CS 6222 Grad Crypto September 4, 2025

Lecture 4: Properties of Computational Indistinguishability

Lecturer: Jack Doerner *Scribe: Sabrina Lopez*

1 Topics Covered

- Useful Lemmas about Computational Indistinguishability
- Pseudorandom Generators Imply P≠NP

Note 1. *If a distinguisher cannot tell the difference between two distributions, then they are indistinguishable. This concept can be formalized as the following definition.*

2 Computational Indistinguishability

Definition 1 (Computational Indistinguishability)**.** *Let $\mathcal{X} = \{X_n\}_{n\in\mathbb{N}}$ and $\mathcal{Y} = \{Y_n\}_{n\in\mathbb{N}}$ be ensembles such that, $\forall n \in \mathbb{N}$, X_n and Y_n are distributions on $\{0,1\}^{\ell(n)}$ for polynomial ℓ. With that in mind, $\mathcal{X}$ and $\mathcal{Y}$ are computationally indistinguishable if and only if, $\forall$ NUPPT (Non-Uniform Probabilistic Polynomial-Time) distinguishers D, there $\exists$ a negligible ε such that, $\forall n \in \mathbb{N}$*

$$|Pr[D(1^n, t) = 1 : t \leftarrow X_n] - Pr[D(1^n, t) = 1 : t \leftarrow Y_n]| < \varepsilon(n)$$

To start explaining the equation, here's the following clarification,

- t represents a random sample from the distribution (e.g., X_n or Y_n), and
- $D(1^n, t) = 1$ represents the event that the distinguisher D, given t and a unary encoding of the security parameter n, outputs 1. An output of 1 does not indicate anything in particular.

With all that said, the equation essentially states that the absolute difference of the probabilities of the distinguisher figuring out that a sample is from one distribution and another distribution is less than negligible ε or simply negligible. This means, for all distinguishers, that they cannot tell which distribution a sample is from at all.

Note 2. *The notation for computation indistinguishability between two ensembles is the following: $\mathcal{X} \approx_c \mathcal{Y}$.*

Note 3. *The definition for computational indistinguishability requires that for some n_0 and every $n > n_0$, the two distributions X_n and Y_n pass all efficient statistical tests that might be used to distinguish them. For example, a statistical test for distinguishing whether a sample comes from the uniform distribution or some other distribution might include:*

- *Checking that there are roughly as many 0 as 1 in the sample.*
- *Checking that each sequence of bits occurs with roughly the same probability.*
- *Checking that given any prefix of a sample, some strategy for guessing the next bit succeeds with probability roughly 1/2.*[1]

Theorem 1 (Computational Indistinguishability is Closed Under NUPPT Post-processing)**.** *If* $\{X_n\}_{n\in\mathbb{N}} \approx_c \{Y_n\}_{n\in\mathbb{N}}$*, then* $\forall$ *NUPPT machines* M*,* $\{M(X_n)\}_{n\in\mathbb{N}} \approx_c \{M(Y_n)\}_{n\in\mathbb{N}}$*.*

Proof. Suppose towards contradiction that there $\exists$ NUPPT D, polynomial p such that $p(n)$ is positive as $n \to \infty$ and

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

for infinitely many $n \in \mathbb{N}$.[2] With that said, let R be a *reduction* such that $R(1^n, u) = D(1^n, M(u))$. Consider the advantage of R in distinguishing $\mathcal{X}$ from $\mathcal{Y}$. For infinitely many $n \in \mathbb{N}$,

$$\begin{aligned}
&|Pr[R(1^n, u) = 1 : u \leftarrow X_n] - Pr[R(1^n, u) = 1 : u \leftarrow Y_n]| \\
&= |Pr[D(1^n, M(u)) = 1 : u \leftarrow X_n] - Pr[D(1^n, M(u)) = 1 : u \leftarrow Y_n]| && \text{by the def. of } R \\
&= |Pr[D(1^n, t) = 1 : t \leftarrow M(X_n)] - Pr[D(1^n, t) = 1 : t \leftarrow M(Y_n)]| && \text{by rearrangement} \\
&\geq \frac{1}{p(n)} && \text{by our supposition}
\end{aligned}$$

This contradicts the computational indistinguishability of $\mathcal{X}$ and $\mathcal{Y}$. Therefore, no such D with a non-negligible distinguishing advantage can exist, and the theorem follows. □

Theorem 2 (Computational Indistinguishability is Transitive)**.** *Let* $\{X^i\}_{i\in[m]}$ *be a sequence of distributions for some constant* m*. If* $\exists$ *any distinguisher* D[3] *and any non-negative constant* ε *such that*

$$|Pr[D(x) = 1 : x \leftarrow X^1] - Pr[D(x) = 1 : x \leftarrow X^m]| \geq \varepsilon \tag{1}$$

then $\exists\ i \in [m-1]$ *such that*

$$|Pr[D(x) = 1 : x \leftarrow X^i] - Pr[D(x) = 1 : x \leftarrow X^{i+1}]| \geq \frac{\varepsilon}{m-1}$$

[1] If this holds for all prefixes, and all strategies, it is known as the *Next-Bit Test*. The next-bit test is complete for all statistical tests [Ps10, Theorem 75.4].

[2] $M(X_n)$ and $M(Y_n)$ represent distributions induced by applying M to samples from distributions X_n and Y_n respectively. Meanwhile, since $p(n)$ is a polynomial, $\frac{1}{p(n)}$ is a non-negligible quantity. This statement communicates that D outputs 1 with non-negligibly greater or lesser probability when given samples from $M(X_n)$ than when given samples from the $M(Y_n)$, or in other words it violates computational indistinguishability, or simply *it distinguishes*.

[3] Not necessarily bounded.

Proof. Let $p_i = Pr[D(x) = 1 : x \leftarrow X_i]$, and suppose toward contradiction that $\forall\ i \in [m-1]$ we have $|p_i - p_{i+1}| < \frac{\varepsilon}{m-1}$. It follows that

$$\sum_{i=1}^{m-1} (|p_i - p_{i+1}|) < (m-1) \cdot \frac{\varepsilon}{m-1} = \varepsilon.$$

The summation on the left hand side can be expanded in the following way:

$$\begin{aligned} & |p_1 - p_2| + |p_2 - p_3| + \cdots + |p_{m-1} - p_m| \\ & \geq |p_1 - p_2 + p_2 - p_3 + \cdots + p_{m-1} - p_m| \qquad \text{by the triangle inequality [Wei25]} \\ & = |p_1 - p_m| \end{aligned}$$

which then implies that

$$|p_1 - p_m| < \varepsilon$$

in contradiction to Equation 1. Therefore, if $|p_1 - p_m| \geq \varepsilon$, then $\exists i \in [m-1]$ such that $|p_i - p_{i+1}| \geq \frac{\varepsilon}{m-1}$. □

Note 4 (On the Uses of Theorem 2)**.** *To help you understand why this theorem is useful, consider the sequence of ensembles* $\{\mathcal{X}^i\}_{i\in[m]}$ *such that* $\forall i \in [m]$, $\mathcal{X}^i = \{X_n^i\}_{n\in\mathbb{N}}$*. If there exists some NUPPD distinguisher* D *and some polynomial* p *such that* $p(n)$ *is positive as* $n \to \infty$ *and for infinitely many* $n \in \mathbb{N}$

$$|Pr[D(x) = 1 : x \leftarrow X_n^1] - Pr[D(x) = 1 : x \leftarrow X_n^m]| \geq \frac{1}{p(n)}$$

then by Theorem 2, for infinitely many $n \in \mathbb{N}$ *there exists some* $i_n \in [m-1]$ *such that*

$$|Pr[D(x) = 1 : x \leftarrow X_n^{i_n}] - Pr[D(x) = 1 : x \leftarrow X_n^{i_n+1}]| \geq \frac{1}{(m-1) \cdot p(n)}.$$

Since $1/p(n)$ *is non-negligible and* m *is constant,* $1/((m-1) \cdot p(n))$ *is also non-negligible, and therefore if* D *can distinguish* $\mathcal{X}^1$ *from* $\mathcal{X}^m$ *then there exists some* $i \in [m-1]$ *such that* D *can distinguish* $\mathcal{X}^i$ *from* $\mathcal{X}^{i+1}$*.*

Corollary 1. *If* $\mathcal{X} \approx_c \mathcal{Y}$ *and* $\mathcal{Y} \approx_c \mathcal{Z}$*, then* $\mathcal{X} \approx_c \mathcal{Z}$*.*

In other words, if no efficient distinguisher or algorithm can tell the difference between $\mathcal{X}$ and $\mathcal{Y}$ or $\mathcal{Y}$ and $\mathcal{Z}$, then none can tell the difference between $\mathcal{X}$ and $\mathcal{Z}$.

Theorem 3 (Prediction Lemma)**.** *Let* ℓ *be a polynomial and let* $\mathcal{X}^b = \{X_n^b\}_{n\in\mathbb{N}}$ *for* $b = \{0,1\}$ *be defined such that* X_n^b *is a distribution on* $\{0,1\}^{\ell(n)}$*.* $\mathcal{X}^0 \approx_c \mathcal{X}^1$ *if and only if* $\forall$ *NUPPT prediction algorithms* A*,* $\exists$ *some negligible function* ε *such that* $\forall n \in \mathbb{N}$*,*

$$\left| Pr[(A(1^n, t) = b : b \leftarrow \{0,1\}, t \leftarrow X_n^b] - \frac{1}{2} \right| < \varepsilon(n) \tag{2}$$

Proof. We can see that the *if* direction of the theorem holds by contraposition: if there exists some A satisfying Equation 2, then it trivially distinguishes $\mathcal{X}^0$ from $\mathcal{X}^1$. The remainder of the proof deals with the *only if* direction by contraposition; i.e. we will show that if there exists any NUPPT D that distinguishes $\mathcal{X}^0$ from $\mathcal{X}^1$ with non-negligible advantage, then there exists some A violating Equation 2.

Suppose without loss of generality[4] that $\exists$ a NUPPT distinguisher D and a non-negligible function μ such that

$$|Pr[D(1^n, t) : t \leftarrow X_n^1] - Pr[D(1^n, t) : t \leftarrow X_n^0]| > \mu(n) \tag{3}$$

and consider what happens if we use D to predict whether a sample came from $\mathcal{X}^0$ or $\mathcal{X}^1$:

$$\begin{aligned}
&Pr[D(1^n, t) = b : b \leftarrow \{0,1\}, t \leftarrow X_n^b] \\
&= \frac{1}{2}(Pr[D(1^n, t) = 1 : t \leftarrow X_n^1] + Pr[D(1^n, t) \neq 1 : t \leftarrow X_n^0]) \\
&= \frac{1}{2}(Pr[D(1^n, t) = 1 : t \leftarrow X_n^1] + 1 - Pr[D(1^n, t) = 1 : t \leftarrow X_n^0]) \\
&\frac{1}{2} + \frac{1}{2}(Pr[D(1^n, t) : t \leftarrow X_n^1] - Pr[D(1^n, t) : t \leftarrow X_n^0]) \\
&> \frac{1}{2} + \frac{\mu(n)}{2} \qquad \text{by plugging in Eqn. 3}
\end{aligned}$$

Note that the prediction advantage $\frac{\mu(n)}{2}$ is non-negligible, since $\mu(n)$ is. □

Note 5 (On the Meaning of Theorem 3)**.** *One way to read this theorem is that there is an algorithm to tell with non-negligible advantage which of two distributions a sample came from if and only if there is an algorithm that distinguishes the distributions with non-negligible advantage, or: good distinguishers imply good predictors and vice versa.*

3 Pseudo-random Generator

Definition 2 (Pseudorandom Generator (PRG))**.** *Let U_n be the uniform distribution on $\{0,1\}^n$ and let ℓ be a polynomial. The function $G : \{0,1\}^n \to \{0,1\}^{\ell(n)}$ is a PRG if:*

- $\ell(n) > n$ [5]
- *G is deterministic and runs in polynomial time*
- $\{G(x) : x \leftarrow U_n\}_{n\in\mathbb{N}} \approx_c \{U_{\ell(n)}\}_{n\in\mathbb{N}}$

[4] If instead there exists D' such that

$$|Pr[D'(1^n, t) : t \leftarrow X_n^0] - Pr[D'(1^n, t) : t \leftarrow X_n^1]| > \mu(n)$$

then we can construct D from D' by inverting the output.

[5] G expands its input to be larger than n

Theorem 4. *If there* $\exists$ *a PRG, then* $\mathsf{P} \neq \mathsf{NP}$.

Proof. Given a PRG $G : \{0,1\}^n \to \{0,1\}^{\ell(n)}$, let language $L = \text{image}(G) = \{G(x) : x \in \{0,1\}^*\}$, $\forall\ y \in L$, $\exists$ a witness x such that $G(x) = y$. G efficiently verifies membership in L given a witness, and thus $L \in \mathsf{NP}$. Suppose towards contradiction that $L \in \mathsf{P}$. By the definition of polynomial-time-recognizable languages, $\exists$ a polynomial-time algorithm A such that $A(y) = 1 \iff y \in L$. It follows $\forall n \in \mathbb{N}$ that

$$Pr[A(G(x)) = 1 : x \leftarrow \{0,1\}^n] = 1$$

and

$$Pr[A(y) = 1 : y \leftarrow (\{0,1\}^{\ell(n)} \setminus \{G(x) : x \in \{0,1\}^n\})] = 0$$

which contradicts the PRG security of G. Therefore, $L \notin \mathsf{P}$ and $\mathsf{P} \neq \mathsf{NP}$. $\square$

References

[Ps10] Rafael Pass and abhi shelat. A course in cryptography. `https://www.cs.cornell.edu/courses/cs4830/2010fa/lecnotes.pdf`, 2010.

[Wei25] Eric W. Weisstein. Triangle inequality. `https://mathworld.wolfram.com/TriangleInequality.html`, 2025.