Class may move to CS 1211.

Goal: Build function $f: \{0,1\}^n \rightarrow \{0,1\}$

that "looks random".

**Distinguishability**

Define: Distributions $D$ and $D'$ are $\epsilon$-statistically indistinguishable if for all algorithms $A = \ldots$

$$\text{Adv}_A = |\Pr[A(x) = 1; x \sim D] - \Pr[A(x) = 1; x \sim D']| \leq \epsilon$$

- always outputs 1, $\text{Adv}_A = 0$
- is a random bit, $\text{Adv}_A = 0$

**Example:**

<table>
<thead>
<tr>
<th>$D$</th>
<th>0: 100%</th>
<th>1: 0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D'$</td>
<td>0: 50%</td>
<td>1: 50%</td>
</tr>
</tbody>
</table>

$A(0) = 0$

$A(1) = 1$

$\text{Adv} A = |0 - .5| = .5$

**Lemma:** If $D$ and $D'$ are $\epsilon$-statistically indistinguishable, then

$$|D(x) - D'(x)| \leq 2\epsilon$$
Proof: consider adversary \( A(x) = \begin{cases} 0 & \text{if } D(x) > D'(x) \\ 1 & \text{if } \text{true} \end{cases} \)

\[
\text{Adv } A = \left| \Pr[A(x) = 1; x \in D] - \Pr[A(x) = 1; x \in D'] \right|
\]

\[
= \left| \sum_{x \in D(x)} D(x) - \sum_{x \in D'(x)} D'(x) \right|
\]

\[
= \sum_{x \in D(x) \setminus D'(x)} D(x) - \sum_{x \in D'(x) \setminus D(x)} D'(x)
\]

\[
= \sum_{x \in D(x) \setminus D'(x)} D(x) - D'(x)
\]

\[
\leq \epsilon
\]

By symmetry

\[
\sum_{x \in D'(x) \setminus D(x)} D'(x) - D(x) \leq \epsilon
\]

\[
\Rightarrow \sum_{x \in D(x)} D'(x) - D(x) \leq 2\epsilon
\]

Defn: Distributions \( D \) and \( D' \) are \((+\epsilon)\)-computationally indistinguishable iff for all algorithms \( A \) executing in less than time \( +t \),

\[
\text{Adv } A \leq \epsilon
\]

Ex: Suppose \( D \) is uniform distribution on graphs with 1,000 nodes and containing a Hamiltonian cycle.

\( D' \) is uniform distribution on graphs with 1,000 nodes without a Hamiltonian cycle.
Conj. D and D' are \(A(100, -01)\). \(A_{11} = \) computationally indistinguishable. 

Notation: \(D^+ \oplus D' = A_{11} \oplus A_{11}

Lemma: (Data Processing Inequality)

If \(D, D'\) are \((+, E)\) compressed and \(f\) is any function executing in time \(t\), then \(f(D), f(D')\)

Proof: (by contrapositive)

Assume \(A\) has \(Adv_A^E = E\) and \(A\) runs in time \(\leq t - t''\). Then \(B = A \circ f\) can distinguish \(D\) and \(D'\) with probability \(E\). B's time is \(t\).

\[
\begin{align*}
0, 1 & \quad \text{(for input)} \\
\text{X} & \quad \text{(output of} \ A \text{)} \\
\text{X} & \quad \text{(output of} \ f(D) \text{)} \\
\text{X} & \quad \text{(output of} \ f(D') \text{)}
\end{align*}
\]

Lemma: \(I \oplus D \oplus D'\) and \(D' \oplus D''\) then \(D \oplus D''\)

Proof: For any \(A\) running in time \(t\),

\[
Adv_A^E = \left| \Pr[A(D) = 1] - \Pr[A(D') = 1] \right|
\]

\[
= \left| \Pr[A(D) = 1] - \Pr[A(D') = 1] \\
+ \Pr[A(D') = 1] - \Pr[A(D'') = 1] \right|
\]
\[ \leq \Pr[a(D) = 1] - \Pr[a(D') = 1] + \Pr[a(D') = 1] - \Pr[a(D'') = 1] \]

\[ = \text{Adv}^a_0 A + \text{Adv}^a_0 A \leq \varepsilon_1 + \varepsilon_2 \]

**Defn:** \( G: \Sigma^0, 13^t \rightarrow \Sigma^0, 13^t \) is a 

\( (+, E) \) Pseudo-random generator (PRG) 

if \( G \circ U \in \Sigma^0 \)

Given a \( L \)-bit string \( x \)

Suppose try to compute \( k \) s.t. \( G(k) = x \)

1. **Success!** Guess \( x \) is from \( G \)
2. **Failure!** Guess \( x \) is uniform.

**Thm:** If \( G_1: \Sigma^0, 13^t \rightarrow \Sigma^0, 13^t \) is \((+, E_1)\) PRG 

and \( G_2: \Sigma^0, 13^t \rightarrow \Sigma^0, 13^m \) is \((+, E_2)\) PRG 

and \( G_2 \) runs in time \( +^t \), then 

\( G_2 \circ G_1 \) is a \((+-^t, E, +E_2)\) PRG

**Proof:** By def., \( U \in \Sigma^0 \), \( G_2 \circ U \)

1. By DPI, \( G_2 \circ U \in \Sigma^0 \), \( G_2 \circ G_1 \circ U \)
2. By def., \( G_2 \circ U \in \Sigma^0 \), \( G_1 \circ U \)
3. By obvious, \( G_2 \circ U \in \Sigma^0 \), \( G_2 \circ G_1 \circ U \)
4. By trans., \( U \in \Sigma^0 \), \( G_2 \circ G_1 \circ U \)
Thm: If $G_i: E_0, 1^L_i \rightarrow E_0, 1^L_i$ is $(+, e_i)$ PRG
and $G_a: E_0, 1^L_a \rightarrow E_0, 1^L_a$ is $(+, e_a)$ PRG
and $G_a$ runs in time $t_a$, then:

$G \parallel G_a$ is a $(\min (+, t_a, t_a + t_2), e_i + e_a)$ PRG

Proof: By def, $G \circ U_{e_i} \overset{t_i}{\rightarrow} G_a \circ U_{e_a} \overset{t_2}{\rightarrow} U_{e_a}
\overset{t_a}{\rightarrow} U_{e_a} \overset{t_2}{\rightarrow} U_{e_a} \overset{t_a}{\rightarrow} U_{e_a}$

By DPI

$f(x) = x \parallel G_a \circ U$