COM S 6810 Theory of Computing January 20, 2009

Lecture 21: pseudo-random Generators

Instructor: Rafael Pass *Scribe: Matt Weinberg*

1 Definitions

Definition 1 *An **Ensemble**, $\{X_n\}_{n\in\mathcal{N}}$ is a sequence of probability distributions.*

Definition 2 *Let $\{X_n\}$ and $\{Y_n\}$ be two ensembles, where X_n and Y_n are probability distributions over $\{0,1\}^{l(n)}$ for some polynomial l. Then we say $\{X_n\}$ and $\{Y_n\}$ are **indistinguishable** if $(\forall PPT,\ D)(\exists$ negligible $\epsilon)(\forall n \in \mathcal{N})(Pr[t \leftarrow X_n, D(1^n,t)=1] - Pr[t \leftarrow Y_n, D(1^n,t)=1] \leq \epsilon(n))$.*

This is just used to formalize what it means for two probability distributions to look the same. Two distributions look the same if no PPT can tell them apart with decent probability (remember that negligible functions are smaller than every polynomial).

Definition 3 *A function, $G : \{0,1\}^* \rightarrow \{0,1\}^*$ is called a **pseudo-random Generator** if it satisfies the following three properties:*

***(1) Efficiency:** G can be computed in polynomial time.*

***(2) Expansion:** $l(|x|) = |G(x)|$ is a well-defined function, and $l(x) > |x|, \forall x$*

***(3) pseudo-randomness:** $\{s \leftarrow \{0,1\}^n, G(s)\}$ is indistinguishable from $\{t \leftarrow \{0,1\}^{l(n)}\}$*

Efficiency means that such a PRG could actually be used in a PPT. Expansion means that it is useful in the sense that it grants us more random bits than we started with. pseudo-randomness means that the output is basically random when viewed by a PPT because no PPT can tell the difference between the output of the PRG and the actual uniform distribution.

Remember that a predicate, $b : \{0,1\}^n \rightarrow \{0,1\}$ is a **hard-core bit** for f if b is efficiently computable and $\forall$PPT $A, \exists$ negligible ϵ such that $Pr[x \leftarrow \{0,1\}^n, A(1^n, f(x)) = b(x)] \leq 1/2 + \epsilon(n)$.

2 Existence of PRGs

As a corollary of the Goldreich-Levin Theorem, $\exists OWP \rightarrow \exists OWP$ with hard-core predicate. Call such a OWP, f, and call the hard-core predicate h.

Now we will use such a OWP to define a PRG. Let $G(s) = f(s)||h(s)$, where $||$ denotes concatenation. Then we have the following claim:

Proposition 1 *If $f(s)$ is a OWP and $h(s)$ is a hard-core predicate of f, then $G(s) = f(s)||h(s)$ is a PRG.*

Proof. Clearly efficiency and expansion are satisfied. Now we'll check pseudo-randomness. Say pseudo-randomness were not satisfied, then $\exists$ PPT, D, such that for infinitely many n, D can distinguish $\{s \leftarrow U_n, G(s)\}$ and $\{U_{n+1}\}$ with probability $\geq 1/p(n)$. Then for these n, and the PPT, D, it is true that:

$$Pr[s \leftarrow \{0,1\}^n, D(1^n, G(s)) = 1] - Pr[s \leftarrow \{0,1\}^{n+1}, D(1^n, s) = 1] \geq 1/p(n)$$

just by the definition of distinguishable. Now consider $G'(s) = f(s)||(1 - h(s))$. Then it is clear that $U_{n+1} = \frac{1}{2}G(s) + \frac{1}{2}G'(s)$ because f is a OWP. So for all $x \in U_{n+1}$, x is of the form $G(s)$ or $G'(s)$ for exactly one s, and exactly one of G or G'.

Now it is clear that if D can distinguish U_{n+1} and $G(s)$, then D can distinguish $G(s)$ and $G'(s)$ because we could use the same distinguisher.

Now denote $G_0 = G$, $G_1 = G'$.

Proposition 2 $Pr[b \leftarrow \{0,1\}, s \leftarrow \{0,1\}^n, D(1^n, G_b(s)) = b] \geq 1/2 + \frac{1}{2p(n)}$

This is the same $p(n)$ as above.

Proof. This seems trivial, because D can distinghuish between G_0 and G_1, but just because D can distinguish G_0 and G_1 doesnt mean D can actually figure out which distribution G_b is. So to prove this proposition, we w.l.o.g. assume that $Pr[s \leftarrow \{0,1\}^n, G'(s) = 1] > Pr[s \leftarrow \{0,1\}^n, G(s) = 1]$ for infinitely many n. This assumption is made w.l.o.g. because if some distinguisher D does not satisfy this, then we can just consider D', where $D'(1^n, f(s)) = 1 - D(1^n, f(s))$. Now we can write the following equations:

$$\begin{aligned}
&Pr[b \leftarrow \{0,1\}, s \leftarrow \{0,1\}^n, D(1^n, G_b(s)) = b] \\
&= 1/2 Pr[s \leftarrow \{0,1\}^n, D(1^n, G_0(s)) = 0] + 1/2 Pr[s \leftarrow \{0,1\}^n, D(1^n, G_1(s)) = 1] \\
&= 1/2[1 - Pr[s \leftarrow \{0,1\}^n, D(1^n, G_0(s)) = 1] + Pr[s \leftarrow \{0,1\}^n, D(1^n, G_1(s)) = 1]] \\
&= 1/2 + 1/2[Pr[s \leftarrow \{0,1\}^n, D(1^n, G_1(s)) = 1] - Pr[s \leftarrow \{0,1\}^n, D(1^n, G_0(s)) = 1] \\
&\geq 1/2 + \frac{1}{2p(n}
\end{aligned}$$

Note that these equations hold for the infinitely many n such that D can distinguish between $G(s)$ and $G'(s)$ and $Pr[s \leftarrow \{0,1\}^n, G'(s) = 1] > Pr[s \leftarrow \{0,1\}^n, G(s) = 1]$. So now the proposition is true.

Now using the above proposition, we consider the following algorithm, $A(1^n, y)$:

(1) Uniformly pick $b \leftarrow \{0,1\}$.

(2) Let $d \leftarrow D(1^n, y||b)$

(3) If d = 0, output b. If d = 1, output 1-b.

Then $A(1^n, y)$ can predict $h(x)$ with probability $1/2 + \frac{1}{2p(n)}$ by the proposition, because for either b, there is a $1/2 + \frac{1}{2p(n)}$ chance that $d = h(x)$. So this is a contradiction because we assumed that $h(x)$ was a hard-core bit for $f(x)$, so it must be the case that $\{s \leftarrow U_n, G(s)\}$ and $\{s \leftarrow U_n, G'(s)\}$ are indistinguishable, and by contrapositive, so are $\{s \leftarrow U_n, G(s)\}$ and $\{U_{n+1}\}$. So pseudo-randomness must be satisfied, and $G(s)$ is in fact a PRG.