COM S 6830 – Cryptography	Tuesday, September 22, 2009

Lecture 8: Computational Indistinguishability and Pseudorandomness

Instructor: Rafael Pass *Scribe: Chin Isradisaikul*

In this lecture we introduce the notion of computational indistinguishability and discuss some properties of indistinguishability, including closure under efficient operations, transitivity, and the relationship between distinguishing and predicting. We then define pseudorandomness and consider the completeness of the next-bit test for pseudorandomness.

1 Computational Indistinguishability

1.1 Definitions

We begin with definitions related to indistinguishability, along with some discussions based on these definitions.

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

As a shorthand, we write $\{X_n\}_n$ to denote $\{X_n\}_{n\in\mathbb{N}}$. We will use this shorthand throughout this handout.

Definition 2 Let $\{X_n\}_n$ and $\{Y_n\}_n$ be ensembles, where X_n's and Y_n's are probability distributions over $\{0,1\}^{\ell(n)}$ for some polynomial $\ell(n)$. We say that $\{X_n\}_n$ and $\{Y_n\}_n$ are *computationally indistinguishable* if for all nuPPT $\mathcal{D}$, there exists a negligible function ε such that for every $n \in \mathbb{N}$,

$$|\Pr[t \leftarrow X_n : \mathcal{D}(t) = 1] - \Pr[t \leftarrow Y_n : \mathcal{D}(t) = 1]| \leq \varepsilon(n).$$

That is, an nuPPT decider $\mathcal{D}$ cannot tell apart a sample from X_n and Y_n. Note that we can consider only deciders (which output only one bit) because we can define a boolean function from the output of any nuPPT algorithm—whose output is polynomial in n—to 0 or 1 and evaluate this function in polynomial time. Another way to understand this is to use $\mathcal{D}$ to determine whether an attack or a simulation succeeds, given input t.

Notation: We write $\{X_n\}_n \approx \{Y_n\}_n$ to denote that $\{X_n\}_n$ and $\{Y_n\}_n$ are indistinguishable.

For applications, if $\{X_n\}_n$ and $\{Y_n\}_n$ are indistinguishable, we can use Y_n instead of X_n (and vice versa) because no efficient machine will be able to tell the difference.

From this definition of indistinguishability, it may seem that any two indistinguishable probability distributions should be really close to each other, but in fact there *are* disjoint probability distributions which are also indistinguishable, assuming the existence of one-way functions and the Discrete Log Assumption.

1.2 Properties of Computational Indistinguishability

1.2.1 Closure Under Efficient Operations

First, computational indistinguishability is preserved under *efficient* operations. (An analogy: An object that has closure should look the same with or without sunglasses on, where the sunglasses are operations.) We introduce this property in the following lemma.

Lemma 3 If $\{X_n\}_n$ is indistinguishable from $\{Y_n\}_n$ and M is an nuPPT, then

$$\{M(X_n)\}_n \approx \{M(Y_n)\}_n ,$$

where $\{M(X_n)\}_n$ denotes $\{t \leftarrow X_n : M(t)\}_{n \in \mathbb{N}}$ and similarly for $\{M(Y_n)\}_n$.

Note that the M is the operation in consideration. Before we begin the proof, observe that this lemma holds under *any* operation for identical distributions, but the lemma stated holds only under *efficient* operations if distributions are not identical. Now we prove this lemma.

Proof: Assume for a contradiction that there exists an nuPPT $\mathcal{D}$ and a polynomial p such that for infinitely many n, $\mathcal{D}$ distinguishes $\{M(X_n)\}_n$ and $\{M(Y_n)\}_n$ with probability $\frac{1}{p(n)}$. That is,

$$|\Pr[t \leftarrow M(X_n) : \mathcal{D}(t) = 1] - \Pr[t \leftarrow M(Y_n) : \mathcal{D}(t) = 1]| \geq \frac{1}{p(n)}. \tag{1}$$

We can rewrite Equation 1 as

$$|\Pr[t \leftarrow X_n : \mathcal{D}(M(t)) = 1] - \Pr[t \leftarrow Y_n : \mathcal{D}(M(t)) = 1]| \geq \frac{1}{p(n)}.$$

But this means that we have an nuPPT $\mathcal{D}'(t) = \mathcal{D}(M(t))$ which distinguishes $\{X_n\}_n$ and $\{Y_n\}_n$ with probability $\frac{1}{p(n)}$ for infinitely many n, contradicting the assumption that $\{X_n\}_n \approx \{Y_n\}_n$. □

Remark: For the lemma, we need M to be nuPPT in order for $\mathcal{D}'$ in the proof to be nuPPT.

1.2.2 Transitivity

Instead of proving transitivity directly, we prove the contrapositive of transitivity in the following lemma. Then transitivity is simply a corollary of this lemma.

Lemma 4 (*Hybrid Lemma*) Let $X_1, \ldots, X_m$ be a sequence of probability distributions. Suppose there exists a distinguisher $\mathcal{D}$ that distinguishes X_1 and X_m with probability ε. Then there exists $i \in [1, m-1]$ such that $\mathcal{D}$ distinguishes X_i and X_{i+1} with probability $\frac{\varepsilon}{m}$.

Proof: We will use the triangle inequality which states that $|a+b| \leq |a| + |b|$ in this proof. In general,

$$\left|\sum_{i=1}^{k} x_i\right| \leq \sum_{i=1}^{k} |x_i|.$$

Suppose that $\mathcal{D}$ distinguishes X_1 and X_m with probability ε. Then

$$|\Pr[t \leftarrow X_1 : \mathcal{D}(t) = 1] - \Pr[t \leftarrow X_m : \mathcal{D}(t) = 1]| \geq \frac{1}{p(n)}. \tag{2}$$

Let $g_i = \Pr[t \leftarrow X_i : \mathcal{D}(t) = 1]$. Then Equation 2 can be written as $|g_1 - g_m|$. Now,

$$\begin{aligned}
\varepsilon &\leq |g_1 - g_m| \\
&= |g_1 - g_2 + g_2 - g_3 + \cdots + g_{m-1} - g_m| \\
&= \left|\sum_{i=1}^{m-1} g_i - g_{i+1}\right| \\
&\leq \sum_{i=1}^{m-1} |g_i - g_{i+1}| \qquad \text{(by the triangle inequality).}
\end{aligned}$$

That is, the sum of $m-1$ absolute values must exceed ε. Hence, there must exist some i such that

$$|g_i - g_{i+1}| \geq \frac{\varepsilon}{m-1} > \frac{\varepsilon}{m}$$

(otherwise the sum would be less than ε). But $|g_i - g_{i+1}|$ is exactly the probability that $\mathcal{D}$ distinguishes X_i and X_{i+1}. Therefore, $\mathcal{D}$ distinguishes X_i and X_{i+1} with probability ε/m, as required. □

If there are polynomially many distributions, the Hybrid Lemma holds. If there are more-than-polynomially many distributions, however, it might be the case that all distributions X_i and X_{i+1} are indistinguishable but X_1 and X_m *are* distinguishable! (This may be explored in the next homework assignment.)

Example 5 Assume $\{X_n\}_n \approx \{Y_n\}_n \approx \{Z_n\}_n$. Furthermore, assume that X_n, Y_n, Z_n can be efficiently sampled. That is, $\{X_n\}_n = \{M(1^n)\}_n$, where M is an nuPPT, and similarly for Y_n and Z_n. Show that $\{X_n Y_n\}_n$ is indistinguishable from $\{Z_n Z_n\}_n$, i.e.,

$$\{a \leftarrow X_n; b \leftarrow Y_n : ab\}_n \approx \{a \leftarrow Z_n; b \leftarrow Z_n : ab\}_n .$$

Solution: Let H_1 (X_1 in Hybrid Lemma) be $X_n Y_n$, $H_2 = X_n Z_n$, and $H_3 = Z_n Z_n$. By Hybrid Lemma, if there exists an nuPPT $\mathcal{D}$ that distinguishes H_1 and H_3, then $\mathcal{D}$ must distinguish H_1 and H_2 or H_2 and H_3. Now, if $\mathcal{D}$ distinguishes H_1 and H_2, then let $M(t) = X \leftarrow X_n$ and output Xt, so that $M(Y_n) = H_1$. Similarly, $M(Z_n) = H_2$. Because, $\{Y_n\}_n \approx \{Z_n\}_n$, it follows that $H_1 \approx H_2$ (by closure). Therefore, it must be the case that $\mathcal{D}$ distinguishes H_2 and

H_3. But because $\{X_n\}_n \approx \{Z_n\}_n$, it follows that $\{X_n Z_n\}_n \approx \{Z_n Z_n\}_n$. We finally arrive at a contradiction. Hence, there does not exist an nuPPT that distinguishes H_1 and H_3, i.e., $\{X_n Y_n\}_n \approx \{Z_n Z_n\}_n$.

Note that for this problem, Y_n and Z_n need not be efficiently sampled.

1.2.3 Prediction vs Distinguishability

The next property of indistinguishability is rather intuitive: If two distributions are indistinguishable, it should be difficult to tell which distribution a sample comes from. Once again we will state the contrapositive of this as the following lemma.

Lemma 6 (*Prediction Lemma*) Let $\{X_n^0\}_n$ and $\{X_n^1\}_n$ be ensembles of distributions where X_n^0 and X_n^1 are probability distributions over $\{0,1\}^{\ell(n)}$ for some polynomial $\ell(n)$. Also, let $\mathcal{D}$ be an nuPPT machine that distinguishes $\{X_n^0\}_n$ and $\{X_n^1\}_n$ with probability $\mu(n)$ for infinitely many n's. Then there exists an nuPPT $\mathcal{A}$ such that for infinitely many n's,

$$\Pr[b \leftarrow \{0,1\}\,; t \leftarrow X_n^b : \mathcal{A}(t) = b] \geq \frac{1}{2} + \frac{\mu(n)}{2}.$$

Remark: It is easy to see that prediction implies distinguishability.

Proof: Consider an nuPPT $\mathcal{D}$ that distinguishes $\{X_n^0\}_n$ and $\{X_n^1\}_n$ with probability $\mu(n)$ for infinitely many n's. Without loss of generality, we may assume that for infinitely many n's,

$$\Pr[t \leftarrow X_n^1 : \mathcal{D}(t) = 1] - \Pr[t \leftarrow X_n^0 : \mathcal{D}(t) = 1] \geq \mu(n).$$

Note that we dropped the absolute value. If the difference is negative, consider $\mathcal{D}'(t) = 1 - \mathcal{D}(t)$ instead (or, identically, switch the role of the two distributions).

We show that $\mathcal{D}$ is in fact a predictor, i.e., $\mathcal{A} = \mathcal{D}$. It follows that

$$\begin{aligned}
&\Pr[b \leftarrow \{0,1\}\,; t \leftarrow X_n^b : \mathcal{D}(t) = b] \\
&= \Pr[b = 1]\Pr[t \leftarrow X_n^1 : \mathcal{D}(t) = 1] + \Pr[b = 0]\Pr[t \leftarrow X_n^0 : \mathcal{D}(t) = 0] \\
&= \frac{1}{2}\Pr[t \leftarrow X_n^1 : \mathcal{D}(t) = 1] + \frac{1}{2}\Pr[t \leftarrow X_n^0 : \mathcal{D}(t) = 0] \\
&= \frac{1}{2}\Pr[t \leftarrow X_n^1 : \mathcal{D}(t) = 1] + \frac{1}{2}\big(1 - \Pr[t \leftarrow X_n^0 : \mathcal{D}(t) = 1]\big) \\
&= \frac{1}{2}\big(\Pr[t \leftarrow X_n^1 : \mathcal{D}(t) = 1] - \Pr[t \leftarrow X_n^0 : \mathcal{D}(t) = 1]\big) + \frac{1}{2} \\
&\geq \frac{1}{2}\mu(n) + \frac{1}{2} = \frac{1}{2} + \frac{\mu(n)}{2},
\end{aligned}$$

as required. □

2 Pseudorandomness

With the notion of computational indistinguishability, we are now ready to discuss pseudorandomness.

2.1 Definitions

Definition 7 Let $\{X_n\}_n$ be an ensemble of distributions, where X_n is a probability distribution over $\{0,1\}^{\ell(n)}$, where ℓ is polynomial. Then $\{X_n\}_n$ is *pseudorandom* if $\{X_n\}_n \approx \{U_n\}_n$, where $U_n = \{t \leftarrow \{0,1\}^n\}$ is the uniform distribution.

Before we continue with pseudorandomness and how to construct a pseudorandom generator, we show that the statistical test for randomness (discussed in the previous lecture)—that given a prefix of a bit sequence, one should not be able to predict the next bit—is complete.

Definition 8 An ensemble of distributions $\{X_n\}_n$, where X_n is over $\{0,1\}^{m(n)}$ is said to *pass the next-bit test* if for all nuPPT $\mathcal{A}$, there exists a negligible function ε such that for all $n \in \mathbb{N}$ and for all $i \in [m(n)]$,

$$\Pr[t \leftarrow X_n : \mathcal{A}(t_{0\to i}) = t_{i+1}] \leq \frac{1}{2} + \varepsilon(n),$$

where $t_{0\to i}$ denotes all bits from the beginning to position i and t_{i+1} denotes the $(i+1)^{\text{st}}$ bit.

Observe that the probability is $\frac{1}{2}$ for blind guessing. That is, this definition tells us that a distribution passes the next-bit test if no nuPPT can do much better than random guessing.

2.2 Completeness of the Next-Bit Test

Surprisingly, this simple next-bit test is complete. In other words, the next-bit test is sufficient to determine whether an ensemble is pseudorandom. We illustrate this in the following theorem due to Yao.

Theorem 9 [Yao] An ensemble $\{X_n\}_n$ passes the next-bit test if and only if it is pseudorandom.

Proof: ($\Leftarrow$) Prove by contradiction. Suppose that $\{X_n\}_n$ is not pseudorandom. Then there is an nuPPT $\mathcal{D}$ that distinguishes $\{X_n\}_n$ from the uniform distribution. Then for infinitely many n's,

$$|\Pr[t \leftarrow X_n : \mathcal{D}(t) = 1] - \Pr[t \leftarrow U_n : \mathcal{D}(t) = 1]| \geq \frac{1}{p(n)}$$

for some polynomial p. In particular, we can treat t as $t_{0\to i}$. Since $\Pr[t \leftarrow U_n : \mathcal{D}(t) = 1] = \frac{1}{2}$ by definition of uniform distribution, it follows that

$$\begin{aligned} |\Pr[t \leftarrow X_n : \mathcal{D}(t) = 1] - \Pr[t \leftarrow U_n : \mathcal{D}(t) = 1]| &= \left|\Pr[t \leftarrow X_n : \mathcal{D}(t) = 1] - \frac{1}{2}\right| \\ &\geq \frac{1}{p(n)}. \end{aligned}$$

That is,

$$\begin{aligned} \Pr[t \leftarrow X_n : \mathcal{D}(t) = 1] - \frac{1}{2} &\geq \frac{1}{p(n)} \\ \Pr[t \leftarrow X_n : \mathcal{D}(t) = 1] &\geq \frac{1}{2} + \frac{1}{p(n)} \end{aligned}$$

or

$$\begin{aligned} -\Pr[t \leftarrow X_n : \mathcal{D}(t) = 1] + \frac{1}{2} &\geq \frac{1}{p(n)} \\ \Pr[t \leftarrow X_n : \mathcal{D}(t) = 1] &\leq \frac{1}{2} - \frac{1}{p(n)} \\ 1 - \Pr[t \leftarrow X_n : \mathcal{D}(t) = 0] &\leq \frac{1}{2} - \frac{1}{p(n)} \\ \Pr[t \leftarrow X_n : \mathcal{D}(t) = 0] &\geq \frac{1}{2} + \frac{1}{p(n)} \end{aligned}$$

In any case, we come to a conclusion that $\Pr[t \leftarrow X_n : \mathcal{D}(t_{0\to i}) = t_{i+1}] \geq \frac{1}{2} + \frac{1}{p(n)}$, contradicting the assumption that $\{X_n\}_n$ passes the next-bit test. Therefore, if $\{X_n\}_n$ passes the next bit test, then it must be pseudorandom.

($\Rightarrow$) [This direction of the proof will be discussed in the next lecture.]