LAKE FOREST COLLEGE

Senior Thesis

The Probabilistic Method

by

Melanie Tian

Dec. 7, 2021

The report of the investigation undertaken as a Senior Thesis, to carry one course of credit in the Department of Mathematics.

Davis Schneiderman
Krebs Provost and Dean of the Faculty

Enrique Treviño, Chairperson

Arthur Bousquet

Michael M. Kash

Contents

Abstract

We adopt a problem solving based approach to introduce the probabilistic method in combinatorics, with Ramsey Numbers and Ramsey Theory as frequent guest objects.

In Section 2 we introduce the spirit of the method, then present the *ninefree* problem (Dartmouth Math 108 Winter 2008 HW1 #5) as a quintessential example.

In Section 3 we introduce the Ramsey Numbers and some elementary bounds.

In Section 4 we introduce the Ramsey Theory and some famous results like Schur's and Van der waerden's.

In Section 5 we introduce the probablistic method with both contest problems and "real math" research results.

In Section 6 we explore the upper bounds of $R(k, k)$ using tools like alterations and Lovász Local Lemma. We also present a brief history of research results.

Motivation

(Personal, not mathematical.)

Professor Treviño expressed his love for the“ninefree” problem (Dartmouth Math 108 Topics in Combinatorics, Winter 2008, Problem Set 1 #5) in a zealous manner early this morning in his office after we exchanged our views on the appreciation of Spider-man. Then I vaguely recall there is indeed some problem has this ninefree shape in one of the linear algebra midterms he created, though that problem is too easy to be fun. Then I went on a Metra train to downtown Chicago later that day like a normal 21-year-old who has a social life. But when I came back to campus around ten that night, I felt a mysterious urge to google ninefree matrix. Then a PDF solution of that problem came up for some reason, and I have zero self-control so I read the whole thing.

I was first appalled at how smooth and brief the reading experience was, because at my intelligence level usually reading “hard math” requires active mental preparation. Then I was convinced how powerful the probabilistic method is, since every move feels “too good to be true”. However, the process has a nostalgia touch, since I have seen problems using expected value to force the existence of something “above average”. It even has a similar spirit like pigeonhole, you try to be all fair and square and even and such. Also, there is this satisfying and virtuous feeling of solving (ok, watching people solving) a problem in an elegant and potentially “cosmopolitan” way.

So I was all wired up by that problem (solution) and this thesis writing endeavor began as an impulse move at 22:30 (with no coffee/drugs) on Wednesday September 8 (and the first draft was done on Friday night).

Acknowledgment

I would like to thank professor Enrique Treviño and every professor in the math department at Lake Forest College, who praised me too much, supported me too often, made my best memories here, and opened once unthinkable possibilities for my future.

I would like to thank my highschool math teacher Yuanyuan Xu (徐媛媛), who believed in me when I didn't believe in myself. Without her, there is less than ϵ ($\forall\epsilon > 0$) probability that I didn't give up math. She is my hero.

Probability bridges these two realms. The wild, woolly world we know is what makes probability necessary. The calm, measured world we can never quite reach is what makes probability possible. As mortal creatures, we'll never set foot in the land of eternity — but probability can offer us a glimpse. [12]

1 Introduction

We will write in the form that we present a problem, then present one or more solutions, then another problem, then solution(s), etc. The main reasons for adopting this approach are as following.

(i) The author is very comfortable with this approach, and has fond memories being taught math in this approach.

(ii) Each problem serves as a good motivation for the technique(s) in the solution(s).

(iii) Each problem serves as a good motivation for keeping reading, because you might want to be able to solve it.

(iv) For more discussions about this approach, we invite you to read Timothy Gowers's The Two Cultures of Mathematics [5].

We assume the reader knows basic graph theory. If not, there is an appendix at the end explaining some basic graph theory definitions.

2 Linearity of Expectation

Let's start with some well-known old toy problems that use linearity of expectation to be solved to get a hold of the spirit of the techniques.

Problem 1. Consider a complete graph K_{3n+1}, $n \in \mathbb{N}$, every vertex is connected to exactly n red edges, n blue edges, and n green edges. Prove there exists a triangle such that all three sides have different colors.

Solution 1. We consider the number of "pairs of edges of different colors sharing a vertex". For every vertex, there are $3n^2$ of those (pick a color combi-

nation then pick the 2 edges), so there are $3n^2(3n+1)$ of them in total.
The expected value of the number of those pairs that every triangle has

$$\frac{3n^2(3n+1)}{\binom{3n+1}{3}} = \frac{6n}{3n-1} > 2.$$

So some triangle must have 3 of them.

Problem 2 (HMMT 2006). At a nursery, 2006 babies sit in a circle. Suddenly, each baby randomly pokes either the baby to its left or to its right. What is the expected value of the number of unpoked babies?

Solution 2. [3] For all x_i, $i \in \{1, 2, ..., 2006\}$, we denote $x_i = 1$ if baby $\#i$ is unpoked and $x_i = 0$ otherwise. We want the expected value of $\sum x_i$.
For every baby, the probability of being unpoked is $(\frac{1}{2})^2 = \frac{1}{4}$. By linearity of expectation,

$$\mathbb{E}[x_1 + x_2 + ... + x_{2006}] = \mathbb{E}[x_1] + \mathbb{E}[x_2] + ... + \mathbb{E}[x_{2006}] = 2006 \times \frac{1}{4} = \frac{1003}{2}$$

is exactly the expected number of unpoked babies since we defined x_i as an indicator function.

Problem 3 (IMC 2002). Two hundred students participated in a mathematical contest. They had 6 problems to solve. It is known that each problem was correctly solved by at least 120 participants. Prove that there must be two participants such that every problem was solved by at least one of these two students..

Solution 3. [3] For every contestant, the probability of them missing a problem is at most $1 - \frac{120}{200} = \frac{2}{5}$. So the probability for a pair of contestants both missing a problem is at most $(\frac{2}{5})^2 = \frac{4}{25}$. Therefore, the expected number of problems a pair of contestants both miss is at most $6 \times \frac{4}{25} < 1$. So there exists such a pair of contestants we want.

Now let's get to the star of tonight.

Problem 4 (Dartmouth Math108 2008). [4] Let $A = (a_{ij})$ be an $n \times n$ matrix with all $a_{ij} \in \{0, 1\}$. We call A *ninefree* if there is no 3×3 submatrix consisting of all ones. (Note: the rows and columns of a submatrix don't need to be consecutive.) Let $f(n)$ denote the maximal number of ones in a ninefree $n \times n$ matrix. Give a lower bound on $f(n)$ by the following argument: Consider a random matrix in which each coefficient is one with independent probability p. Now for each 3×3 submatrix consisting of all ones change one of the ones to a zero. This gives a ninefree matrix. Give a precise theorem giving a lower bound for $f(n)$.

Solution 4. [6] Let $x_{i,j}$ be the random variable for the i, j-entry of A. Define

$$X = \sum_{i,j} x_{i,j}.$$

The expected number of 1's in A is $\mathbb{E}[X]$.

$$\mathbb{E}[X] = \sum_{i,j} \mathbb{E}[X_{i,j}] = n^2 p.$$

There are $\binom{n}{3}\binom{n}{3}$ 3×3 submatrices. Let $y_i = 1$ if submatrix i is all 1's and $y_i = 0$ otherwise. Define

$$Y = \sum_i y_i.$$

We have

$$\mathbb{E}[Y] = \sum \mathbb{E}[y_i] = \binom{n}{3}\binom{n}{3}p^9 < \frac{n^6}{36}p^9.$$

Then

$$\mathbb{E}[X - Y] = \mathbb{E}[X] - \mathbb{E}[Y] > n^2 p - \frac{n^6}{36}p^9.$$

We want to maximize this. We will differentiate it in terms of p and set it to 0 to get

$$n^2 - \frac{9n^6}{36}p^8 = 0$$
$$p = \sqrt[8]{4}n^{-\frac{1}{2}}.$$

So we have our lower bound as

$$f(n) > \mathbb{E}[X] - \mathbb{E}[Y] = \sqrt[8]{4}n^{\frac{3}{2}} - \frac{\sqrt[8]{4}}{9}n^{\frac{3}{2}} = \frac{8}{9}\sqrt[8]{4}n^{\frac{3}{2}}.$$

Notice how the last problem is more sophisticated in the sense that we introduced the alterations method (we "remove" obstructions) and so on and we warn you that certain techniques in it will be recurrent.

3 Ramsey Numbers

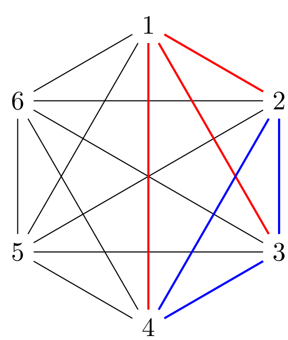
Now let's get sidetracked a little bit and bring out a famous object, the Ramsey Numbers.

We follow the long standing tradition around the world by introducing them with the following classic problem.

Problem 5. Prove that among any 6 people in the world, either there are 3 of them who all know each other, or there are 3 of them such that no one knows each other.

Solution 5. It translates to proving there exists a monochromatic triangle in a 2-colored K_6.

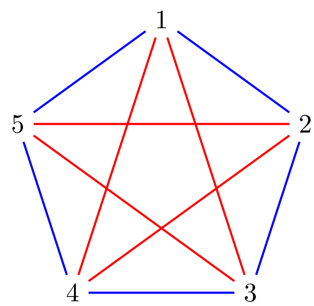
Assume the opposite. We denote the edge between vertex i and j as e_{ij}. Without loss of generality, let e_{12}, e_{13}, e_{14} be red (by pigeonhole principle, among the 5 edges from v_1, 3 of them are of the same color), now e_{23}, e_{34}, e_{24} all have to be blue (otherwise there will be red triangles). Then, we have a blue $\triangle v_2v_3v_4$, contradiction.



But this only proved $R(3,3) \leq 6$, in order to show $R(3,3) = 6$, we need to give a construction of K_5 such that there is no monochromatic triangle.

Problem 6. Give a construction of K_5 such that there is no monochromatic triangle.

Solution 6.



Now we have $R(3,3) = 6$.

It might seem like these kind of problems are easy and bashy, or at least doable. For example, for $R(3,3,3) \leq 17$, it's simply just "repeat the process".

Problem 7 (IMO 1964 P4). Seventeen people correspond by mail with one another - each one with all the rest. In their letters only three different topics are discussed. Each pair of correspondents deals with only one of these topics. Prove that there are at least three people who write to each other about the same topic.

Solution 7. Among the 16 edges from v_1, 6 of them are of the same color, say color 1. If among these 6 vertices, there are edges of color 1 then we are done. Otherwise there are only 2 colors left, by problem 5 we are done.

Now it is time for the formal definition of Ramsey Numbers. Although even if you have not seen it before, by the examples above you should have already guessed it.

Theorem 1 (Ramsey). $\forall\, s, t \in \mathbb{N}$, $\exists\, n \in \mathbb{N}$ such that any 2-colored K_n contains either a red K_s or a blue K_t.
The least such n is the Ramsey Number $R(s, t)$.
(The fact that $R(s, t)$ even exists will be proved by problem 8.)

Currently we know $R(4, 4) = 18$, but for $n \geq 5$, we only have (not so tight) bounds for some $R(n, n)$, like $R(5, 5) \in [43, 48]$, $R(10, 10) \in [798, 23556]$ [15].

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, we should marshal all our computers and all our mathematicians and attempt to find the value. But suppose, instead, that they asked for $R(6,6)$, we should attempt to destroy the aliens.
–Paul Erdős

But we could try to observe some elementary bounds now, grabbing some low-hanging fruit.

Problem 8. Prove (i) $R(1, n) = n$.
(ii) $R(s, t) \leq R(s - 1, t) + R(s, t - 1)$.

Solution 8. (i) Either the edges are all blue (red), or there is an blue (red) edge.

(ii) (Because man these are a lot of vertices.)

Pick a vertex, it has either $\geq R(s-1,t)$ red neighbors or $\geq R(s,t-1)$ blue neighbors. The former allows us to build a red K_s or a blue K_t, the latter allows us to build a blue K_t or a red K_s.

Problem 9. [10] Prove for $s,t \geq 2$,

$$R(s,t) \leq \binom{s+t-2}{s-1}.$$

Solution 9. We have

$$\binom{(s-1)+t-2}{s-2} + \binom{s+(t-1)-2}{s-1} = \binom{s+t-2}{s-1}$$

by Pascal's Triangle.

Since $R(2,2) = 2 \leq \binom{2}{1}$, by induction we are done.

Just by the way, about K_6, we could achieve a "better" result by just analyzing it a little bit closer. Also it has a somewhat similar taste as "the probabilistic method".

Problem 10. [10] Prove there exists 2 monochromatic triangles in a 2-colored K_6.

Solution 10. We call a pair of edges which share a vertex and have the same color as a *good pair*. (Like in Problem 1.)

Let the number of monochromatic triangles be x,

so the number of non-monochromatic triangles is $\binom{6}{3} - x = 20 - x$.

Count the number of good pairs.

For each monochromatic triangle, there are 3. For each non-monochromatic triangle, there is 1 by pigeonhole.

So in total there are $3x + 20 - x = 20 + 2x$ good pairs.

Now consider a vertex which has r red edges, the number of good pairs from this vertex is

$$\binom{r}{2} + \binom{5-r}{2} \geq \binom{3}{2} + \binom{2}{2} = 4.$$

So we have

$$20 + 2x \geq 6 \times 4$$
$$x \geq 2.$$

4 Ramsey Theory

Complete disorder is impossible.

– T. S. Motzkin

We consider problems like: Color a large structure however you like it, then there must be some small substructure which is monochromatic.

Problem 11 (IMO 1978 P6). An international society has its members from six different countries. The list of members has 1978 names, numbered $1, 2, \ldots, 1978$. Prove that there is at least one member whose number is the sum of the numbers of two (not necessarily distinct) members from his own country.

Solution 11.1. Consider members $1, 2, ..., 989$. Say they belong to $k \in [1, 6]$ countries. By pigeonhole, there exists a country which has at least $\lfloor \frac{988}{k} \rfloor + 1$

members. Let them be $a_1, a_2, ..., a_{\lfloor \frac{988}{k} \rfloor + 1}$, where $a_i \in [1, 989]$.

Notice we pick 989 because that is half of 1978, so $a_i + a_j \in (1, 1978)$.

As long as we have

$$\binom{\lfloor \frac{988}{k} \rfloor + 1}{2} > 1978 - \left(\left\lfloor \frac{988}{k} \right\rfloor + 1 \right), \quad (*)$$

then we could guarantee some of those sums will have to fall into its own country.

We checked that $(*)$ does hold for $k = 1, 2, 3, 4, 5, 6$.

Solution 11.2. [8] We partition $\{1, 2, ..., 1978\}$ into 6 color classes $C_1, ..., C_6$. Consider a graph K_{1978}. For the edge connecting vertex i and j, we color it color r if $|i - j| \in C_r$.

Theorem 2. [7] A k-colored graph $K_{\lfloor ek! \rfloor + 1}$ contains a monochromatic triangle for $k \geq 2$.

Now since $1978 > \lfloor e6! \rfloor + 1$, there exists $i, j, k \in \{1, 2, ..., 1978\}$ such that $|i - j|$, $|j - k|$, $|k - i|$ belong to the same color class. So we have some two from the same country sum to the third.

About this $x + y = z$ structure, there is a big open problem (which would be too ambitious to attack here).

Theorem 3 (Schur). $\forall r \in \mathbb{N}$, $\exists n \in \mathbb{N}$ such that every r-colored $\{1, 2, ..., n\}$ contains a monochromatic triple (x, y, z) such that $x + y = z$.

The least such n is the Schur's number $S(r)$.

The only known values are $S(1) = 2$, $S(2) = 5$, $S(3) = 14$, $S(4) = 45$, $S(5) = 161$. The proof that $S(5) = 161$ was announced in 2017 and took up 2 petabytes of space [16].

What if we change $x + y = z$ into $x + y = 2z$?

Theorem 4 (Van der Waerden). $\forall\, r, k \in \mathbb{N}$, $\exists\, n \in \mathbb{N}$ such that every r-colored $\{1, 2, ..., n\}$ contains a monochromatic arithmetic progression with length at least k.
The least such n is the Van der Waerden number $W(r, k)$.

Theorem 4 (Timothy Gowers).

$$W(r,k) \leq 2^{2^{r^{2^{2^{k+9}}}}}.$$

Which is the best upper bound currently known. Word on the street that this is related to this (much much easier) problem (and Timothy Gowers went to this IMO himself and got a perfect score).

Problem 12 (IMO 1981 P6). The function $f(x, y)$ satisfies
(1) $f(0, y) = y + 1$,
(2) $f(x + 1, 0) = f(x, 1)$,
(3) $f(x + 1, y + 1) = f(x, f(x + 1, y))$,
$\forall x, y \in \mathbb{N}_0$. Determine $f(4, 1981)$.

Solution 12. [2] We have $f(1, 0) = f(0, 1) = 2$.
Then $f(1, y + 1) = f(0, f(1, y)) = f(1, y) + 1$.
So $f(1, y) = y + 2$ by induction.
Similarly, $f(2, 0) = f(1, 1) = 3$ and $f(2, y+1) = f(2, y)+2$ gives $f(2, y) = 2y+3$.
Continue in the same fashion,

$f(3,0)+3=8$, $f(3,y+1)+3=2(f(3,y)+3)$, we get $f(3,y)+3=2^{y+3}$.
$f(4,0)+3=2^{2^2}$, $f(4,y)+3=2^{f(4,y)+3}$.
So $f(4,1981)=2^{2^{\cdot^{\cdot^{2}}}}-3$ where there are 1984 twos.

However, these might lead you to think that it's hard to not have any arithmetic progression. After all, *complete disorder is impossible*, right?

Problem 13 (IMO 1983 P5). Is it possible to choose 1983 distinct positive integers, all less than or equal to 10^5, no three of which are consecutive terms of an arithmetic progression? Justify your answer.

Solution 13. [8] For all $i \in \mathbb{N}$ we do the following: write i in base 2 but treat it like a base 3 number and call it a_i.
We claim that the set $\{a_1, \ldots, a_{1983}\}$ works.
Because if there exists $2a_i = a_j + a_k$, first $2a_i$ only has 0s and 2s as its digits, and for any 0, a_j and a_k must have 0 in that place as well, and for any 2, a_j and a_k must have 1 in that place, since we are in the base 3 world. So $a_j = a_k$, contradiction.
Notice $1983 = 11110111111_2$, $a_{1983} = 11110111111_3 < \frac{3^{11}}{2} < 10^5$ so the answer is yes.

Anyone who has not seen this problem (I guess this is a famous example for binary constructions, but still) but happens to have heard of Van der Waerden's might guess the answer is NO!!! within 10 seconds and then get stuck forever... .

5 Probabilistic Method

The problems in this section showcase the versatility of using probabilistic methods to attack problems in extremal combinatorics, graph theory, number theory, algebra and analysis.

Problem 14 (All-Russian 1996). In the Duma there are 1600 delegates, who have formed 16000 committees of 80 persons each. Prove that one can find two committees having no fewer than four common members.

Solution 14.1. Assume every 2 committees share ≤ 3 people.
We count the number of triples (i, C_1, C_2) where person i is on both C_1 and C_2 in two ways. Say person i is on a_i many committees.
On one hand, The number of such triples is

$$n = \sum \binom{a_i}{2} \geq 1600 \binom{\frac{\sum a_i}{1600}}{2} = 1600 \binom{\frac{80 \times 16000}{1600}}{2} = 1600 \binom{800}{2}.$$

On the other hand, since every 2 committees share at most 3 people

$$n \leq 3 \binom{16000}{2}.$$

However, $1600\binom{800}{2} > 3\binom{16000}{2}$, contradiction.

Solution 14.2. [11] We randomly choose 2 committees. Let X be the number of persons on both chosen committees, notice $X = \sum X_i$, where X_i is the $\{0, 1\}$ random variable telling whether the i-th person is on both committees.
By linearity of expectation, $E[X] = \sum E[X_i]$.

Say person i is on a_i many committees.

$$E[X_i] = \frac{\binom{a_i}{2}}{\binom{16000}{2}},$$

$$E[X] \geq 1600 \frac{\binom{\frac{16000 \times 80}{1600}}{2}}{\binom{16000}{2}} = 1600 \frac{800 \times 799}{16000 \times 15999} \approx 4$$

which is a little bit less than 4. So some choice would have $X \geq 4$ since it should be an integer.

Most people would agree solution 2 is "more natural" and the triple we considered in solution 1 is somewhat coming out of nowhere (though some might argue once you have seen enough problems everything is natural).

Problem 15 (MOP 2007). In an $n \times n$ array, each of the numbers $1, 2, ..., n$ appears exactly n times. Show that there is a row or a column in the array with at least $\sqrt{n}$ distinct numbers.

Solution 15. [11] Choose a random row or column. Let X be the number of distinct numbers in it.

Then $X = \sum I_i$, where I_i is the indicator variable of i appearing.

We have $\mathbb{E}[I_i] = \mathbb{P}[I_i = 1]$.

To have a lower bound, we consider the worst case scenario that it clustered in some $\sqrt{n} \times \sqrt{n}$. (This is indeed the worst case scenario because, if it's not in a $\sqrt{n} \times \sqrt{n}$ cluster, we first move everything to the top left corner by shifting rows and/or columns, then without loss of generality we assume there are a rows and b columns where $a > \sqrt{n}$ and $b \leq \sqrt{n}$. We claim $a + b \geq 2\sqrt{n}$. Because to minimize $a + b$, we fill out each of the a rows one by one, and we have $ab \geq n$. Now by AM-GM, $a + b \geq 2\sqrt{ab} \geq 2\sqrt{n}$.)

So $\mathbb{P}[I_i = 1] \geq \frac{2\sqrt{n}}{2n} = \frac{1}{\sqrt{n}}$.
So $\mathbb{E}[X] \geq n \times \frac{1}{\sqrt{n}} = \sqrt{n}$.

Problem 16 (MOP 2008). Let $a, b, c \in \mathbb{R}^+$, assume that for all $n \in \mathbb{Z}$,

$$\lfloor an \rfloor + \lfloor bn \rfloor = \lfloor cn \rfloor.$$

Prove that at least one of a, b, c is an integer.

Solution 16. [11] Assume none of them is an integer. Dividing by n and taking the limit as $n \to \infty$ gives $a + b = c$. So we also have

$$\{an\} + \{bn\} = \{cn\} \quad (*)$$

We know that if x is irrational, $\{xn\}$ is equidistributed [14] over $(0, 1)$. So if we choose n randomly from $\{1, 2, ..., N\}$ when $N \to \infty$, $E[\{xn\}] \to \frac{1}{2}$.
If x is rational and $x = \frac{p}{q}$ (with p, q in lowest terms), then $E[\{xn\}] \to \frac{q-1}{2q} \in [\frac{1}{4}, \frac{1}{2})$.
($E[\{xn\}] \to \frac{q-1}{2q}$ because when $(p, q) = 1$, from p to $(q-1)p$ we run through all possible remainders of q, so the average is $\frac{1+\cdots+(q-1)}{q} \times \frac{1}{q} = \frac{q-1}{2q}$.)
So $E[\{xn\}] \to$ something $\in [\frac{1}{4}, \frac{1}{2}]$ if x is not an integer.
So consider the limits, the only way to achieve (*) is to have $E[\{an\}]$ and $E[\{bn\}]$ $\to \frac{1}{4}$ and $E[\{xn\}] \to \frac{1}{2}$. So a, b is rational and c is irrational, contradiction.

Problem 17 (Crossing Lemma). Draw a graph with V vertices and E edges in the plane, for $E \geq 4V$, there are at least $\frac{E^3}{64V^2}$ pairs of crossing edges.

Solution 17. [11] It is well known that for planar graphs $E \leq 3V - 6$ holds (by Euler's V + F - E = 2).

So at least the crossing number $\geq E-(3V-6) > E-3V$. (*)

Take a graph with c crossings, sample vertices randomly with probability p.

Now V becomes pV, E becomes p^2E, c becomes p^4c.

But $(*)$ still holds so

$$p^4c > p^2E - 3pV.$$

We get $c \geq \frac{E^3}{64V^2}$ by choosing $p = \frac{4V}{E}$ (which shouldn't be too surprising since $E \geq 4V$ is there for a reason).

Problem 18 (Erdős, 1965). Prove that every set A of nonzero integers has a subset S such that $|S| > \frac{|A|}{3}$ and S is sum-free.

Solution 18. Pick a prime $p = 3k+2$ and p is greater than twice the maximum absolute value of any element in A.

Let $B = \{k+1, ..., 2k+1\}$, which is sum-free mod p.

Then randomly pick $x \in \{1, 2, ..., p-1\}$ and let C be the set of each element of A multiplied by x mod p.

For each element, the probability of falling into B is $\frac{|B|}{p-1} > \frac{1}{3}$. So the expected number of elements falling into B is $> \frac{|A|}{3}$, so such S exists.

Problem 19 (Alon-Spencer, Theorem 3.2.1). Prove that every n-vertex graph with $\frac{nd}{2}$ edges has a subset U of pairwise nonadjacent vertices of size $|U| \geq \frac{n}{2d}$.

Solution 19. [1] We use the alterations method. Pick a random subset U_0 by taking each element independently with probability p. Then for every edge that survives to U_0, delete one of its endpoints, and call the result U. This has expected size

$$E[U] = np - (\frac{nd}{2})p^2$$

We choose $p = \frac{1}{d}$ and get $E[U] = \frac{n}{2d}$.

Problem 20 (Alon-Spencer, Exercise 2.9). [1] Suppose that every vertex of a n-vertex bipartite graph is given a personalized list of $> \log_2 n$ possible colors. Prove that it is possible to give each vertex a color from its list such that no two adjacent vertices receive the same color.

Solution 20. Let the bipartition of the vertex set be $V_1 \cup V_2$. Let X be the total set of all the colors that ever appear in any list. For each color, flip a fair coin.

The idea is that we are done if for all $v \in V_1$, we can choose a color that flips a head, and for all $v \in V_2$, we can choose a color that flips a tail.

Let N be the number of vertices that fails such. For each vertex, the probability of failure

$$< \frac{1}{2^{\log_2 n}} = \frac{1}{n}.$$

By linearity of expectation $E[N] < 1$ so it is possible that $N = 0$.

6 $R(k, k)$ Bounds

In this section we gradually develop better and better bounds for $R(k, k)$.

Let's start with an "easy" construction first.

Problem 21 (Erdős, 1947). Prove $R(k, k) > n$ if

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

Solution 21. [9] We randomly color K_n. For any set S of k vertices, let A_S be the event that S is monochromatic. The probability of A_S is

$$\frac{2}{2^{\binom{k}{2}}} = 2^{1-\binom{k}{2}}$$

for any S. Because there are 2 good colorings and $2^{\binom{k}{2}}$ total colorings. There are $\binom{n}{k}$ such S, so as long as

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

there exists a coloring of K_n with no monochromatic K_k.

The next problem uses a technique that appeared in problem 4 before.

Problem 22 (Alterations). [9] Prove

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

Solution 22. We randomly color K_n. For any set S of k vertices, this time let A_S be the indicator variable ($A_S = 1$ if S is monochromatic and 0 otherwise). For each S,

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

so the expected number of monochromatic K_k is

$$\binom{n}{k} 2^{1-\binom{k}{2}} \quad (*),$$

Now if we delete one vertex from each monochromatic k-clique, we delete at most (*) many vertices.

So we have an expected

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

vertices with no monochromatic k-clique.

The next improvement comes from a well-known and powerful lemma often refered to as LLL.

Problem 23 (Lovász Local Lemma). [9] Prove $R(k,k) > n$ if

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

Solution 23.

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

We randomly color K_n. For any set I of k vertices, let A_I be the indicator variable. Note that A_S and A_T are independent unless they share 2 vertices (an edge). For each S, there are at most $\binom{k}{2}\binom{n}{k-2}$ choices for S to have $|R \cap S| \geq 2$. So by Lovász Local Lemma we have a positive probability no A_I occurs as long as

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

Hall of Fame

Over the last a few decades, people are fond of setting world records for the upper bound. Here we present some of them [13].

Erdős and Szekeres, 1935.

$$R(s+1,t+1) \le \binom{s+t}{s}.$$

Rödl, 1980s, unpublished. For some $c > 0$,

$$R(s+1,t+1) \le \frac{\binom{s+t}{s}}{c\log^c(s+t)}.$$

Thomason, 1988. For some $A > 0$ when $s > t$,

$$R(s+1,t+1) \le s^{\frac{-t}{2s}+\frac{A}{\sqrt{\log s}}}\binom{s+t}{s}.$$

Conlon, 2009. For some $c > 0$,

$$R(k+1,k+1) \le k^{\frac{-c\log k}{\log\log k}}\binom{2k}{k}.$$

Ashwin Sah (PhD student since Fall 2020), May 2020.

For some $c > 0$ when $k \ge 3$,

$$R(k+1,k+1) \le e^{-c(\log k)^2}\binom{2k}{k}.$$

望峰息心，窥谷忘反。

Appendix. Abbreviations and Glossary

Note: Qualifications for the glossary are based on professor recommendations.

HMMT: Harvard-MIT Math Tournament

IMC: International Mathematics Competition for University Students

IMO: International Mathematical Olympiad

MOP: Mathematical Olympiad Summer Program

Graph (finite simple graph): Consider a bunch of vertices, the only legal operation is to pick two vertices and connect them with an edge (if they are not already connected). Anything generated by legal operations is a graph.

Complete graph K_n: Given a set of n vertices, everything that could possibly be connected between them is connected.

Triangle: Three vertices and three edges. Every pair of these three vertices are connected by one of the three edges.

Monochromatic: All of the edges have the same color.

k-colored: Every edge is colored with one of the k colors.

AM-GM: Arithmetic Mean-Geometric Mean Inequality.

Sum-free: Set A is sum-free if the equation $a + b = c$ has no solution with $a, b, c \in A$.

References

[1] Noga Alon and Joel H. Spencer. *The probabilistic method*. Second. Wiley-Interscience Series in Discrete Mathematics and Optimization. With an appendix on the life and work of Paul Erdős. Wiley-Interscience [John Wiley & Sons], New York, 2000, pp. xviii+301. ISBN: 0-471-37046-0. DOI: 10.1002/0471722154. URL: https://doi.org/10.1002/0471722154.

[2] AOPS. *1981 IMO Problems/Problem 6*. https://artofproblemsolving.com/wiki/index.php/1981_IMO_Problems/Problem_6. Accessed: 2021-11-29.

[3] E. Chen. *The OTIS Excerpts*. https://web.evanchen.cc/textbooks/OTIS-Excerpts.pdf. Accessed: 2021-11-29.

[4] S. Elizalde. *Math 108. Topics in combinatorics: The probabilistic method. Assignment 1.* https://math.dartmouth.edu/archive/m108w08/public_html/ass1.pdf. Accessed: 2021-11-29.

[5] T. Gowers. *The Two Cultures of Mathematics*. https://www.dpmms.cam.ac.uk/~wtg10/2cultures.pdf. Accessed: 2021-11-29.

[6] F. C. Graham. *Math 262A Midterm Solutions*. http://math.ucsd.edu/~fan/math262/midsol.pdf. Accessed: 2021-11-29.

[7] Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer. *Ramsey theory*. Wiley Series in Discrete Mathematics and Optimization. Paperback edition of the second (1990) edition [MR1044995]. John Wiley & Sons, Inc., Hoboken, NJ, 2013, pp. xiv+196. ISBN: 978-1-118-79966-6.

[8] B. Green. *Ramsey Theory and the IMO*. The Mathematical Gazette, vol. 86, no. 506, 2002, pp. 204–207. JSTOR, https://doi.org/10.2307/3621841. Accessed: 2021-11-29.

[9] A. Lin. *The Probabilistic Method in Combinatorics Lecture Notes*. https://ocw.mit.edu/courses/mathematics/18-218-probabilistic-method-in-combinatorics-spring-2019/lecture-notes/MIT18_218S19_full_notes.pdf. Accessed: 2021-11-29.

[10] P. Loh. *Graph Theory*. https://www.math.cmu.edu/~ploh/docs/math/mop2008/graph-theory-soln.pdf. Accessed: 2021-11-29.

[11] P. Loh. *Probabilistic Methods in Combinatorics*. https://www.math.cmu.edu/~ploh/docs/math/mop2009/prob-comb.pdf. Accessed: 2021-11-29.

[12] B. Orlin. *Math with bad drawings : illuminating the ideas that shape our reality*. Running Press Book Publishers, 2018.

[13] Ashwin Sah. *Diagonal Ramsey via effective quasirandomness*. 2020. arXiv: 2005.09251 [math.CO].

[14] H. Weyl. *Ueber die Gleichverteilung von Zahlen mod. Eins*. Math. Ann. 77 (3): 313–352. doi:10.1007/BF01475864. S2CID 123470919.

[15] Wikipedia. *Ramsey's theorem*. https://en.wikipedia.org/wiki/Ramsey\%27s_theorem. Accessed: 2021-11-29.

[16] Wikipedia. *Schur's theorem*. https://en.wikipedia.org/wiki/Schur\%27s_theorem. Accessed: 2021-11-29.