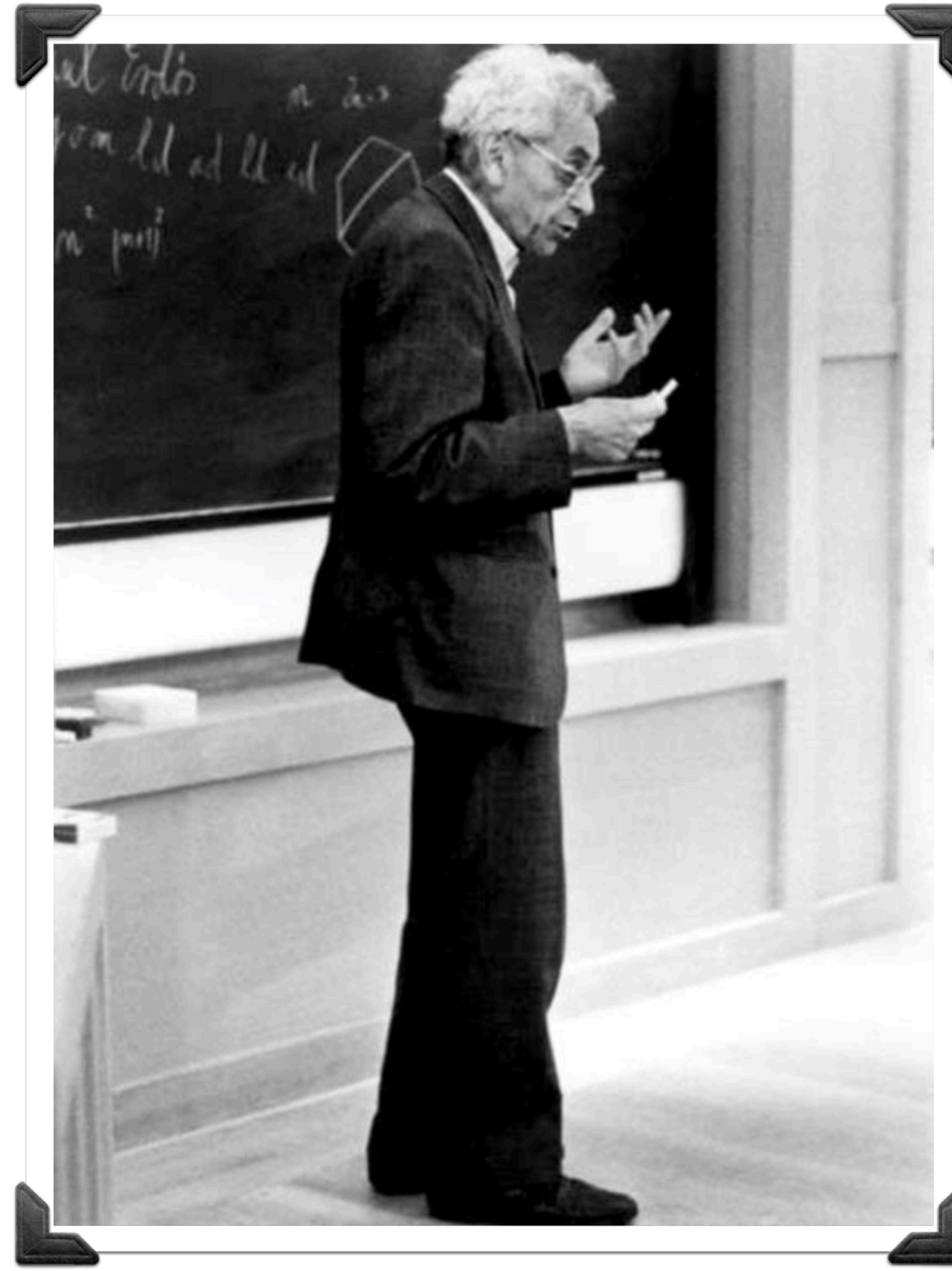
Combinatorics

The Probabilistic Method

尹一通 Nanjing University, 2023 Spring

The Probabilistic Method

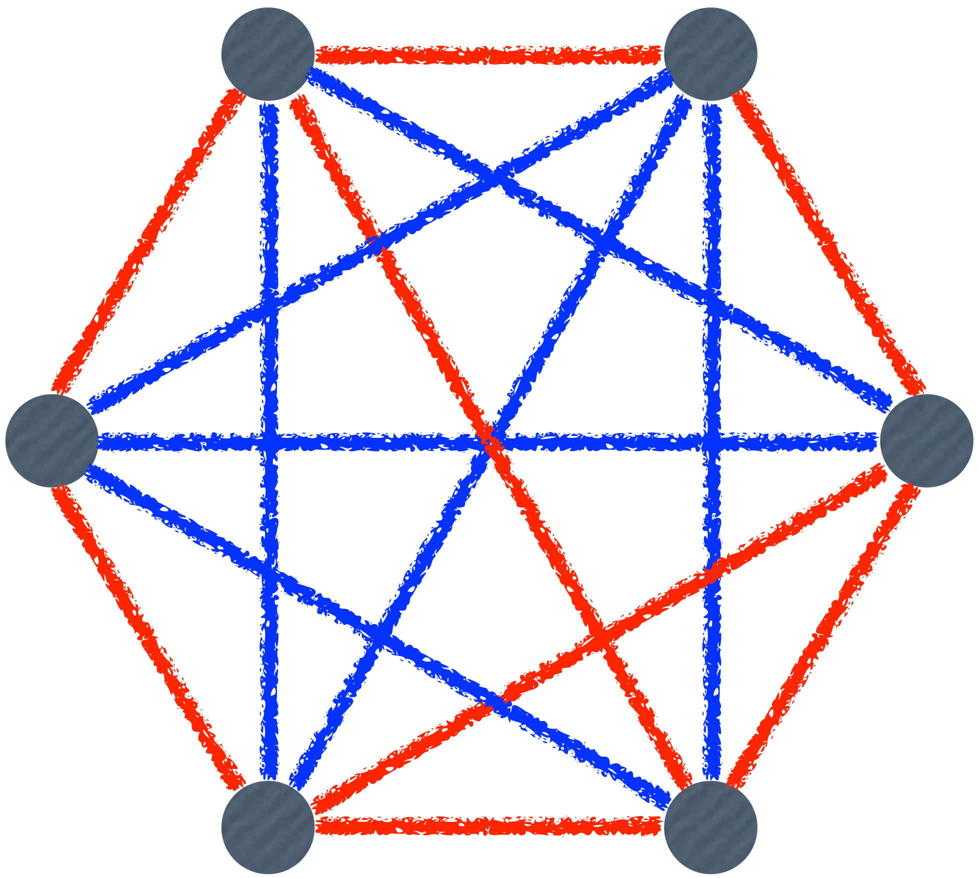


Paul Erdős
(1913-1996)

Ramsey Number

"In any party of six people, either at least three of them are mutual strangers or at least three of them are mutual acquaintances"

- For any edge-2-coloring of K_6, there is a ***monochromatic*** K_3.



Ramsey Theorem

If $n \geq R(k, k)$, for any edge-2-coloring of K_n, there is a monochromatic K_k.

Ramsey number: $R(k, k)$

Theorem (Erdős 1947)

If $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$ then it is possible to color the edges of K_n with 2 colors so that there is no monochromatic K_k subgraph.

Each edge $e \in K_n$ is colored
- with prob 1/2
- with prob 1/2

For any K_k subgraph:

$$\Pr[\text{the } K_k \text{ is monochromatic}] = \Pr[{\color{red}K_k} \text{ or } {\color{blue}K_k}]$$
$$= 2^{1-\binom{k}{2}}$$

Theorem (Erdős 1947)

If $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$ then it is possible to color the edges of K_n with 2 colors so that there is no monochromatic K_k subgraph.

Each edge $e \in K_n$ is colored $\begin{cases} \text{with prob } 1/2 \\ \text{with prob } 1/2 \end{cases}$

$$\Pr[\exists K_k \text{ is monochromatic}] \leq \binom{n}{k} 2^{1-\binom{k}{2}} < 1$$

$$\Longrightarrow \Pr[\text{no } K_k \text{ is monochromatic}] > 0$$

$\Longrightarrow$ $\exists$ a 2-coloring of edges of K_n without monochromatic K_k

Tournament

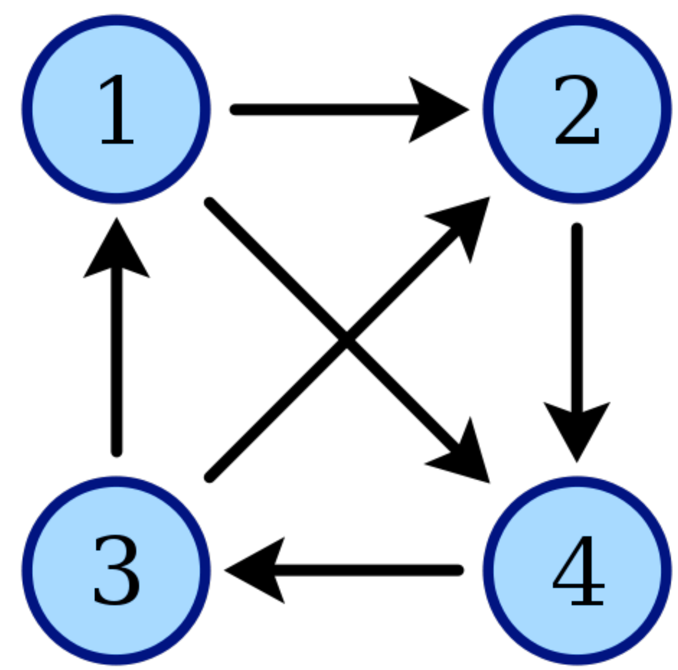
$T(V, E)$

n players, each pair has a match.

$u \to v$ iff u beats v.

k-paradoxical:

For every k-subset S of V, there is a player in $V \backslash S$ who beats all players in S.



"*Does there exist a k-paradoxical tournament for every finite k?*"

Theorem (Erdős 1963)

If $\binom{n}{k}\left(1-2^{-k}\right)^{n-k} < 1$ then there is a k-paradoxical tournament of n players.

Pick a random tournament T on n players $[n]$.

Fixed any $S \in \binom{[n]}{k}$

Event A_S: no player in $V \setminus S$ beat all players in S.

$$\Pr[A_S] = \left(1-2^{-k}\right)^{n-k}$$

Theorem (Erdős 1963)
If $\binom{n}{k}\left(1-2^{-k}\right)^{n-k} < 1$ then there is a k-paradoxical tournament of n players.

Pick a random tournament T on n players $[n]$.

Event A_S: no player in $V \setminus S$ beat all players in S.

$$\forall S \in \binom{[n]}{k}: \quad \Pr[A_S] = \left(1-2^{-k}\right)^{n-k}$$

$$\Pr\left[\bigvee_{S\in\binom{[n]}{k}} A_S\right] \le \sum_{S\in\binom{[n]}{k}} (1-2^{-k})^{n-k} < 1$$

Theorem (Erdős 1963)

If $\binom{n}{k}\left(1-2^{-k}\right)^{n-k} < 1$ then there is a k-paradoxical tournament of n players.

Pick a random tournament T on n players $[n]$.

Event A_S: no player in $V\backslash S$ beat all players in S.

$$\Pr\left[\bigvee_{S\in\binom{[n]}{k}} A_S\right] < 1$$

$$\Pr[T \text{ is } k\text{-paradoxical}] = 1 - \Pr\left[\bigvee_{S\in\binom{[n]}{k}} A_S\right] > 0$$

Theorem (Erdős 1963)

If $\binom{n}{k}\left(1-2^{-k}\right)^{n-k} < 1$ then there is a k-paradoxical tournament of n players.

Pick a random tournament T on n players $[n]$.

$$\Pr[T \text{ is } k\text{-paradoxical}] \ > 0$$

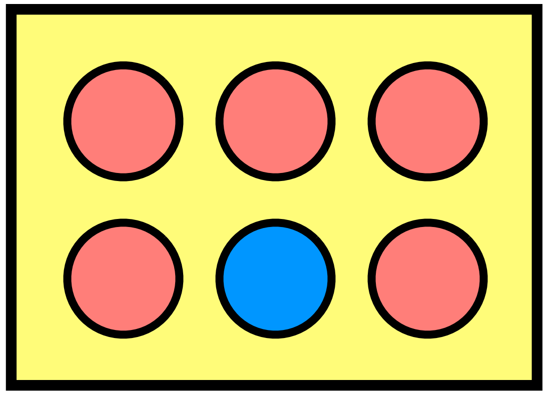
There is a k-paradoxical tournament on n players.

The Probabilistic Method

- Pick random ball from a box,

  Pr[the ball is blue]>0.

  $\Rightarrow$ There is a blue ball.



- Define a probability space Ω, and a property P:

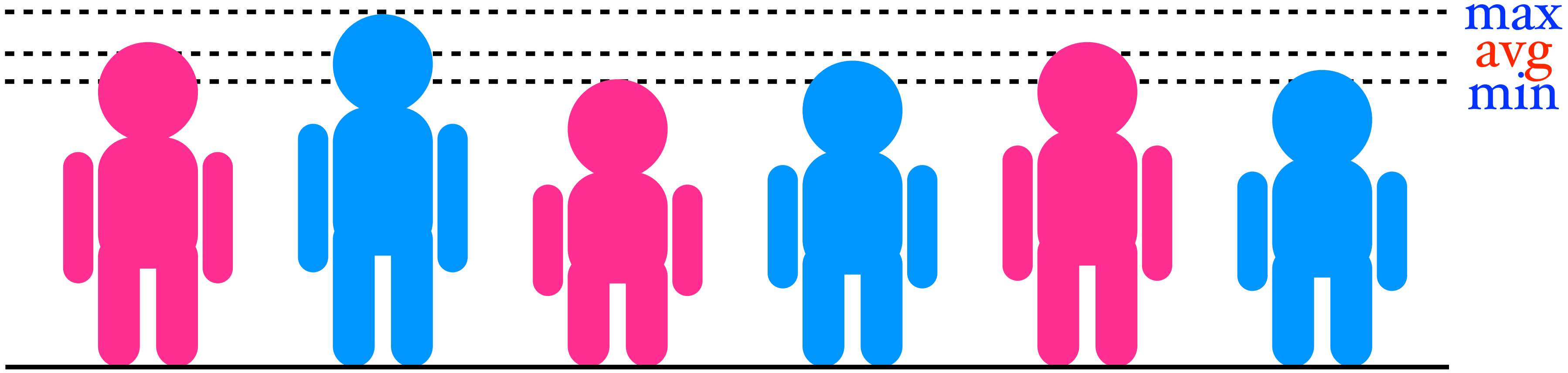
$$\Pr_x[P(x)] > 0$$

$\Longrightarrow \exists$ a sample $x \in \Omega$ with property P.

Averaging Principle

- Average height of the students in class is l.

  $\Rightarrow$ There is a student of height $\geq l$ ($\leq l$)

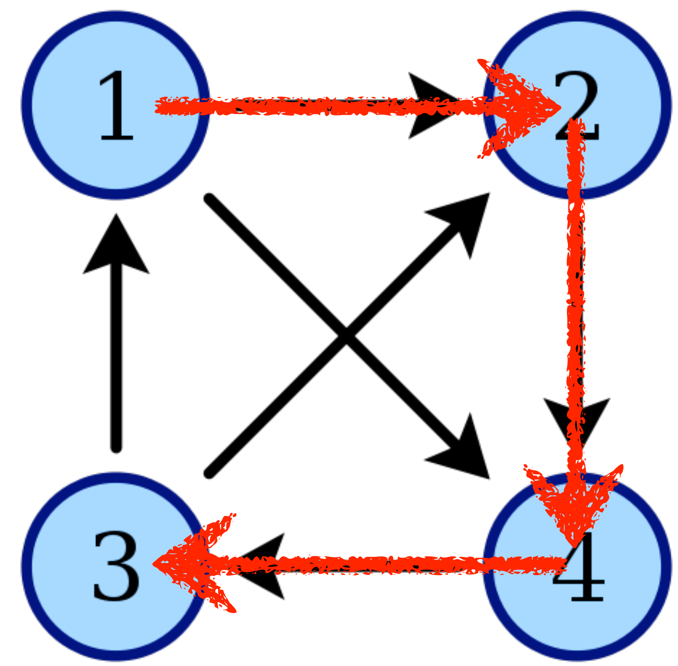


- For a random variable X,
  - $\exists\, x \leq \mathrm{E}[X]$, such that $X = x$ is possible;
  - $\exists\, x \geq \mathrm{E}[X]$, such that $X = x$ is possible.

Hamiltonian Paths in Tournament

Hamiltonian path:

a path visiting every vertex *exactly* once.



Theorem (Szele 1943)

There is a tournament on n players with at least $n!2^{-(n-1)}$ Hamiltonian paths.

Theorem (Szele 1943)

There is a tournament on n players with at least $n!2^{-(n-1)}$ Hamiltonian paths.

Pick a random tournament T on n players $[n]$.

For every permutation π of $[n]$,

$$X_\pi = \begin{cases} 1 & \pi \text{ is a Hamiltonian path} \\ 0 & \pi \text{ is } \textit{not} \text{ a Hamiltonian path} \end{cases}$$

Hamiltonian paths: $\quad X = \sum_{\pi} X_\pi$

$$\mathrm{E}[X_\pi] = \Pr[X_\pi = 1] = 2^{-(n-1)}$$

Theorem (Szele 1943)

There is a tournament on n players with at least $n!2^{-(n-1)}$ Hamiltonian paths.

Pick a random tournament T on n players $[n]$.

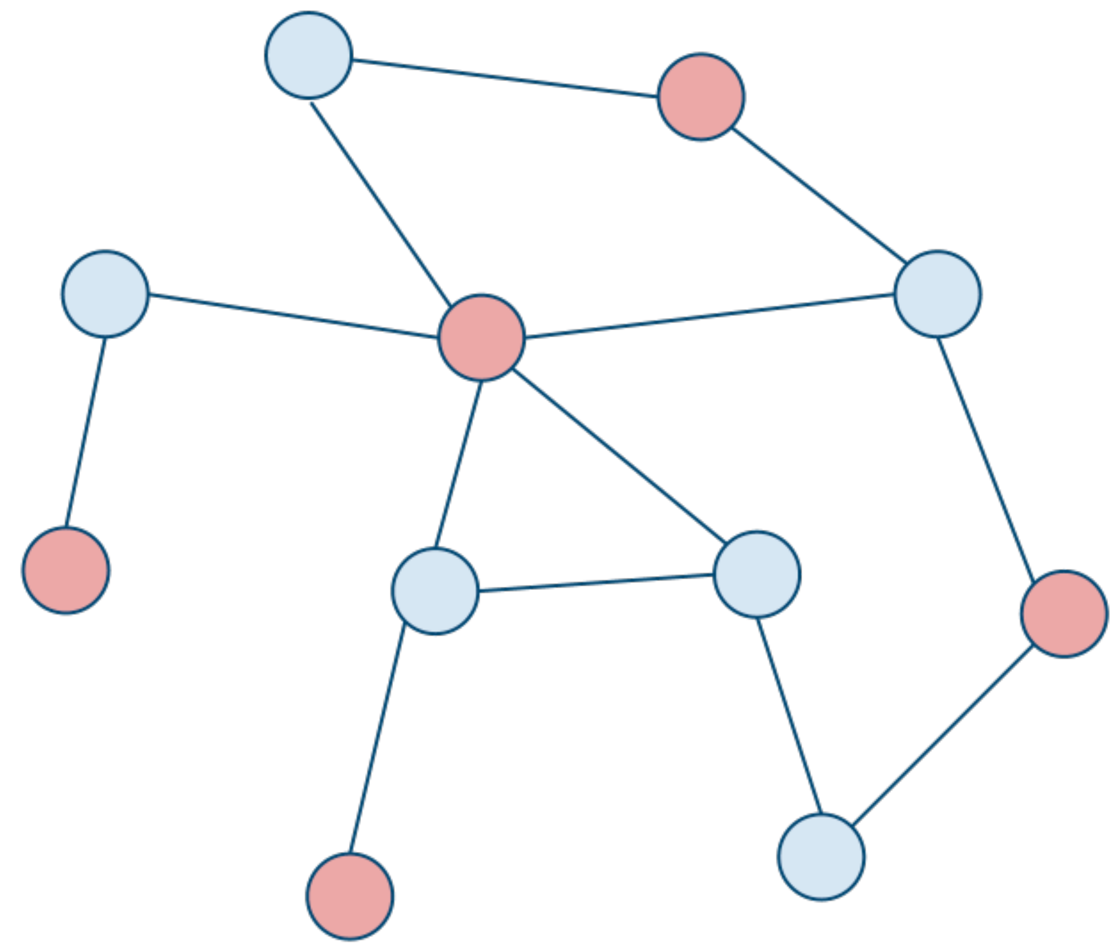
\# Hamiltonian paths: $X = \sum_{\pi} X_\pi$

$$\mathrm{E}[X_\pi] = \Pr[X_\pi = 1] = 2^{-(n-1)}$$

$$\mathrm{E}[X] = \sum_{\pi} \mathrm{E}[X_\pi] = n!2^{-(n-1)}$$

Large Independent Set

- Graph $G(V, E)$
- independent set $S \subseteq V$
  - no adjacent vertices in S
- max independent set is **NP**-hard



Theorem: G has n vertices and m edges

$\Rightarrow$ $\exists$ an independent set S of size $\frac{n^2}{4m}$

- Draw a random independent set $S \subseteq V$ (How?)
  - each $v \in V$ is selected into a random set R independently with probability p (to be fixed later)
  - for every $uv \in E$: delete one of u, v from R if $u, v \in R$
  - the resulting set is an independent set S
- Show that $\mathbf{E}[\,|S|\,] \geq \dfrac{n^2}{4m}$

$G(V,E)$: n vertices, m edges

1. sample a random R : each vertex is chosen *independently* with probability p

2. modify R to S : independent set!

$\forall uv \in E$ if $u, v \in R$

delete one of u, v from R

Y: # of edges in R $Y = \sum_{uv \in E} Y_{uv}$ $Y_{uv} = \begin{cases} 1 & u, v \in S \\ 0 & \text{o.w.} \end{cases}$

$$\mathbf{E}[\,|S|\,] \geq \mathbf{E}[\,|R| - Y] = \mathbf{E}[\,|R|\,] - \mathbf{E}[Y]$$

$$\mathbf{E}[\,|R|\,] = np \qquad \mathbf{E}[Y] = \sum_{uv \in E} \mathbf{E}[Y_{uv}] = mp^2$$

$G(V, E)$: n vertices, m edges

1. sample a random R :   each vertex is chosen *independently* with probability p

2. modify R to S :   independent set!

$\forall uv \in E$ if $u, v \in R$

delete one of u, v from R

$$\mathbf{E}[\,|S|\,] \geq np - mp^2 \;=\; \frac{n^2}{4m}$$

when $p = \dfrac{n}{2m}$

$G(V,E)$: n vertices, m edges

average degree $d = \frac{2m}{n}$

random independent set S:

$$\mathbf{E}[\,|S|\,] \geq \frac{n^2}{4m} = \frac{n}{2d}$$

Theorem: G has n vertices and m edges

$\Rightarrow$ $\exists$ an independent set S of size $\frac{n^2}{4m}$

Theorem: G has n vertices and m edges

$\Rightarrow$ $\exists$ an independent set S of size $\dfrac{n^2}{2m+n}$

- Draw a random independent set $S \subseteq V$
  - each $v \in V$ draws a real number $r_v \in [0,1]$ uniform and independent at random
  - each $v \in V$ joins S iff r_v is local maximal within the neighborhood of v
  - S must be an independent set
- $\forall v \in V$: $\Pr[v \in S] = \dfrac{1}{d_v + 1} \implies \mathbf{E}[\,|S|\,] = \sum_{v \in V} \dfrac{1}{d_v + 1}$

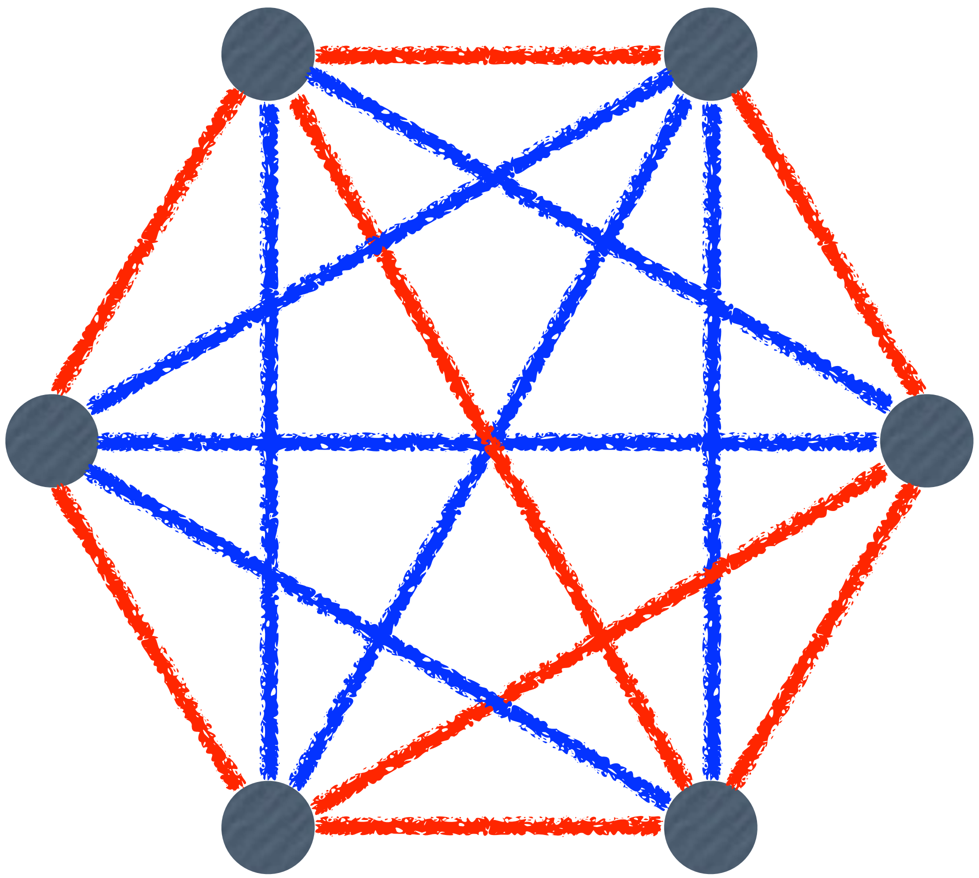
(Cauchy-Schwarz) $\geq \dfrac{n^2}{2m+n}$

Lovász Local Lemma

Ramsey Number

"In any party of six people, either at least three of them are mutual strangers or at least three of them are mutual acquaintances"

- For any edge-2-coloring of K_6, there is a ***monochromatic*** K_3.



Ramsey Theorem

If $n \geq R(k, k)$, for any edge-2-coloring of K_n, there is a monochromatic K_k.

Ramsey number: $R(k, k)$

$$R(k,k) > ?$$

"∃ a 2-coloring of K_n with no monochromatic K_k."

The Probabilistic Method:

a random 2-coloring of K_n

$\forall S \in \binom{[n]}{k}$

event A_S: S is a monochromatic K_k

To prove:

$$\Pr\left[\bigwedge_{S\in\binom{[n]}{k}} \overline{A_S}\right] > 0$$

Dependency!

Lovász Sieve

- Bad events: $A_1, A_2, \ldots, A_n$
- None of the bad events occurs:

$$\Pr\left[\bigwedge_{i=1}^{n} \overline{A_i}\right]$$

- The probabilistic method: being good is possible

$$\Pr\left[\bigwedge_{i=1}^{n} \overline{A_i}\right] > 0$$

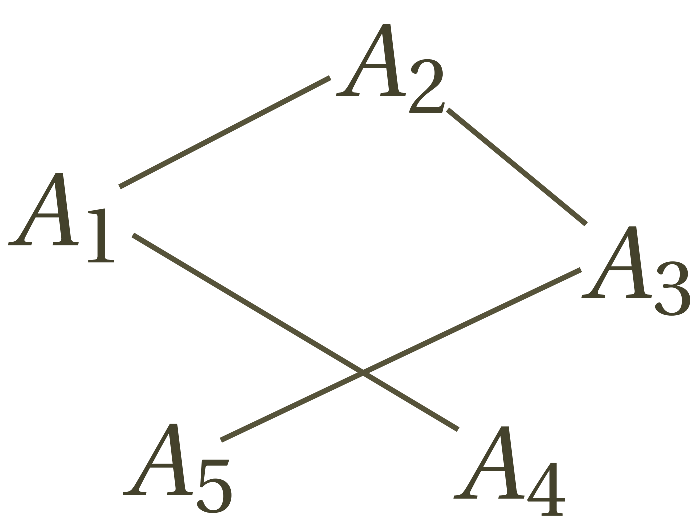
events: $A_1, A_2, \ldots, A_n$

dependency graph: $D(V, E)$

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

$ij \in E \Longleftrightarrow A_i$ and A_j are dependent

d : max degree of dependency graph



$A_1(X_1, X_4)$ $\quad$ $A_4(X_4)$

$A_2(X_1, X_2)$ $\quad$ $A_5(X_3)$

$A_3(X_2, X_3)$

$X_1, \ldots, X_4$ mutually independent

events: $A_1, A_2, \ldots, A_n$

d : max degree of dependency graph

Lovász Local Lemma

- $\forall i,\ \Pr[A_i] \le p$
- $\mathrm{e}p(d+1) \le 1$

$$\Rightarrow \Pr\left[\bigwedge_{i=1}^{n} \overline{A_i}\right] > 0$$

General Lovász Local Lemma

$\exists x_1, \ldots, x_n \in [0,1)$

$\forall i, \Pr[A_i] \le x_i \prod_{j\sim i}(1-x_j)$

$$\Rightarrow \Pr\left[\bigwedge_{i=1}^{n} \overline{A_i}\right] \ge \prod_{i=1}^{n}(1-x_i)$$

$$R(k,k) \geq n$$

"∃ a 2-coloring of K_n with no monochromatic K_k."

a random 2-coloring of K_n:

$\forall \{u, v\} \in K_n$, *uniformly* and *independently* $\begin{cases} uv \\ uv \end{cases}$

$\forall S \in \binom{[n]}{k}$ event A_S: S is a monochromatic K_k

$$\Pr[A_S] = 2 \cdot 2^{-\binom{k}{2}} = 2^{1-\binom{k}{2}}$$

A_S, A_T dependent $\Longleftrightarrow$ $|S \cap T| \geq 2$

max degree of dependency graph $d \leq \binom{k}{2}\binom{n}{k-2}$

To prove: $\Pr\left[\bigwedge_{S \in \binom{[n]}{k}} \overline{A_S}\right] > 0$

Lovász Local Lemma

- $\forall i,\ \Pr[A_i] \le p$
- $ep(d+1) \le 1$

$\Rightarrow$ $\Pr\left[\bigwedge_{i=1}^{n} \overline{A_i}\right] > 0$

$$\left.\begin{aligned} \Pr[A_S] &= 2^{1-\binom{k}{2}} \\ d &\le \binom{k}{2}\binom{n}{k-2} \end{aligned}\right\} \Rightarrow$$

for some $n = ck2^{k/2}$

with constant c

$e2^{1-\binom{k}{2}}(d+1) \le 1$

To prove: $\Pr\left[\bigwedge_{S\in\binom{[n]}{k}} \overline{A_S}\right] > 0$

$$R(k,k) \ge n = \Omega(k2^{k/2})$$

events: $A_1, A_2, \ldots, A_n$

General Lovász Local Lemma

$\exists x_1, \ldots, x_n \in [0,1)$

$\forall i, \Pr[A_i] \le x_i \prod_{j \sim i}(1 - x_j)$ $\Rightarrow$ $\Pr\left[\bigwedge_{i=1}^{n} \overline{A_i}\right] \ge \prod_{i=1}^{n}(1 - x_i)$

$$\Pr\left[\bigwedge_{i=1}^{n} \overline{A_i}\right] = \prod_{i=1}^{n} \Pr\left[\overline{A_i} \,\middle|\, \bigwedge_{j=1}^{i-1} \overline{A_j}\right] = \prod_{i=1}^{n}\left(1 - \Pr\left[A_i \,\middle|\, \bigwedge_{j=1}^{i-1} \overline{A_j}\right]\right)$$

Lemma For any $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$,

$$\Pr\left[\bigwedge_{i=1}^{n} \mathcal{E}_i\right] = \prod_{k=1}^{n} \Pr\left[\mathcal{E}_k \mid \bigwedge_{i<k} \mathcal{E}_i\right].$$

proof:

$$\Pr\left[\mathcal{E}_n \,\middle|\, \bigwedge_{i=1}^{n-1} \mathcal{E}_i\right] = \frac{\Pr\left[\bigwedge_{i=1}^{n} \mathcal{E}_i\right]}{\Pr\left[\bigwedge_{i=1}^{n-1} \mathcal{E}_i\right]}$$

recursion!

events: $A_1, A_2, \ldots, A_n$

General Lovász Local Lemma

$\exists x_1, \ldots, x_n \in [0,1)$

$\forall i, \Pr[A_i] \le x_i \prod_{j \sim i} (1 - x_j)$

$\Rightarrow \Pr\left[\bigwedge_{i=1}^{n} \overline{A_i}\right] \ge \prod_{i=1}^{n} (1 - x_i)$

I.H.

$$\Pr\left[A_{i_1} \mid \overline{A_{i_2}} \cdots \overline{A_{i_m}}\right] \le x_{i_1} \text{ for any } \{i_1, \ldots, i_m\}$$

induction on m:

$m = 1$, trivial

events: $A_1, A_2, \ldots, A_n$

$$\exists x_1, \ldots, x_n \in [0,1)$$
$$\forall i, \Pr[A_i] \le x_i \prod_{j \sim i} (1 - x_j)$$

I.H. $\Pr\left[A_{i_1} \mid \overline{A_{i_2}} \cdots \overline{A_{i_m}}\right] \le x_{i_1}$ for any $\{i_1, \ldots, i_m\}$

suppose i_1 adjacent to $i_2, \ldots, i_k$

$$\Pr\left[A_{i_1} \mid \overline{A_{i_2}} \cdots \overline{A_{i_m}}\right] = \frac{\Pr\left[A_{i_1} \overline{A_{i_2}} \cdots \overline{A_{i_k}} \mid \overline{A_{i_{k+1}}} \cdots \overline{A_{i_m}}\right]}{\Pr\left[\overline{A_{i_2}} \cdots \overline{A_{i_k}} \mid \overline{A_{i_{k+1}}} \cdots \overline{A_{i_m}}\right]}$$

$$\le \Pr\left[A_{i_1} \mid \overline{A_{i_{k+1}}} \cdots \overline{A_{i_m}}\right] = \Pr\left[A_{i_1}\right] \le x_{i_1} \prod_{j=2}^{k} (1 - x_{i_j})$$

$$= \prod_{j=2}^{k} \Pr\left[\overline{A_{i_j}} \mid \overline{A_{i_{j+1}}} \cdots \overline{A_{i_m}}\right] = \prod_{j=2}^{k} \left(1 - \Pr\left[A_{i_j} \mid \overline{A_{i_{j+1}}} \cdots \overline{A_{i_m}}\right]\right)$$

I.H. $\ge \prod_{j=2}^{k} (1 - x_{i_j})$

events: $A_1, A_2, \ldots, A_n$

General Lovász Local Lemma

$\exists x_1, \ldots, x_n \in [0,1)$

$\forall i, \Pr[A_i] \leq x_i \prod_{j \sim i} (1 - x_j)$ $\Rightarrow$ $\Pr\left[\bigwedge_{i=1}^{n} \overline{A_i}\right] \geq \prod_{i=1}^{n} (1 - x_i)$

$\Pr\left[A_{i_1} \mid \overline{A_{i_2}} \cdots \overline{A_{i_m}}\right] \leq x_{i_1}$ for any $\{i_1, \ldots, i_m\}$

$$\Pr\left[\bigwedge_{i=1}^{n} \overline{A_i}\right] = \prod_{i=1}^{n} \Pr\left[\overline{A_i} \,\middle|\, \bigwedge_{j=1}^{i-1} \overline{A_j}\right] = \prod_{i=1}^{n} \left(1 - \Pr\left[A_i \,\middle|\, \bigwedge_{j=1}^{i-1} \overline{A_j}\right]\right)$$

$$\geq \prod_{i=1}^{n} (1 - x_i) > 0$$

events: $A_1, A_2, \ldots, A_n$

d : max degree of dependency graph

Lovász Local Lemma

- $\forall i,\ \Pr[A_i] \le p$
- $\mathrm{e}p(d+1) \le 1$

$$\Rightarrow \Pr\left[\bigwedge_{i=1}^{n} \overline{A_i}\right] > 0$$

General Lovász Local Lemma

$\exists x_1, \ldots, x_n \in [0,1)$

$\forall i, \Pr[A_i] \le x_i \prod_{j\sim i}(1-x_j)$

$$\Rightarrow \Pr\left[\bigwedge_{i=1}^{n} \overline{A_i}\right] \ge \prod_{i=1}^{n}(1-x_i)$$

Constraint Satisfaction Problem (CSP)

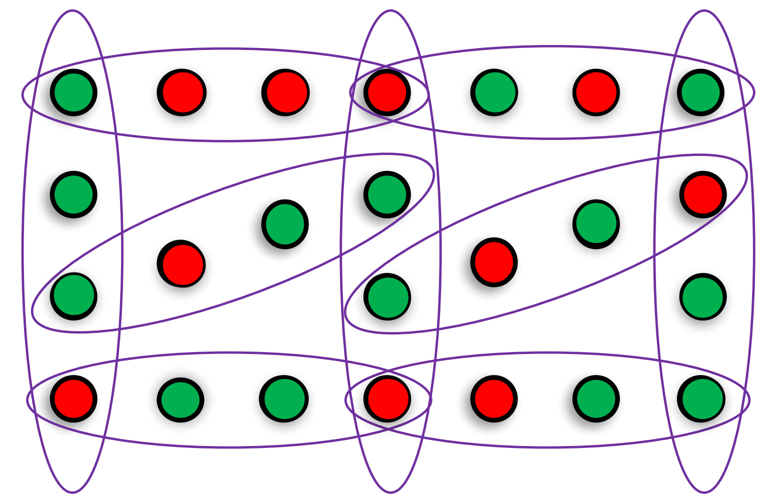
- **Variables:** $x_1, \ldots, x_n \in [q]$
- (local) **Constraints:** $C_1, \ldots, C_m$
  - each C_i is defined on a subset $\mathsf{vbl}(C_i)$ of variables

$$C_i : [q]^{\mathsf{vbl}(C_i)} \to \{\mathtt{True}, \mathtt{False}\}$$

- Any $\boldsymbol{x} \in [q]^n$ is a **CSP solution** if it satisfies all $C_1, \ldots, C_m$
- **Examples:**
  - k-CNF, (hyper)graph coloring, set cover, unique games…
  - vertex cover, independent set, matching, perfect matching, …

Hypergraph Coloring

- k-uniform hypergraph $H = (V, E)$:
  - V is vertex set, $E \subseteq \binom{V}{k}$ is set of hyperedges



- degree of vertex $v \in V$: # of hyperedges $e \ni v$
- proper q-coloring of H:
  - $f : V \to [q]$ such that no hyperedge is *monochromatic*

$$\forall e \in E, \quad |f(e)| > 1$$

Theorem: For any k-uniform hypergraph H of max-degree Δ,

$$\Delta \leq \frac{q^{k-1}}{\mathrm{e}k} \implies H \text{ is } q\text{-colorable}$$

$k \geq \log_q \Delta + \log_q \log_q \Delta + O(1)$

Hypergraph Coloring

Theorem: For any k-uniform hypergraph H of max-degree Δ,

$$\Delta \le \frac{q^{k-1}}{\mathrm{e}k} \implies H \text{ is } q\text{-colorable}$$

- Uniformly and independently color each $v \in V$ a random color $\in [q]$
- Bad event A_e for each hyperedge $e \in E \subseteq \binom{V}{k}$: e is monochromatic
  - $\Pr[A_e] \le p = q^{1-k}$
- Dependency degree for bad events $d \le k(\Delta - 1)$
  - $\Delta \le \frac{q^{k-1}}{\mathrm{e}k} \implies \mathrm{e}p(d+1) \le 1$ Apply LLL