PROOF COMPLEXITY AS A COMPUTATIONAL LENS

LECTURE 5

Recap

Clique-colouring formula hard for resolution

"There exists an n-vertex graph G such that
(a) G has m-clique
(b) G is $(m-1)$-colourable"

Set $m = n^{\delta}$

Proof technique: Given resolution refutation, use FEASIBLE (MONOTONE) INTERPOLATION to extract monotone circuit C that given graph G determines whether (a) or (b) applies

Circuit size = Θ(resolution proof length)

But can prove there are no small monotone circuits for this problem

Hence no short resolution refutations either

Question

What if we fix graph G?

i.e., consider fixed assignment $\rho : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ to edge indicator variables?

GRAPH COLOURING

Strong lower bounds for resolution and polynomial calculus – will see this later in the course

CLIQUE

Much less well understood

CLIQUE PROBLEM

Given undirected n-vertex graph G, does G contain a k-clique?
(i.e., k vertices $v_1, v_2, \ldots, v_k \in V(G)$ such that $(v_i, v_j) \in E(G)$ for all $1 \le i < j \le k$)

Trivial algorithm in time $\binom{n}{k} n^{O(1)}$
Check all subsets of vertices of size k
Can do slightly better with algebraic techniques: $O(n^{\omega k/3})$ where $\omega \le 2.373$ is matrix multiplication exponent
But $n^{O(k)}$ upper bound best known

Believed to be hard
- even to meaningfully approximate size of largest clique
- even on average for random graphs

$n^{\Omega(k)}$ believed to be correct lower bound
True if SAT problem requires exponential time (Exponential Time Hypothesis or ETH)

$n^{\Omega(k)} \iff \exists \delta > 0$ such that running time is $\Omega(n^{\delta k})$ ← independent of k

But in practice there are very good algorithms for computing max cliques

What about unconditional results?

In circuit complexity, $n^{\Omega(k)}$ lower bounds for

- bounded-depth circuits
- monotone circuits

Can we prove lower bounds for conflict-driven clause learning (CDCL) SAT solvers?

For state-of-the-art max clique algorithms?

What we will talk about today and next time

THEOREM (very informal)
For $k \ll \sqrt{n}$ it holds for suitably sampled random n-vertex graphs G that
(a) G does not have k-clique a.a.s.
(b) REGULAR resolution requires length $n^{\Omega(k)}$ to prove this a.a.s.

Recall that a sequence of events $\{E_n\}_{n=1}^{\infty}$ happens ASYMPTOTICALLY ALMOST SURELY if

$$\lim_{n\to\infty} \Pr[E_n] = 1.$$

COROLLARY
~~Essentially all~~ Well, at least Many state-of-the-art max clique algorithms require time $n^{\Omega(k)}$

Corollary follows since reasoning in many max clique algorithms is captured by regular resolution

(This requires an argument that we will ignore. Also, there is a 2^k-loss in reasoning power, but this is dominated by the $n^{\Omega(k)}$ lower bound)

~~~~~~~~~~

But we still have no good bounds for general resolution and CDCL

~~~~~~~~~~

Need to make precise what we mean by

(i) "random" graph

(ii) "regular" resolution

(iii) "prove" – which formula is refuted?

~~~~~~~~~~

(i) Erdős-Rényi random graphs $G(n,p)$

n-vertex undirected graph; think of $V=[n]$

For every potential edge (i,j), $i<j$, flip independent coin and include edge with probability p

For $X \subseteq V$, $|X|=k$,

$$\Pr[X \text{ forms clique}] = p^{\binom{k}{2}}$$
~~~~~~~~~~

Let N_k = # k-cliques in random graph

$$\mathbb{E}[N_k] = \binom{n}{k} p^{\binom{k}{2}} \qquad (*)$$

by linearity of expectation

For $p = n^{-\frac{2}{k-1}}$ get $\mathbb{E}[N_k] \approx 1$

Markov's inequality says that

$$\Pr[N_k \geq 1] \leq \mathbb{E}[N_k]$$

Fix constant $\eta > 1$ and set

$$p = n^{-\frac{2\eta}{k-1}} \qquad (**)$$

Then

$$\Pr[N_k \geq 1] \leq \mathbb{E}[N_k] \to 0 \quad \text{as } n \to \infty$$

$G \sim G(n,p)$ very unlikely to contain k-clique (but will contain many cliques of size $\Omega(k)$)

This is the random graph distribution that we will study.

(ii) Regular resolution

General resolution: Proof DAG without restrictions

Tree-like resolution: Proof DAG is binary tree

Equivalent to decision tree for falsified clause search problem:

- Internal nodes query variables
- Edges labelled by values to queried variables
- Each leaf labelled by clause C falsified by total value assignment defined by path to leaf

Regular resolution: DAG-like proof
But on each path every variable resolved over at most once
(once variable eliminated, not reintroduced again)

Regular resolution is exponentially stronger than tree-like resolution
(E.g., for XORified pebbling formulas, which we will see later in the course)

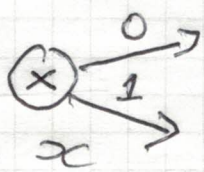
Regular seems "morally close" to general resolution
(Why re-introduce variables that have been eliminated?!)
Has been conjectured to be as strong as general resolution for some classes of formulas

But regular resolution is exponentially weaker than general resolution (but separating formulas are hard to find and are quite contrived)

Regular resolution = Read-once branching program (ROBP)

Read-once branching program (ROBP) for CNF formula F

Directed acyclic graph (DAG)
One source node (no incoming edges)
= start node
Every sink node (no outgoing edges) labelled by clause $C \in F$
Size = # nodes

Every non-sink node 

- labelled by variable x
- 2 outgoing edges labelled 0 and 1

"Node queries x and takes edge with label agreeing with value of x"

Any assignment α to $Vars(F)$ defines source-to-sink path

Branching program solves (falsified clause search problem for) F if $\forall \alpha$ reach sink labelled by C such that $\alpha(C) = 0$.

Read-once Along any path, query any variable at most once

General resolution proof $\Rightarrow$ general branching program
- Make $\bot$ start node
- flip direction of all edges
- Let resolvent node query variable (x) resolved over
- Label edge to premise by value $\{0,1\}$ falsifying literal over x in premise clause

So lower bounds for branching program size $\Rightarrow$ lower bounds for resolution length

Unfortunately, general branching programs much more powerful than resolution...

Regular resolution proof ⇒ Read-once branching program

By definition of regularity, every variable queried at most once

In fact, regular resolution ⇔ ROBP, but we only need the easy direction (The harder direction tells us that this is a promising approach, though)

(iii) Clique formula

Basically clique colouring formula but with restriction to edge indicator variables But with some constraints added to avoid formula being hard for "stupid reasons"

Variables $x_{v,i}$ = "v is ith member of clique"

$v \in V = \{1, 2, \ldots, n\} = [n]$

$i \in \{1, 2, \ldots, k\} = [k]$

CLIQUE AXIOMS	$\bigvee_{v \in V} x_{v,i}$	$i \in [k]$
EDGE AXIOMS	$\bar{x}_{u,i} \vee \bar{x}_{v,j}$	$i \neq j \in [k]$ $(u,v) \notin E$
FUNCTIONALITY AXIOMS	$\bar{x}_{u,i} \vee \bar{x}_{v,i}$	$i \in [k]$ $u \neq v \in V$
ORDERING AXIOMS	$\bar{x}_{u,i} \vee \bar{x}_{v,j}$	$u < v \in V$ $(u,v) \in E$ $i > j \in [k]$

"if u & v are both in clique and $u < v$, then clique index of u < clique index of v"

THEOREM [Atserias, Bonacina, de Rezende, Lauria, Nordström, Razborov]

For any constant $\varepsilon > 0$, $\eta > 1$, if $G \sim G(n, n^{-2\eta/(k-1)})$ for $k \leq n^{1/4-\varepsilon}$ then asymptotically almost surely G does not contain a k-clique but any regular resolution refutation of the k-clique formula for G has length $n^{\Omega(k)}$.

Ordering axioms are important — but we will cheat a bit and ignore them for simplicity of exposition

All key ideas still there, and even without ordering axioms the lower bound is highly nontrivial and interesting

Will do slightly simpler proof from conference paper [ABdRLNR '18]. Full details in journal version [ABdRLNR '21].

PREVIOUSLY KNOWN

TREE-LIKE RESOLUTION

Upper bound $n^{O(k)}$

Proof sketch: Build decision tree that tests all $\binom{n}{k}$ clique candidates (and for each candidate finds missing edge)

Lower bound $n^{\Omega(k)}$ for Erdős-Rényi random graphs [Beyersdorff, Galesi, Lauria '13]

Prosecutor-Defendant game w/o forgetting (= decision tree)

Defendant commits to k'-clique ($\binom{K}{k'}$) for $k' = \Omega(k)$

When asked about $x_{v,i}$, $v \in V$, say no/0 if $v \notin K$. Say yes/1 if $v \in K$ and not already answered 1 for $x_{v,j}$, $j \neq i$

When Prosecutor wins, has identified clique $n^{\Omega(k)}$ different cliques $\Rightarrow$ lower bound

Same lower-bound argument works for balanced, complete $(k-1)$-partite graphs

GENERAL RESOLUTION

[Beame, Impagliazzo, Sabharwal '07]

Lower bound $\exp(\Omega(n^{\varepsilon}))$, $\varepsilon > 0$, for $n^{5/6} \ll k \leq n/3$ for very dense Erdős-Rényi graphs

ε depends on density. For $k = n/3$, get $\exp(\Omega(n))$

But not right dependence on k

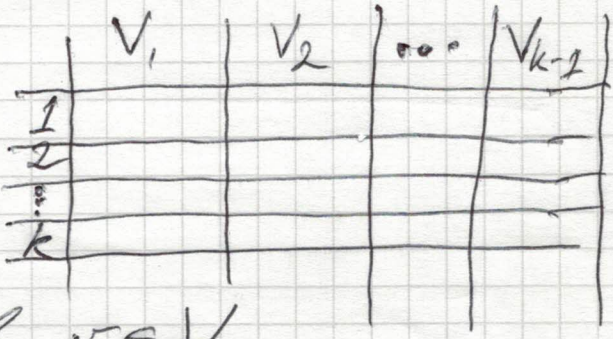
REGULAR RESOLUTION

CX

Complete $(k-1)$-partite graphs
(and any $(k-1)$-colourable graphs)
are EASY – regular resolution refutation
in length $\leq \boxed{2^k k^2 n^2}$ [BGL13]

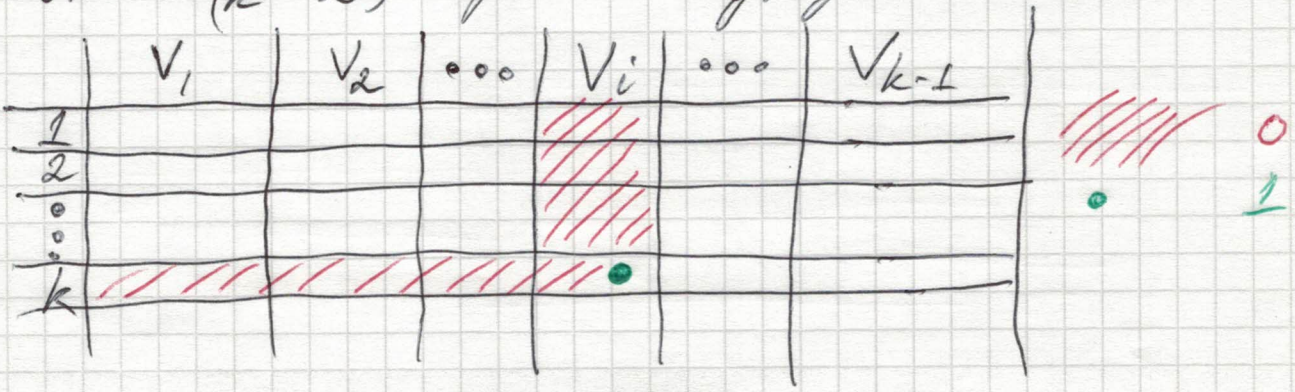
$$V = V_1 \cup V_2 \cup \cdots \cup V_{k-1}$$

$$E = \bigcup_{1 \leq i < j \leq k-1} V_i \times V_j$$



Prosecutor strategy:

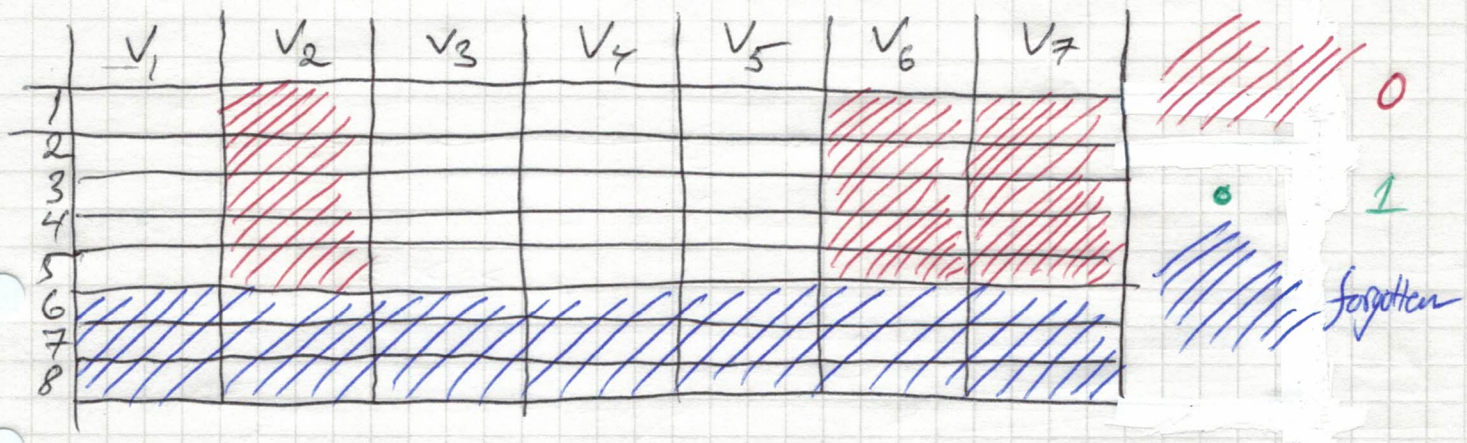
Ask $x_{v,k}$ for all $v \in V$
All answers $= 0 \Rightarrow$ win (clique axiom)
Suppose $x_{v,k} = 1$ for $v \in V_i$
Ask $x_{u,\ell}$ for all $u \in V_i$, $\ell \leq k-1$
All answers 0, since otherwise violate edge axiom
Forget all variables $x_{v,k}$
Now we have $(k-1)$-clique problem
on $(k-2)$-partite graph



r-th round:

Looking for $(k+1-r)$-clique in $(k-r)$ possible graph

For purposes of illustration, fix $k=8$, $r=4$



$(r-1)$ blocks V_i already ruled out

$\binom{k}{r-1}$ possible starting records

(1) Ask $x_{v,k+1-r}$ for all (remaining) $v \in V$

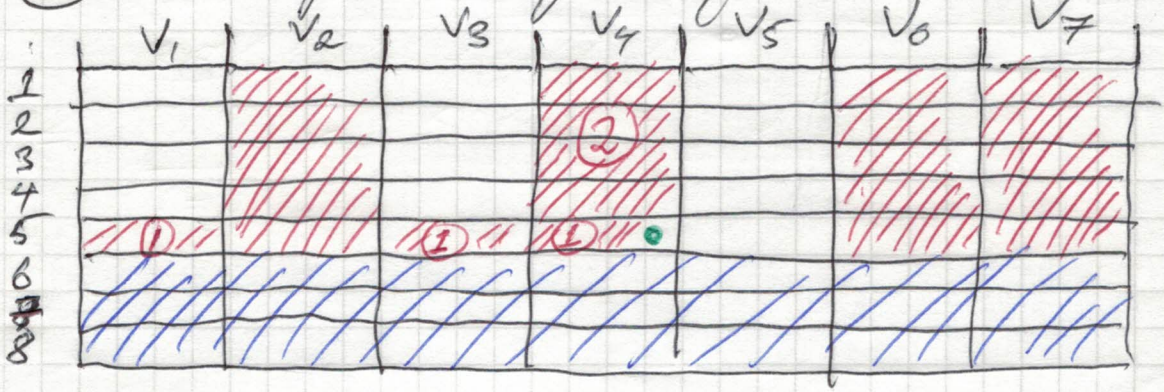
$\leq n$ queries

until find a 1 in some V_i.

(2) Ask for all $u \in V_i$ and all $\ell \leq k-r$ about $x_{u,\ell}$

$\leq k \cdot n$ queries

(3) Forget everything about variables $x_{v,k+1,r}$



Total # records something like

$\leq \sum_{r=0}^{k} \binom{k}{r} k \cdot n^2 = 2^k \cdot k \cdot n^2$ (missing factor k?)

Note that we only use clique and edge axioms — no functionality or ordering axioms

Balanced, complete $(k-1)$-partite graphs (i.e., all V_i satisfy $|V_i| = \lceil \frac{V}{k-1} \rceil$ or $|V_i| = \lfloor \frac{|V|}{k-1} \rfloor$) good testbed for lower bound ideas.

If a promising lower bound idea identifies a property that holds for $(k-2)$-partite graphs, then it is no longer so promising...

Proof idea used for PHP and Tseitin

BOTTLENECK COUNTING

(1) Define random walk R_α from $\perp$ in proof DAG

(2) Show that all random walks R_α have to pass through special "bottleneck node" w (informative record)

(3) Prove that for any fixed bottleneck node w

$$\Pr[R_\alpha \text{ passes through } w] \leq \text{small}$$

(4) But

$$1 = \Pr[R_\alpha \text{ passes through } \underline{\text{some}}\ w]$$

$$\leq \sum_{\substack{\text{bottle-} \\ \text{necks} \\ w}} \Pr[R_\alpha \text{ passes through fixed } w]$$

$$\leq (\#\ \text{bottlenecks}) \cdot \text{small}$$

so $(\#\ \text{bottlenecks}) = \text{large}$

Twists:

- Typically, small $= \exp(-n^{\delta})$ but we can't get better than $n^{-\Omega(k)}$
- We will need pair of bottleneck nodes

Phrase proof in terms of ROBPs

ROBP path $\Leftrightarrow$ partial truth value assignment
(because of read-once property)

Branching program: NODES a, b, c
Graph $G = (V, E)$ VERTICES u, v, w

$\beta(a)$ = maximal partial assignment contained in every path from source to a
= all variable queries & answers obtained along every path to a
"ASSIGNMENTS REMEMBERED" at a

Sourcepath = path from source to some node
Node a queries x: $x(a) = x$
If on path α x is queried and edge $(x) \xrightarrow{b}$ taken, say "α SETS x TO b"

If $\exists$ sourcepath α to a such that
- α sets x
- $\beta(a)$ doesn't contain x

then we say "x IS FORGOTTEN AT a"

OBSERVATION 2
If x forgotten at a and $\exists$ path $a \rightsquigarrow b$, then x forgotten at b

Proof Use read-once property
$\exists \alpha$: source $\rightsquigarrow a$ not setting x
$\exists \alpha'$: source $\rightsquigarrow a$ setting x
If $\exists \gamma : a \rightsquigarrow b$ queries x, then composed path $\alpha' \circ \gamma$ violates read-once property ☐

Focus now on clique formula

If $x_{v,i}$ forgotten at a for some $v \in V$, say "INDEX i FORGOTTEN AT a"

If source path α ends at clique axiom $\bigvee_{v \in V} x_{v,i}$ (which α thus falsifies), say "PATH α RULES OUT INDEX i"

OBSERVATION 2

If source path α rules out index i at sink b, then i is not forgotten at b.

Proof All source paths α ending at b must have queried all $x_{v,i}$, $v \in V$ and have received answers 0, so no $x_{v,i}$ can be forgotten ☐

Fix from now on arbitrary ROBP Π solving k-clique formula for $G = (V, E)$ without k-clique (and omitting ordering axioms). Want to show Π has size $n^{\Omega(k)}$ a.a.s. if $G \sim G\left(n, n^{-2/(k-1)}\right)$

Define RANDOM DISTRIBUTION 2 OVER PATHS α by following process, starting with (current node a) := source

(a) If $\beta(a) \cup \{X(a) \mapsto 1\}$ falsifies edge or functionality axiom, take 0-edge from a

(b) If $X(a) = x_{v,i}$ and i forgotten at a, take 0-edge from a — set $x_{v,i} = 0$

(c) Otherwise, flip $n^{-\delta}$-biased coin, $\delta > 0$

$X(a) = x_{v,i} \mapsto 1$ with probability $n^{-\delta}$

$x_{v,i} \mapsto 0$ otherwise

(a) + (b) "FORCED CHOICES" always setting $x_{v,i} = 0$

(c) "FREE CHOICE / COIN FLIP" but still highly likely to set $x_{v,i} = 0$

OBSERVATION 3

(i) Any $\alpha \sim \mathcal{D}$ ends by ruling out some index i (never falsifies functionality or edge axiom)

(ii) Any $\alpha \sim \mathcal{D}$ sets at most k variables $x_{v,i}$ to 1.

Proof Straightforward exercise using definition of random path distribution $\mathcal{D}$.

FIRST PROOF IDEA (WON'T WORK)

Given any $\alpha \sim \mathcal{D}$, define WAYPOINTS (for $t \in \mathbb{N}^+$ fixed)

$a_0 :=$ source

$a_{i+1} :=$ first time after a_i we have $\lceil k/t \rceil$ additional assignments $x_{v,i} \mapsto 1$ on α or sink of path, whichever comes first

Suppose α rules out i

Let $V_i = \{ v \mid x_{v,i} \mapsto 0$ between a_{i-1} and $a_i \}$

$V = V_1 \cup V_2 \cup \dots \cup V_{t'}$ for $t' \leq t$

Must exist j such that $|V_j| \geq n/t$

i.e., between a_{j-1} and a_j we have $\geq n/t$ assignments $x_{v,i} = 0$

ROBP Π makes lots of progress in ruling out vertices on subpath (a_{j-1}, a_j)

Call such a pair (a_{j-1}, a_j) USEFUL*

We just proved:

OBSERVATION 4

Any path $\alpha \sim \mathcal{D}$ passes through a useful* pair (a, b) with probability 1.

Now we want to fix any pair (a,b) of nodes in Π and argue that

$$\Pr_{\alpha \sim \mathcal{D}}\left[(a,b) \text{ useful* for } \alpha\right] \leq n^{-\Omega(k)}$$

If we can do this, then we're done, since

$$1 = \Pr_{\alpha \sim \mathcal{D}}\left[\exists (a,b) \in \Pi^2 \text{ useful* for } \alpha\right]$$

$$\leq \sum_{(a,b)\in\Pi^2} \Pr_{\alpha}\left[\text{fixed } (a,b) \text{ useful* for } \alpha\right]$$

$$\leq |\Pi|^2 \cdot n^{-\Omega(k)}$$

so $|\Pi| = n^{\Omega(k)}$

Little bit of notation overload

$\alpha \sim \mathcal{D}$:
α sampled according to $\mathcal{D}$
α is path in support of $\mathcal{D}$

OBSERVATION 5 (CRUCIAL)

Suppose $\exists$ path $\alpha \sim \mathcal{D}$ such that

(i) α passes through a and b

(ii) α assigns $x_{v,i} = 0$ for $v \in V_j$

(iii) α rules out index i

Then any path α' through a and b has to set $x_{v,i} = 0$ for $v \in V_j$

Proof Exercise. (Use read-once property)

HOPEFUL CLAIM 6 (BROKEN)

For any $(a,b) \in \Pi^2$ it holds that

$$\Pr_{\alpha \sim D}[(a,b) \text{ useful}^* \text{ for } \alpha] \leq n^{-\Omega(k)}$$

Attempt at proof Case analysis

(1) If (a,b) not useful for any path
$\Rightarrow$ probability 0

(2) Suppose $\beta(a)$ contains $\Omega(k)$ assignments $x_{v,j} \mapsto 1$. Note that 1-assignments are never forced, but are free (and unlikely) coin flips. Then

$$\Pr_{\alpha \sim D}[\alpha \text{ passes through } a] \leq n^{-\delta k}$$

(By read-once property, only have one chance for every $x_{v,j} \mapsto 1$ in $\beta(a)$.)

(3) Relatively few 1s in $\beta(a)$: Then Π knows almost nothing about the candidate clique found up until this point

Yet between a and b makes $\geq n/t$ assignments $x_{v,i} \mapsto 0$. Some of these might be forced, to be sure. But a polynomial fraction should be free coin flips, say # free choices n^{δ} for $\gamma \gg \delta$

$$\Pr_{\alpha}[\text{all these coin flips yield } 0] = (1 - n^{-\delta})^{n^{\gamma}}$$
$$= \exp(-n^{\gamma-\delta}), \text{ so great } \textcircled{=}$$

CLB VI
END

Except case 3 is broken.
We could have essentially all choices forced
Sanity check:
- Did we use so far that G_1 is not $(k-1)$-partite?
- Did we use any nontrivial property of G_1?

Can make this approach work for Erdős–Rényi random graphs and regular resolution / ROBPs
But need
- to use nontrivial properties of random graphs
- to have better notion of "useful pair" (without the asterisk)

Will have to wait until next lecture.