CS59200-PRS: Pseudorandomness Aug 25, 2025

Lecture 1: Introduction to Pseudorandomness

Lecturer: Wei Zhan *Scribe: Xiuyu Ye*

In this class, we are interested in the following three questions.

1. ***What*** is pseudorandomness?
2. ***How*** to achieve pseudorandomness?
3. ***Why*** are we interested in studying pseudorandomness?
   - Cryptographic applications.
   - Derandomization.

1 Definitions

We begin with a "dictionary" definition of pseudorandomness.

Pseudorandomness describes an object that looks random but uses less randomness to construct.

The questions are what all the underlining phrases mean. To give it a more formal treatment, we have the following generic definition.

Definition 1 (Pseudorandomness and PRG)**.** *Let $\mathcal{D}$ be a target distribution and $\mathcal{S}$ be a source distribution. We say G is* pseudorandom *against a set $\mathcal{A}$ consisting of functions $A\colon \mathrm{supp}(\mathcal{D}) \to [0,1]$, if there exists small $\varepsilon > 0$ such that for all $A \in \mathcal{A}$,*

$$\left| \mathop{\mathbb{E}}_{s\sim\mathcal{S}}[A(G(s))] - \mathop{\mathbb{E}}_{r\sim\mathcal{D}}[A(r)] \right| \leq \varepsilon.$$

By specifying $\mathcal{D}, \mathcal{S}, \mathcal{A}$ and ε we get a class of pseudorandom objects. In particular, when $\mathcal{D}$ is uniform over $\{0,1\}^n$ and $\mathcal{S}$ is uniform over $\{0,1\}^\ell$ for some $\ell = \ell(n)$ depending on n, we say G is a pseudorandom generator (PRG) *against $\mathcal{A}$.*

Here are some terminologies. In the above definition, we call each $A \in \mathcal{A}$ a *distinguisher*, and we say that G *ε-fools* the distinguisher A. We call $s \sim \mathcal{S}$ a seed, and in the PRG case $\ell(n)$ is the seed length. When we later talk about asymptotics, we should always think of $n \to \infty$ and G is actually a family of PRGs $\{G_n\}$.

2 Examples

2.1 Question 1

Given the PRG $G\colon \{0,1\}^{\ell(n)} \to \{0,1\}^n$, construct a binary distinguisher $A\colon \{0,1\}^n \to \{0,1\}$ as simple as possible that is ***not*** "fooled" by G, i.e. $\left|\mathop{\mathbb{E}}_{s\sim\mathcal{S}}[A(G(s))] - \mathop{\mathbb{E}}_{r\sim\mathcal{D}}[A(r)]\right| > \varepsilon$.

We define the distinguisher as follows.

$$A(x) \ := \ \begin{cases} 1 & \text{if } x \in \text{range}(G) \\ 0 & \text{otherwise} \end{cases} .$$

Then,

$$\mathop{\mathbb{E}}_{r\sim\mathcal{D}}[A(r)] = \frac{|\text{range}(G)|}{2^n} = 2^{\ell - n},$$
$$\mathop{\mathbb{E}}_{s\sim\mathcal{S}}[A(G(s))] = 1,$$

the distinguishing advantage is large.

The above distinguisher A runs in $O(2^{\ell(n)} \cdot n)$ time with oracle access to a PRG G (denote as $A \in \mathsf{TIME}^G(2^{\ell(n)} \cdot n)$). If $\ell(n) = O(\log n)$ and $G \in \mathsf{TIME}(n^{O(1)})$ (G is a polynomial time computable function), then G cannot "fool" P (the set of all languages computable in deterministic polynomial time). In other words,polynomial-time PRGs with logarithmic seed length cannot fool all polynomial-time distinguishers. The contrapositive states that if $G \in \mathsf{TIME}(n^{O(1)})$ ε-fools P with any constant $\varepsilon < 1$, then $\ell(n) = \omega(\log n)$.

2.2 Question 2

Consider the reverse direction, where we are given the set of distinguishers $\mathcal{A}$ and we want to construct the PRG G as simple as possible. What is the smallest seed length $\ell(n)$ that fools every distinguisher in $\mathcal{A}$?

1. $|\mathcal{A}| = 1$. Say $\mathcal{A} = \{A\}$.

   For example, consider a binary distinguisher $A\colon \{0,1\}^n \to \{0,1\}$ that outputs 1 for m out of the 2^n bit-strings, that is, $\mathop{\mathbb{E}}_{r\sim\mathcal{D}}[A(r)] = m/2^n$. To achieve $\mathop{\mathbb{E}}_{s\sim\mathcal{S}}[A(G(s))] \approx \mathop{\mathbb{E}}_{r\sim\mathcal{D}}[A(r)]$, we want G to map k out of the 2^ℓ bit-strings to something in $A^{-1}(1)$. Therefore we need $\forall m \in \{1, 2, \ldots, 2^n\}, \exists k \in \{1, 2, \ldots, 2^\ell\}$ such that $\left|\frac{m}{2^n} - \frac{k}{2^\ell}\right| \le \varepsilon$. The smallest seed length to fool this class of distinguisher is

   $$\ell \ge \lceil \log(1/\varepsilon) \rceil - 1.$$

2. $\mathcal{A} = \{\text{all boolean functions } \{0,1\}^n \to \{0,1\}\}$.

   $$\ell(n) = n.$$

   This is because for any $\ell(n) < n$, the distinguisher in Section 2.1 serves as a counterexample.

3. Generic $\mathcal{A}$.

   Consider a random function $G\colon \{0,1\}^{\ell(n)} \to \{0,1\}^n$ where each output is uniformly and independently drawn. Then for any $s \in \{0,1\}^{\ell(n)}$, $G(s)$ also looks random and $\mathop{\mathbb{E}}_{G}[A(G(s))] = \mathop{\mathbb{E}}_{r}[A(r)]$. Hence, through Hoeffding bound and union bound, we get

   $$\Pr_G\left[\forall A \in \mathcal{A},\ \left|\mathop{\mathbb{E}}_{s\sim\mathcal{S}}[A(G(s))] - \mathop{\mathbb{E}}_{r\sim\mathcal{D}}[A(r)]\right| \le \varepsilon\right] \ge 1 - 2\cdot\exp\left(-2^{\ell}\cdot\varepsilon^2\right)\cdot|\mathcal{A}|\,.$$

   That meanings, when

   $$\ell = \log\log(|\mathcal{A}|) + 2\log(1/\varepsilon) + O(1)$$

   there exists a function $G\colon \{0,1\}^{\ell(n)} \to \{0,1\}^n$ that ε-fools every $A \in \mathcal{A}$. However, this PRG is not explicit as we do not know how it is actually constructed.

4. $\mathcal{A} = \{\text{all size } K \text{ Boolean fan-in-2 circuits}\}$, K is the number of gates in the circuit.

   Note that $|\mathcal{A}| = 2^{O(K\log(K))}$. Take any $K = 2^{\omega(\log n)}$, we know that for every $A \in \mathsf{P}$ there exists $N \in \mathbb{N}$, such that the computation of A with input length n is captured by circuits in $\mathcal{A}$ for all $n \ge N$. By the probabilistic bound above, there exists a PRG against $\mathcal{A}$ (and thus against P) with $\varepsilon = 1/K$ and seed length $O(\log K)$.

   We call a PRG that ε-fools P with some $\varepsilon = \text{negl}(n)$ (smaller than every inverse-polynomial) a *cryptographic PRG*. It means that for every $\ell(n) = \omega(\log n)$, there exists a cryptographic PRG with seed length $\omega(\log n)$ (which is also necessary from Section 2.1). However, the construction is again not explicit, and the to construct an explicit (polynomial-time computable) cryptographic PRG, even with seed length $n-1$, is an open question.

3 Next Time: MAX-CUT

Given a graph $G = (V,E)$, find labeling $r(v) \in \{0,1\}$ of vertices V that maximizes the size of the cut according this labeling, that is

$$\text{maximize} \sum_{(i,j)\in E} \mathbb{1}_{r(i)\neq r(j)}.$$

We will look at a randomized approximation algorithm and derandomize the construction.