Notes for Lecture 6

Last time, we gave concrete and asymptotic definitions for one-way permutation, hard-core predicate, indistinguishable distributions, and pseudorandom generator.

Today, we will show that if there is a permutation with a hard-core predicate, then there is a pseudo-random generator. Precisely:

Theorem 1 *If $B : \{0,1\}^n \to \{0,1\}$ is (S,ϵ)-hard core for $p : \{0,1\}^n \to \{0,1\}^n$, then*

$$G(\underbrace{x}_{n}) = \underbrace{p(x)}_{n}, \underbrace{B(x)}_{1}$$

is (S,ϵ)-pseudorandom.

Equivalently, we will show: If there exists a distinguisher circuit D of size $\leq S$ such that

$$\left| \mathop{\mathbf{Pr}}_{x}[D(p(x), B(x)) = 1] - \mathop{\mathbf{Pr}}_{x,r}[D(p(x), r) = 1] \right| \geq \epsilon \ , \tag{1}$$

then $\exists$ C of size $\leq S$: $\mathbf{Pr}_x[C(p(x)) = B(x)] \geq \frac{1}{2} + \epsilon$.

Proof: Without loss of generality, we may assume the distinguishing difference in Eq. (1) is positive – otherwise use $\bar{D}$ (recall that NOT gates aren't counted towards the circuit size). We give two (equivalent) constructions. The first is optimal and simpler, but perhaps less intuitive.

1. Input z $(= p(x))$
   Pick $b \in \{0,1\}$ at random
   if $D(z,b) = 1$
   then output b
   else output 1-b

   Let $A_b(z)$ be the output of the *algorithm* (not yet a circuit) on input z, random choice b. Then

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{x,b}[A_b(p(x)) = B(x)] &= \tfrac{1}{2}\left(\mathop{\mathbf{Pr}}_{x}[A_B(p) = B] + \mathop{\mathbf{Pr}}_{x}[A_{\bar{B}}(p) = B]\right) \\
&= \tfrac{1}{2}\left(\mathop{\mathbf{Pr}}_{x}[D(p,B) = 1] + \mathop{\mathbf{Pr}}_{x}[D(p,\bar{B}) = 0]\right) \\
&\geq \tfrac{1}{2} + \epsilon \ .
\end{aligned}$$

   Here the first equality is from averaging over the cases $b = B(x)$ and $b = \overline{B(x)}$. The second equality is from the definition of the algorithm: $A_B(p) = B \Leftrightarrow D(p,B) = 1$, and $A_{\bar{B}}(p) = B \Leftrightarrow D(p,\bar{B}) = 0$. The final inequality came from substituting $\mathbf{Pr}_x[D(p,\bar{B}) = 0] = 1 - \mathbf{Pr}_x[D(p,\bar{B}) = 1]$ and using

$$\begin{aligned}
\epsilon &\leq \mathop{\mathbf{Pr}}_{x}[D(p,B) = 1] - \mathop{\mathbf{Pr}}_{x,r}[D(p,r) = 1] \\
&= \mathop{\mathbf{Pr}}_{x}[D(p,B) = 1] - \tfrac{1}{2}\left(\mathop{\mathbf{Pr}}_{x}[D(p,0) = 1] + \mathop{\mathbf{Pr}}_{x}[D(p,1) = 1]\right) \\
&= \tfrac{1}{2}\left(\mathop{\mathbf{Pr}}_{x}[D(p,B) = 1] - \mathop{\mathbf{Pr}}_{x}[D(p,\bar{B}) = 1]\right) \ .
\end{aligned}$$

To get a circuit from this algorithm, note that there exists a fixed $b_0 \in \{0,1\}$ so $\mathbf{Pr}_x[A_{b_0}(p(x)) = B(x)] \geq \frac{1}{2} + \epsilon$. The circuit for $A_0(z)$ is $\overline{D(z,0)}$, and the circuit for $A_1(z)$ is $D(z,1)$. In either case, the size is at most S.

2.

Input z $(= p(x))$			
Compute	$D(z,0)$	$D(z,1)$	
case :	0	0	**output** random bit
	0	1	**output** 1
	1	0	**output** 0
	1	1	**output** random bit

This algorithm is equivalent to the first algorithm because the random bit, call it b, can be chosen before computing $D(z,0)$ and $D(z,1)$. If we then output $1-b$ on case $(0,0)$, and b on case $(1,1)$, then the output is always determined by only $D(z,b)$; evaluating $D(z,1-b)$ is unnecessary. Regardless, we shall give a separate analysis.

Define the four disjoint events E_{00}, E_{11}, E_c, E_w according to the the four possibilities for $(D(z,0), D(z,1))$: either $(0,0)$, $(1,1)$, $(\bar{B}, B)$ or $(B, \bar{B})$ respectively. That is, $E_{00} \equiv \{x : D(p(x),0) = 0, D(p(x),1) = 0\}$ and similarly for the other events. Using these definitions, we get

$$\begin{aligned} \epsilon &\leq \mathbf{Pr}[D(p,B)=1] - \mathbf{Pr}[D(p,r)=1] \\ &= (\mathbf{Pr}[E_c] + \mathbf{Pr}[E_{11}]) - \left(\mathbf{Pr}[E_{11}] - \tfrac{1}{2}\mathbf{Pr}[E_c] - \tfrac{1}{2}\mathbf{Pr}[E_w]\right) \\ &= \tfrac{1}{2}\left(\mathbf{Pr}[E_c] - \mathbf{Pr}[E_w]\right) \ , \end{aligned}$$

and therefore the algorithm is correct with probability

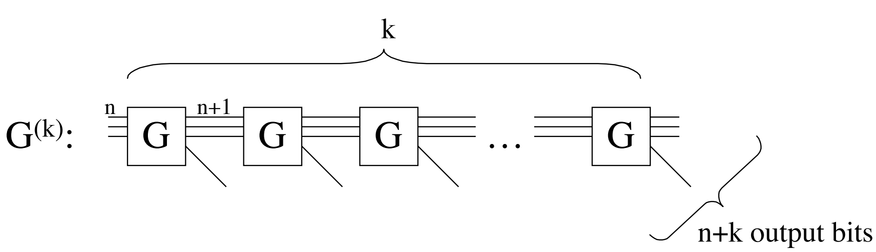
$$\begin{aligned} \mathbf{Pr}[\text{correct}] &= \tfrac{1}{2}\mathbf{Pr}[E_{00}] + \tfrac{1}{2}\mathbf{Pr}[E_{11}] + \mathbf{Pr}[E_c] \\ &\geq \tfrac{1}{2}\left(\mathbf{Pr}[E_{00}] + \mathbf{Pr}[E_{11}] + \mathbf{Pr}[E_c] + \mathbf{Pr}[E_w]\right) + \epsilon \\ &= \tfrac{1}{2} \cdot 1 + \epsilon \ . \end{aligned}$$

□

We have shown how given an n-bit permutation with a hard-core predicate, we get a pseudo-random generator $\{0,1\}^n \to \{0,1\}^{n+1}$ with the same security parameters. Now how can we get a PRG with longer stretch?

How to get a longer stretch

For $G : \{0,1\}^n \to \{0,1\}^{n+1}$, define $G^{(k)} : \{0,1\}^n \to \{0,1\}^{n+k}$ by composing G on its n of its output bits k times sequentially. The extra output bit from each round, together with the $n+1$ bits output from the last round, form the output of $G^{(k)}$:



Theorem 2 *If G is (S,ϵ)-pseudorandom and computable by a circuit of size t, then $G^{(k)}$ is $(S - O(tk), k\epsilon)$-pseudorandom.*

Notice that the circuit size security parameter decreases in addition to the ϵ parameter increasing. This reflects that in our proof by contradiction, given a distinguisher for $G^{(k)}$ we will build a distinguisher for G essentially by adding on the computation of at most k rounds of G. As an aside, it is certainly important that G be efficiently computable. Say for example that $f : \{0,1\}^n \to \{0,1\}$ satisfies that for all circuits C of size $\leq S$, $\mathbf{Pr}[C(x) = f(x)] \leq \frac{1}{2} + \epsilon$. Then $x \mapsto x, f(x)$ is (S,ϵ)-pseudorandom. But applying this construction would result in the first k bits being constant, certainly not random-looking.

Proof: Say we have D of size S such that $\left|\mathbf{Pr}[D(G^{(k)}(x) = 1] - \mathbf{Pr}[D(r) = 1]\right| \geq \epsilon$. We want to show that there is a C of size $\leq S + O(tk)$ such that $|\mathbf{Pr}[C(G(x)) = 1] - \mathbf{Pr}[C(r) = 1]| \geq \epsilon$. The argument is by the standard hybrid technique. Assume for simplicity that G is a PRG of the form we constructed earlier today: $G(x) = (p(x), B(x))$. Define the following distributions $H_0, \ldots, H_k$:

$$\begin{array}{lllllll}
H_0: & B(x), & B(p(x)), & B(p^{(2)}(x)), & \ldots, & G(p^{(k-1)}(x)) \\
H_1: & r_1, & B(p(x)), & B(p^{(2)}(x)), & \ldots, & G(p^{(k-1)}(x)) \\
H_2: & r_1, & r_2, & B(p^{(2)}(x)), & \ldots, & G(p^{(k-1)}(x)) \\
\vdots & & & & & \\
H_k: & r_1, & r_2, & r_3, & \ldots, & r_k, \; p^{(k)}(x)
\end{array}$$

In each case, x is chosen at random from $\{0,1\}^n$ and independently $r_1, \ldots, r_k$ each at random from $\{0,1\}$. Then $H_0 \sim$ output of $G^{(k)}$, while $H_k \sim$ the uniform distribution. Since

$$\epsilon \leq \mathbf{Pr}[D(H_0) = 1] - \mathbf{Pr}[D(H_k) = 1] = \sum_{i=0}^{k-1} \big(\mathbf{Pr}[D(H_i) = 1] - \mathbf{Pr}[D_(H_{i+1} = 1]\big)$$

(where we have w.l.o.g. assumed $\mathbf{Pr}[D(H_0) = 1] > \mathbf{Pr}[D(H_k) = 1]$ and telescoped the sum), there exists an i such that $\mathbf{Pr}[D(H_i) = 1] - \mathbf{Pr}[D(H_{i+1}) = 1] \geq \epsilon/k$.

We will finish the proof next time. But the basic idea is that our G-distinguisher C will construct either the distribution H_i or H_{i+1} (depending on whether its input is from G or is truly random) using at most k computations of G. It then feeds this distribution over $\{0,1\}^{n+k}$ to the $G^{(k)}$-distinguisher D. $\square$