**Advanced Topics in Combinatorics** **Summer 2018**

# Lecture 3: Erdős, Graph Coloring, and the Probabilistic Method

*Lecturer: Kavish Gandhi*

## 3.1 Introduction

The probabilistic method is based on a very simple idea: if the probability of an event is positive, it must be possible! The method was pioneered by and used frequently by Paul Erdős, a Hungarian mathematician who wrote almost 1500 papers. The goal of this lecture is to learn about the probabilistic method and see how Erdős applied it to seemingly unrelated topics like graph theory, number theory, and group theory.

## 3.2 The Pigeonhole Principle

The proof of the pigeonhole principle, a basic tool in combinatorics, using the probabilistic method gives a good, simple sense of how the method is used in general.

**Theorem 3.1** (The Pigeonhole Principle)**.** *Given $n$ boxes and $n+1$ objects, if all the objects are put into a box, then there exists a box with more than one object.*

This is a very intuitive theorem, and it is most often proved using a simple proof by contradiction. The probabilistic method is more complicated, but it illustrates the general idea of the probabilistic method very well. We start with the following definition of a random variable.

**Definition 3.2.** A random variable $X$ is a mapping $\Omega \to \mathbb{R}$ where $\Omega$ is a probability space.

The above may look complicated, but it's actually a very simple concept. A probability space consists of a set of outcomes and the random variable is assigning each of these outcome's "values." For example, the results of a coin flip form a probability space with the outcomes "heads" and "tails," with a possible corresponding random variable mapping "heads" to 1 and "tails" to 0. Another probability space could be the result of a dice roll, and the corresponding random variable could be the number rolled. There is no need for these specific random variables, however, and the mapping can be totally arbitrary. For the coin flip example, one could create a trivial random variable where both outcomes map to 1 or create a random variable where "heads" maps to $\pi$ and "tails" maps to $e$.

We now introduce the expected value of a random variable.

**Definition 3.3.** The expected value $\mathbb{E}[X]$ of a random variable $X$ is given by

$$\mathbb{E}[X] = \sum_{x \in \Omega} X(x)p(x)$$

The expected value is a weighted average of the values the random variable achieves with the weights being the probabilities. For the coin flip random variable where "heads" maps to 1 and "tails" maps to 0, we have an expected value of $1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = \frac{1}{2}$. For the rolls of a six-sided die, the expected value of the roll is

$$1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{7}{2}$$

Now we have all the tools we need to prove the Pigeonhole Principle probabilistically.

*Proof of Theorem 3.1.* Suppose we have split the $n+1$ objects among the $n$ bins. We pick a random bin with equal probability and let our probability space $\Omega$ be the $n$ bins and we define a random variable $X$ to be the number of objects in some bin. Then since the probability of picking any bin is $\frac{1}{n}$ we have

$$\mathbb{E}[X] = \sum_{x \in \Omega} X(x)p(x) = \frac{1}{n}\sum_{x \in \Omega} X(x) = \frac{n+1}{n}$$

Now suppose with probability 1, $X$ takes values at most 1. Then we have

$$\mathbb{E}[X] = \sum_{x \in \Omega} X(x)p(x) \leq \frac{1}{n}\sum_{x \in \Omega} X(x) = 1$$

Thus, $X$ takes values at most 1 with probability less than 1 so $X$ takes values greater than 1 with a positive probability. Since the probability of picking a bin with more than one object always has positive probability, it must be possible to pick such a box regardless of how the objects are distributed. □

This captures the general gist of the probabilistic method. One must first define an appropriate probability space and then prove that the probability of a desired outcome is positive, which means it must occur for some object, as in this case, or some circumstance.

If you can, see if you can prove the more general form of the Pigeonhole Principle using the same technique.

**Problem 3.2.1.** *Use the probabilistic method to prove the more general form of the Pigeonhole Principle: Given $n$ boxes and $kn+1$ objects where $k$ is a natural number, there exists a box with more than $k$ objects regardless of how the objects are put into the boxes.*

The pigeonhole principle has a number of applications; we'll present a classic example.

**Proposition 3.4.** *At a party, there are at least two people who have the same number of friends.*

*Proof.* Let the number of people be $n$. First, assume that there is someone, say person $X$, that has no friends; then the maximum number of people that anyone can be friends with is $n-2$, since $X$ is not friends with anyone. Thus, if we let our objects be the $n$ people and our boxes be the possible number of friends, ranging from 0 to $n-2$, by the Pigeonhole principle, since there are $n-1$ boxes, there is a box with more than object, or a number where more than one person has that number of friends.

If everyone has at least one friend, then the minimum number of people anyone can be friends with is 1 and the maximum is $n-1$. Again, with a quick application of the pigeonhole principle, since there are $n$ people and $n-1$ boxes, we must have some number where more than one person has that number of friends. □

## 3.3 A Quick Theorem in Ramsey Theory

We first start off with a well-known problem on edge colorings, which, as it turns out, was one of the first problems in Ramsey Theory, a now burgeoning field of combinatorics concerned with patterns that emerge in large systems.

**Proposition 3.5.** *In the complete graph with 6 vertices, any red-blue coloring of the edges will create a triangle which is either all red or all blue.*

*Proof.* Suppose this were not the case. Consider some vertex $v$; it has edges to 5 other vertices. By the pigeonhole principle, at least 3 of these are either red or blue; without loss of generality, let this be red. Let these three vertices be $x, y, z$. Note that none of $(x, y)$, $(y, z)$, $(x, z)$ can be red, as this would complete a triangle with $v$, since the edge from each of these to $v$ is red. Thus, they must all be blue; however, this creates a blue triangle $(x, y, z)$, a contradiction. Thus, we are done. □

This is the smallest such number for which this is true; to see this, consider the 2-coloring of the complete graph $K_5$ shown below, in which there are no monochromatic triangles.

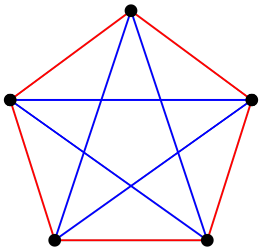

Figure 3.1: A 2-coloring of $K_5$ with no monochromatic triangles.

In 1947, Erdős used the probabilistic method to prove one of the first major results in Ramsey Theory.

**Theorem 3.6.** *Let $R(k, k)$ be the minimum size of a complete graph such that when each edge is colored either red or blue in which there is guaranteed to be $k$ vertices such that edges between them are either all blue or all red. Then we have $2^{\frac{k}{2}} < R(k, k)$.*

At the time, this result was really astonishing because many expected $R(k, k)$ to grow polynomially, but this theorem shows that it grows much faster. Even cooler is that this theorem, which was proved just in the last century, can easily be proved using elementary techniques and the probabilistic method.

We introduce an simple inequality which we will use in the proof.

**Theorem 3.7** (Boole's Inequality)**.** *Given a finite set of events (sets of good outcomes), $\{A_1, \ldots, A_n\}$, we have*

$$\mathbb{P}\left[\bigcup A_i\right] \leq \sum \mathbb{P}[A_i]$$

This is easily visualized with a Venn diagram. If none of the events intersect, then there is equality. If they do overlap, then the overall area covered will be less than the sum of the individual areas. This theorem can be generalized so that the set of events is countably infinite, but the proof of that is beyond the scope of this class and it is not necessary for the proof.

*Proof of Theorem 3.6.* We will prove the theorem by contradiction. Let $r = R(k, k) \leq 2^{\frac{k}{2}}$ and $k \geq 4$. Consider the set of all edge-colorings of the complete graph with $r$ vertices and pick one randomly. Our probability space will be the possible colorings one can pick. We will show that there is a positive probability that one picks a coloring with a blue or red complete subgraph of $k$ vertices. First, pick $k$ of the $r$ vertices, of which there are $\binom{r}{k}$ ways to do so. The probability that a random coloring will make this set of $k$ vertices have all red edges or all blue edges is $2 \cdot \frac{1}{2^{\binom{k}{2}}}$. Then by Boole's Inequality, we have that the overall probability of picking a coloring which has a red or blue complete subgraph of $k$ vertices is at most

$$\binom{r}{k} \frac{2}{2^{\binom{k}{2}}} < \frac{2^{\frac{k^2}{2}+1}}{k!\, 2^{\frac{k^2-k}{2}}} = \frac{2^{\frac{k+2}{2}}}{k!} < 1,$$

where the second inequality comes from the fact that $\binom{r}{k} = \frac{r(r-1)\cdots(r-k+1)}{k!} < \frac{r^k}{k!}$. Then there is a positive probability of picking a graph where no such subgraph exists so there must exist a coloring where we cannot find a blue or red complete subgraph of $k$ vertices. Thus, we have that $R(k,k) > 2^{\frac{k}{2}}$. □

## 3.4 Large Chromatic Numbers

Erdős also proved in 1959 an astonishing result about graph colorings, also provable with the probabilistic method.

**Definition 3.8.** The **chromatic number** of a graph is the minimum number of colors required to color the vertices of a graph such that no two adjacent vertices share the same color.

**Theorem 3.9.** *For any $g$ and $k$, there exists a graph with only cycles of length at least $g$ such that the chromatic number of $G$ is at least $k$.*

This theorem states that we can find a graph such that all the cycles are arbitrarily large and the chromatic number is arbitrarily large. One small example of this is the Grötzch graph, which is triangle-free, yet the chromatic number is 4. This graph is pictured below.

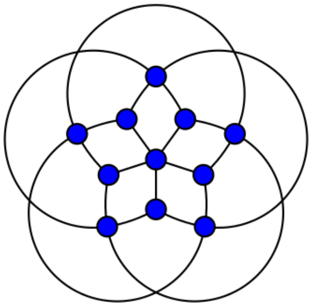

Figure 3.2: The Grötzch graph, which has no triangles yet has chromatic number 4.

In particular, the theorem makes finding $k$-colorings much harder, since one cannot just use cycle length to identify bounds on the chromatic number. Theorems like these have huge impacts on computer science, as $k$-coloring is known to be an NP-complete problem. Coloring also pops up in the schedulers for tasks in computers, since each core of a processor can only do one thing at a time, even if this computation occurs very quickly. As a result, colorability is useful for determining the optimal scheduling of tasks.

Before we begin the proof of the theorem, we first state an important inequality in probability.

**Theorem 3.10** (Markov's Inequality)**.** *If $X$ is a nonnegative random variable and $a > 0$, then*

$$\mathbb{P}[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$$

*Proof.* This follows from the fact that

$$\mathbb{E}[X] = \sum_{x\in\Omega} X(x)p(x) \geq \sum_{x\geq a} X(x)p(x) \geq \sum_{x\geq a} ap(x) = a\mathbb{P}[X \geq a]$$

□

*Proof of Theorem 3.9.* We create a random graph on $n$ vertices, where each edge has a probability $p = n^{\frac{1}{g}-1}$ of existing in this graph.

We first show probabilistically that there exists a graph such that no more than $\frac{n}{2}$ cycles have length less than $g$. The number of cycles of length $i$ in the complete graph $K_n$ is given by

$$\frac{n!}{2\cdot i\cdot (n-i)!} \leq \frac{n^i}{2},$$

since there are $n(n-1)(n-2)\cdots(n-i+1)$ ways to pick the $i$ vertices in the cycle, but this is overcounting by a factor of $i$, since any of the $i$ vertices can be the start point, and a factor of 2, since the cycles can go in either direction.

Given this, we have from Markov's inequality that, if $X$ is the random variable representing the number of cycles,

$$\mathbb{P}\left[X > \frac{n}{2}\right] \leq \frac{2}{n}\mathbb{E}[X] \leq \frac{1}{n}\sum_{i=3}^{g-1} p^i n^i = \frac{1}{n}\sum_{i=3}^{g-1} n^{\frac{i}{g}} \leq \frac{g}{n}n^{-\frac{1}{g}}$$

which goes to 0 as $n \to \infty$. Since this is less than 1, there must exist some graph where $X \leq \frac{n}{2}$.

We can also prove probabilistically that there are no independent sets (sets of vertices such that none of them are connected with edges) of size more than $\lceil\frac{n}{2k}\rceil$. In particular, if we let $y = \lceil\frac{n}{2k}\rceil$ and $Y$ be the random variable corresponding to the size of the maximum independent set,

$$P[Y \geq y] \leq \binom{n}{y}(1-p)^{\frac{y(y-1)}{2}} \leq n^y e^{-\frac{py(y-1)}{2}},$$

where the first inequality comes from the fact that lack of an edge between two vertices has probability $1-p$, and the second inequality comes from the fact that $1-p < e^{-p}$ by the Taylor expansion of $e^x = 1+x+\frac{x^2}{2!}+\cdots$. This also goes to 0 as $n \to \infty$, which implies our desired result by the probabilistic method.

Then there exists an $n$ such that the probability of finding such a graph with fewer than $\frac{n}{2}$ cycles of length less than $g$ and no independent sets of size more than $\lceil\frac{n}{2k}\rceil$. Then we can create a new graph of size $n'$ where $n' > \frac{n}{2}$ by removing one vertex from each of the cycles of length less than $g$. This subgraph will have no independent sets of size more than $\lceil\frac{n'}{k}\rceil$, which implies that its chromatic number must be at least $k$, and all its cycles have length at least $g$, since we removed all of the cycles with smaller length. □

## 3.5 The Erdös-Ko Rado Theorem

Now, we will use the probabilistic method to prove the Erdös-Ko Rado Theorem, which concerns intersecting families of sets.

**Definition 3.11.** A family $F = \{A_1, A_2, \ldots, A_x\}$ of sets is intersecting if for all $i \neq j$, $|A_i \cap A_j| \neq 0$.

An obvious question to ask about such a family is the following: what is the maximum size of $F$, given the sets $A_i$ are a subset of some large set $Y$ of size $n$? In particular, we are concerned with this maximum when all of the $A_1, \ldots, A_x$ are of equal size $k$. In this case, one construction that does a good job including many sets is the following: pick one element to be in all of the sets, and then choose the remaining $k-1$ elements among all of the possibilities. This gives $\binom{n-1}{k-1}$ sets $A_i$. The Erdös-Ko Rado Theorem shows that this is maximal.

**Theorem 3.12.** *If $|Y|= n, n \geq 2k$, and $F$ is an intersecting family of subsets $A_1, \ldots, A_x$ of $Y$ such that $|A_j|= k$ for all $1 \leq j \leq x$, then $x \leq \binom{n-1}{k-1}$.*

Our proof will be modeled off of that of Alon and Spencer. In particular, we first prove the following lemma.

**Lemma 3.13.** *Consider $Y$ an $n$-element set $\{y_1, \ldots, y_n\}$, and consider some ordering of these elements specified by a permutation $\pi$ $\{y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(n)}\}$. Define $A_s^\pi = \{y_{\pi(s)}, y_{\pi(s+1)}, \ldots, y_{\pi(s+k-1)}\}$, where we take sums modulo $n$ so this is meaningful. Then, if $n \geq 2k$, any intersecting family $F$ of subsets of size $k$ contains at most $k$ of the sets $A_s^\pi$.*

*Proof.* Suppose that $A_i^\pi \in F$. Any other $A_j^\pi \in F$ must be one of $\{A_{i-k+1}^\pi, \ldots, A_{i+k-1}^\pi\}$, and in particular, we can divide these sets into pairs $(A_{i-k+1}^\pi, A_{i+1}^\pi), (A_{i-k+2}^\pi, A_{i+2}^\pi), \ldots, (A_{i-1}^\pi, A_{i+k-1}^\pi)$, such that at most one set from each pair can appear in $F$, because they do not intersect (since each set is of size $k$ and has elements that appear consecutively). This proves the desired result, since there are $k-1$ such pairs. □

Now, we prove Theorem 3.12.

*Proof.* Using our lemma, if we choose $\pi$ and $s$ randomly, our set $A_s^\pi$ has an equal probability of being any $k$-element set of $Y$, and from our lemma we have that $\mathbb{P}[A_s^\pi \in F] \leq \frac{k}{n}$, since in the worst case it is always among the sets specified by our permutation. But notice that the probability of a random $k$-element set being in $F$ is exactly $\frac{|F|}{\binom{n}{k}}$, so we have that

$$\frac{|F|}{\binom{n}{k}} \leq \frac{k}{n} \Rightarrow |F| \leq \frac{n!\,(k)}{n(n-k)!\,(k!\,)} = \frac{(n-1)!}{(k-1)!\,(n-k)!} = \binom{n-1}{k-1},$$

as desired. □

The use of the probabilistic method here is a little more subtle, but it still does appear. Isn't this a fascinating result?

## 3.6 Practice and Challenge Problems

We include only competition problems this time; some more basic problems can be found online.

### 3.6.1 Competition Problems

**Problem 3.6.1** (USAMO 2012 Problem 2). *A circle is divided into 432 congruent arcs by 432 points. The points are colored in four colors such that some 108 points are colored red, some 108 points are colored green, some 108 points are colored blue, and the remaining 108 points are colored yellow. Prove that one can choose three points of each color in such a way that the four triangles formed by the chosen points of the same color are congruent.*

**Problem 3.6.2** (Leningrad Math Olympiad 1987). *Let $A_1, \dots, A_n$ be subsets of $\{1, \dots, M\}$, and suppose that none of the $A_i$ are subsets of each other. For each index $i$, let $a_i = |A_i|$. Prove that*

$$\sum_{i=1}^{n} \frac{1}{\binom{M}{a_i}} \leq 1$$

**Problem 3.6.3** (APMO 1998). *Let $F$ be the set of all $n$-tuples $(A_1, A_2, \dots, A_n)$ where each $A_i$, $i = 1, 2, \dots, n$ is a subset of $\{1, 2, \dots, 1998\}$. Let $|A|$ denote the number of elements of the set $A$. Compute*

$$\sum_{(A_1,\dots,A_n)\in F} |A_1 \cup A_2 \cup \cdots \cup A_n|.$$

### 3.6.2 Research Problems

This is not necessarily a problem related to research currently being done, but is a famous result in graph theory that can be solved, in a certain interpretation, using the probabilistic method. See if you can solve it!

**Problem 3.6.4** (Turán's Theorem). *Let $G$ be a graph with $n$ vertices with no set of $r + 1$ vertices all connected to one another. What is the maximum number of edges in $G$?*