Pseudorandom generators from general one-way functions III

15-859I
Spring 2003

Review:

- Our goal is to construct a PRG from any OWF
- A False Entropy Generator is a function $f:\{0,1\}^n \to \{0,1\}^{\ell(n)}$ that has $f(U_n)$ computationally indistinguishable from some ptc ensemble D_n: $\{0,1\}^{\ell(n)}$ where $H(D) > H(f(U))$.
- Using universal hash functions and product distributions, we can construct a PRG from a F.E.G. (4 pages from [HILL99])

Review: f ' construction

- Let $f:\{0,1\}^n \to \{0,1\}^{\ell(n)}$ be a one-way function, and let $h: \{0,1\}^{p(n)} \times \{0,1\}^n \to \{0,1\}^{n+\lceil \log 2n \rceil}$ be a universal hash function. Define

  $f'(x,i,r) = (f(x), h_r(x)|_{1\ldots i+\lceil \log 2n \rceil}, i, r)$
- Let $Y \leftarrow U_n$, then when $I < \tilde{\mathbf{D}}_f(f(X))$, we will have $(f'(X,I,R),Y,X \bullet Y) \cong (f'(X,I,R),Y,U_1)$.
- To formalize, define two sets:
  - $T = \{(x,i) : x \in \{0,1\}^n, i \in \{0,\ldots, \tilde{\mathbf{D}}_f(f(x))\}\}$
  - $T^C = \{(x,i) : x \in \{0,1\}^n, i \in \{\tilde{\mathbf{D}}_f(f(x))+1, \ldots, n-1\}\}$

Review: FEG Construction

- Let $k(n) \geq 125n^3$, $I \in_U \{0,\ldots,n-1\}$, and define

  $p_n = \Pr[I \leq \tilde{\mathbf{D}}_f(f(x))]$

  $m(n) = k(n)p_n - 2k(n)^{2/3}$
- Let $X', Y' \leftarrow U_{nk(n)}, I' \in_U \{0,\ldots,n-1\}^{k(n)}$, $R' \leftarrow U_{k(n)p(n)}$, $Z \leftarrow U_{m(n)}$.
- Let $h' : \{0,1\}^{p'(n)} \times \{0,1\}^{k(n)} \to \{0,1\}^{m(n)}$ be a universal hash function, and $V \leftarrow U_{p'(n)}$
- Define $g(p_n,X',Y',I',R',V) =$

  $(h'_V(X' \bullet Y'), f'^{k(n)}(X',I',R'), V, Y')$

Review: Main Theorem

- False Entropy Theorem: g is a mildly nonuniform false entropy generator.
- Proof: Delayed…
- Main Theorem: If there exists a one-way function, then there exists a pseudorandom generator.
- Proof: Compose previous theorems: False Entropy Theorem, FEG → (mildly nonuniform) PEG theorem, PEG → PRG theorem, mildly nonuniform PRG → PRG theorem.
- We're done! Oh wait, that pesky False entropy theorem…

Review: False Entropy Theorem

- Proof: Consider the distributions:

  D $= g(p_n,X',Y',I',R',V)$ and

  E $= (Z, f'^{k(n)}(X',I',R'), V, Y')$

  Lemma 1: $H(E) \geq H(D) + 10n^2$.

  Lemma 2: $D \cong E$

  Thus, g is a false entropy generator given p_n. We will show in the proof of lemma 2 that it is OK to use a value ρ with $p_n \leq \rho \leq p_n+1/n$. Therefore we only need log n bits of advice. So g is a mildly nonuniform false entropy generator. QED

Lemma 2: $D \cong E$

- Recall:
  - D = $h_V(X' \bullet Y'), f'^{k(n)}(X',I',R'), V, Y'$
  - E = $(Z, f'^{k(n)}(X',I',R'), V, Y')$
- Another way to describe D:
  - For each j, choose $C_j=1$ with probability p_n
  - When $C_j = 1$, choose $(X_j',I_j') \in T$, else $(X_j',I_j') \in T^C$
- Define the distribution D':
  - Same as D, except when $C_j = 1$ replace j^{th} input to h $(X'_j \bullet Y'_j)$ by $B_j \leftarrow U_1$.

Lemma 2 intuition…

- Notice that by the Leftover Hash Lemma, $L_1(D',E) \le 2^{-k(n)^{1/3}} = 2^{-5n}$, so $D' \cong E$.
- Intuitively, in D' we just replace $X_j' \bullet Y_j'$ by B_j when $(X_j',I_j') \in T$; and we have already shown that in this case $X_j' \bullet Y_j' \cong B_j$. So we would expect $D \cong D'$, giving $D \cong E$.
- The hybrid argument fails, however, because we can't efficiently sample from D'

Hybrid argument for $D \cong D'$

- Suppose we have A such that
  $\Pr[A(D)=1] - \Pr[A(D')=1] = \delta(n)$
- Define the hybrid distributions $F^{(j)}$ so that $F^{(j)}$ is distributed identically to D' up to position j and D afterwards, i.e., $F^{(j)}$ is chosen like D except that for $i \le j$, when $C_i=1$ we replace $X_i' \bullet Y_i'$ by B_i. Thus $F^{(0)} = D$, $F^{(k(n))} = D'$
- If $J \in_U \{1,\ldots,k(n)\}$, then we have that
  $E_J[\, A(F^{(J-1)}) - A(F^{(J)})\,] = \delta(n)/k(n)$

How to fix our Hybrid argument?

- Notice that when $C_j = 0$, A has no advantage, yet when $C_j=1$ A has significant advantage.
- So A "knows" when an element $W \in T$, given $f'(W,R)$.
- We will take advantage of this to build hybrid distributions which are "close" to $F^{(j)}$ allowing us to get by the problem.
- This is the last 4 technical pages of [HILL99]

New Hybrids…

- We will define two sets of hybrid distributions, $E^{(j)}, D^{(j)}$ for $j \in \{0,\ldots,k(n)\}$.
- We will have $E^{(0)} = E$, $D^{(0)} = D$, and $E^{(k(n))} \approx D^{(k(n))}$.
- Define $\delta^{(j)} = \Pr[A(D^{(j)}) = 1] - \Pr[A(E^{(j)})=1]$.
- Then $\delta^{(0)} = \delta(n)$ and $\delta^{(k(n))} \approx 0$
- We will also have: $E_J[\delta^{(J-1)} - \delta^{(J)}] \ge \delta(n)/k(n)$
- This will allow us to (indirectly) invert f' later.

Definition of $D^{(j)}$, $E^{(j)}$

- Define parameters:
  - $\rho = \delta(n)/16k(n)$
  - $\tau = 64n^2/\rho$
- Define: $D^{(0)} = D$; $E^{(0)} = E$; $B \leftarrow U_{k(n)}$.
- Suppose $D^{(j-1)}$ is defined. Then to sample from $D^{(j)}$:
  - Choose $c_j \in \{0,1\}$ so that $\Pr[c_j=1] = p_n$
  - Sample $x_m \leftarrow U_n$, $i_m \in_U \{1\ldots n\}$, let $w_m = (x_m,i_m)$, $1 \le m \le \tau$.

$D^{(j)}$ and $E^{(j)}$ continued…

- Define $D^{(j-1)}(c_j,w_m)$ to be the same as $D^{(j-1)}$ except that (X_j',I_j') is fixed to w_m and the j^{th} input bit of h' is set to $x_m \bullet Y_j'$ if $c_j=0$ and B_j otherwise.
- Define $E^{(j-1)}(w_m)$ to be the same as $E^{(j-1)}$ except (X_j',I_j') is fixed to w_m.
- Define $\delta^{(j-1)}(c_j,w_m) = \Pr[A(D^{(j-1)}(c_j,w_m)) = 1] - \Pr[A(E^{(j-1)}(w_m)) = 1]$.

Sampling from $D^{(j)}$ and $E^{(j)}$…

- Use A and draw $O(n/\rho^2)$ samples from $D^{(j-1)}(c_j,w_m)$, $E^{(j-1)}(w_m)$ to get an estimate $\Delta^{(j-1)}(c_j,w_m)$ such that
  $\Pr[|\Delta^{(j-1)}(c_j,w_m) - \delta^{(j-1)}(c_j,w_m)| > \rho] \leq 2^{-n}$
  (i.e., take average over $O(n/\rho^2)$ samples)
- Let $\mu \in \{1,\ldots,\tau\}$ be such that $\Delta^{(j-1)}(c_j,w_\mu)$ is maximized.
- Define $D^{(j)} = D^{(j-1)}(c_j, w_\mu)$, $E^{(j)} = E^{(j-1)}(w_\mu)$

Using our hybrids

- Define $D^{(j)}(w,r,b,y)$ to be $D^{(j)}$ with $f'(X'_{j+1},Y'_{j+1},R'_{j+1})$ replaced by $f'(w,r)$, the j+1 input bit to h' replaced by b, and Y'_{j+1} replaced by y; Same for $E^{(j)}(w,r,y)$.
- Define $M^A(f'(w,r),b,y)$ =
  - Choose $j \in_U \{0,\ldots,k(n)-1\}$
  - Draw $d \leftarrow D^{(j)}(w,r,b,y)$, $e \leftarrow E^{(j)}(w,r,y)$, $b' \leftarrow U_1$.
  - If A(d) = A(e), output b'; else output A(d).

Hybrid claim

- Hybrid Claim: if A distinguishes D and E with probability $\delta(n)$, M^A distinguishes $f'(W,R),X \bullet Y,Y$ from $f'(W,R),B,Y$ with probability at least $\delta(n)/16k(n)$
- (Hang in there… only 2pp left!)

Proof of Hybrid claim

- $\Pr[M(f(w,r),b,y) = 1] =$
  $\frac{1}{2}\Pr[A(D^{(j)}(w,r,b,y)) = A(E^{(j)}(w,r,y)]$
  $+ \Pr[A(D^{(j)}(w,r,b,y)) = 1 \;\&\; A(E^{(j)}(w,r,y)=0)]$
  $= \frac{1}{2}\Pr[A(D^{(j)}(w,r,b,y) = 1) \;\&\; A(E^{(j)}(w,r,y)) = 1] +$
  $\frac{1}{2}\Pr[A(D^{(j)}(w,r,b,y) = 0) \;\&\; A(E^{(j)}(w,r,y)) = 0] +$
  $\Pr[A(D^{(j)}(w,r,b,y)) = 1 \;\&\; A(E^{(j)}(w,r,y)) = 0]$
  $= \frac{1}{2} + \frac{1}{2}(E[A(D^{(j)}(w,r,b,y)] - E[A(E^{(j)}(w,r,y)])$
  $\equiv \frac{1}{2} + \frac{1}{2}(d(j,w,r,b,y) - e(j,w,r,y))$

Proof, con't…

- Notice that:
  - $E[d(j,w,R,x \bullet Y,Y) - e(j,w,R,Y)] = \delta^{(j)}(0,w)$
  - $E[d(j,w,R,B,Y) - e(j,w,R,Y)] = \delta^{(j)}(1,w)$
- Define $\varepsilon^{(j)} = E[\delta^{(j)}(0,W) - \delta^{(j)}(1,W)]$
- Then the advantage of M^A is:
  $E[M^A(f'(W,R),X \bullet Y,Y)] - E[M^A(f'(W,R),B,Y)] =$
  $E[\delta^{(j)}(0,W)/2] - E[\delta^{(j)}(1,W)/2] = E_j[\varepsilon^{(j)}]/2$
- So we just need to show that
  $E_j[\varepsilon^{(j)}] \geq \delta(n)/8k(n)$

Alternatively…

- Alternatively we can show that
  $E[\Sigma_j \varepsilon^{(j)}] \geq 2\rho k(n)$
- We will prove this by showing that:
  (a) $E[\delta^{(k(n))}] \leq 2^{-n+1}$
  (b) $E[\delta^{(j)} - \delta^{(j+1)}] \leq \varepsilon^{(j)} + 4\rho$
- This will give us:

$$\begin{aligned} 8\rho k(n) &= \delta(n)/2 \\ &< \delta(n) - E[\delta^{k(n)}] \\ &= \Sigma_j E[\delta^{(j)} - \delta^{(j+1)}] \\ &\leq 4k(n)\rho + E[\Sigma_j \varepsilon^{(j)}]. \end{aligned}$$

Proof of (a) $E[\delta^{(k(n))}] \leq 2^{-n+1}$

- Notice that $E^{(k(n))}$ and $D^{(k(n))}$ are identical except that the first m(n) bits of $E^{(k(n))}$ are Z and the first m(n) bits of $D^{(k(n))}$ are the output of h'.
- But H_R(input to h' | rest of $D^{(k(n))}$) $\geq \Sigma_j c_j$.
- A Chernoff bound gives us that with probability at least $1-2^{-n}$,
  $\Sigma_j c_j \geq k(n)p_n - k(n)^{2/3} = m(n) + k(n)^{2/3}$
- When this is true, we get from the Leftover hash lemma that $L_1(D^{(k(n))},E^{(k(n))}) \leq 2^{-k(n)^{2/3}/2} < 2^{-n}$.
- This gives us $E[\delta^{(k(n))}] \leq 2^{-n+1}$.

Proof of (b) $E[\delta^{(j)} - \delta^{(j+1)}] \leq \varepsilon^{(j)} + 4\rho$

- Recall that $W \in_U T$. Define $W^C \in_U T^C$.
- Then since the j+1 input to h' in $D^{(j)}$ is always $X'_{j+1}\bullet Y'_{j+1}$, we have

$$\begin{aligned} \delta^{(j)} &= p_n E[\delta^{(j)}(0,W)] + (1-p_n)E[\delta^{(j)}(0,W^C)] \\ &= p_n E[\delta^{(j)}(1,W)] + p_n(E[\delta^{(j)}(0,W)] - E[\delta^{(j)}(1,W)]) + (1-p_n)E[\delta^{(j)}(0,W^C)] \\ &< \varepsilon^{(j)} + p_n E[\delta^{(j)}(1,W)] + (1-p_n)E[\delta^{(j)}(0,W^C)] \end{aligned}$$

- We will complete the proof by showing that
  $E[\delta^{(j+1)}] + 4\rho \geq p_n E[\delta^{(j)}(1,W)] + (1-p_n)E[\delta^{(j)}(0,W^C)]$.

To show: $E[\delta^{(j+1)}]+4\rho \geq p_n E[\delta^{(j)}(1,W)]+(1-p_n)E[\delta^{(j)}(0,W^C)]$

- A Chernoff Bound gives us that with probability at least $1-2^{-n}$, for stage j, at least n/ρ of the w_m are in T and at least n/ρ of the w_m are in T^C.
- Thus with probability at least $1-2^{-n}$, we have:
  $\max_m \{\delta^{(j)}(c,w_m)\} \geq \max\{E[\delta^{(j)}(c,W)],E[\delta^{(j)}(c,W^C)]\} - \rho$
- Also recall that with probability at least $1-2^{-n}$, we have $|\Delta^{(j)}(c,w_m) - \delta^{(j)}(c,w_m)| < \rho$

To show: $E[\delta^{(j+1)}]+4\rho \geq p_n E[\delta^{(j)}(1,W)]+(1-p_n)E[\delta^{(j)}(0,W^C)]$

So

$$\begin{aligned} \delta^{(j)}(c,w_\mu) &\geq \Delta^{(j)}(c,w_\mu) - \rho \\ &= \max_m \{\Delta^{(j)}(c,w_m)\} - \rho \\ &\geq \max_m \{\delta^{(j)}(c,w_m)\} - 2\rho \\ &\geq \max\{E[\delta^{(j)}(c,W)],E[\delta^{(j)}(c,W^C)]\} - 3\rho \end{aligned}$$

With probability at least $1 - 3\cdot 2^{-n}$. Thus:

$$\begin{aligned} E[\delta^{(j+1)}(c)] &= E[\delta^{(j)}(c,w_\mu)] \\ &\geq \max\{E[\delta^{(j)}(c,W)],E[\delta^{(j)}(c,W^C)]\} - 4\rho \end{aligned}$$

Giving the required inequality.

So we are done

- This completes the proof that A distinguishes f'(w,r),x•y,y from f'(w,r),b,y .
- Thus completing the proof that a F.E. Generator can be constructed from any one-way function…
- HUGE issue: suppose we compose the various constructions to get a pseudorandom generator. Then to get inputs to f of size n, the inputs to the resulting generator will have size n^{34}. [HILL99]

Open problem

- Now we don't actually require all of the intermediate product distributions… [HILL99] claim that the same techniques can chip it down to inputs of size n^8.
- Open problem: construct a pseudorandom generator from any one-way function f such that the security of f on inputs of size n is related to the security of g on inputs of size n^2 or n^3.