6.842 Randomness and Computation 2017-09-06

Lecture 1

Lecturer: Ronitt Rubinfeld *Scribe: Ray Hua Wu*

1 The Probabilistic Method

Show an object exists by showing that it probably exists, specifically that it occurs with probability greater than 0. (An object exists either with probability 0 or 1, so if it exists with probability greater than 0, its probability is 1.)

Descartes: "I think, therefore I am."

Erdos: "I probably exist, therefore I am."

1.1 Example 1: Proper n-Coloring

Consider a set S, with subsets $S_1, ..., S_m$ all of the same size l. Can we n-color the elements of S such that no set is monochromatic?

Let's look specifically at proper 2-coloring, that is, coloring with just two colors.

With S containing three elements a, b, and c, with $S_1 = \{a, b\}$, $S_2 = \{a, c\}$, and $S_3 = \{b, c\}$, there is no solution. From the perspective of any one element, both other elements (with which a subset is shared) must be a different color, but there is only one other color, so they're the same color, making the subset containing just them monochromatic.

With larger sets and larger number of elements, the problem could become more unclear.

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

Proof If we color each entity in S either red or blue randomly, then for each subset, monochromaticity occurs when either all were colored red or all were colored blue. As there are l elements in each subset, the probability the subset is monochromatic is $\frac{1}{2^l} + \frac{1}{2^l} = \frac{1}{2^{l-1}}$. As there are m such subsets, the sum of the probabilities for the subsets is $\frac{m}{2^{l-1}}$. By the union bound, the probability that there exists a monochromatic subset is at most this quantity, since the union of the events of monochromaticity for each subset is the event of monochromaticity in at least one of the subsets. If $m < 2^{l-1}$, then $\frac{m}{2^{l-1}} < 1$, so the probability there is no monochromatic subset is positive, and thus by the probabilistic method there is necessarily a coloring of the elements that is a proper 2-coloring. ■

Note that this proof is nonconstructive: despite proving our theorem, we have not been equipped with a method to generate a proper 2-coloring given a problem instance.

In general, Proper n-Coloring is an NP-complete problem.

1.2 Example 2: Sum-free Sets

Definition 2 *A sum-free set is a set of positive integers A such that $\nexists a_1, a_2, a_3 \in A$ where $a_1 + a_2 = a_3$.*

Theorem 3 (Erdos) *For any subset B of n positive integers, $\exists A \subset B$ such that A is sum-free and $|A| > \frac{n}{3}$*

Proof Without loss of generality, let b_n be the largest element in B. Find a prime $p > 2b_n$ such that $p \equiv 2 \pmod 3$. Define an integer k such that $p = 3k + 2$. Let C be $k + 1, ..., 2k + 1$. The following are true about C:

- $C \subset \mathbb{Z}_p^*$
- $\frac{|C|}{p-1} = \frac{k+1}{3k+1} > \frac{1}{3}$

- C is sum-free
- C is sum-free under addition in $\mathbb{Z}_p^*$, that is, it is sum-free mod$p$

C is sum-free under addition because twice its smallest element is greater than its largest element ($2(k+1) > 2k+1$), and C is sum-free under addition in $\mathbb{Z}_p^*$ because $(2k+1)+(2k+1) = 4k+2 \equiv k \pmod{3k+2}$, and if the highest numbers can't add up to be high enough to be $k+1 \pmod{3k+2}$, neither can lower numbers.

Construct an A_x with elements from B such that $xb_i \bmod p \in C$, where x is a randomly chosen element of $\mathbb{Z}_p^*$. We claim that there exists some x for which A_x both is sum-free and has more than $\frac{n}{3}$ elements, and choose this as our A satisfying the theorem.

Suppose A_x isn't actually sum-free. Then, $\exists b_i, b_j, b_k \in A_x$ such that $b_i + b_j = b_k$. If so, then $xb_i + xb_j = xb_k$, and thus also $xb_i + xb_j \equiv xb_k \pmod{p}$, but this is not possible, because $xb_i, xb_j, xb_k \in C$, and C is sum-free and sum-free mod$p$.

$\forall i, \forall y \in \mathbb{Z}_p^*$, a unique $x \in \mathbb{Z}_p^*$ satisfies $y \equiv xb_i \pmod{p}$. Thus, the probability y maps to b_i is $\frac{1}{p-1}$. Thus, the expectation of $xb_i \text{mod} p \in C$ is $\frac{|C|}{p-1}$, which we established above as greater than $\frac{1}{3}$. Thus, the expectation on $|A_x|$ is greater than $\frac{n}{3}$. Define indicator $\sigma_i(x) = 1$ if $xb_i \bmod p \in C$ and 0 otherwise. Then, $E_x[\sigma_i(x)] = Pr_x[\sigma_i(x) = 1] = \frac{|C|}{p-1} > \frac{1}{3}$ and $E_x[|A_x|] = E_x[\sum_i \sigma_i(x)] = \sum_i E_x[\sigma_i(x)] > \frac{n}{3}$. Thus there must be some x such that $|A_x| > \frac{n}{3}$. ■

2 Brief Preview: Lovász Local Lemma

We will discuss the Lovász Local Lemma when the class next meets. We are motivated to attempt to improve our methods of combining small results because the union bound is actually really weak, and often we can make at least partial assumptions of independence.

If the probabilities of our "bad events" (ones we want to avoid) are independent and nontrivial, the latter term meaning that for none of the events is the probability 1, then the probability of the union of the bad events is less than 1, allowing us to prove the avoidability of the bad events.

The Lovász Local Lemma says that if for all bad events A_i, the probability is less than p, and the dependency digraph has degree $\leq d$, where $ep(d+1) \leq 1$, then the probability of no bad events is greater than 0.

Note that the number of bad events itself does not appear in the inequality.