Lecturer: Prof. Yufei Zhao Notes by: Andrew Lin

Definition 0.1 (Asymptotic notation)

Given functions or sequences $f, g > 0$ (usually of some parameter $n \to \infty$), the notation in each bullet point below are considered equivalent:

- $f \lesssim g$, $f = O(g)$, $g = \Omega(f)$, $f \le Cg$ (for some constant C);
- $f \ll g$, $f = o(g)$, $\frac{f}{g} \to 0$, $g = \omega(f)$.
- $f \asymp g$, $f = \Theta(g)$, $g \lesssim f \lesssim g$.
- $f \sim g$, $\frac{f}{g} \to 1$, $f = (1 + o(1))g$.

Some event holds *with high probability* if its probability is $1 - o(1)$.

Warning: analytic number theorists like to use the Vinogradov notation, where $f \ll g$ means $f = O(g)$ instead of $f = o(g)$ as we do. In particular, $100 \ll 1$ is correct in Vinogradov notation.

1 Introduction to the probabilistic method

In combinatorics and other fields of math, we often wish to show existence of some mathematical object. One clever way to do this is to try to construct this object randomly and then show that we succeed with positive probability.

Proposition 1.1

Every edge $G = (V, E)$ with vertices V and edges E contains a bipartite subgraph with at least $\frac{|E|}{2}$ edges.

Proof. We can form a bipartite graph by partitioning the vertices into two groups. Randomly color each vertex either white or black (making the white and black sets the two groups), and include only the edges between a white and a black edge in a new graph G'. Since all vertices are colored independently at random, each edge is included in G' with probability $\frac{1}{2}$. Thus, we have an average of $\frac{|E|}{2}$ edges in our graph by linearity of expectation, and this means that at least one coloring will work. □

This class will introduce a variety of methods to solve these types of problems, and we'll start with a survey of those techniques.

1.1 The Ramsey numbers

Definition 1.2

Let the **Ramsey number** $R(k, \ell)$ be the smallest n such that if we color the edges of K_n (the complete graph on n vertices) red or blue, we always have a K_k that is all red or a K_ℓ that is all blue.

Theorem 1.3 (Ramsey, 1929)

For any integers k, ℓ, $R(k, \ell)$ is finite.

One way to do this is to use the recurrence inequality

$$R(r, s) \le R(r - 1, s) + R(r, s - 1)$$

by picking an arbitrary vertex v and partitioning the remaining vertices by the color of their edge to v.

Theorem 1.4 (Erdős, 1947)
We have $R(k,k) > n$ for all

$$\binom{n}{k} 2^{1-\binom{k}{2}} < 1.$$

In other words, for any n that satisfies this inequality, we can color K_n with no monochromatic K_k.

Proof. Color the edges of K_n randomly. Given any set R of k vertices, let A_R be the event where R is monochromatic (all $\binom{k}{2}$ edges are the same color). The probability A_R occurs for any given R is $2^{1-\binom{k}{2}}$, since there are only 2 ways to color R, and thus the total probability that K_n is monochromatic is

$$\Pr\left[\bigcup_{R \in \binom{[n]}{k}} A_R\right]$$

and we can "union bound" this: the total probability is at most the sum of the probabilities of the independent events, so

$$\Pr(\text{monochromatic}) \le \sum_R \Pr(A_R) = \binom{n}{k} 2^{1-\binom{k}{2}},$$

and as long as this is less than 1, there is a positive probability that no monochromatic coloring exists, and thus $R(k,k) > n$. □

Fact 1.5
We can optimize Theorem 1.4 with Stirling's formula to find that

$$R(k,k) > \left(\frac{1}{e\sqrt{2} + o(1)}\right) k 2^{k/2},$$

where the $o(1)$ term goes to 0 as $k \to \infty$.

This is a lower bound on the Ramsey numbers. It turns out we can also get an upper bound

$$R(s,s) \le \left(\frac{1}{4\sqrt{\pi}} + o(1)\right) \frac{4^s}{\sqrt{s}}.$$

Currently, this is basically the best we can do: it is still an open problem to make the bases of the exponents tighter than $\sqrt{2}$ and 4.

Remark. *Because the name is Hungarian, the "s" in Erdős is pronounced as "sh," while "sz" is actually pronounced "s."*

1.2 Alterations

We can almost immediately improve our previous bound by a bit.

Proposition 1.6
For all k, n, we have

$$R(k,k) > n - \binom{n}{k} 2^{1-\binom{k}{2}}.$$

Proof. As before, color the edges of K_n randomly. This time, let A_R be the **indicator variable** for a set R of k vertices. (This means that A_R is equal to 1 if R is monochromatic and 0 otherwise.) The expected value of each A_R is just the probability that R is monochromatic, which is $2^{1-\binom{k}{2}}$, so the expected number of monochromatic K_ks is the sum of all A_Rs, which is

$$\mathbb{E}[X] = \binom{n}{k} 2^{1-\binom{k}{2}}.$$

Now delete one vertex from each monochromatic k-clique: we delete X vertices at most (possibly with repeats), so now we have an expected

$$n - \binom{n}{k} 2^{1-\binom{k}{2}}$$

vertices. But this graph has all monochromatic k-cliques removed, and thus there exists a graph with at least this many vertices and no monochromatic k-clique. □

Fact 1.7

Using the same optimization with Stirling's formula on Proposition 1.6,

$$R(k,k) > \left(\frac{1}{e} + o(1)\right) k 2^{k/2},$$

which is better than the result above by a factor of 2.

Both of these proofs are interesting, because although we now know a graph exists, we can't actually construct such an example easily!

1.3 Lovász Local Lemma

We're going to discuss some methods in this class beyond just picking things randomly: here's one of them. Let's say that we are trying to avoid a bunch of bad events $E_1, E_2, \cdots, E_n$ simultaneously. There's two main ways we know how to avoid them:

- All the probabilities are small, and there aren't too many of them. In particular, if the total sum of probabilities is at most 1, we always have a positive chance of success.
- If all the events are independent, then the probability of success is just the product of individual avoidances.

Theorem 1.8 (Lovász Local Lemma)

Let $E_1, \cdots, E_n$ be events each with probability at most p, where each event E_i is mutually independent of all other E_js except at most d of them. If $ep(d+1) \le 1$, then there is a positive probability that no E_i occurs.

Corollary 1.9 (Spencer, 1975)

We have $R(k,k) > n$ if

$$e\left(\binom{k}{2}\binom{n}{k-2} + 1\right) 2^{1-\binom{k}{2}} \le 1.$$

Proof. Randomly color all the edges, and again let A_R be the indicator variable for a subset R of k vertices forming a monochromatic clique. Note that all A_R and A_S are mutually independent unless they share an edge, meaning $|R \cap S| \ge 2$. For each given R, there are at most $\binom{k}{2}\binom{n}{k-2}$ choices for S, since we pick 2 vertices to share with R

and then pick the rest however we'd like. Now by Lovász Local Lemma, we have a positive probability no A_R occurs as long as

$$ep(d+1) = e\left(\binom{k}{2}\binom{n}{k-2}+1\right)2^{1-\binom{k}{2}} \leq 1.$$

□

Fact 1.10

This time, optimizing n in Corollary 1.9 yields

$$R(k,k) > \left(\frac{\sqrt{2}}{e}+o(1)\right)k2^{k/2}.$$

1.4 Set systems

Let $\mathcal{F}$ be a collection of subsets of $[n] = \{1, 2, \cdots, n\}$ (there are a total of 2^n subsets to put in $\mathcal{F}$). We call this an **antichain** if there is no set in $\mathcal{F}$ that is contained in another one.

Our question: what is the largest possible antichain? One thing we can do is to only use subsets of a fixed size k, since no set can be contained in another. This means we can at least get $\binom{n}{\lfloor n/2 \rfloor}$, the largest binomial coefficient. It turns out that this is the best bound:

Theorem 1.11 (Sperner, 1928)

If $\mathcal{F}$ is an antichain of subsets of $[n]$, then it has size at most $\binom{n}{\lfloor n/2 \rfloor}$.

To show this, we'll prove a more slightly general result:

Theorem 1.12

For any antichain $\mathcal{F}$ of the subsets of $[n]$,

$$\sum_{A\in\mathcal{F}}\binom{n}{|A|}^{-1} \leq 1.$$

This implies the result above, because it is a weighted sum where each weight $\binom{n}{|A|}$ is at most $\binom{n}{\lfloor n/2 \rfloor}$ (and the central binomial coefficients are largest).

Proof. Fix a random permutation σ of $[n]$. Associated with this permutation, we have a chain

$$\varnothing \subseteq \{\sigma(1)\} \subseteq \{\sigma(1), \sigma(2) \subseteq \cdots \subseteq \{\sigma(1), \cdots, \sigma(n)\} = [n].$$

Each subset A has probability $P_A = \binom{n}{|A|}^{-1}$ of appearing in such a chain, since each $|A|$-element subset has the same chance of appearing. However, no two subsets can appear in the same chain, so the events are disjoint. Thus, the sum of probabilities that A appears in the chain must be at most 1, and thus

$$\sum_{A\in\mathcal{F}} P_A = \sum_{A\in\mathcal{F}}\binom{n}{|A|}^{-1} \leq 1.$$

□

Theorem 1.13 (Bollobás' Two Families Theorem)
Given r-element sets $A_1, \cdots, A_m$ and s-element sets $B_1, \cdots, B_m$, if we know that

$$A_i \cap B_j = \varnothing \quad \text{if and only if} \quad i = j$$

(all A_i and B_j intersect except for $i = j$), then $m \le \binom{r+s}{r}$.

Where's the motivation for this coming from?

Definition 1.14
Given a family of sets $\mathcal{F}$, let a **transversal** T be a set that intersects all $S \in \mathcal{F}$, and let the **transversal number** $\tau(\mathcal{F})$ denote the size of the smallest transversal of $\mathcal{F}$. $\mathcal{F}$ is **τ-critical** if we have $\tau(\mathcal{F} \setminus S) < \tau(\mathcal{F})$ for all $S \in \mathcal{F}$.

Corollary 1.15 (of Theorem 1.13)
An r-uniform τ-critical family of sets $\mathcal{F}$ with $\tau(F) = s + 1$ has size at most $\binom{r+s}{r}$.

Proof. Let our family of sets be $A_1, \cdots, A_m$. $\mathcal{F}$ being τ-critical implies that for any i, we can find a transversal of size s for $\mathcal{F} \setminus A_i$. Letting this be B_i, notice that $A_i \cap B_j = \varnothing \iff i = j$, and thus by Bollobás' Theorem we can find the upper bound stated. □

Here's a slightly more general version of Bollobás' Theorem, which we'll prove now:

Theorem 1.16
Let $A_1, \cdots, A_m, B_1, \cdots, B_m$ be finite sets, such that $A_i \cap B_j = \varnothing$ if and only if $i = j$. Then

$$\sum_{i=1}^{m} \binom{|A_i| + |B_i|}{|A_i|}^{-1} \le 1.$$

Notice that if we make $B_i = [n] \setminus A_i$ for all i, we get Sperner's theorem. Meanwhile, if all A_is have size r and all B_js have size s, we get Bollobás' Two Families Theorem.

Proof. Like in Sperner's theorem, randomly order all elements in the union of all A_i and B_js. For any i, the probability that all of A_i occurs before all of B_i is $\binom{|A_i|+|B_i|}{|A_i|}^{-1}$, and we can't have this happen with two different is in any given ordering, because this would mean that either A_i and B_j are disjoint or A_j and B_i are disjoint. Thus all events of this form are disjoint, and we must have $\sum_{i=1}^{m} \binom{|A_i|+|B_i|}{|A_i|}^{-1} \le 1$, as desired. □

Definition 1.17
A family of sets $\mathcal{F}$ is **intersecting** if $A \cap B \neq \varnothing$ for all $A, B \in \mathcal{F}$.

Note that this does not mean they must all have a common element: for example, $\{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$ is intersecting.

Theorem 1.18 (Erdős-Ko-Rado 1961)
If $n \geq 2k$, then all intersecting families of k-element subsets of $[n] = \{1, 2, \cdots, n\}$ have size at most $\binom{n-1}{k-1}$.

(This can be constructed by having all sets share the element 1, for example.)

Proof. Order the integers $1, 2, \cdots, n$ around a circle randomly. Let a subset $A \subseteq [n]$ be **contiguous** if all elements lie in a contiguous block around the circle. For any subset A with $|A| = k$; the probability it is contiguous is

$$\binom{n}{\binom{n}{k}},$$

(think of picking k of the spots around the circle). So the expected number of contiguous subsets is $|\mathcal{F}|\binom{n}{\binom{n}{k}}$, but if all subsets are intersecting, we can only have k contiguous subsets (here, as long as $n \geq 2k$, all contiguous subsets must pass through a common point, which is why we set up the problem this way). Thus, $|\mathcal{F}|\binom{n}{\binom{n}{k}} \leq k$, and rearranging yields

$$|F| \leq \frac{k}{n}\binom{n}{k} = \binom{n-1}{k-1},$$

as desired. □

1.5 Hypergraph colorings

This is a topic we'll be discussing quite a bit in this class, but the idea is very similar to that of set systems.

Definition 1.19
A ***k*-uniform hypergraph** $H(V, E)$ has a (finite) set of vertices V and a set of edges E, each of which is a k-element subset of V. H is ***r*-colorable** if we can color V with r colors such that no edge is monochromatic (that is, not all the vertices in an edge have the same color).

(Regular graphs are 2-uniform hypergraphs.) Let $m(k)$ to be the minimum number of edges in a k-uniform hypergraph that isn't 2-colorable.

Example 1.20
A triangle is not 2-colorable, so $m(2) = 3$. The Fano plane is not 2-colorable if we interpret lines as edges, so $m(3) = 7$ (any smaller example can be checked).

These quickly become hard to calculate, though: $m(4) = 23$, but $m(5)$ is actually currently unknown.

Theorem 1.21
A k-uniform hypergraph with fewer than 2^{k-1} edges is 2-colorable.

Proof. Color each vertex randomly; each edge has probability 2^{1-k} of being monochromatic, since all k vertices need to be one color or the other. Thus, if we have less than 2^{k-1} edges, the expected number of monochromatic edges is less than 1, so there is a way to 2-color the hypergraph successfully. □

To date, we have the bounds (which are reasonably close to each other)

$$m(k) \geq 2^k \sqrt{\frac{k}{\log k}} \quad \text{and} \quad m(k) = O(k^2 2^k).$$

How do we show the upper bound? We can restate it as follows:

Problem 1.22
Construct a k-uniform hypergraph with $O(k^2 2^k)$ edges that is not k-colorable.

Solution. Start with a set of vertices V where $|V| = n$, and let H be the hypergraph constructed by choosing m edges $S_1, S_2, \cdots, S_m$ at random. For any coloring of the vertices $\chi : V \to$ red, blue, the event $A(\chi)$ refers to H containing no monochromatic edges. Then our goal is to pick m, n so that

$$\sum_{\chi} \Pr(A_i) < 1,$$

because this means there is a graph H that cannot be properly colored regardless of which χ we pick.

A coloring χ that colors a vertices red and b vertices blue has a given S_i monochromatic with probability

$$\frac{\binom{a}{k} + \binom{b}{k}}{\binom{n}{k}} \geq \frac{2\binom{n/2}{k}}{\binom{n}{k}}$$

(since there are $\binom{n}{k}$ total sets of vertices and $\binom{a}{k} + \binom{b}{k}$ of them are monochromatic). Further bounding, this is

$$\geq 2\left(\frac{n/2 - k + 1}{n - k + 1}\right)^k = 2^{-k+1}\left(1 - \frac{k-1}{n-k+1}\right)^k \geq c2^{-k}$$

where we pick $n = k^2$ so that we can have

$$2\left(1 - \frac{k-1}{n-k+1}\right)^k \geq c,$$

a constant. So now the probability that we have a proper coloring (which means no S_i is monochromatic) is at most (looking at all S_is now)

$$(1 - c2^{-k})^m \leq e^{-c2^{-km}},$$

since we chose our S_is randomly (possibly with replacement), and then $1 + x \leq e^x$ for all x. Therefore, if we sum over all χ, we have

$$\sum_{\chi} e^{-c2^{-km}} = 2^n e^{-c2^{-km}} < 1$$

for some value of $m = O(k^2 2^k)$, as desired. □

Now that we have a sampling of some preliminary techniques, we'll begin examining them in more detail in the next few chapters!