6.842 Randomness and Computation May 7, 2008

Lecture 24

Lecturer: Ronitt Rubinfeld *Scribe: Sam McVeety*

1 Last Time

Definition 1 *$X = x_1 \dots x_n$ is* ***next bit unpredictable*** *if $\forall$ ppt P, $\exists$ negligible $\epsilon(n)$ such that*

$$\Pr_{x,i,\text{coins}_P}[P(x_1, \dots, x_{i-1}) = x_i] \leq \frac{1}{2} + \epsilon(n)$$

Theorem 2 *X is pseudorandom iff X is next bit unpredictable (n.b.u.)*

Last time, we proved the forward direction of the theorem. Today, we will finish the proof.

2 Finishing the Proof

Proof
$\Rightarrow$ (last time)
$\Leftarrow$ (today)

Again, we will prove the contrapositive. Assume that X is not pseudorandom. This implies that $\exists$ ppt T such that

$$|\Pr[T(X) = 1] - \Pr[T(U) = 1]| > \frac{1}{n^k}$$

Without loss of generality, assume that that the value in $|\cdot|$ is > 0, otherwise, just use $\overline{T}$. (Note: When the test outputs one, it is guessing that the output came from the pseudorandom generator.)

We will use a hybrid argument, which makes use of a series of sequences that differ from the previous sequence in only one place. This may look familiar, as it resembles the canonical path argument:

$$\begin{aligned} U = D_0 &= U_1 \dots U_n \\ D_1 &= X_1 U_2 \dots U_n \\ D_2 &= X_1 X_2 U_3 \dots U_n \\ &\vdots \\ X = D_n &= X_1 X_2 \dots X_n \end{aligned}$$

We know that $\Pr[T(D_n) = 1] - \Pr[T(D_0) = 1] > \frac{1}{n^k}$. Also,

$$\frac{1}{n^k} < \Pr[T(D_n) = 1] - \Pr[T(D_0) = 1] = \sum_{i=1}^{n} (\Pr[T(D_i) = 1] - \Pr[D_{i-1} = 1])$$

So that there must exist some pair of neighboring sequences in the above list such that the test T should have a difference of $\frac{1}{n} \cdot \frac{1}{n^k} = \frac{1}{n^{k+1}}$ in performance between them, or in other words, some i such that

$$\Pr[T(D_i) = 1] - \Pr[T(D_{i+1}) = 1] \geq \frac{1}{n^{k+1}}$$

The next bit predictor will work in the following way: next bit predictor P_i (for specific i) takes input $X_1 \dots X_{i-1}$. To construct $P(X_1 \dots X_{i-1})$:

1. Choose $U_i \dots U_n \in \{0,1\}^{n-i}$
2. $b \leftarrow T(X_1 \dots X_{i-1} U_i \dots U_n)$ (b is the output of T on a given sequence)
3. if $b = 1$ output U_i. Else output $\overline{U_i}$.

Observation 3 *Note that $P(X_1 \dots X_{i-1}) = X_i$ iff $((b = 1, U_i = X_i)$ or $(b = 0, U_i \neq X_i))$ (probability 1/2 each)*

$$\Pr[\underbrace{P(X_1 \dots X_{i-1}) = X_i}_{*}] = \frac{1}{2}[\Pr[* \,|\, U_i = X_i] + \Pr[* \,|\, U_i \neq X_i]]$$

$$= \frac{1}{2}[\Pr[b = 1 \,|\, U_i = X_i] + \Pr[b = 0 \,|\, U_i \neq X_i]]$$

$$= \frac{1}{2} + \frac{1}{2}[\underbrace{\Pr[T(X_1 \dots X_i U_{i+1} \dots U_n) = 1]}_{\text{(a)}} - \underbrace{\Pr[T(X_1 \dots X_{i-1} \overline{X}_i U_{i+1} \dots U_n) = 1]}_{\text{(b)}}]$$

$$= \frac{1}{2} + \frac{1}{2}[\underbrace{\Pr[T(X_1 \dots X_i U_{i+1} \dots U_n) = 1]}_{\text{(a)}} - \underbrace{\Pr[T(X_1 \dots X_{i-1} U_i U_{i+1} \dots U_n) = 1]}_{\text{(c)}}]$$

$$+ \frac{1}{2}[\underbrace{\Pr[T(X_1 \dots X_{i-1} U_i U_{i+1} \dots U_n) = 1]}_{\text{(c)}} - \underbrace{\Pr[T(X_1 \dots X_{i-1} \overline{X}_i U_{i+1} \dots U_n) = 1]}_{\text{(b)}}]$$

Fact 4 $(c) = \frac{(a)+(b)}{2}$

We use the claim to substitute for (b). From above, we know that $(a) - (c) \geq \frac{1}{n^{k+1}}$ for some i. Accordingly, the entire expression has the following lower bound:

$$\Pr[P(X_1 \dots X_{i-1}) = X_i] \geq \frac{1}{2} + \frac{1}{2}\frac{1}{n^{k+1}} + \frac{1}{2}\frac{1}{n^{k+1}} = \frac{1}{2} + \frac{1}{n^{k+1}}$$

So X is not next bit unpredictable. ■

Notice that a corollary of this theorem is that order does not matter in next bit unpredictable sequences, because order does not matter in pseudorandom sequences, and we have just shown that the definitions are equivalent.

Definition 5 *f is a **One-way Function** (OWF) if*

1. *f is compute in deterministic polynomial time*
2. *$\forall$ ppt A $\exists$ negligible $\epsilon(n)$ such that $\forall n$ large enough, $\Pr_{x,\text{coins}_A}[A(f(x)) \in f^{-1}(f(x))] \leq \epsilon(n)$*

Definition 6 *f is a **One-way Permutation** (a bijective one-way function) if*

1. *f is compute in deterministic polynomial time*
2. *$\forall$ ppt A $\exists$ negligible $\epsilon(n)$ such that $\forall n$ large enough, $\Pr_{x,\text{coins}_A}[A(f(x)) = x] \leq \epsilon(n)$*

3 Creating PRGs

3.1 Candidate OWFs

- $f(x, y) = x \cdot y$ (worst-case complexity)
- The RSA function: $f_{m,e}(x) = x^e \mod m$
- Raven's function: $f_m(x) = x^2 \mod m$
- Discrete Log: $f_{p,g}(x) = g^x \mod p$

Why do we care? Because we need OWFs to have pseudorandom generators.

Theorem 7 *$f : \{0, 1\}^n \to \{0, 1\}^{2n}$ is a PRG $\Rightarrow$ f is one-way*

Proof We will prove the contrapositive. Assume that f is not one way, that is, that there $\exists$ ppt A such that

$$\Pr_x[A(f(x)) \in f^{-1}(f(x))] \geq \frac{1}{n^k}$$

Then we can create $T(y)$, which runs $A(y)$ and outputs 1 if success, 0 otherwise.

$$\Pr_x[T(f(x)) = 1] \geq \frac{1}{n^k}$$
$$\Pr[T(U) = 1] \leq \frac{2^n}{2^{2n}} \leq \frac{1}{2^n}$$

This distinguishes the output of f from the uniform distribution, and so f is not pseudorandom. ■

Theorem 8 *(Hastad, Impagliazzo, Levin, Luby) Efficient PRGs exist $\Leftrightarrow$ one-way functions exist.*

We will not prove this theorem in class, because the proof is intricate and 40 pages long. Instead, we will prove a significantly easier result. Historically, the latter result was all that we had for quite a long time.

Theorem 9 *Efficient PRGs exist $\Leftrightarrow$ one-way permutations exist.*

We will look at a few approaches to converting a OWP into a PRG, settling on hardcore bits as the most viable candidate.

3.2 Candidate PRGs

Let f be a OWP. What if we input x to our PRG and output $f(x) \circ x$? This does not work because f is calculable in polynomial time, so we can have a statistical test that always passes $f(x) \circ x$ but almost never passes $y \circ x$ for random y, distinguishing it from random.

Another approach would be to use a single bit of the output. This may be a viable approach; several papers have been written to this effect.

The approach that we will actually explore is $f(x) \circ b(x)$, where b is a hardcore bit of f.

Definition 10 *$b : \{0, 1\}^* \to \{0, 1\}$ is a* ***Hardcore Bit*** *(HCB) for a one-way function f if $\forall$ ppt A, $\exists$ negligible $\epsilon(n)$ such that $\Pr_{x \in \{0,1\}^l}[A(f(x)) = b(x)] \leq \frac{1}{2} + \epsilon(l)$.*

Theorem 11 *If b is a hardcore bit for a one-way permutation f, then $G = f(x) \circ b(x)$ is a pseudorandom generator.*

Proof $x \in_R U_l \Rightarrow f(x) \in_R U_l$ since f is a permutation, so $f(x)$ is next bit unpredictable. $b(U_l)$ is unpredictable from $f(U_l)$ since it is hardcore. Thus, $f(x) \circ b(x)$ is pseudorandom. ■

3.3 Another Candidate HCB

For p, q prime, $\equiv 3 \mod 4$, let $N = pq, x \in \mathbb{Z}_N^*$. $f(x) = x^2 \mod N$ is one-way? Repeat a lot: output lsb(x) , $x \leftarrow x^2 \mod N$.

3.4 Next Time

An idea for next time is to take x, l bits long, and apply f to x, record $b(x)$, apply $f^2(x)$, record $b(f(x))$, up to $f^k(x)$, recording $b(f^{k-1}(x))$. Finally, output the values $b(f^*(x))$, which we label as $\sigma_0, \sigma_1, \ldots, \sigma_m$. We will show that this sequence is pseudorandom, so the order does not matter.