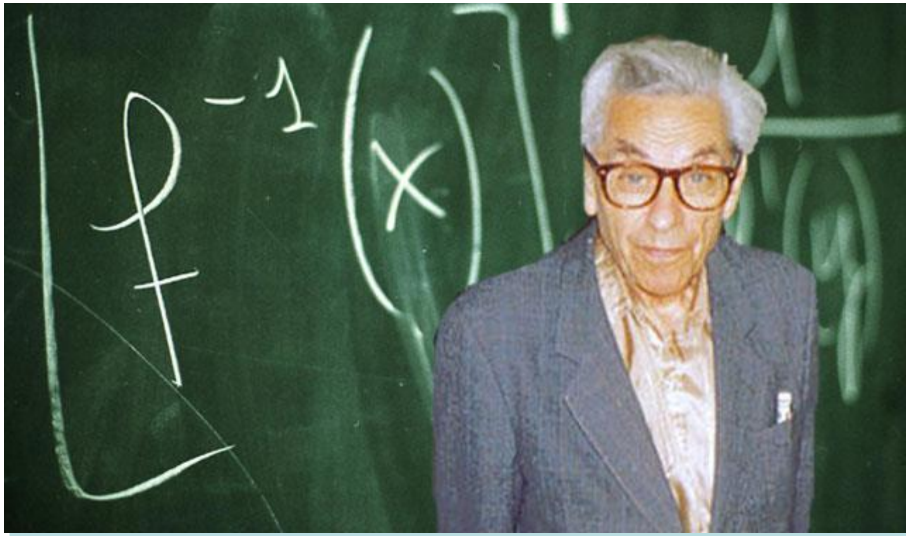
The Probabilistic Method

longhuan@sjtu.edu.cn

The probabilistic method



Paul Erdős (26 March 1913 – 20 September 1996)

Hungarian mathematician. Erdős published more papers than any other mathematician in history, working with hundreds of collaborators. He worked on problems in combinatorics, graph theory, number theory, classical analysis, approximation theory, set theory, and probability theory.

- The probabilistic method is a nonconstructive method, primarily used in combinatorics and pioneered by Paul Erdős.
- *For proving the existence of a prescribed kind of mathematical object. It works by showing that if one randomly chooses objects from a specified class, the probability that the result is of the prescribed kind is more than zero.*

Basic Counting Argument

The Expectation Argument

Lovasz Local Lemma

1. Cards Shuffling

- Consider a new deck of 52 cards. We will shuffle the cards by so-called dovetail shuffling (a.k.a. ‘riffle’).
- Is 4 rounds of dovetail shuffling enough to yield a random order of the cards?

$$\binom{52}{26}^4 < 52!$$

$$\frac{3\log_2 n}{2} + \theta$$

1. Cards Shuffling

- Consider a new deck of 52 cards. We will shuffle the cards by so-called dovetail shuffling (a.k.a. ‘riffle’).
- Is 4 rounds of dovetail shuffling enough to yield a random order of the cards?

$$\binom{52}{26}^4 < 52!$$

The Annals of Applied Probability
1992, Vol. 2, No. 2, 294–313

TRAILING THE DOVETAIL SHUFFLE TO ITS LAIR

By Dave Bayer[1] and Persi Diaconis[2]

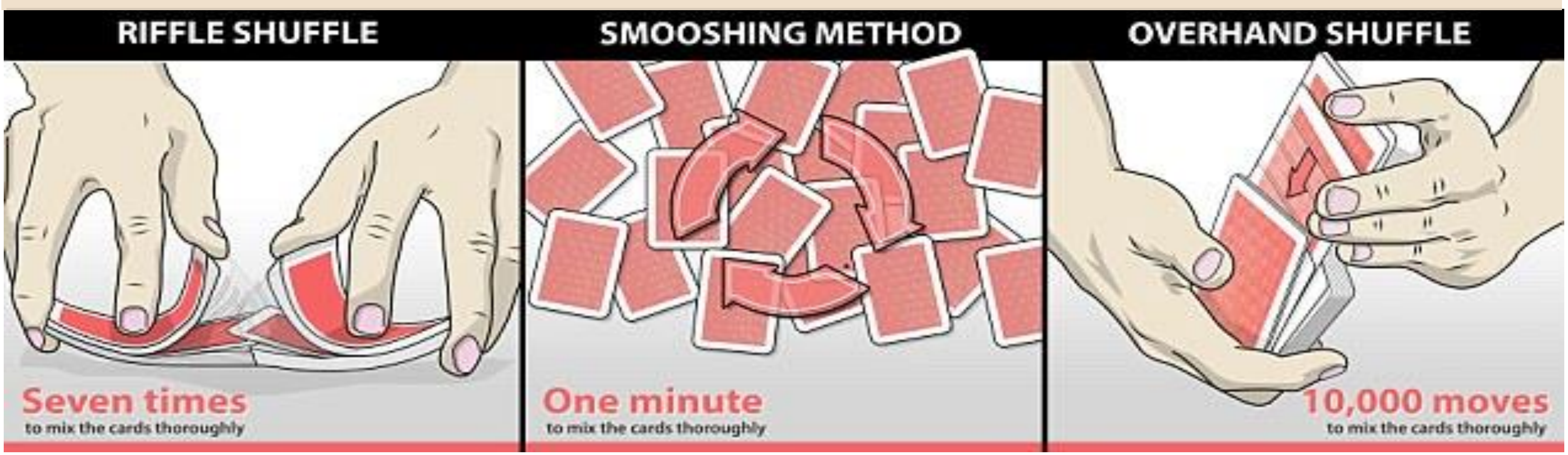
Columbia University and Harvard University

We analyze the most commonly used method for shuffling cards. The main result is a simple expression for the chance of any arrangement after any number of shuffles. This is used to give sharp bounds on the approach to randomness: $\frac{3}{2}\log_2 n + \theta$ shuffles are necessary and sufficient to mix up n cards.

Key ingredients are the analysis of a card trick and the determination of the idempotents of a natural commutative subalgebra in the symmetric group algebra.

How to shuffle cards like a pro: Mathematician shows why the 'riffle' technique is more effective than the flashy 'overhand'

- A Stanford University mathematician compared shuffling techniques
- Dealers using a 'riffle' shuffle need to repeat the process seven times to get a random pack of cards, said Peri Diaconis
- This technique involves cutting a deck and shuffling the halves together
- Whereas 'overhand' needs to be repeated 10,000 times to get same results
- The 'smooshing' or wash method takes one minute to randomise cards



2. Difficult Boolean Functions

- n variable Boolean functions:

$$f: \{0,1\}^n \to \{0,1\}.$$

- Logical formula in n variables:
  - Symbols: $x_1, x_2, \dots, x_n$;
  - Parenthesis：(,);
  - Logical connectives: $\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$, $\neg$;

Proposition. There exists a Boolean function of n variables that cannot be defined by any formula with fewer than $2^n / \log_2(n+8)$ symbols.

Proposition. There exists a Boolean of n variables that cannot be defined by any formula with fewer than $2^n / \log_2(n+8)$ symbols.

- Proof:

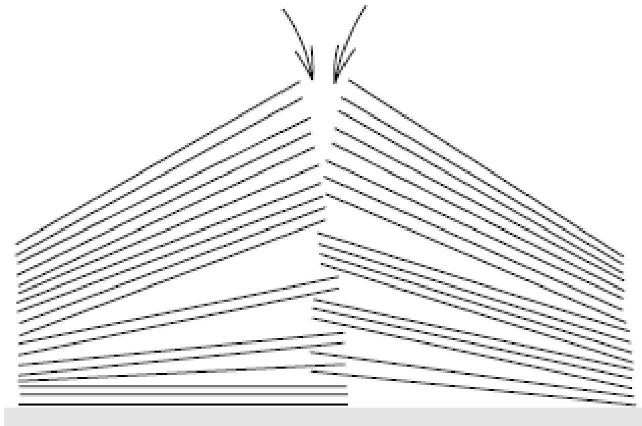
The number of all Boolean functions of n variables: $= 2^{2^n}$

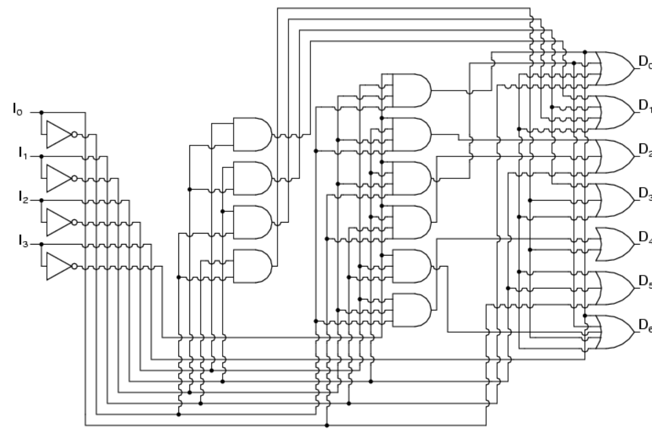
The number of formulas in n variables written by at most m symbols is: $\leq (n+8)^m$

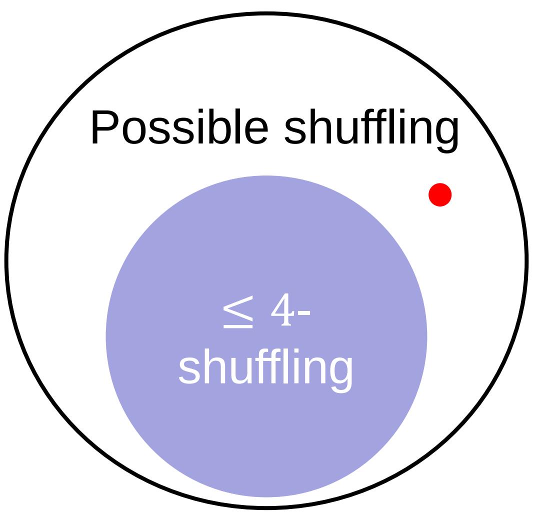
Complications will emerge when: $2^{2^n} > (n+8)^m$

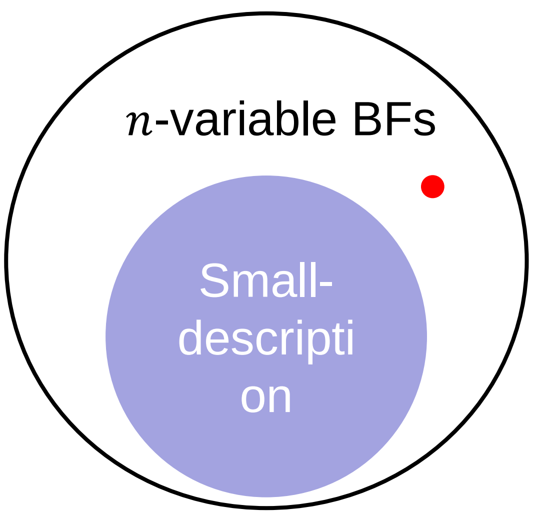
$$m < 2^n / \log_2(n+8)$$

The existence of certain objects

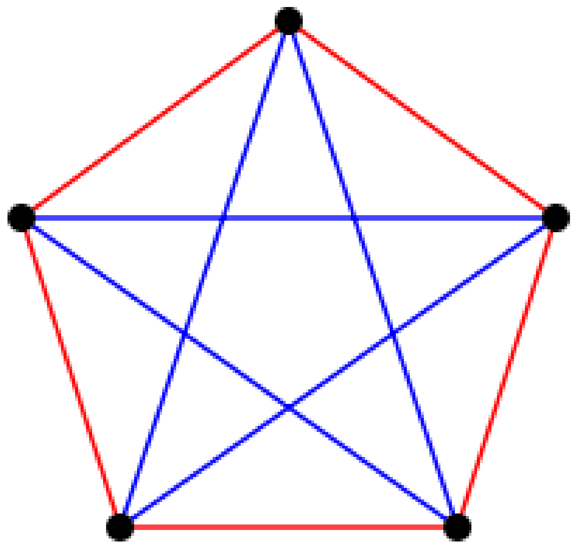


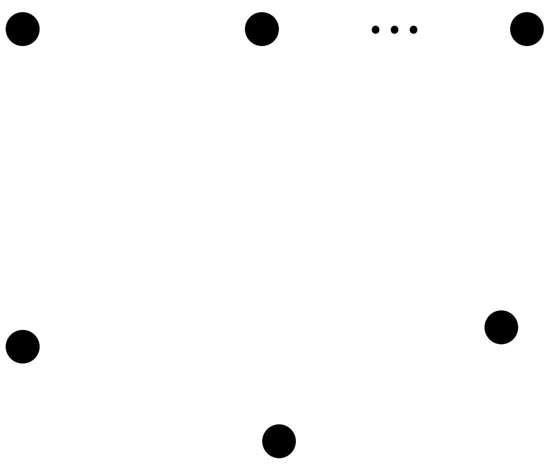






3. Edge Coloring (a.k.a. Ramsey number $R(k,k)$)

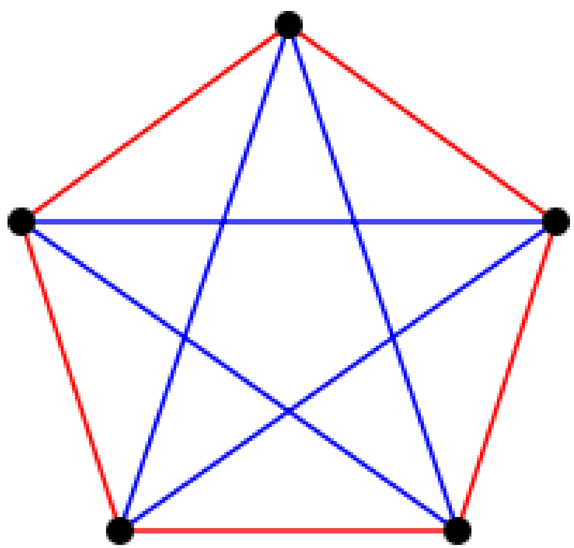


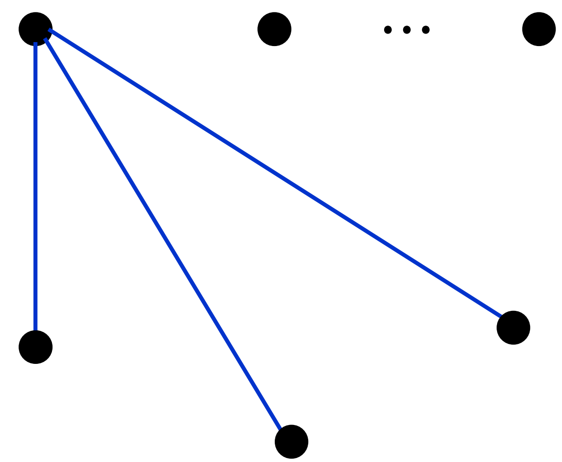


A Ramsey Number, $n = R(r, b)$, is the smallest integer n such that the 2-colored graph K_n, using the colors red and blue for edges, implies ① a red monochromatic subgraph K_r ,or ② a blue monochromatic subgraph K_b.

$$R(3,3) = 6$$

3. Edge Coloring (a.k.a. Ramsey number $R(k,k)$)

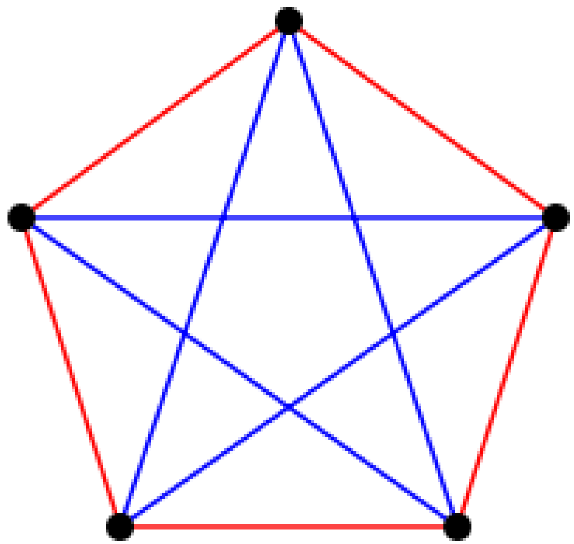


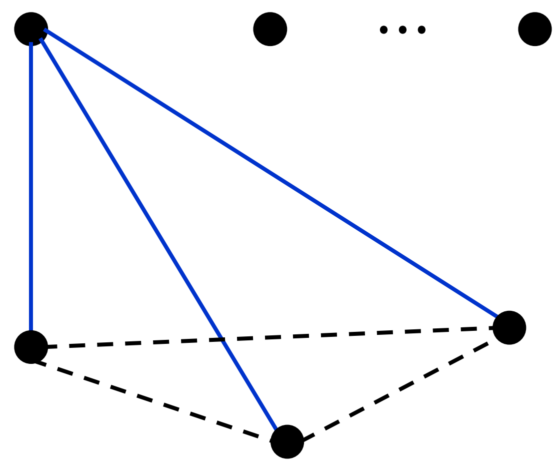


A Ramsey Number, $n = R(r,b)$, is the smallest integer n such that the 2-colored graph K_n, using the colors red and blue for edges, implies ① a red monochromatic subgraph K_r ,or ② a blue monochromatic subgraph K_b.

$$R(3,3) = 6$$

3. Edge Coloring (a.k.a. Ramsey number $R(k,k)$)

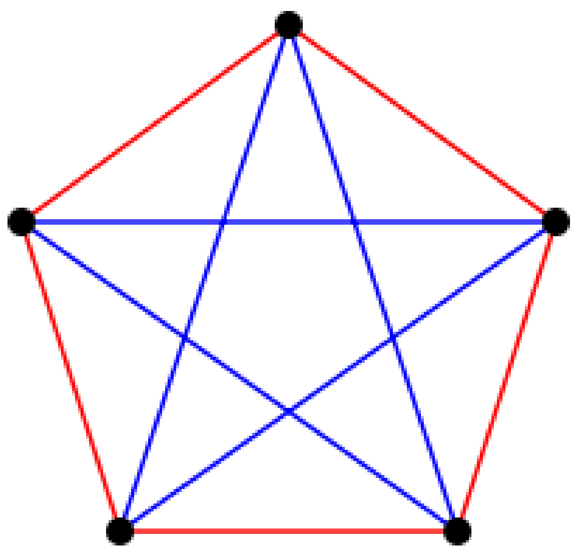


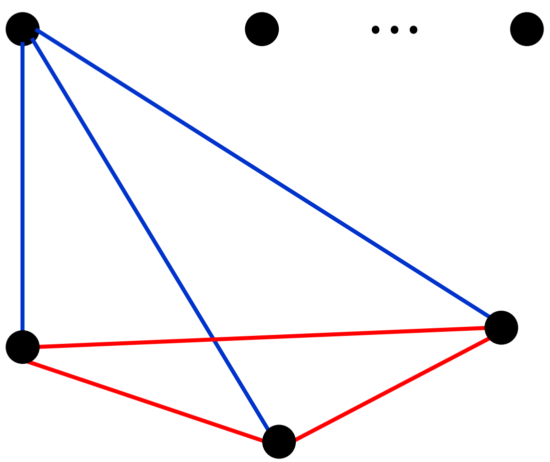


A Ramsey Number, $n = R(r,b)$, is the smallest integer n such that the 2-colored graph K_n, using the colors red and blue for edges, implies ① a red monochromatic subgraph K_r ,or ② a blue monochromatic subgraph K_b.

$$R(3,3) = 6$$

3. Edge Coloring
(a.k.a. Ramsey number $R(k,k)$)

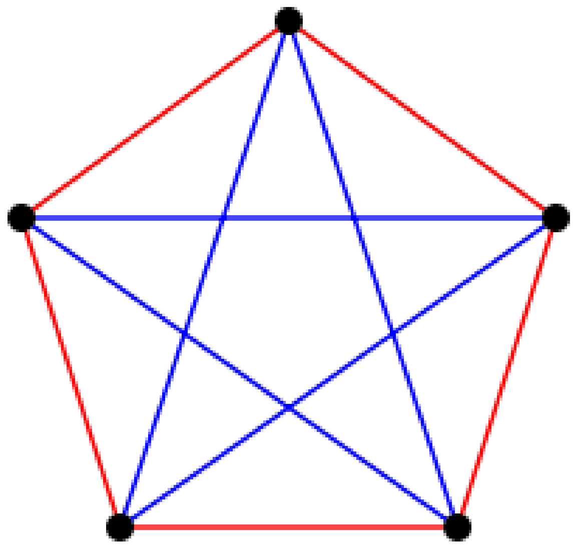


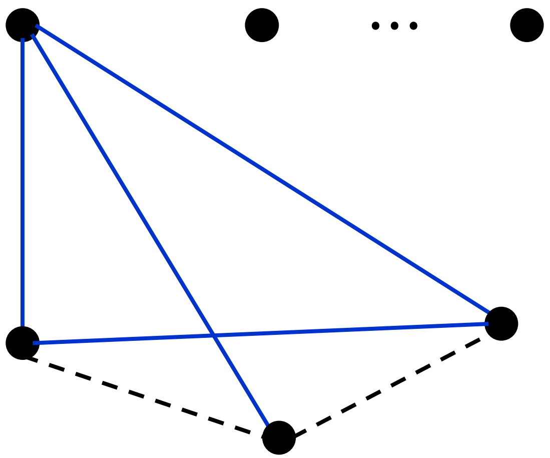


A Ramsey Number, $n = R(r,b)$, is the smallest integer n such that the 2-colored graph K_n, using the colors red and blue for edges, implies ① a red monochromatic subgraph K_r ,or ② a blue monochromatic subgraph K_b.

$R(3,3) = 6$

3. Edge Coloring
(a.k.a. Ramsey number $R(k,k)$)





Values / known bounding ranges for Ramsey numbers $R(r, s)$ (sequence A212954 in the OEIS)

r \ s	1	2	3	4	5	6	7	8	9	10
1	**1**	1	1	1	1	1	1	1	1	1
2		**2**	3	4	5	6	7	8	9	10
3			**6**	9	14	18	23	28	36	40–42
4				**18**	25[7]	36–41	49–61	59[13]–84	73–115	92–149
5					**43–48**	58–87	80–143	101–216	133–316	149[13]–442
6						**102–165**	115[13]–298	134[13]–495	183–780	204–1171
7							**205–540**	217–1031	252–1713	292–2826
8								**282–1870**	329–3583	343–6090
9									**565–6588**	581–12677
10										**798–23556**

Erdős asks us to imagine an alien force, vastly more powerful than us, landing on Earth and demanding the value of $R(5, 5)$ or they will destroy our planet. In that case, he claims, we should marshal all our computers and all our mathematicians and attempt to find the value. But suppose, instead, that they ask for $R(6, 6)$. In that case, he believes, we should attempt to destroy the aliens.

— Joel Spencer

Theorem. If $\binom{n}{k} 2^{-\binom{k}{2}+1} < 1$, then it is possible to color the edges of K_n with two colors so that it has no single-colored (monochromatic) K_k subgraphs.

- Proof.

For each $e = \{u, v\}$

$$\begin{cases} \text{Head:} & f(e) = \text{RED} \\ \text{Tail:} & f(e) = \text{BLUE} \end{cases}$$

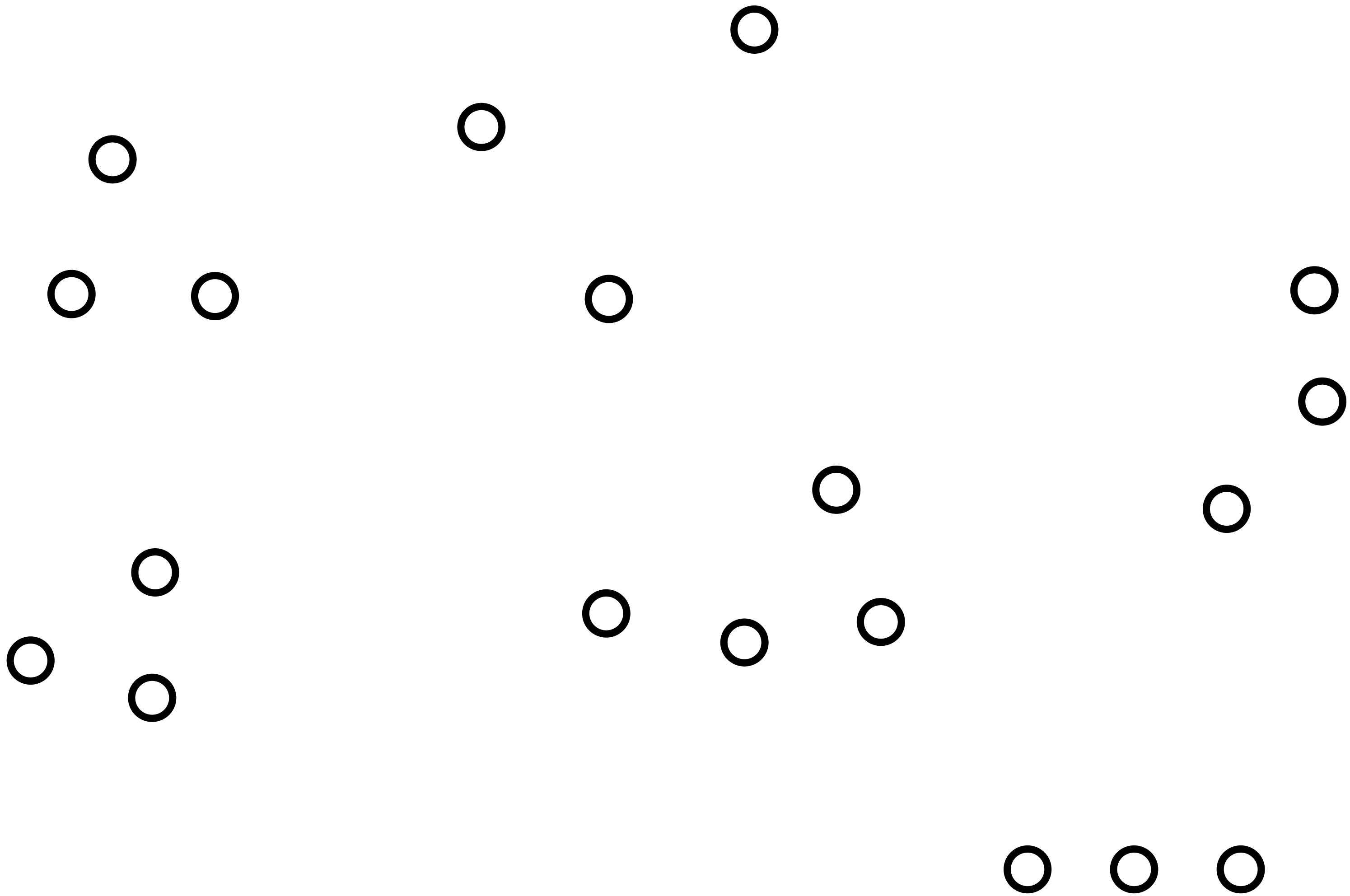
A certain K_k subgraph is monochromatic: $= 2 \cdot \frac{1}{2^{\binom{k}{2}}}$

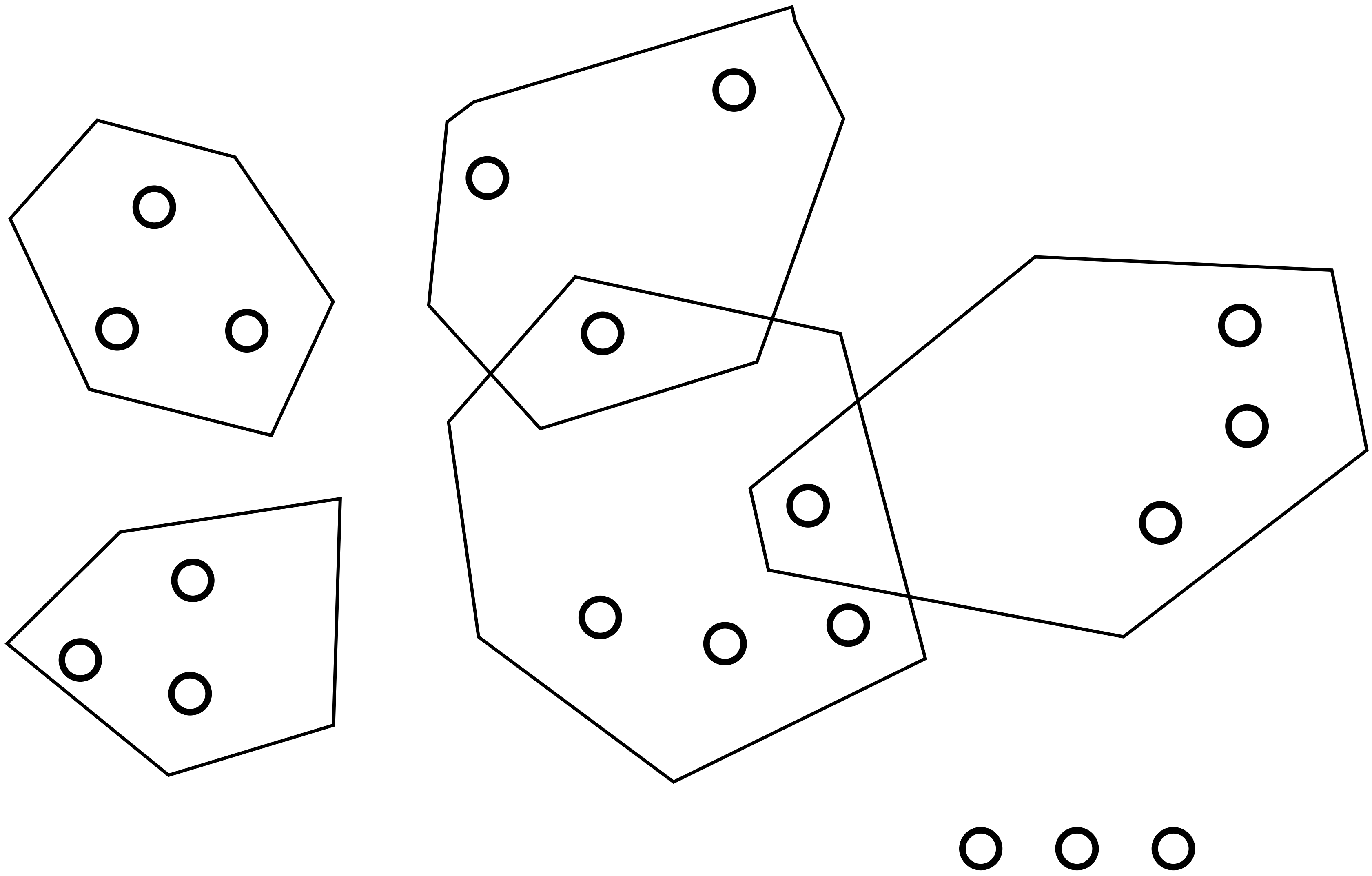
The probability that one of K_k subgraph is monochromatic: $\le \binom{n}{k} \cdot 2 \cdot \frac{1}{2^{\binom{k}{2}}} = \binom{n}{k} 2^{-\binom{k}{2}+1}$

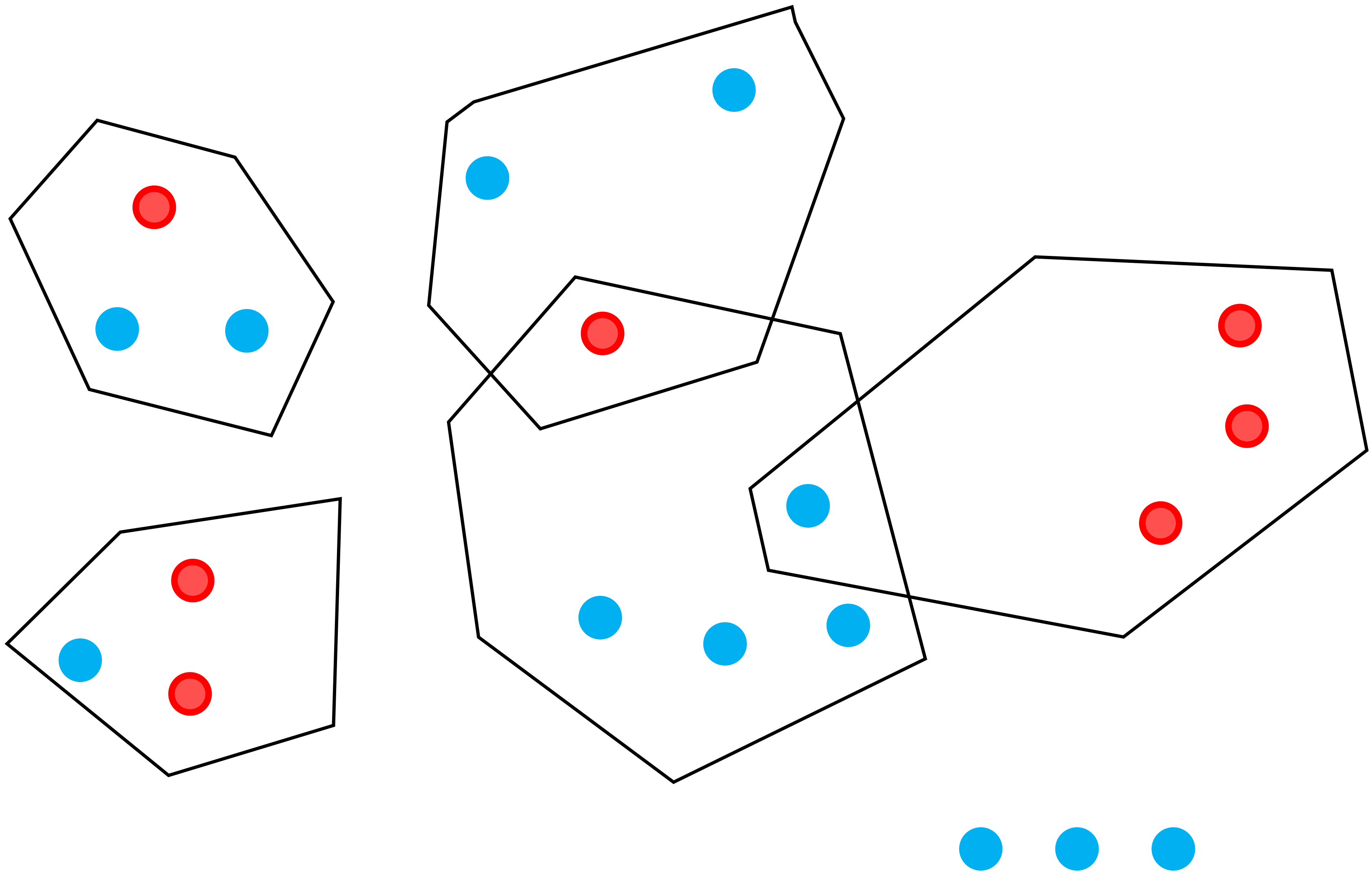
$< \mathbf{1}$

4. Coloring set systems by two colors(*)

- X is a finite set, $M \subseteq P(X)$.
- **Coloring function** $f: X \to \{\mathrm{RED}, \mathrm{BLUE}\}$
- **2-Colorability**. if there is a coloring function such that every $S \in M$ contains points of both colors. Then M is 2-colorable.
- **Example**. $X = \{1,2,3\}, M = \{\{1,2\}, \{1,3\}, \{2,3\}\}$ then M is not 2-colorable.







- X is a finite set, $M \subseteq P(X)$.
- **Coloring function** $f: X \to \{\mathrm{RED}, \mathrm{BLUE}\}$
- **2-Colorability**. if there is a coloring function such that every $S \in M$ contains points of both colors. Then M is 2-colorable.
- $\forall\, S \in M\ (|S| = k)$
- $s(k)$ is the smallest number of sets in a system M (i.e., $|M|$) that is not 2-colorable.
- **Example:** $s(2) = 3$.

Theorem. $s(k) \geq 2^{k-1}$, i.e. any system consisting of fewer than 2^{k-1}sets of size k admits a 2-coloring.

Theorem. $s(k) \geq 2^{k-1}$, i.e. *any* system consisting of fewer than 2^{k-1} sets of size k admits a 2-coloring.

- Proof. $M \subseteq \binom{X}{k}, \ |M| = m$

For each $x \in X$

$$\begin{cases} \text{Head:} & f(x) = \text{RED} \\ \text{Tail:} & f(x) = \text{BLUE} \end{cases}$$

$S \in M$, the probability that S is single-colored is: $\frac{1}{2^k} + \frac{1}{2^k} = 2^{1-\mathrm{k}}$

The probability that at least one of the m sets in M is <u>*monochromatic (single-color)*</u> is: $\leq m \cdot 2^{1-\mathrm{k}}$

If $m < 2^{k-1}$ the probability is strictly less than 1.

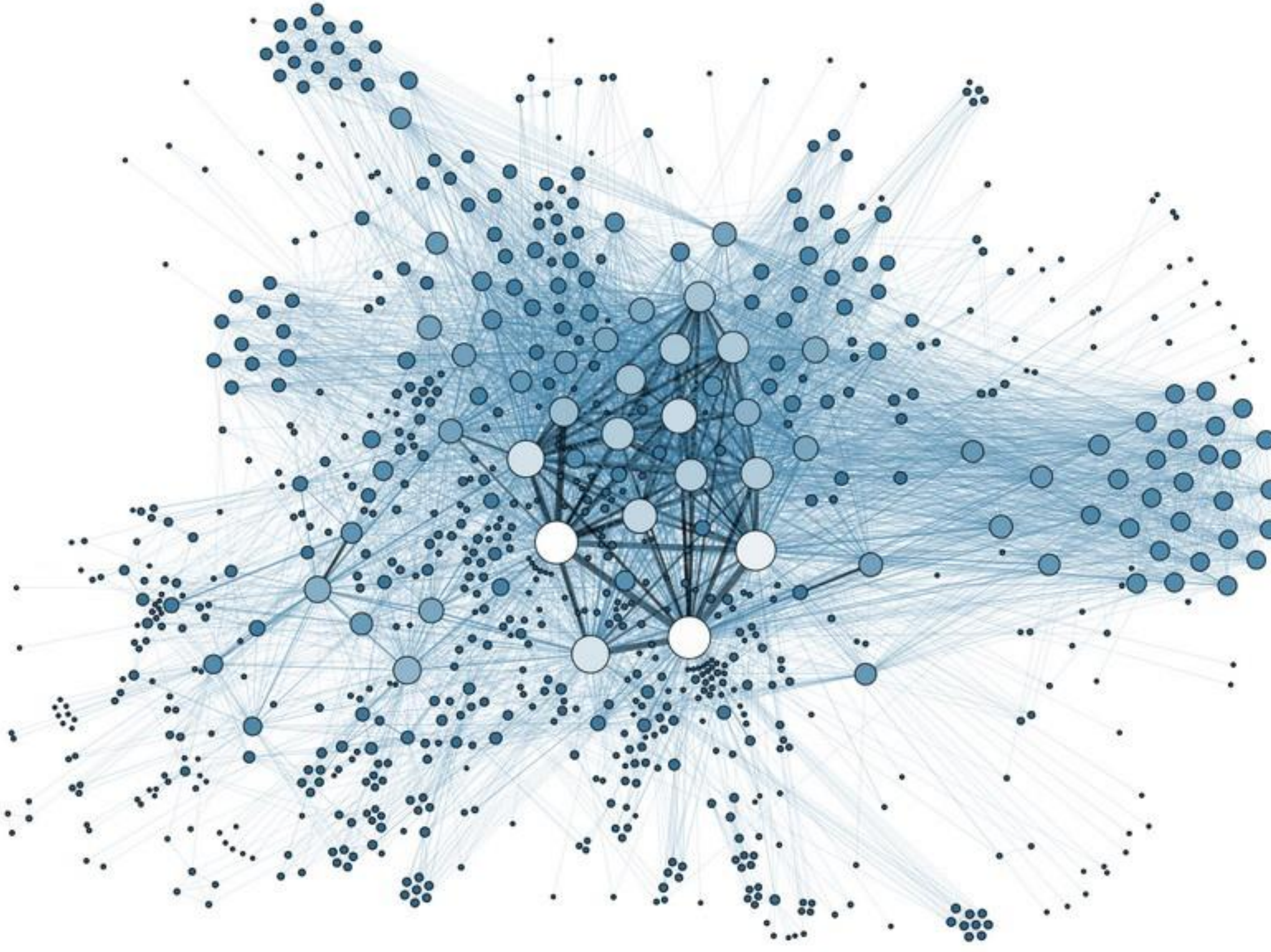
Some M is 2-colorable. $\therefore s(k) \geq 2^{k-1}$.

Basic Counting Argument

The Expectation Argument

Lovasz Local Lemma

1. *Dense* Partition



Theorem. Let G be a graph with an even number, $2n$, of vertices and with $m > 0$ edges. Then the set $V = V(G)$ can be divided into two disjoint n-element subsets A and B in such a way that more than $\frac{m}{2}$ edges go between A and B.

Proof. Randomly choose n vertex to form set A.
Then $B = V \backslash \mathrm{A}$.

For any edge $e = \{u, v\}$, the probability of e being lying 'across' A and B is: $\frac{2\binom{2n-2}{n-1}}{\binom{2n}{n}} = \frac{n}{2n-1} > \frac{1}{2}$

$|E(G)| = m$, the expectation of the number of edges lying 'across' : $E\big(C(A, B)\big) = m \cdot \frac{n}{2n-1} > \frac{m}{2}$

There must exist a choice of A with more than half of the edges going across.

A Las Vegas algorithm for finding an partition

Let $p = \Pr\left(C(A,B) \geq \frac{m}{2}\right)$,

$$\frac{m}{2} < E\left(C(A,B)\right) = \sum_{i \leq \frac{m}{2}-1} i \cdot \Pr(C(A,B) = i) + \sum_{i \geq \frac{m}{2}} i \cdot \Pr(C(A,B) = i)$$

$$\leq (1-p)\left(\frac{m}{2} - 1\right) + pm$$

$$\therefore \ p \geq \frac{1}{\frac{m}{2} + 1}$$

The expected number of samples before finding a cut with value at least $m/2$ is therefore just $\frac{m}{2} + 1$.

Sample and testing.

Derandomization using conditional expectation

Placing the vertices deterministically, in an arbitrary order $v_1, v_2, \dots, v_n$.

For each v_i, define $x_i \in \{A, B\}$ to be the set where v_i is placed.

$\boldsymbol{E}[C(A,B) | \, x_1, x_2, \dots, x_k]$ is the conditional expectation of the value of the cut given the location $x_1, x_2, \dots, x_k$ of the first k vertices.

We can always place the next vertex so that

$$\boldsymbol{E}[C(A,B) | \, x_1, x_2, \dots, x_k] \le \boldsymbol{E}[C(A,B) | \, x_1, x_2, \dots, x_k, x_{k+1}]$$

Then

$$\frac{m}{2} \le \boldsymbol{E}[C(A,B)] = \boldsymbol{E}[C(A,B) | \, x_1] \le \boldsymbol{E}[C(A,B) | \, x_1, x_2, \dots, x_n]$$

To get $\boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k] \le \boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k, x_{k+1}]$

Consider placing v_{k+1} in A or B with equal probability $\frac{1}{2}$. Let Y_{k+1} be a random variable representing the set where v_{k+1} is places. Then

$$\boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k] = \frac{1}{2}\boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k, Y_{k+1} = A] + \frac{1}{2}\boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k, Y_{k+1} = B]$$

Therefore,

$$\boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k] \le \max\begin{pmatrix}\boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k, Y_{k+1} = A], \\ \boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k, Y_{k+1} = B]\end{pmatrix}$$

Therefore, we just need to decide which of $\boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k, Y_{k+1} = A]$ and $\boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k, Y_{k+1} = B]$ is larger. And then set Y_{k+1} accordingly.

$\boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k, Y_{k+1} = Z]$ where $Z \in \{A, B\}$, is the number of edges ① crossing the cut whose endpoints are both among the first $k+1$ vertices, plus ② half of the remaining edges.

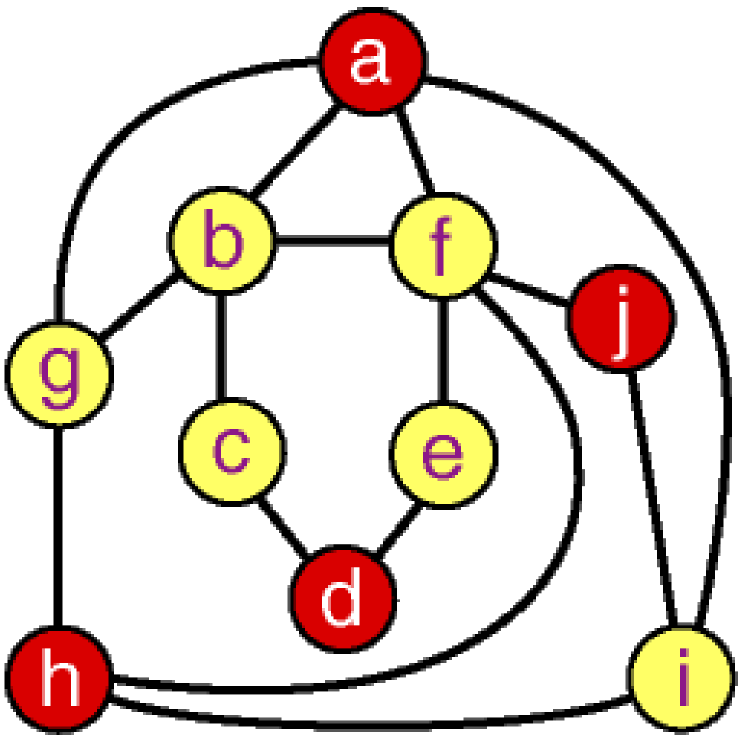
$$\boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k] \le \boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k, x_{k+1}]$$

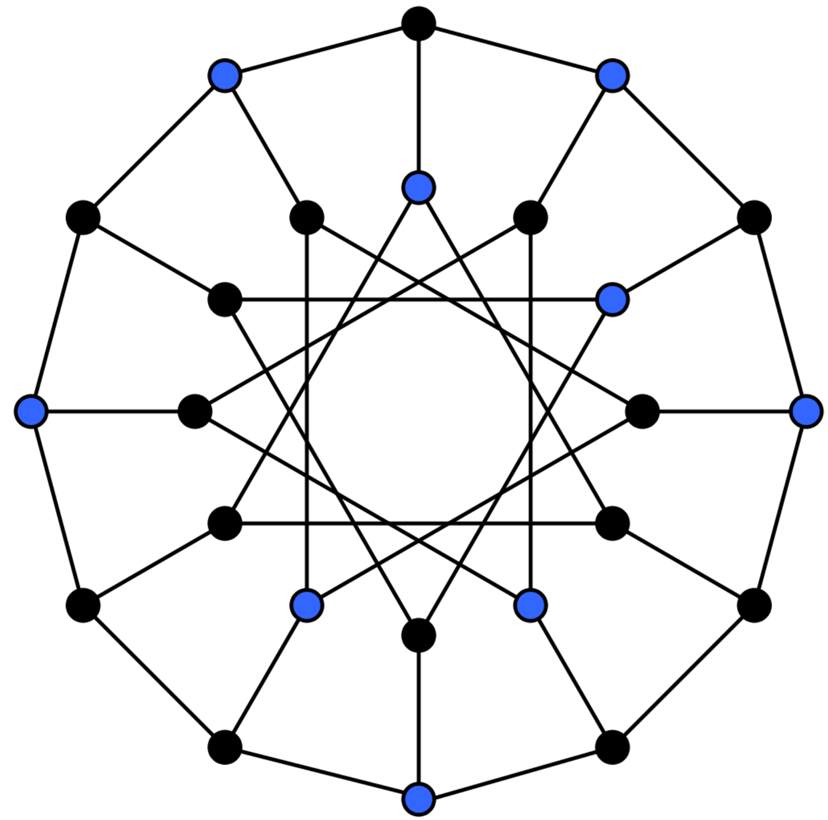
Thus, the larger of $\boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k, Y_{k+1} = A]$ and $\boldsymbol{E}[C(A,B)|\, x_1, x_2, \dots, x_k, Y_{k+1} = B]$ is simply determined by whether v_{k+1} has more neighbors in A or B.

The derandomized algorithm:

- Take the vertices in some order;
- Place the first vertex arbitrarily in A.
- Place each successive vertex to maximize the number of edges crossing the cut. (Equivalently, place each vertex on the side with fewer neighbors.)

2. Independent set





Theorem. (Tur$\grave{a}$n's theorem). For any graph G on n vertices, we have $\alpha(G) \geq \frac{n^2}{2|E(G)|+n}$.
where $\alpha(G)$ denotes the size of the largest independent set of vertices in the graph G.

Lemma. For any graph G, we have

$$\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{\deg_G(v) + 1}.$$

Lemma. For any graph G, we have

$$\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{\deg_G(v) + 1}.$$

- **Proof.** $V = \{1,2, \dots, n\}$

Randomly pick a permutation $\pi\colon \mathrm{V} \to V$,

$M \stackrel{\text{def}}{=} M(\pi) \subseteq V; M = \{v \mid \forall u\ (\{u, v\} \in E(G) \to \pi(u) > \pi(v))\}$,

$M(\pi)$ is an independent set in G, $\therefore$ for any $\pi, |M(\pi)| \leq \alpha(G)$.

A_v: the event "$v \in M(\pi)$"

$$P(A_v) = \frac{1}{1 + |N_v|} = \frac{1}{\deg_G(v) + 1}$$

$$\alpha(G) \geq E(|M|) = \sum_{v \in V} E[I_{A_v}] = \sum_{v \in V} P(A_v) = \sum_{v \in V} \frac{1}{\deg_G(v) + 1}$$

Lemma. For any graph G, we have

$$\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{\deg_G(v) + 1}.$$

Theorem. (Turàn's theorem). For any graph G on n vertices, we have $\alpha(G) \geq \frac{n^2}{2|E(G)|+n}$.
where $\alpha(G)$ denotes the size of the largest independent set of vertices in the graph G.

$$\sum_{v \in V(G)} \frac{1}{\deg_G(v)+1}$$

will be minimal, when $d_1 = d_2 = \cdots = d_n = \frac{2|E(G)|}{n}$.

3. Maximum Satisfaction

- Logical formula:
  $(x_1 \vee \overline{x_2} \vee \overline{x_3}) \wedge (\overline{x_1} \vee \overline{x_3}) \wedge (x_1 \vee x_2 \vee x_4) \wedge (x_4 \vee \overline{x_3}) \wedge (x_4 \vee \overline{x_1})$
- SAT is NP-hard
- MAXSAT: Given a SAT formula, satisfying as many clauses as possible.

Theorem. Given a set of m clauses, let k_i be the number of literals in the ith clause for $i = 1, \dots, m$. Let $k = \min_{1 \le i \le m} k_i$. Then there is a truth assignment that satisfies at least

$$\sum_{i=1}^{m} \left(1 - 2^{-k_i}\right) \ge m\left(1 - 2^{-k}\right).$$

- **Proof**

Assign values independently and uniformly at random to the variables.

The probability that the ith clause with k_i literals is satisfied is $1 - 2^{-k_i}$

The expected number of satisfied clauses is $\sum_{i=1}^{m} \left(1 - 2^{-k_i}\right) \ge m\left(1 - 2^{-k}\right).$

Basic Counting Argument

The Expectation Argument

Lovasz Local Lemma

László Lovász (Hungarian: born March 9, 1948) is a Hungarian mathematician and professor emeritus at Eötvös Loránd University, best known for his work in combinatorics, for which he was awarded the 2021 Abel Prize jointly with Avi Wigderson. He was the president of the International Mathematical Union from 2007 to 2010 and the president of the Hungarian Academy of Sciences from 2014 to 2020.

Lovász in 2017

László Lovász - Wikipedia

- $E_1, E_2, \ldots, E_n$ is a set of bad events.
- The probability that none of the bad events occurs is

$$\Pr\left(\bigcap_{i=1}^{n} \overline{E}_i\right)$$

- Mutual independence is rare in real applications.
- What if the dependency is limited.

Mutually independent of a set

- Event F is mutually independent of the events $F_1, F_2, \ldots, F_n$ if, for any subset $I \subseteq [1, n]$:

$$\Pr\left(F \mid \cap_{j \in I} F_j\right) = \Pr(F)$$

- **Dependency graph.** for a set of events $E_1, E_2, \ldots, E_n$, define graph $G = (V, E)$ such that $V = \{1,2, \ldots, n\}$ and, for $i = 1, \ldots, n$, event E_i is mutually independent of the events $\left\{E_j \mid (i, j) \notin E\right\}$.

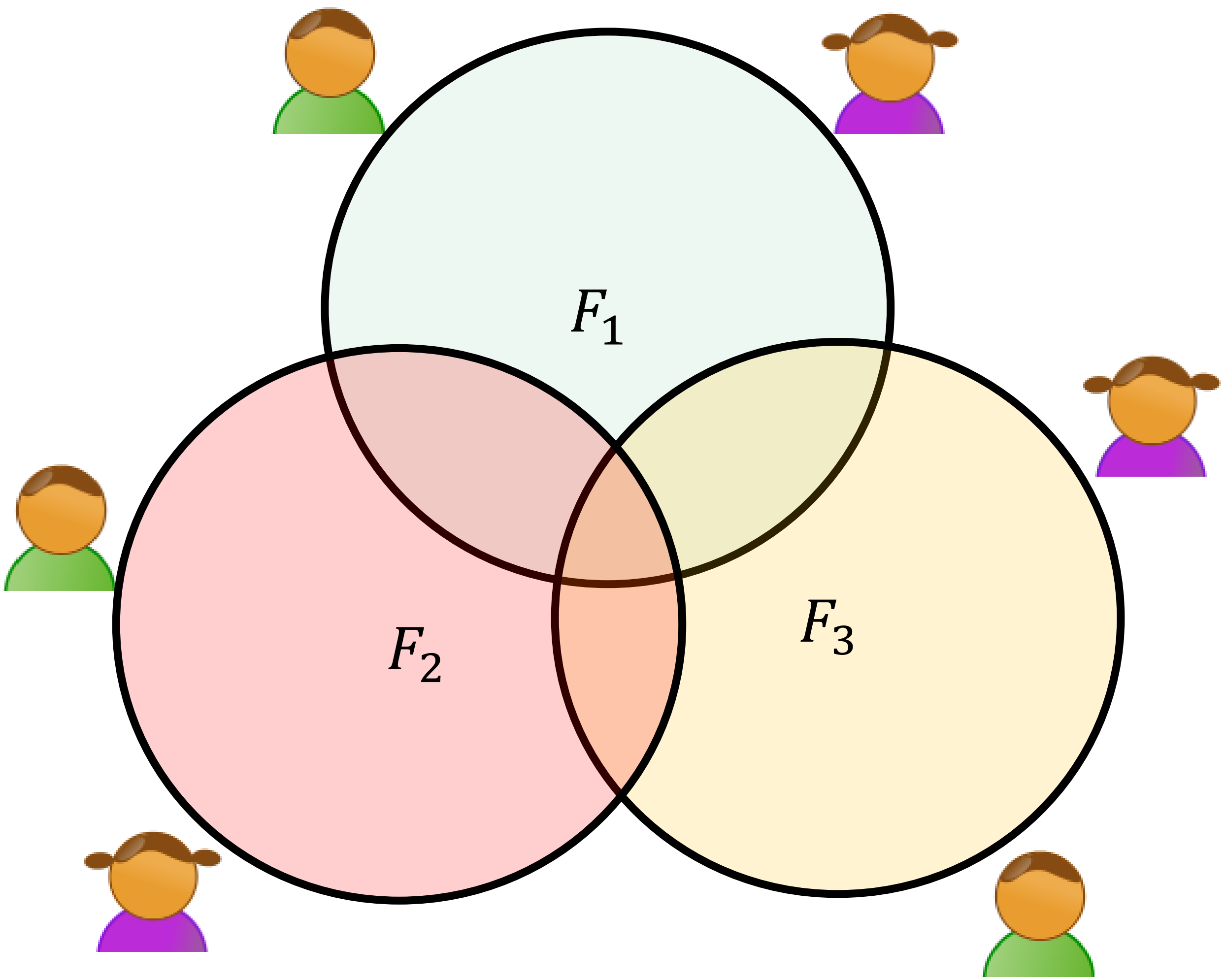
Theorem[Lovasz Local Lemma]:

Let $E_1, E_2, \ldots, E_n$ be a set of events, and assume that the following holds:

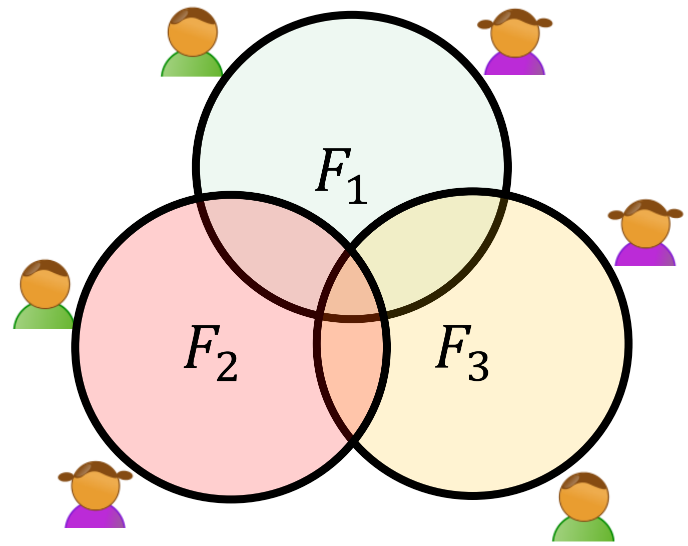
1. For all i, $\Pr(E_i) \le p$;
2. The degree of the dependency graph given by $E_1, E_2, \ldots, E_n$ is bounded by d;
3. $4dp \le 1$.

Then $\Pr(\bigcap_{i=1}^{n} \overline{E}_i) > 0$.

Application 1: Edge-disjoint path



- Scenario
  - n pairs of users need to communicate using edge-disjoint paths on a given network.
  - Each pair $i = 1, \ldots, n$ can choose a path from a collection F_i of m path (i.e. $|F_i| = m$).



Theorem: If any path in F_i shares edges with no more than k paths in F_j, where $i \neq j$ and $\frac{8nk}{m} < 1$, then there is a way to choose n edge-disjoint paths connecting the n pairs.

Theorem: If any path in F_i shares edges with no more than k paths in F_j, where $i \neq j$ and $\frac{8nk}{m} \leq 1$, then there is a way to choose n edge-disjoint paths connecting the n pairs.

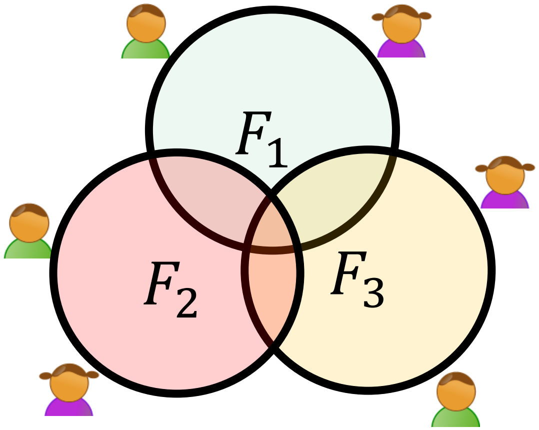
Proof. Each pair i chooses a path independently and uniformly at random from F_i.

$E_{i,j}$: the event that the path chosen by pairs i and j share at least one edge.

Obviously, $p = \Pr(E_{i,j}) \leq \frac{k}{m}$,

Dependency graph, $d < 2n$.

$$4dp < \frac{8nk}{m} \leq 1$$



$\therefore \Pr\left(\bigcap_{i \neq j} \overline{E_{i,j}}\right) > 0$ by Lovasz local lemma.

Application 2: Satisfiability

- If no variable in a k –SAT formula appears in more than $T = \frac{2^k}{4k}$ clauses, then the formula has a satisfying assignment.
- **Proof.**
  - E_i: the ith clause is not satisfied.
  - $p = 2^{-k}$, $d \leq k \cdot T \leq 2^{k-2}$

Theorem[Lovasz Local Lemma]: Let $E_1, E_2, \dots, E_n$ be a set of events, and assume that the following holds:

1. For all i, $\Pr(E_i) \le p$;
2. The degree of the dependency graph given by $E_1, E_2, \dots, E_n$ is bounded by d;
3. $4dp \le 1$.

Then $\Pr(\bigcap_{i=1}^{n} \bar{E}_i) > 0$.

Proof.

$$Pr\left(\bigcap_{i=1}^{n} \bar{E}_i\right) = \prod_{i=1}^{n} \Pr\left(\bar{E}_i \middle| \cap_{j=1}^{i-1} \bar{E}_j\right)$$

$$= \prod_{i=1}^{n} \left(1 - \Pr\left(E_i \middle| \cap_{j=1}^{i-1} \bar{E}_j\right)\right)$$

$$\ge \prod_{i=1}^{n} (1 - 2p) > 0$$

Let $S \subset \{1, \dots, n\}$. We prove that for all $s = 0, \dots, n-1$, if $|S| \le s$, then for all $k \notin S$:

$$\Pr\left(E_k \middle| \cap_{j \in S} \bar{E}_j\right) \le 2p$$

Let $S \subset \{1, \ldots, n\}$. We prove that for all $s = 0, \ldots, n-1$, if $|S| \leq s$, then for all $k \notin S$: $\Pr\left(E_k \middle| \cap_{j \in S} \bar{E}_j\right) \leq 2p$.

Proof. (by induction on s)

Base case: $s = 0$, the results holds for the assumption $\Pr(E_k) \leq p$;

Inductive step: $s > 0$, first we show $\Pr\left(\cap_{j \in S} \bar{E}_j\right) > 0$

$s = 1$: it is true for $\Pr\left(\bar{E}_j\right) \geq 1 - p > 0$

$s > 1$: w.l.o.g. $S = \{1,2, \ldots, s\}$, then

$$Pr\left(\cap_{j \in S} \bar{E}_j\right) = \prod_{i=1}^{s} \Pr\left(\bar{E}_i \middle| \cap_{j=1}^{i-1} \bar{E}_j\right)$$

$$= \prod_{i=1}^{s} \left(1 - \Pr\left(E_i \middle| \cap_{j=1}^{i-1} \bar{E}_j\right)\right)$$

$$\geq \prod_{i=1}^{s} (1 - 2p) > 0$$ by I.H.

Let $S \subset \{1, \dots, n\}$. We prove that for all $s = 0, \dots, n-1$, if $|S| \le s$, then for all $k \notin S$: $\Pr\left(E_k \middle| \bigcap_{j \in S} \bar{E}_j\right) \le 2p$.

Proof. (by induction on s)

Base case: $s = 0$, the results holds for the assumption $\Pr(E_k) \le p$;

Inductive step: $s > 0$, $\boxed{\text{we know } \Pr\left(\bigcap_{j \in S} \bar{E}_j\right) > 0}$

Let $S_1 = \{j \in S \mid (k, j) \in E\}$, and $S_2 = S - S_1$,

Case 1: $S_2 = S$, (i.e. $S_1 = \emptyset$)

then E_k is mutually independent of $\bar{E}_i$, $i \in S$, and

$\Pr\left(E_k \middle| \bigcap_{j \in S} \bar{E}_i\right) = \Pr(E_k) \le p \le 2p$ holds.

Case 2: $|S_2| < s$.

Let $S \subset \{1, \dots, n\}$. We prove that for all $s = 0, \dots, n-1$, if $|S| \le s$, then for all $k \notin S$: $\Pr\left(E_k \middle| \cap_{j \in S} \bar{E}_j\right) \le 2p$.

Proof. (by induction on s)

Inductive step: $s > 0$, we know $\Pr\left(\cap_{j \in S} \bar{E}_j\right) > 0$

Let $S_1 = \{j \in S \mid (k, j) \in E\}$, and $S_2 = S - S_1$,

Case 2: $|S_2| < s$. Let $F_S = \cap_{j \in S} \bar{E}_j$, $F_{S_1} = \cap_{j \in S_1} \bar{E}_j$, $F_{S_2} = \cap_{j \in S_2} \bar{E}_j$

Obviously, $F_S = F_{S_1} \cap F_{S_2}$

$$\Pr(E_k | F_S) = \frac{\Pr(E_k \cap F_S)}{\Pr(F_S)}$$

$$= \frac{\Pr\left(E_k \cap F_{S_1} \cap F_{S_2}\right)}{\Pr(F_{S_1} \cap F_{S_2})} = \frac{\Pr\left(E_k \cap F_{S_1} \middle| F_{S_2}\right) \Pr(F_{S_2})}{\Pr\left(F_{S_1} \middle| F_{S_2}\right) \Pr(F_{S_2})}$$

$$= \frac{\Pr\left(E_k \cap F_{S_1} \middle| F_{S_2}\right)}{\Pr\left(F_{S_1} \middle| F_{S_2}\right)}$$

Let $S \subset \{1, \dots, n\}$. We prove that for all $s = 0, \dots, n-1$, if $|S| \leq s$, then for all $k \notin S$: $\Pr\left(E_k \middle| \bigcap_{j \in S} \bar{E}_j\right) \leq 2p$.

Proof. (by induction on s)

Inductive step: $s > 0$, we know $\Pr\left(\bigcap_{j \in S} \bar{E}_j\right) > 0$

Let $S_1 = \{j \in S \mid (k, j) \in E\}$, and $S_2 = S - S_1$,

Case 2: $|S_2| < s$. Let $F_S = \bigcap_{j \in S} \bar{E}_j$, $F_{S_1} = \bigcap_{j \in S_1} \bar{E}_j$, $F_{S_2} = \bigcap_{j \in S_2} \bar{E}_j$, $F_S = F_{S_1} \cap F_{S_2}$

$$\Pr(E_k | F_S) = \frac{\Pr\left(E_k \cap F_{S_1} \middle| F_{S_2}\right)}{\Pr\left(F_{S_1} \middle| F_{S_2}\right)}$$

Let $S \subset \{1, \ldots, n\}$. We prove that for all $s = 0, \ldots, n-1$, if $|S| \leq s$, then for all $k \notin S$: $\Pr\left(E_k \middle| \cap_{j \in S} \bar{E}_j\right) \leq 2p$.

Proof. (by induction on s)

Inductive step: $s > 0$, we know $\Pr\left(\cap_{j \in S} \bar{E}_j\right) > 0$

Let $S_1 = \{j \in S \mid (k, j) \in E\}$, and $S_2 = S - S_1$,

Case 2: $|S_2| < s$. Let $F_S = \cap_{j \in S} \bar{E}_j$, $F_{S_1} = \cap_{j \in S_1} \bar{E}_j$, $F_{S_2} = \cap_{j \in S_2} \bar{E}_j$, $F_S = F_{S_1} \cap F_{S_2}$

$$\Pr(E_k | F_S) = \frac{\Pr\left(E_k \cap F_{S_1} \middle| F_{S_2}\right)}{\Pr\left(F_{S_1} \middle| F_{S_2}\right)}$$

$$\Pr\left(E_k \cap F_{S_1} \middle| F_{S_2}\right) \leq \Pr\left(E_k \middle| F_{S_2}\right)$$

$= \Pr(E_k) \leq p$ by assumption.

Let $S \subset \{1, \ldots, n\}$. We prove that for all $s = 0, \ldots, n-1$, if $|S| \le s$, then for all $k \notin S$: $\Pr\left(E_k \middle| \cap_{j \in S} \bar{E}_j\right) \le 2p$.

Proof. (by induction on s)

Inductive step: $s > 0$, we know $\Pr\left(\cap_{j\in S} \bar{E}_j\right) > 0$

Let $S_1 = \{j \in S \mid (k, j) \in E\}$, and $S_2 = S - S_1$,

Case 2: $|S_2| < s$. Let $F_S = \cap_{j\in S} \bar{E}_j$, $F_{S_1} = \cap_{j\in S_1} \bar{E}_j$, $F_{S_2} = \cap_{j\in S_2} \bar{E}_j$, $F_S = F_{S_1} \cap F_{S_2}$

$$\Pr(E_k | F_S) = \frac{\Pr\left(E_k \cap F_{S_1} \middle| F_{S_2}\right)}{\Pr\left(F_{S_1} \middle| F_{S_2}\right)}$$

$$\begin{aligned}
\Pr\left(F_{S_1} \middle| F_{S_2}\right) &= \Pr\left(\cap_{i\in S_1} \bar{E}_i \middle| \cap_{j\in S_2} \bar{E}_j\right) \\
&\ge 1 - \sum_{i \in S_1} \Pr\left(E_i \middle| \cap_{j\in S_2} \bar{E}_j\right) \\
&\ge 1 - \textstyle\sum_{i\in S_1} 2p \text{ by I.H.} \\
&\ge 1 - 2pd \ge \frac{1}{2}
\end{aligned}$$

Let $S \subset \{1, \dots, n\}$. We prove that for all $s = 0, \dots, n-1$, if $|S| \le s$, then for all $k \notin S$: $\Pr\left(E_k \middle| \cap_{j\in S} \bar{E_j}\right) \le 2p$

Proof. (by induction on s)

Inductive step: $s > 0$, we know $\Pr\left(\cap_{j\in S} \bar{E_j}\right) > 0$

Let $S_1 = \{j \in S \mid (k, j) \in E\}$, and $S_2 = S - S_1$,

Case 2: $|S_2| < s$. Let $F_S = \cap_{j\in S} \bar{E_j}$, $F_{S_1} = \cap_{j\in S_1} \bar{E_j}$, $F_{S_2} = \cap_{j\in S_2} \bar{E_j}$, $F_S = F_{S_1} \cap F_{S_2}$

$$\Pr(E_k | F_S) = \frac{\Pr\left(E_k \cap F_{S_1} \middle| F_{S_2}\right)}{\Pr\left(F_{S_1} \middle| F_{S_2}\right)} \quad \begin{matrix} \le p \\ \ge \frac{1}{2} \end{matrix}$$

$$\le 2p$$

Lovasz Local Lemma: The General Form

Theorem: Let $E_1, \ldots, E_n$ be a set of events in an arbitrary probability space, and let $G = (V, E)$ be the dependency graph for these events. Assume there exist $x_1, \ldots, x_n \in [0,1)$ such that, for all $i \leq i \leq n$,

$$\Pr(E_i) \leq x_i \cdot \prod_{(i,j) \in E} (1 - x_j)$$

Then

$$Pr\left(\bigcap_{i=1}^{n} \bar{E}_i\right) \geq \prod_{i=1}^{n} (1 - x_i)\,.$$