1 Recap

1.1 Ensemble of distributions

An ensemble of distributions $\{X_n\}_{n \in \mathbb{N}}$ ($\{X_n\}$ for short) is a sequence of probability distributions $X_1, X_2, \ldots$

1.2 Computational indistinguishability

Two ensembles of distributions $\{X_n\}$ and $\{Y_n\}$ are said to be computationally indistinguishable ($\{X_n\} \approx \{Y_n\}$) if:

$$\forall \text{nuPPT } D \exists \text{neg } \epsilon \forall n \in \mathbb{N} \left| \Pr[t \leftarrow X_n : D(1^m, t) = 1] - \Pr[t \leftarrow X_n : D(1^m, t) = 1] \right| \leq \epsilon(n)$$

1.3 Two properties of (in)distinguishability

1.3.1 Closure under efficient operations

$\{X_n\} \approx \{Y_n\}$ implies $\{M(X_n)\} \approx \{M(Y_n)\}$ for any nuPPT $M$.

1.3.2 Hybrid lemma

For a sequence of probability distributions $X_1, X_2, \ldots, X_m$, if there exists a machine $D$ that distinguishes $X_1$ from $X_m$ with probability $\epsilon$, then there exists some $i$ such that $D$ distinguishes $X_i$ from $X_{i+1}$ with probability at least $\frac{\epsilon}{m}$.
2 Prediction lemma

A third property of distinguishability, the prediction lemma, intuitively states that if you can distinguish two distributions, then you should be able to guess which distribution an arbitrary sample came from as well.

**Lemma 1** For ensembles \( \{X^0_n\} \) and \( \{X^1_n\} \) where each \( X^0_n \) and \( X^1_n \) is a distribution over \( \{0,1\}^m(n) \) for some polynomial \( m \), let \( D \) be a nuPPT that distinguishes \( \{X^0_n\} \) and \( \{X^1_n\} \) with probability \( \mu(n) \) for infinitely many \( n \). Then there exists an nuPPT \( A \) such that for infinitely many \( n \),

\[
\Pr[b \leftarrow \{0,1\}, t \leftarrow X^b_n : A(t) = b] \geq \frac{1}{2} + \frac{\mu(n)}{2}
\]

**Proof.** Assume without loss of generality that \( D \) outputs 1 with higher probability when getting a sample from \( X^1_n \). This assumption is safe because either \( D(t) \) or \( 1 - D(t) \) will work for infinitely many \( n \), or alternatively, because \( D \) can accept additional information about whether to invert its output as a nuPPT.

We’ll show that \( D \) actually satisfies the above conditions for \( A \), so \( D \) is a predictor:

\[
\Pr[b \leftarrow \{0,1\} : t \leftarrow X^b_n : D(t) = b] = \frac{1}{2} \Pr[t \leftarrow X^1_n : D(t) = 1] + \frac{1}{2} \Pr[t \leftarrow X^0_n : D(t) \neq 1] \\
= \frac{1}{2} \Pr[t \leftarrow X^1_n : D(t) = 1] + \frac{1}{2}(1 - \Pr[t \leftarrow X^0_n : D(t) = 1]) \\
= \frac{1}{2} + \frac{1}{2}(\Pr[t \leftarrow X^1_n : D(t) = 1] - \Pr[t \leftarrow X^0_n : D(t) = 1]) \\
= \frac{1}{2} + \frac{\mu(n)}{2}
\]

Note that there are no restrictions on \( \mu \), but the predictor \( A \) will have some special properties if \( \mu \) is polynomial. This will be discussed in a later lecture.

3 Pseudorandomness

With these three lemmas, we can define pseudorandomness as indistinguishability from the uniform distribution: \( \{X_n\} \) is pseudorandom if \( \{X_n\} \approx \{U_{m(n)}\} \), where \( X_n \) is over \( \{0,1\}^m(n) \), \( m \) is polynomial, and \( U \) is the uniform distribution.

We can show that this definition of pseudorandomness is equivalent to passing the next-bit test using Yao’s theorem.
3.1 Next-bit test

**Definition 1** An ensemble \( \{X_n\} \) over \( \{0, 1\}^{m(n)} \) passes the **next-bit test** if

\[
\forall \text{nuPPT} \ A \ \exists \text{neg} \ \forall n \in \mathbb{N}, i \in [0, 1, \ldots, m(n) - 1] \Pr[t \leftarrow X_n : A(1^n, t_{0\rightarrow i}) = t_{i+1}] \leq \frac{1}{2} + \epsilon(n)
\]

where \( t_{0\rightarrow i} \) denotes the first \( i + 1 \) bits of \( t \).

Intuitively, a prefix of a sample of \( \{X_n\} \) cannot be used to predict the next bit in the sample with high probability.

3.2 Half of Yao’s theorem

**Theorem 2** Any ensemble \( \{X_n\} \) over \( \{0, 1\}^{m(n)} \) that passes the next-bit test is pseudorandom.

**Proof.** The proof will proceed by contradiction, as usual. Assume for the sake of contradiction that there exists a \( D \) distinguishing \( \{X_n\} \) from \( \{U_{m(n)}\} \) with probability \( \frac{1}{p(n)} \) for polynomial \( p \), so \( \{X_n\} \) is not pseudorandom. We will use this \( D \) to predict the next bit of any sample from \( \{X_n\} \).

Consider the hybrid distributions \( H^i_n = \{l \leftarrow X_n, r \leftarrow U_{m(n)} : l_{0\rightarrow i} || r_{i+1 \rightarrow m(n)}\} \). The first \( i \) bits of \( H^i_n \) come from \( X_n \), while the rest are uniformly random (so we can generate them ourselves).

\( H^0_n = U_{m(n)} \) and \( H^{m(n)}_n = X_n \), with each \( H^i_n \) in between injecting \( i \) bits from \( X_n \). By our assumption, \( \{X_n\} \) is distinguishable from \( \{U_{m(n)}\} \), so for infinitely many \( n \), \( H^{m(n)}_n = X_n \) is distinguishable by \( D \) from \( H^0_n = U_{m(n)} \) with probability at least \( \frac{1}{p(n)} \).

Applying the hybrid lemma, there exists some \( i \) such that \( D \) distinguishes \( H^i_n \) from \( H^{i+1}_n \) for each of these \( n \). The only difference between these distributions is that bit \( i + 1 \) of \( H^i_n \) is uniformly random, whereas bit \( i + 1 \) of \( H^{i+1}_n \) is drawn from \( X_n \).

Define another hybrid \( \tilde{H}^{i+1}_n = \{l \leftarrow X_n, r \leftarrow U_{m(n)} : l_{0\rightarrow i} || 1 - l_{i+1} || r_{i+2 \rightarrow m}\} \). \( \tilde{H}^{i+1}_n \) is exactly \( H^{i+1}_n \) with bit \( i + 1 \) flipped; if \( D \) can distinguish \( H^i_n \) from \( H^{i+1}_n \), then it can certainly distinguish \( H^{i+1}_n \) from \( \tilde{H}^{i+1}_n \). We can show this with some algebra:

\[
|\Pr[t \leftarrow H^{i+1}_n : D(t) = 1] - \Pr[t \leftarrow H^i_n : D(t) = 1]| = |\Pr[t \leftarrow H^{i+1}_n : D(t) = 1] - \frac{1}{2} \Pr[t \leftarrow H^{i+1}_n : D(t) = 1] - \frac{1}{2} \Pr[t \leftarrow \tilde{H}^{i+1}_n : D(t) = 1]|.
\]
= \left| \frac{1}{2} \Pr[t \leftarrow H_n^{i+1} : D(t) = 1] - \frac{1}{2} \Pr[t \leftarrow \tilde{H}_n^{i+1} : D(t) = 1] \right|

\geq \frac{1}{p(n)m(n)} \text{ by assumption on } D \text{ and hybrid lemma}

where the second line follows because $H_n^i$ can be expressed as $\frac{1}{2}H_n^{i+1} + \frac{1}{2}\tilde{H}_n^{i+1}$ (making bit $i + 1$ uniformly random).

Now that we can distinguish $H_n^{i+1}$ and $\tilde{H}_n^{i+1}$, we can apply the prediction lemma to show that there exists a nuPPT $A$ for infinitely many $n \in \mathbb{N}$ such that

$$\Pr[b \leftarrow \{0, 1\}, t \leftarrow H_n^{b,i+1} : A(t) = b] \geq \frac{1}{2} + \frac{1}{p(n)m(n)}$$

where $H_n^{0,i+1} = \tilde{H}_n^{i+1}$ and $H_n^{1,i+1} = H_n^{i+1}$.

$A$ can be used to construct a predictor for bit $i + 1$ of infinitely many $X_n$: construct $A'(1^n, y)$ (where $y$ is an $i + 1$ bit prefix of a sample of some such $X_n$) to pick a random $r \leftarrow \{0, 1\}^{m(n) - i}$. If $A(y||r) = 1$, then $A'$ outputs $r_i$, otherwise if $A(y||r) = 0$, then $A'$ outputs $1 - r_i$. $A'$ effectively predicts bit $i + 1$ of $\{X_n\}$ because it satisfies

$$\Pr[t \leftarrow X_n : A'(1^n, t_{0 \rightarrow i}) = t_{i+1}] \geq \frac{1}{2} + \frac{1}{p(n)m(n)}$$

for infinitely many $n$. ■