Lecture 09: Next-bit Unpredictability

Indistinguishability

Consider two distributions X and Y over the sample space Ω. The distributions X and Y are ε-indistinguishable from each other if:

- For all algorithms $\mathcal{A}\colon \Omega \to \{0,1\}$ the following holds

$$\left|\Pr[\mathcal{A}(X) = 1] - \Pr[\mathcal{A}(Y) = 1]\right| \leqslant \varepsilon$$

- This is also equivalent to the following: For all algorithms $\mathcal{A}\colon \Omega \to \{0,1\}$ the following holds

$$\mathrm{SD}\left(\mathcal{A}(X), \mathcal{A}(Y)\right) \leqslant \varepsilon$$

- Represented by $X \approx_\varepsilon Y$
- Think: Why are these equivalent?

Computational Indistinguishable

Definition

The distributions X and Y over the sample space Ω are ε-computationally indistinguishable, if: For all efficient algorithms $\mathcal{A}\colon \Omega \to \{0,1\}$ the following holds

$$\mathrm{SD}\left(\mathcal{A}(X), \mathcal{A}(Y)\right) \leqslant \varepsilon$$

Represented by $X \approx_{\varepsilon}^{(c)} Y$

Negligible Functions

Definition (Negligible)

A function $\varepsilon\colon \mathbb{N} \to \mathbb{R}$ is a negligible function if for all constant c, (eventually) we have $\varepsilon(n) \leqslant 1/n^c$.

- The term "eventually" can be ignored for this class
- Intuition: A negligible function is smaller than all inverse-polynomial functions
- Negligible function examples: $\frac{1}{2^n}, \frac{1}{2^{\sqrt{n}}}, \frac{1}{2^{\log^2 n}}$
- Non-negligible function examples: $\frac{1}{n^{100}}, \frac{1}{2^{\log n}}$

Definition (Pseudo-random Generators)

An efficient function $G\colon \{0,1\}^n \to \{0,1\}^{n+\ell}$ is a pseudo-random generator, if $\ell \geqslant 1$ and there exists a negligible ε such that

$$\mathrm{SD}\left(G(U_{\{0,1\}^n}), U_{\{0,1\}^{n+\ell}}\right) \leqslant \varepsilon$$

- We write ε instead of $\varepsilon(n)$ for brevity
- Intuition: Any computationally bounded adversary cannot distinguish the output of G from a uniform distribution

Next-bit Unpredictability

A string $y_1 \dots y_i$ sampled according to a distribution X has next-bit unpredictability if the following is satisfied.
Consider the game between an honest challenger and an arbitrary efficient adversary $\mathcal{A}$

- The honest challenge $\mathcal{H}$ samples $\equiv y_1 \dots y_i \sim X$. Let $\alpha = y_1 \dots y_{i-1}$ and sample $b \xleftarrow{\$} \{0,1\}$. If $b = 0$, define $\beta = y_i$; otherwise define $\beta \xleftarrow{\$} \{0,1\}$. Send (α, β) to the adversary $\mathcal{A}$.
- The adversary replies back with a bit $\widetilde{b}$.
- The adversary wins the game if $b = \widetilde{b}$. The honest challenger $\mathcal{H}$ sets $z = 1$ if $b = \widetilde{b}$; otherwise $z = 0$. Output z.
- The advantage of $\mathcal{A}$ is negligible, i.e. there is a negligible function ε such that $\left|\Pr[z=1] - \frac{1}{2}\right| \leqslant \varepsilon$

Next-bit Unpredictable PRG

- Let X be a distribution over the sample space $\{0,1\}^n$
- We use $X_{\leqslant i}$ to represent the distribution of first i-bits of X
- In particular, $G(U_n)_{\leqslant i}$ represents the distribution of the first i-bits of the output of the function G with n-bit uniformly random string as input.

Definition (Next-bit Unpredictable PRG)

An efficient function $G\colon \{0,1\}^n \to \{0,1\}^{n+\ell}$, for $\ell \geqslant 1$, is a next-bit unpredictable PRG if, for all $i \in \{1,\dots,n+\ell\}$, the distribution $G(U_n)_{\leqslant i}$ is next-bit unpredictable.

Equivalence of PRG Definitions

Theorem

For an efficient function $G\colon \{0,1\}^n \to \{0,1\}^{n+\ell}$, where $\ell \geqslant 1$, the following two statements are equivalent:

1. *For arbitrary efficient $\mathcal{A}$, there exists a negligible ε such that* $\mathrm{SD}\left(\mathcal{A}(G(U_{\{0,1\}^n})), \mathcal{A}(U_{\{0,1\}^{n+\ell}})\right) \leqslant \varepsilon$
2. *For every $i \in \{1, \dots, n+\ell\}$ the distribution $G(U_{\{0,1\}^n})_{\leqslant i}$ is next-bit unpredictable*

- To show this theorem, we will have to show "(1) $\implies$ (2)" and "(2) $\implies$ (1)"

$(1) \implies (2)$

- This direction should be easy. We expect that a pseudo-random string should definitely satisfy the condition that "given its first $(i-1)$-bits the next bit is completely unpredictable" (because uniformly random bits have this property as well)
- We will prove the contrapositive: "$\neg(2) \implies \neg(1)$"
- Mathematically, $\neg(2)$ gives us: There exists i such that $G(U_{\{0,1\}^n})_{\leqslant i}$ is next-bit predictable. That is, there exists an adversary $\mathcal{A}^*$ that has advantage $1/n^c$ in the next-bit unpredictability experiment.
- Target Mathematical Statement: We need to show $\neg(1)$. This is equivalent to constructing an adversary $\widetilde{\mathcal{A}}$ such that $\mathrm{SD}\left(\widetilde{\mathcal{A}}(G(U_{\{0,1\}^n})), \widetilde{\mathcal{A}}(U_{\{0,1\}^{n+\ell}})\right)$ is at least an inverse polynomial.

$(1) \implies (2)$

Adversary $\widetilde{\mathcal{A}}$ construction: On input $y_1 \dots y_{n+\ell}$ the adversary $\widetilde{\mathcal{A}}$ does the following:

- Define $\alpha = y_1 \dots y_{i-1}$
- Pick $b \xleftarrow{\$} \{0,1\}$
- If $b = 0$, set $\beta = y_i$; otherwise $\beta \xleftarrow{\$} \{0,1\}$.
- Send (α, β) to $\mathcal{A}^*$
- Receive $\widetilde{b}$ from $\mathcal{A}^*$
- Output $z = 1$ if $b = \widetilde{b}$; otherwise $z = 0$.

(1) $\implies$ (2)

Suppose $y_1 \dots y_{n+\ell}$ is sampled from $U_{n+\ell}$

- For every fixed value of α, the value of β is a uniform independent random bit irrespective of whether $b = 0$ or $b = 1$
- So, in this case, $\Pr[z = 1] = 1/2$
- The output of $\widetilde{\mathcal{A}}$ is identical to the $U_{\{0,1\}}$ random variable
- Equivalently $\widetilde{\mathcal{A}}(U_{\{0,1\}^{n+\ell}}) = U_{\{0,1\}}$

$(1) \implies (2)$

Suppose $y_1 \dots y_{n+\ell}$ is sampled from $G(U_n)$

- In this case, the output of $\widetilde{\mathcal{A}}$ is identical to the output of the next-bit unpredictability experiment between the honest challenger $\mathcal{H}$ and the adversary $\mathcal{A}^*$
- In this case, we know that the advantage of $\mathcal{A}^*$ is $1/n^c$
- That is, we know that $\Pr[z = 1] = \frac{1}{2} + 1/n^c$
- The output of $\widetilde{\mathcal{A}}$ is a distribution that outputs 0 with probability $\frac{1}{2} - 1/n^c$ and outputs 1 with probability $\frac{1}{2} + 1/n^c$. For brevity we will call it the $(\frac{1}{2} - 1/n^c, \frac{1}{2} + 1/n^c)$ distribution.
- Equivalently $\widetilde{\mathcal{A}}(G(U_{\{0,1\}^n})) = (\frac{1}{2} - 1/n^c, \frac{1}{2} + 1/n^c)$

Now, we have

$$\begin{aligned}\mathrm{SD}\left(\widetilde{\mathcal{A}}(G(U_{\{0,1\}^n})), \widetilde{\mathcal{A}}(U_{\{0,1\}^{n+\ell}})\right) &= \mathrm{SD}\left(\left(\frac{1}{2} - 1/n^c, \frac{1}{2} + 1/n^c\right), U_{\{0,1\}}\right) \\ &= 1/n^c\end{aligned}$$

So, we have shown that the efficient adversary $\widetilde{\mathcal{A}}$ can distinguish the output of a PRG from a uniform distribution. This completes the proof in one direction.