6.5420 Randomness and Computation	February 5, 2026
Lecture 2	
Lecturer: Ronitt Rubinfeld	*Scribe: Edward Xiong*

Today we apply the probabilistic method to several problems. This method essentially does the following: we construct an object with a random procedure, and show that the probability the random object satisfies our desired properties is nonzero. That implies that such an object exists.

1 Hypergraph Coloring

A hypergraph is analogous to a graph, where instead of edges between pairs of vertices, we have hyperedges that connect any subset of vertices. Formally, a hypergraph contains a set X of elements (vertices), and some hyperedges $S_1, S_2, \ldots, S_m \subseteq X$. In the 2-coloring problem, we ask whether it is possible to color the elements of X red or blue such that none of the hyperedges are monochromatic. On arbitrary hypergraphs, this problem is known to be NP-hard.

Instead, we'll look at the case where the size of all hyperedge are the same and the number of hyperedges is relatively small: letting ℓ be the number of elements $|S_i|$ in each hyperedge and m be the number of hyperedges, we can guarantee that a 2-coloring exists if $m < 2^{\ell-1}$.

Theorem 1. *If $m < 2^{\ell-1}$, there always exists a proper 2-coloring.*

Proof. Consider any hypergraph X with hyperedges $S_1, \ldots, S_m$. We randomly color each element $x \in X$ red or blue, independently and uniformly at random. Then for any hyperedge S_i, there is a $1/2^\ell$ probability that all $x \in S_i$ are randomly colored red, and $1/2^\ell$ probability that all $x \in S_i$ are randomly colored blue so the probability that S_i is monochromatic is

$$\Pr[S_i \text{ monochromatic}] = \frac{1}{2^{\ell-1}}.$$

Now by taking the union bound over all m hyperedges, the probability that at least one S_i is monochromatic is at most

$$\Pr[\text{at least one } S_i \text{ monochromatic}] \leq \sum_{i=1}^{m} \Pr[S_i \text{ monochromatic}] = \frac{m}{2^{\ell-1}}.$$

Thus if $m < 2^{\ell-1}$, this probability is strictly less than 1. Equivalently, the probability that there are no monochromatic edges is positive and our random coloring is legal; this implies that there exists at least one legal 2-coloring of X. $\square$

Note that this proof is not constructive; on any given input, we know that a legal 2-coloring exists, but we don't know what that coloring is or how to find it. If we make the bound slightly looser and consider graphs on which $m < 2^{\ell-2}$, then observe that in our proof, the probability of having a legal coloring is greater than 1/2. Then it is possible to construct a legal coloring in expected constant time by randomly sampling a coloring and checking whether any edges are monochromatic.

2 Dominating Set

We look at another example of the probabilistic method. On a graph $G = (V, E)$, a subset of vertices $U \subseteq V$ is a *dominating set* if every other vertex $v \in V \setminus U$ is adjacent to at least one vertex in U. Just as in hypergraph coloring, the problem of finding the smallest dominating set is NP-hard. With the probabilistic method, we can derive some upper bounds.

Theorem 2. *If G has minimum degree Δ then G has a dominating set of size at most*

$$\frac{4n\ln(4n)}{\Delta+1}.$$

Proof. We randomly construct a candidate set $\hat{U}$ by including each node $v \in V$ independently with probability

$$p = \frac{\ln(4n)}{\Delta+1}.$$

We call a vertex $w \in V$ good if it is in $\hat{U}$ or it is adjacent to a vertex in $\hat{U}$, and call it bad otherwise. It's clear from the definition that $\hat{U}$ is a dominating set if all vertices $w \in V$ are good. Since we included the vertices independently, the probability that any w is bad is at most

$$\Pr[w \text{ bad}] \leq (1-p)\cdot(1-p)^{\Delta} = (1-p)^{\Delta+1}.$$

Taking the union bound of all vertices $w \in V$, the probability that at least one of the vertices is bad is

$$\Pr[\text{at least one bad vertex}] \leq n(1-p)^{\Delta+1} = n\left(1-\frac{\ln 4n}{\Delta+1}\right)^{\Delta+1} \leq ne^{-\ln 4n} = \frac{n}{4n} = \frac{1}{4},$$

where we used the fact that $(1-1/x)^x < e^{-1}$ for all positive x. Thus the probability that our candidate set $\hat{U}$ is a dominating set is at at least $\frac{3}{4}$.

Now recalling our theorem, $\hat{U}$ satisfies the conditions if it is a dominating set and has at most $4np$ vertices. Each vertex is in $\hat{U}$ with probability p, so the expected number of vertices in $\hat{U}$ is np. Applying the Markov bound, there is probability $3/4$ that the number of vertices in $\hat{U}$ is at most $4np$. Thus applying another union bound, the probability that $\hat{U}$ is a dominating set and is small enough is at least $1/2$, so such a set exists. □

In this proof, our random sampling has probability at least $1/2$ of yielding a sufficiently small dominating set, so this also yields an efficient algorithm of generating such a set.

3 Sum-Free Subsets

A set $A \subseteq \mathbb{N}$ of positive integers is *sum-free* if there are no triples $a_1, a_2, a_3 \in A$ such that $a_1+a_2+a_3$. For example, if we have the set $B = \{1, \ldots, n\}$, there is a sum-free subset of size $\lceil n/2 \rceil$ by taking the larger half $A = \{\frac{n+1}{2}, \ldots, n\}$.

Theorem 3. *For any set B of positive integers with size $|B| = n$, there exists a sum-free subset $A \subseteq B$ with size $|A| \geq n/3$.*

Proof. Let b_n be the largest number in B and fix a prime $p > 2b_n$ such that $p \equiv 2 \pmod 3$. In this proof, we work in the integers modulo p, denoted $\mathbb{Z}_p$. We also denote $\mathbb{Z}_p^\times$ to be the multiplicative group $\{1, 2, \ldots, p-1\}$. Importantly, every number in $\mathbb{Z}_p^\times$ has a multiplicative inverse mod p.

We have that $p = 3k+2$ for some integer k. Letting C be the set $\{k+1, k+2, \ldots, 2k+1\}$ covering roughly the middle third of $\mathbb{Z}_p$, we can observe that C is always sum-free in $\mathbb{Z}_p$ since the sum of any two elements ranges from $2k+2$ to $4k+2 \equiv k \pmod{3k+2}$. .

Now for each $x \in \mathbb{Z}_p^\times$, define the set

$$A_x = \{b \in B \colon xb \bmod p \in C\}.$$

That is, A_x contains the elements of b of which multiplication by x maps to C. The set A_x will always be sum-free; otherwise, if there are elements $a_1, a_2, a_3 \in A_x$ such that $a_1 + a_2 = a_3$, this would imply $xa_1 + xa_2 \equiv xa_3 \pmod p$ which is impossible since C is sum-free.

Since all elements in $\mathbb{Z}_p^\times$ have inverses, for each $b \in B$ and $y \in \mathbb{Z}_p^\times$, there is exactly one element $x \in \mathbb{Z}_p^\times$ such that $xb \equiv y \pmod{p}$.This means that if we pick random x, the value of xb is a uniformly random distribution across all elements of $\mathbb{Z}_p^\times$.

This lets us apply the probabilistic method as follows: we pick a uniformly random $x \in \mathbb{Z}_p^\times$. Observing that C contains $k+1$ elements, the probability that any element b is mapped to C is

$$\Pr[xb \in C] = \frac{|C|}{p-1} = \frac{k+1}{3k+2} > 1/3.$$

Thus any $b \in B$ is included in A with probability more than $1/3$, and by linearity of expectation the expected size of A is greater than $n/3$. This means that at least one of the A_x has size greater than $n/3$, and we are done. □