CAS CS 535: Complexity Theory

Lecturer: Mark Bun — Fall 2023

Lecture Notes 15:

Probabilistic Time Classes, Concentration Inequalities, Error Reduction

Reading.

- Arora-Barak § 7.3-7.4

Last time: Randomized Algorithms

1 Probabilistic Time Classes

Definition 1. A probabilistic TM has two transition functions δ_0, δ_1. It computes by applying either δ_0 or δ_1 independently in each time step, each with probability $1/2$.

If M is a probabilistic TM and x is an input, let $M(x)$ denote the random variable where $M(x) = 1$ if M accepts x on a realized sequence of random transitions, and $M(x) = 0$ otherwise.

Let $T_{M,x}$ be the random variable denoting the number of steps M takes on x before halting.

Definition 2 (Runtime for Probabilistic TMs)**.**

- M runs in (worst-case) time $T(n)$ if for all $x \in \{0,1\}^*$, we have $T_{M,x} \leq T(|x|)$ with probability 1.
- M runs in expected time $T(n)$ if for all $x \in \{0,1\}^*$, we have $\mathbb{E}[T_{M,x}] \leq T(|x|)$.

Example 3. The algorithm Find$k$ we discussed last time runs in expected time $O(n)$.

Some Probabilistic Time Classes

Definition 4. $L \in \mathbf{RP}$ if there exists a probabilistic TM running in time $\mathrm{poly}(n)$ such that

$$x \in L \implies \Pr[M(x) = 1] \geq 2/3$$
$$x \notin L \implies \Pr[M(x) = 1] = 0.$$

Definition 5. $L \in \mathbf{coRP}$ if there exists a probabilistic TM running in time $\mathrm{poly}(n)$ such that

$$x \in L \implies \Pr[M(x) = 1] = 1$$
$$x \notin L \implies \Pr[M(x) = 1] \leq 1/3.$$

Example 6. Zero testing for arithmetic formulas over the integers, $\mathsf{ZEROF}_{\mathbb{Z}} \in \mathbf{coRP}$.

Two-sided error:

In general, probabilistic algorithms may make errors on both YES instances and NO instances of a problem.

Definition 7. $L \in \mathbf{BPP}$ if there exists a probabilistic TM running in time $\mathrm{poly}(n)$ such that

$$x \in L \implies \Pr[M(x) = 1] \geq 2/3$$
$$x \notin L \implies \Pr[M(x) = 1] \leq 1/3.$$

Some Relationships Between Classes

- $\mathbf{RP} \subseteq \mathbf{BPP}$
- $\mathbf{coRP} \subseteq \mathbf{BPP}$
- $\mathbf{RP} \subseteq \mathbf{NP}$

The last statement is true because of the following alternative characterization of $\mathbf{NP}$.

Claim 8. *A language $L \in \mathbf{NP}$ if and only if there exists a probabilistic TM M running in time $\mathrm{poly}(n)$ such that*

$$x \in L \implies \Pr[M(x) = 1] > 0$$
$$x \notin L \implies \Pr[M(x) = 1] = 0.$$

Proof. Let L be decided by a poly-time NTM N. One can also interpret N as a PTM. There exists a computational branch on which N accepts iff, as a PTM, it accepts with positive probability. □

Zero error:

Some probabilistic algorithms have "zero-sided error" but may not always terminate within a strict time bound.

Definition 9. $L \in \mathbf{ZPP}$ if there exists a probabilistic TM running in expected time $\mathrm{poly}(n)$ such that

$$x \in L \implies \Pr[M(x) = 1] = 1$$
$$x \notin L \implies \Pr[M(x) = 1] = 0.$$

Example 10. Find$k$ runs in expected time $O(n)$ with zero error (as a search problem).

Despite exploiting randomness in seemingly different ways, there is a very clean relationship between our one-sided error and zero-sided error classes.

Theorem 11. $\mathbf{ZPP} = \mathbf{RP} \cap \mathbf{coRP}$.

The proof of this is a good excuse to introduce a bit of important probabilistic machinery.

2 Concentration Inequalities

Let's start with a simple motivating question. Suppose I toss n fair coins. One expects roughly $n/2$ of those coins to come up heads, and than an absurd majority like $3n/4$ will come up heads with tiny probability. How can we prove this?

First, let's set up the math problem capturing this question. Define random variables $H_1, \ldots, H_n$ via

$$H_i = \begin{cases} 1 & \text{if coin } i \text{ comes up heads} \\ 0 & \text{if coin } i \text{ comes up tails.} \end{cases}$$

Let $H = \sum_{i=1}^n H_i$ denote the number of heads. Then by linearity of expectation,

$$\mathbb{E}[H] = \sum_{i=1}^n \mathbb{E}[H_i] = \frac{n}{2}.$$

Our goal is to estimate $\Pr[H \geq 3n/4]$, and show that this probability is small.

Markov's Inequality: For any random variable $X \geq 0$, and real number $t \geq 0$, we have

$$\Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t}.$$

Proof. For a nonnegative random variable X, we have

$$\mathbb{E}[X] = \int_0^\infty \Pr[X \geq x]\, dx \geq \Pr[X \geq t] \cdot t.$$

Rearranging gives the statement. □

Example 12. Applying Markov's inequality with $X = H, t = 3n/4$, and $\mathbb{E}[H] = n/2$, we get

$$\Pr[H \geq 3n/4] \leq \frac{n/2}{3n/4} \leq 2/3.$$

True, but not that useful.

Chebyshev's Inequality: If X is any real-valued random variable, and $t > 0$ is a real number, then

$$\Pr[|X - \mathbb{E}[X]| > t] \leq \frac{\mathrm{V}ar[X]}{t^2}.$$

Proof. We calculate

$$\begin{aligned} \Pr[|X - \mathbb{E}[X]| > t] &= \Pr[(X - \mathbb{E}[X])^2 > t^2] \\ &\leq \frac{1}{t^2} \cdot \mathbb{E}[(X - \mathbb{E}[X])^2] = \frac{\mathrm{V}ar[X]}{t^2}, \end{aligned}$$

where the inequality is Markov applied to the r.v. $(X - \mathbb{E}[X])^2$. □

Example 13.

$$\begin{aligned} \Pr[H \geq 3n/4] &\leq \Pr[|H - \mathbb{E}[H]| \geq n/4] \\ &\leq \frac{\mathrm{V}ar[H]|}{(n/4)^2} \\ &= \frac{16}{n^2} \sum_{i=1}^n \mathrm{V}ar[H_i] \\ &= \frac{16}{n^2} \cdot \frac{n}{4} = \frac{4}{n}. \end{aligned}$$

Here, the identity $\mathrm{V}ar[H] = \sum_{i=1}^n \mathrm{V}ar[H_i]$ holds because the coin tosses are independent.

Hoeffding's Inequality: Let $X_1, \ldots, X_n$ be independent, $[0,1]$-valued random variables, and let $X = \sum_{i=1}^{n} X_i$. Then for all $t > 0$,

$$\Pr[|X - \mathbb{E}[X]| > t] \leq 2\exp\left(-\frac{2t^2}{n}\right).$$

This inequality is related to the "Chernoff Bound" stated as Corollary A.15 in Arora-Barak. It's technically incomparable – more general in the sense of applying to bounded r.v.'s instead of just Boolean ones, but only controlling additive rather than multiplicative error relative to the mean. (Knock on wood) everything we'll do in this class only requires Hoeffding's inequality.

Example 14.

$$\Pr\left[|H - \frac{n}{2}| \geq n/4\right] \leq 2\exp\left(-\frac{2(n/4)^2}{n}\right) \leq 2e^{-n/8}.$$

3 Proof of Theorem 11

Proof that $\mathbf{ZPP} \subseteq \mathbf{RP} \cap \mathbf{coRP}$. Let $L \in \mathbf{ZPP}$ be decided by a zero-error PTM M running in expected time $p(n)$. We'll first show that $L \in \mathbf{RP}$. Here's the algorithm.
Algorithm $N(x)$
On input x:

1. Run M on x for $3p(|x|)$ steps.
2. If M accepted, accept. If M either rejected or did not yet halt, reject.

Runtime analysis: The algorithm always halts after $O(p(n))$ steps, so it has polynomial worst-case runtime.
Correctness analysis: First suppose $x \in L$. Then by our assumption on M, we have $\mathbb{E}[T_{M,x}] \leq p(|x|)$. By Markov, this implies

$$\Pr[T_{M,x} \geq 3p(|x|)] \leq \frac{1}{3}.$$

Hence, M halts on x with probability at least $2/3$, and if it halts, it correctly accepts by zero-error. Therefore, N accepts x with probability at least $2/3$.

On the other hand, if $x \notin L$, then with probability 1 we have that M either rejects (by zero error) or fails to halt within $3p(|x|)$ steps. In either case, N correctly rejects.

The proof for $\mathbf{ZPP} \subseteq \mathbf{coRP}$ is similar. The only change is that if M fails to halt, we *accept*.

Proof that $\mathbf{RP} \cap \mathbf{coRP} \subseteq \mathbf{ZPP}$. Let $L \in \mathbf{RP} \cap \mathbf{coRP}$ be decided by $p(n)$ time $\mathbf{RP}$ machine M_1 and $\mathbf{coRP}$ machine M_2. We'll argue that the following PTM decides L with zero error.
Algorithm $N(x)$
On input x

Repeat the following indefinitely:

1. Run M_1 on x. If it accepts, accept.
2. Run M_2 on x. If it rejects, reject.

I'll leave the proof of correctness and polynomial expected runtime for your homework.

4 Error Reduction

The constant $2/3$ in the definitions of $\mathbf{RP}$, $\mathbf{coRP}$, and $\mathbf{BPP}$ may seem arbitrary – and that's because they are. These classes don't change when $2/3$ is replaced by any positive constant, in the case of $\mathbf{RP}$ and $\mathbf{coRP}$, or any constant $> 1/2$ in the case of $\mathbf{BPP}$.

In fact, something stronger is true. The definition of $\mathbf{BPP}$ doesn't change even when we replace the error threshold with $1/2 + 1/\mathrm{poly}(n)$ or with $1 - 2^{-\mathrm{poly}(n)}$.

Theorem 15 (Error Reduction for $\mathbf{BPP}$)**.** *Suppose M is a PTM running in time $T(n)$ such that*

$$x \in L \implies \Pr[M(x) = 1] \geq \frac{1}{2} + \varepsilon$$
$$x \notin L \implies \Pr[M(x) = 1] \leq \frac{1}{2} - \varepsilon.$$

We'll abuse notation a bit and write these conditions together as "For all x, we have $\Pr[M(x) = L(x)] \geq \frac{1}{2} + \varepsilon$."

Then there exists a PTM M' running in time $O(\frac{\log(1/\delta)}{\varepsilon^2} \cdot T(n))$ such that $\Pr[M'(x) = L(x)] \geq 1 - \delta$.

That is, one can start with a PTM that beats random guessing by with advantage ε, and boost its success probability to $1 - \delta$, all with a modest blowup in runtime.

Before proving this, let's see some examples of how to use it.

If I start with $\Pr[M(x) = L(x)] \geq$ _____	Then I get $\Pr[M(x) = L(x)] \geq$ _____	Using _____ repetitions of M
$\frac{1}{2} + \frac{1}{n^2}$	$\frac{2}{3}$	$O\left(\frac{\log 3}{(1/n^2)^2}\right) = O(n^4)$
$2/3$	$1 - \frac{1}{n^2}$	$O(\log n)$
$2/3$	$1 - 2^{-10n}$	$O(n)$
$\frac{1}{2} + \frac{1}{n^{100}}$	$1 - 2^{-n^{-100}}$	$O(n^{300})$
$\frac{1}{2} + 2^{-n}$	$1 - 2^{-n}$	$O(n \cdot 2^{2n})$

Proof. Define the algorithm $M'(x)$ as follows.

On input x:

1. Run $M(x)$ independently $k = \frac{\log(2/\delta)}{\varepsilon^2}$ times, producing outputs $b_1, \ldots, b_k$.
2. Out the majority vote of those outputs.

By construction, this has the stated runtime.

To analyze correctness, define the random variables

$$X_i = \begin{cases} 1 & \text{if } b_i = L(x) \\ 0 & \text{otherwise.} \end{cases}$$

Why are these r.v.'s useful? Note that

- Our algorithm outputs the majority vote of the b_i's, which agree with L iff the majority of the X_i's are 1. In other words, our algorithm makes an error iff $X < k/2$ where $X = \sum_{i=1}^{k} X_i$.
- Each X_i individually takes the value 1 with some advantage over random guessing, i.e., $\mathbb{E}[X_i] \geq 1/2 + \varepsilon$. So $\mathbb{E}[X] \geq k/2 + \varepsilon k$.

Putting these observations together and applying Hoeffding's inequality, we have

$$\begin{aligned}\Pr[M'(x) \neq L(x)] &= \Pr[X < k/2] \\ &\leq \Pr[|X - \mathbb{E}[X]| > \varepsilon k] \\ &\leq 2\exp\left(-\frac{2(\varepsilon k)^2}{k}\right) \leq \delta.\end{aligned}$$

□

Theorem 16 (Error Reduction for $\mathbf{RP}$)**.** *If M is a PTM deciding L with one-sided error, i.e.,*

$$\begin{aligned}x \in L &\implies \Pr[M(x) = 1] \geq \varepsilon \\ x \notin L &\implies \Pr[M(x) = 1] = 0,\end{aligned}$$

then the PTM M' obtained by repeating M independently $k = O(\log(1/\delta)/\varepsilon)$ times and accepting if <u>at least one</u> run accepts has one-sided error, and $x \in L \implies \Pr[M'(x) = 1] \geq 1 - \delta$.