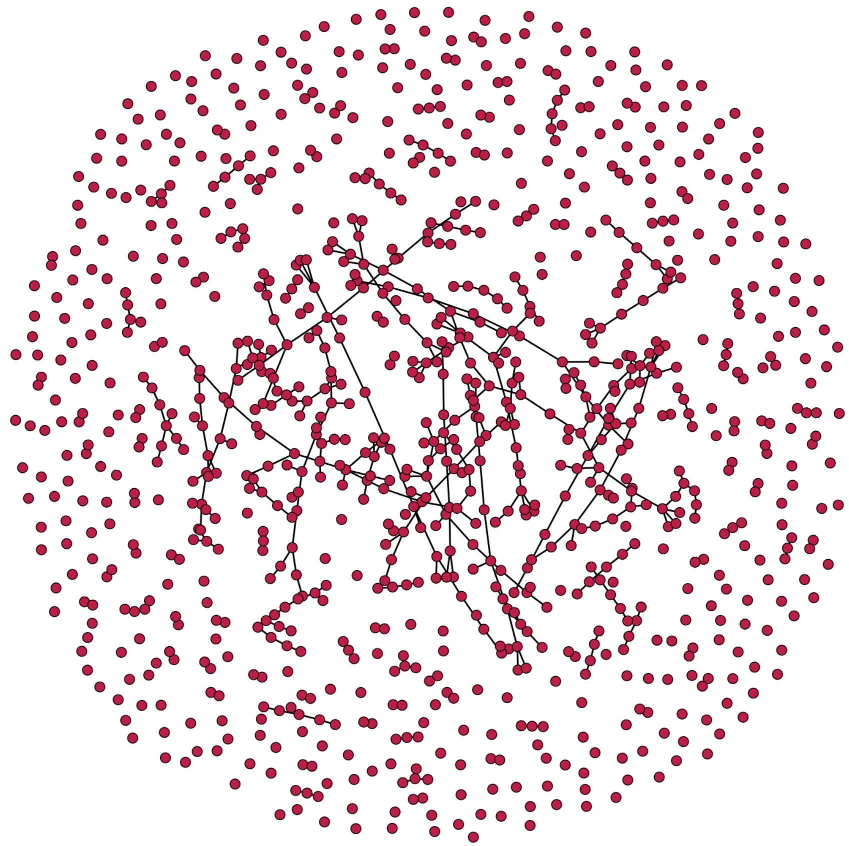
Probability	Algorithms (/ results)
Multiplicative Chernoff Bound	Random graphs & the giant component

Erdös-Rényi graphs $G(n, p)$ (1959)

- n vertices
- each edge sampled independently w/ probability p

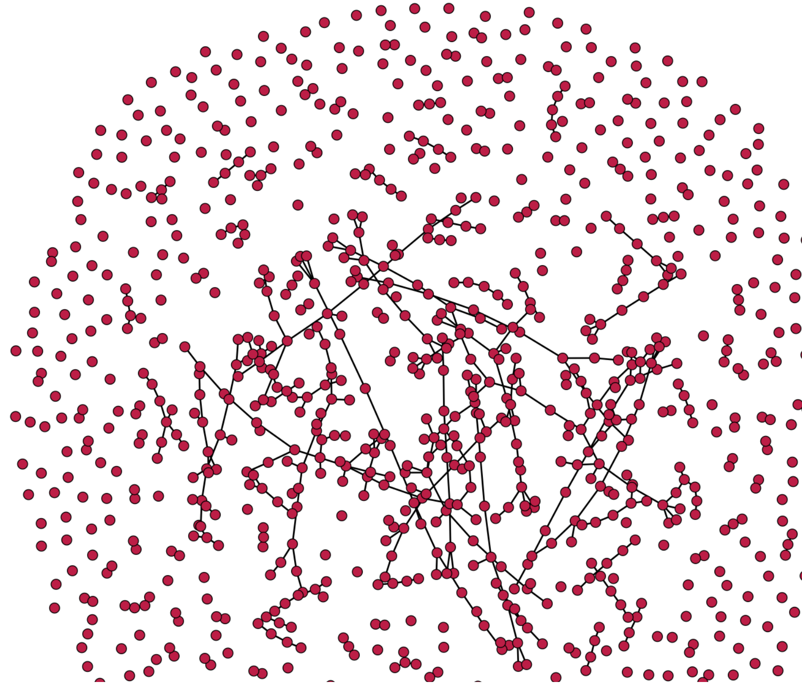
Q: what does such a graph look like?

does it consistently have certain properties?



Erdös-Rényi graphs $G(n, p)$

- n vertices
- each edge sampled independently w/ probability p

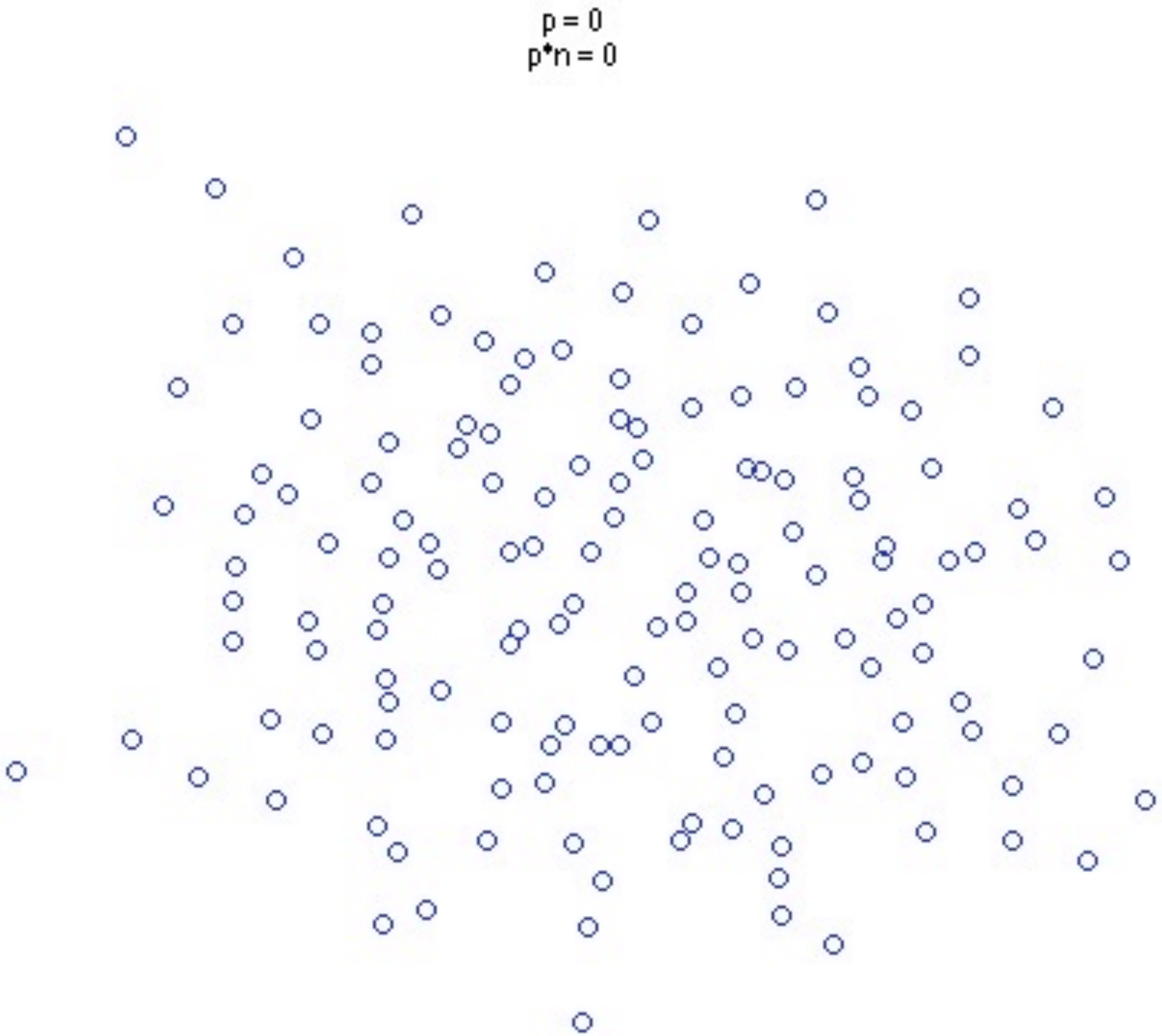


Theorem 5.3. *Consider a random graph $G \sim \mathcal{G}(n,p)$ for $p = c/n$, where c is a constant.*

1. *If $c > 1$, then with high probability, there is exactly one connected component of G with $\Omega(n)$ vertices, and all other components have size $\leq O(\log n)$.*

2. *For $c < 1$, then with high probability, all connected components of G have size $< O(\log n)$.*

"threshold phenomena"



https://www.cs.purdue.edu/homes/dgleich/demos/erdos_renyi/er-150.gif

Theorem 5.3. *Consider a random graph $G \sim \mathcal{G}(n,p)$ for $p = c/n$, where c is a constant.*

1. *If $c > 1$, then with high probability, there is exactly one connected component of G with $\Omega(n)$ vertices, and all other components have size $\leq O(\log n)$.*
2. *For $c < 1$, then with high probability, all connected components of G has size $< O(\log n)$.*

Proof outline ($c>1$)

① Gap theorem let $p=(1+\varepsilon)/n$. w/ high probability, all components have size

"small component" $\leq O\left(\frac{\log n}{\varepsilon^2}\right)$ or "large component" $\geq \Omega(\varepsilon n)$

② Existence of large component

③ Uniqueness of large component

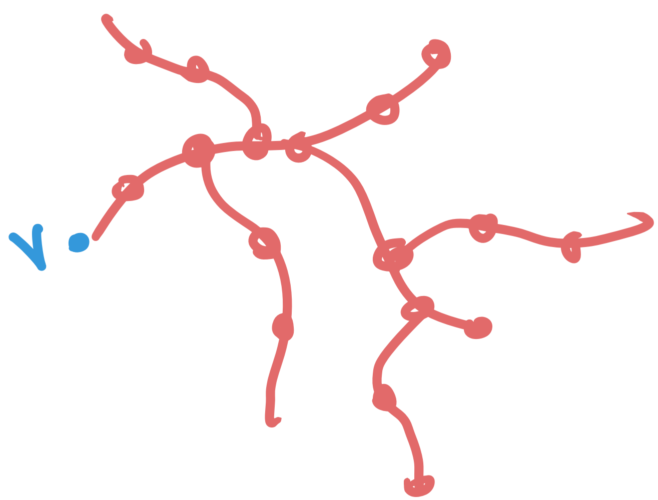
① Gap theorem let $p=(1+\varepsilon)/n$. w/ high probability, all components have size

$$\leq O\left(\frac{\log n}{\varepsilon^2}\right) \quad \text{or} \quad \geq \Omega(\varepsilon n)$$

Setup

fix vertex v.

let $C(v) \subseteq V$ = connected component of v



to analyze $C(v)$, imagine revealing $C(v)$ w/ a search algo

Fix vertex v. $C(v) \subseteq V$ = connected component of v

"Search algo"

$A = \{$vertices known to be in $C(v)\}$ (initially $A = \{v\}$)

$B = \{$explored vertices$\}$ (initially $B = \emptyset$)

① while there is a vertex $v_i \in A \setminus B$

ⓐ add all neighbors of v_i to A

ⓑ add v_i to B

Indexed version of "Search algo"

$A_i = \{$vertices known to be in $C(v)$ after i iter.$\}$ $(A_0 = \{v\})$

$B_i = \{$explored vertices after i iterations$\}$ $(B_0 = \phi)$

1. For $i = 1, 2, \ldots$ until $A_i = B_i$
   a. Let $v_i \in A_{i-1} \setminus B_{i-1}$
   b. $A_i = A_{i-1} \cup N(v_i)$
   c. $B_i = B_{i-1} \cup \{v_i\}$

Observation 1

algo terminates iff $|A_i| = i$

b/c $|B_i| = i$

$B_i \subseteq A_i$

$B_i = A_i \iff |A_i| = i$

$A_i = \{\text{vertices known to be in } C(v) \text{ after } i \text{ iter.}\}$ $(A_0 = \{v\})$

$B_i = \{\text{explored vertices after } i \text{ iterations}\}$ $(B_0 = \phi)$

1. For $i = 1, 2, \ldots$ until $A_i = B_i$
   a. Let $v_i \in A_{i-1} \setminus B_{i-1}$
   b. $A_i = A_{i-1} \cup N(v_i)$
   c. $B_i = B_{i-1} \cup \{v_i\}$

$A_i = \{$vertices known to be in $C(v)$ after i iter.$\}$ $(A_0 = \{v\})$

$B_i = \{$explored vertices after i iterations$\}$ $(B_0 = \phi)$

1. For $i = 1, 2, \dots$ until $A_i = B_i$
   a. Let $v_i \in A_{i-1} \setminus B_{i-1}$
   b. $A_i = A_{i-1} \cup N(v_i)$
   c. $B_i = B_{i-1} \cup \{v_i\}$

Observation 1

algo terminates iff $|A_i| = i$

Observation 2 $A_{i+1} = A_i \cup S_{i+1}$

where S_{i+1} samples each $v \in V \setminus A_i$ independently w/ probability p

Key point: Analysis oblivious to B

Observation 1 algo terminates iff $|A_i| = i$

Observation 2 $A_{i+1} = A_i \cup S_{i+1}$

where S_{i+1} samples each $v \in V \setminus A_i$ independently w/ probability p

<table>
<tr>
<td>

1. For $i = 1, 2, \ldots$ until $A_i = B_i$
   a. Let $v_i \in A_{i-1} \setminus B_{i-1}$
   b. $A_i = A_{i-1} \cup N(v_i)$
   c. $B_i = B_{i-1} \cup \{v_i\}$

</td>
<td>

1. For $i = 1, 2, \ldots$ until $|A_i| = i$
   a. Let S_i randomly sample $V \setminus A_{i-1}$ w/ probability p
   b. $A_i = A_{i-1} \cup S_i$

</td>
</tr>
</table>

(equivalent)

what is the distribution of:

① Gap theorem let $p = (1+\varepsilon)/n$. w/ high probability, all components have size $\leq O\left(\frac{\log n}{\varepsilon^2}\right)$ or $\geq \Omega(\varepsilon n)$

Multiplicative Chernoff bound

let $X_1, \dots, X_n \in [0,1]$ be independent

$\mu = E[X_1 + \dots + X_n]$ $\varepsilon \in (0,1)$

$P[X_1 + \dots + X_n \leq (1-\varepsilon)\mu] \leq e^{-\varepsilon^2 \mu / 2}$

1. For $i = 1, 2, \dots$ until $|A_i| = i$
   a. Let S_i randomly sample $V \setminus A_{i-1}$ w/ probability p
   b. $A_i = A_{i-1} \cup S_i$

$|A_1| - 1$ $E[|A_1|] = 1 + p(n-1)$

$p[x \notin A_2] = (1-p)^2$

$|A_2| - 1 =$

$|A_2| - 1 = \#$ heads w/ $n-1$ cointosses

prob[heads] $= 1 - (1-p)^2$

$|A_3| - 1$

$|A_3| - 1 = 1 - (1-p)^3$

$|A_i| - 1$

$|A_i| - 1 = 1 - (1-p)^i$

let $i \le \frac{\varepsilon n}{4}$.

Lemma 1. $E[|A_i|] \ge (1+\frac{\varepsilon}{2})i$

$$E[|A_i|-i] = (1-(1-p)^i)\cdot(n-1)$$

$$(1-p)^i \le e^{-pi}$$

$$E[\sim] \ge \left(1-e^{-\frac{(1+\varepsilon)i}{n}}\right)\cdot(n-1)$$

$$\ge \left[\frac{(1+\varepsilon)i}{n} - \frac{\left(\frac{(1+\varepsilon)i}{n}\right)^2}{2}\right](n-1)$$

$$\ge \left(1+\varepsilon - \frac{(1+\varepsilon)^2\varepsilon}{8}\right)\frac{i}{n}(n-1)$$

$$\left(\frac{(1+\varepsilon)i}{n}\right)^2 \le \frac{(1+\varepsilon)^2\varepsilon i}{4n}$$

$$\ge \left(1+\frac{3\varepsilon}{4}\right)\frac{i}{n}(n-1)$$

① Gap theorem let $p=(1+\varepsilon)/n$. w/ high probability, all components have size $\le O(\frac{\log n}{\varepsilon^2})$ or $\ge \Omega(\varepsilon n)$

Multiplicative Chernoff bound

let $X_1,\dots,X_n \in [0,1]$ be independent

$\mu = E[X_1+\dots+X_n]$ $\quad \varepsilon \in (0,1)$

$$P[X_1+\dots+X_n \le (1-\varepsilon)\mu] \le e^{-\varepsilon^2\mu/2}$$

1. For $i=1,2,\dots$ until $|A_i|=i$
   a. Let S_i randomly sample $V\setminus A_{i-1}$ w/ probability p
   b. $A_i = A_{i-1}\cup S_i$

$$p=\frac{1+\varepsilon}{n}$$

$$1+x \le e^x \le 1+x+\frac{x^2}{2}$$

$x \le 0$

let $i \le \frac{\varepsilon n}{4}$. $\ge (1+\frac{\varepsilon}{2}) \cdot i$

① Gap theorem let $p = (1+\varepsilon)/n$. w/ high probability, all components have size $\le O(\frac{\log n}{\varepsilon^2})$ or $\ge \Omega(\varepsilon n)$

Multiplicative Chernoff bound

let $X_1, \ldots, X_n \in [0,1]$ be independent

$\mu = E[X_1 + \cdots + X_n]$ $\varepsilon \in (0,1)$

$P[X_1 + \cdots + X_n \le (1-\varepsilon)\mu] \le e^{-\varepsilon^2 \mu/2}$

1. For $i = 1, 2, \ldots$ until $|A_i| = i$
   a. Let S_i randomly sample $V \setminus A_{i-1}$ w/ probability p
   b. $A_i = A_{i-1} \cup S_i$

Lemma 1. $E[|A_i|] \ge (1+\underset{2}{\varepsilon})i$

Lemma 2. $\Leftarrow 16 \frac{\log n}{\varepsilon^2}$

$$P[|A_i| \le i] \le e^{-\varepsilon^2 i/8}$$

$$\le \frac{1}{n^2}$$

$$A_i = \sum_v X_v$$

$$X_v = \begin{cases} 1 & \text{if } v \in A_i \\ 0 & \text{otherwise} \end{cases}$$

$$\mu = E[A_i] \ge (1+\underset{2}{\varepsilon})i$$

$$\Pr[|A_i| \le i] \le \Pr[|A_i| \le (1-\tfrac{\varepsilon}{4})\mu]$$

$$\le e^{-(\frac{\varepsilon}{4})^2 \mu/2}$$

$$= e^{-\varepsilon^2 \mu/32}$$

let $i \le \frac{\varepsilon n}{4}$.

Lemma 1. $E[|A_i|] \ge (1+\varepsilon) i$

Lemma 2.

$$P[|A_i| \le i] \le e^{-\varepsilon^2 i/8}$$

for $i \ge 32 \log n/\varepsilon^2$, $\le \frac{1}{n^4}$

by union bound over $\frac{32\log n}{\varepsilon^2} \le i \le \frac{\varepsilon n}{4}$,
A_i fails for some

$$P\left[\frac{32\log n}{\varepsilon^2} \le i \le \frac{\varepsilon n}{4}\right] \le \frac{1}{n^3}$$

① Gap theorem let $p = (1+\varepsilon)/n$. w/ high probability, all components have size $\le O(\frac{\log n}{\varepsilon^2})$ or $\ge \Omega(\varepsilon n)$

Multiplicative Chernoff bound
let $X_1, \ldots, X_n \in [0,1]$ be independent

$\mu = E[X_1 + \cdots + X_n]$ $\varepsilon \in (0,1)$

$$P[X_1 + \cdots + X_n \le (1-\varepsilon)\mu] \le e^{-\varepsilon^2 \mu/2}$$

1. For $i = 1, 2, \ldots$ until $|A_i| = i$
   a. Let S_i randomly sample $V \setminus A_{i-1}$ w/ probability p
   b. $A_i = A_{i-1} \cup S_i$

Theorem 5.3. *Consider a random graph $G \sim \mathcal{G}(n,p)$ for $p = c/n$, where c is a constant.*

1. *If $c > 1$, then with high probability, there is exactly one connected component of G with $\Omega(n)$ vertices, and all other components have size $\leq O(\log n)$.*
2. *For $c < 1$, then with high probability, all connected components of G has size $< O(\log n)$.*

Proof outline ($c>1$)

① Gap theorem let $p=(1+\varepsilon)/n$. w/ high probability, all components have size

$$\leq O\left(\frac{\log n}{\varepsilon^2}\right) \quad \text{or} \quad \geq \Omega(\varepsilon n)$$

② Existence of large component

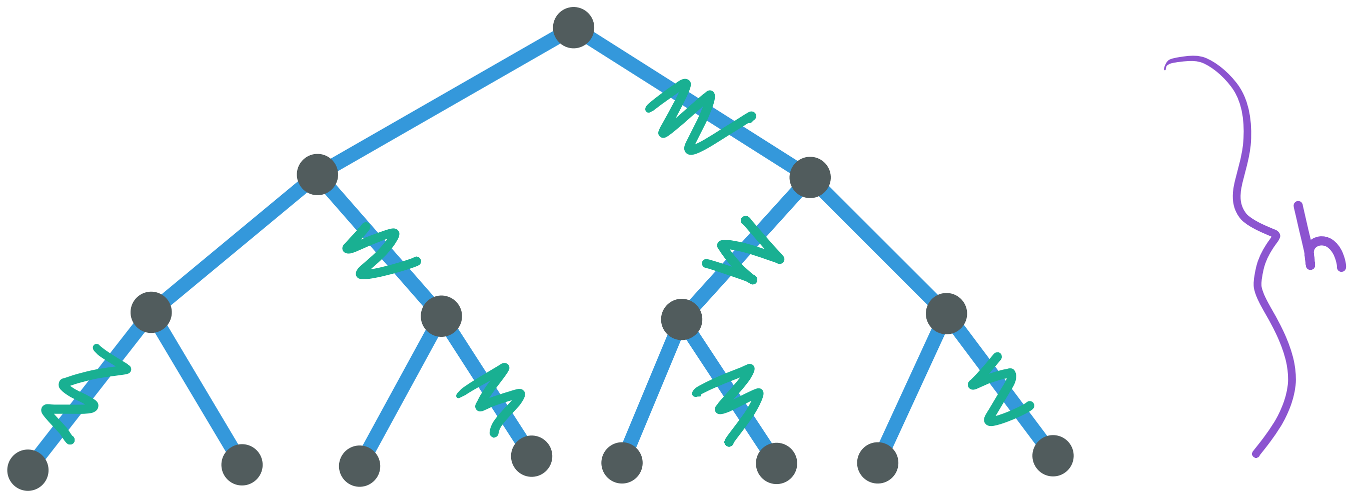
③ Uniqueness of large component

② Existence of large component

Lemma 5.5. *Let $p = (1+\epsilon)/n$ for $\epsilon > 0$. Let $v \in V$ be a vertex. For all $3 \leq h \leq \epsilon n$, with probability at least $1/h$, v has at least $1+h$ vertices in its connected component.*

Lemma 5.5. *Let $p = (1+\epsilon)/n$ for $\epsilon > 0$. Let $v \in V$ be a vertex. For all $3 \leq h \leq \epsilon n$, with probability at least $1/h$, v has at least $1+h$ vertices in its connected component.*

Last time

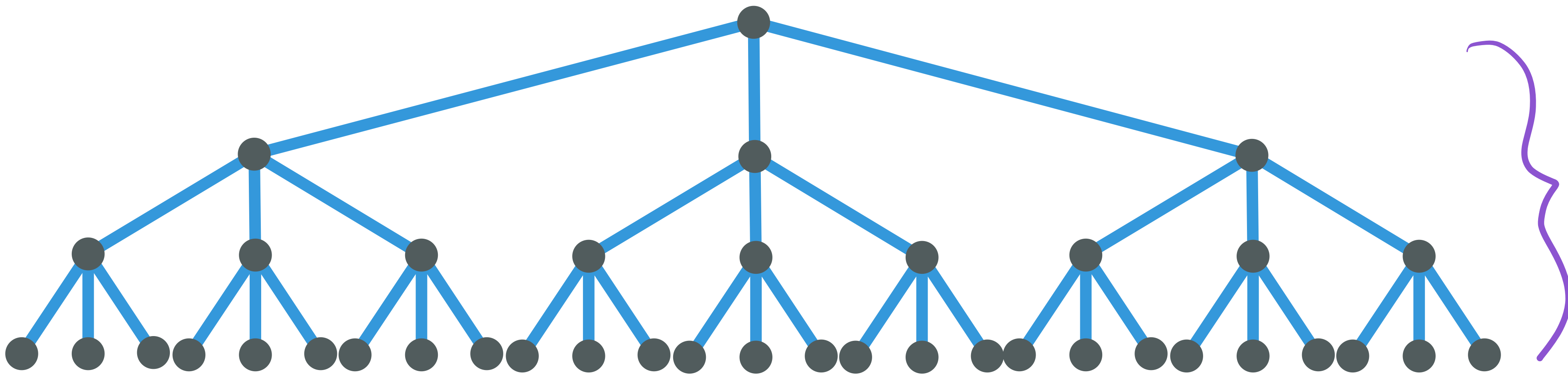


let T = complete binary tree of height h, w/ every edge deleted w/ prob $\leq 1/2$

Fact: $P[\text{some leaf connected to root}] \geq \frac{1}{h}$

"Galton-Watson process"

Lemma 5.5. *Let $p = (1+\epsilon)/n$ for $\epsilon > 0$. Let $v \in V$ be a vertex. For all $3 \le h \le \epsilon n$, with probability at least $1/h$, v has at least $1+h$ vertices in its connected component.*



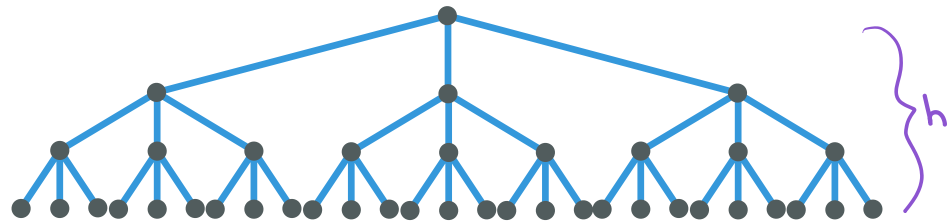
let T = complete k-ary tree of height h,

w/ every edge ~~deleted~~ kept w/ prob $\le 1/k$

Fact: $P[\text{some leaf connected to root}] \ge \frac{1}{h}$

"Galton-Watson process"

Lemma 5.5. *Let $p = (1+\epsilon)/n$ for $\epsilon > 0$. Let $v \in V$ be a vertex. For all $3 \le h \le \epsilon n$, with probability at least $1/h$, v has at least $1 + h$ vertices in its connected component.*



let T = complete k-ary tree of height h,
w/ every edge deleted w/ prob $\le 1/k$

Fact: $P[\text{some leaf connected to root}] \ge \frac{1}{h}$

"Galton-Watson process"

Equivalent process

start w/ population of size 1

Each generation, each individual flips k coins w/ prob. $1/k$ of heads

heads = # children

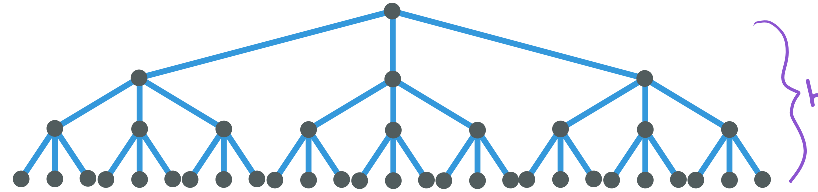
Fact: $P[\text{surviving } h \text{ generations}] \ge \frac{1}{h}$

Lemma 5.5. *Let $p = (1+\epsilon)/n$ for $\epsilon > 0$. Let $v \in V$ be a vertex. For all $3 \le h \le \epsilon n$, with probability at least $1/h$, v has at least $1 + h$ vertices in its connected component.*

Idea:

"couple"

(i.e, relate)

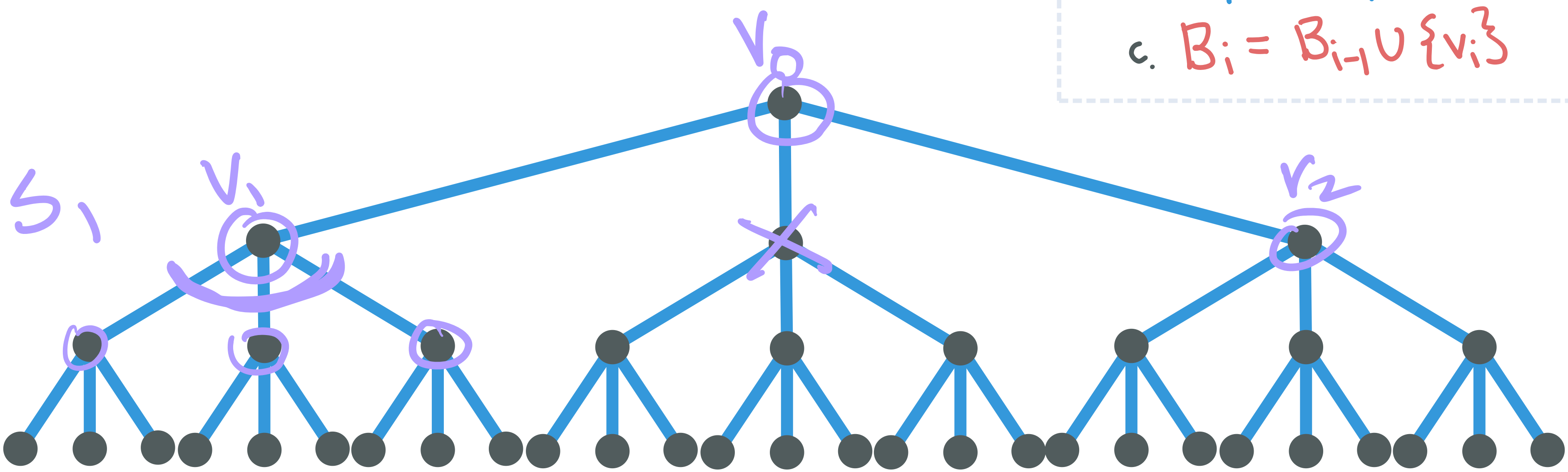


let T = complete k-ary tree of height h,
w/ every edge deleted w/ prob $\le 1/k$

Fact: $P[\text{some leaf connected to root}] \ge \frac{1}{h}$

"Galton-Watson process"

1. For $i = 1, 2, \ldots$ until $A_i = B_i$
   a. Let $v_i \in A_{i-1} \setminus B_{i-1}$
   b. $A_i = A_{i-1} \cup N(v_i)$
   c. $B_i = B_{i-1} \cup \{v_i\}$



Lemma 5.5. *Let $p = (1+\epsilon)/n$ for $\epsilon > 0$. Let $v \in V$ be a vertex. For all $3 \le h \le \epsilon n$, with probability at least $1/h$, v has at least $1+h$ vertices in its connected component.*

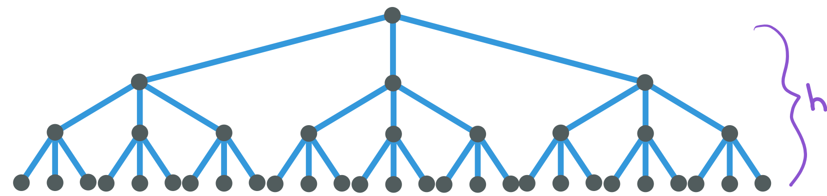
Recall search algo.

Conditional on $|A_{i-1}| \le h$,

A_i / A_{i-1} is like sampling $n - h \ge \frac{n}{1+\epsilon}$ children of some node, each w/ prob. $p = \frac{1+\epsilon}{n}$

Surviving h generations $\Rightarrow$ total population $\ge h \Rightarrow$

$|A_i| > i \quad \forall\, i < h$ (i.e., reach $|A_i| \ge h$)



let T = complete k-ary tree of height h, w/ every edge deleted w/ prob $\le 1/k$

Fact: $P[\text{some leaf connected to root}] \ge \frac{1}{h}$

"Galton-Watson process"

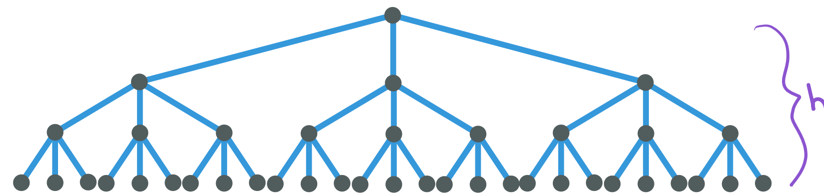
1. For $i = 1, 2, \ldots$ until $A_i = B_i$
   a. Let $v_i \in A_{i-1} \setminus B_{i-1}$
   b. $A_i = A_{i-1} \cup N(v_i)$
   c. $B_i = B_{i-1} \cup \{v_i\}$

Lemma 5.5. *Let $p = (1+\epsilon)/n$ for $\epsilon > 0$. Let $v \in V$ be a vertex. For all $3 \le h \le \epsilon n$, with probability at least $1/h$, v has at least $1+h$ vertices in its connected component.*

Recall search algo.

conditional on $|A_{i-1}| \le h$,

A_i / A_{i-1} is like sampling $n - h \ge \frac{n}{1+\epsilon}$ children of some node, each w/ prob. $p = \frac{1+\epsilon}{n}$



let T = complete k-ary tree of height h, w/ every edge deleted w/ prob $\le 1/k$

Fact: $P[\text{some leaf connected to root}] \ge \frac{1}{h}$

"Galton-Watson process"

1. For $i = 1, 2, \ldots$ until $A_i = B_i$
   a. Let $v_i \in A_{i-1} \setminus B_{i-1}$
   b. $A_i = A_{i-1} \cup N(v_i)$
   c. $B_i = B_{i-1} \cup \{v_i\}$

$$P\left[|A_i| > i \ \forall\, i < h\right] \ge P\left[\begin{matrix}\text{surviving } h \\ \text{generations}\end{matrix}\right] \ge \frac{1}{h}$$

Theorem 5.3. *Consider a random graph $G \sim \mathcal{G}(n,p)$ for $p = c/n$, where c is a constant.*

1. *If $c > 1$, then with high probability, there is exactly one connected component of G with $\Omega(n)$ vertices, and all other components have size $\leq O(\log n)$.*
2. *For $c < 1$, then with high probability, all connected components of G has size $< O(\log n)$.*

Proof outline ($c>1$)

① Gap theorem: let $p=(1+\varepsilon)/n$. w/ high probability, all components have size

$$\leq O\left(\frac{\log n}{\varepsilon^2}\right) \quad \text{or} \quad \geq \Omega(\varepsilon n)$$

② Existence of large component

③ Uniqueness of large component

Imagine sampling $G \sim (n,p)$ as:

① sample $G_1 \sim (n/2, p)$

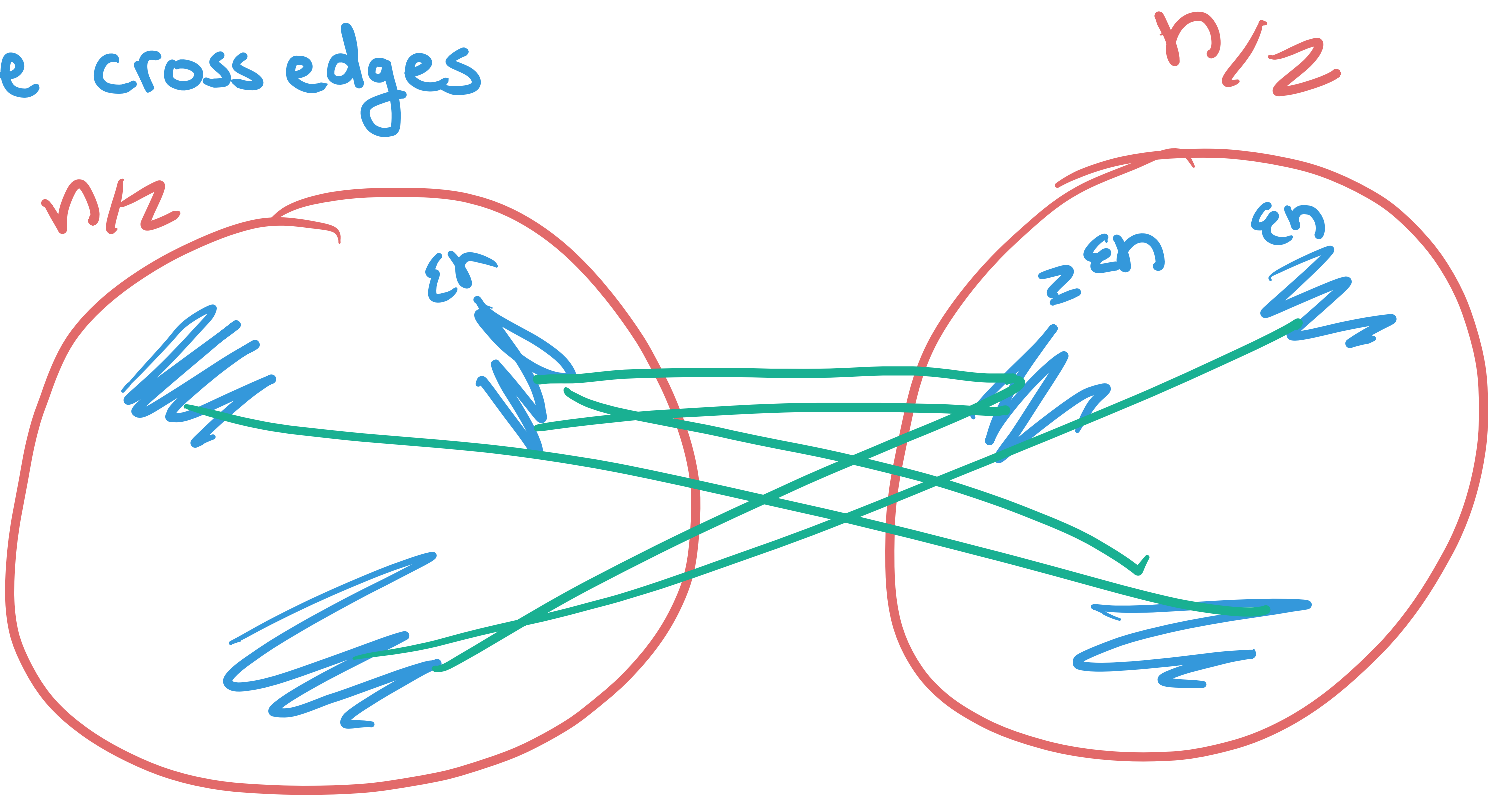
② sample $G_2 \sim (n/2, p)$

③ combine G_1 and G_2

④ sample cross edges

$\varepsilon^2 n^2$ possible

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



Multiplicative Chernoff bound

let $X_1, \cdots, X_n \in [0,1]$ be independent

$\mu = E[X_1 + \cdots + X_n]$

$\varepsilon \in (0,1)$

- $P[X_1 + \cdots + X_n \geq (1+\varepsilon)\mu] \leq e^{-\varepsilon^2 \mu / 3}$
- $P[X_1 + \cdots + X_n \leq (1-\varepsilon)\mu] \leq e^{-\varepsilon^2 \mu / 2}$

(proof at end if time permits)