| | |
|---|---|
| **CS276: Cryptography** | Due date: October 3, 2017 |

# Problem Set 2

*Instructor: Alessandro Chiesa* — *GSI: Benjamin Caulfield*

## Problem 1

Two ensembles $X = \{X_k\}_{k\in\mathbb{N}}$ and $Y = \{Y_k\}_{k\in\mathbb{N}}$ are statistically indistinguishable, denoted $X \simeq Y$, if for all positive constants $c$ and sufficiently large $k$,

$$\frac{1}{2}\sum_{\alpha\in\{0,1\}^k}\Big|\Pr\big[X_k=\alpha\big]-\Pr\big[Y_k=\alpha\big]\Big| < \frac{1}{k^c}\ .$$

1. Prove that if $X$ and $Y$ are statistically indistinguishable, then they are computationally indistinguishable.
2. Show that there exist two ensembles $X$ and $Y$ that are computationally indistinguishable but *not* statistically indistinguishable. (Do not use any computational assumption!)

## Problem 2

Let $G$ be a pseudorandom generator with expansion factor $\ell$ and let $h$ be any (not necessarily polynomial-time computable) length-preserving permutation over $\{0,1\}^*$. (The *expansion factor* of a pseudorandom generator $G$ is a positive polynomial $\ell$ such that $|G(x)| = \ell(k)$ for all $x \in \{0,1\}^k$ and $k \in \mathbb{N}$.)

**1)** Is it always the case that $\{s \leftarrow \{0,1\}^k : h(G(s))\}$ and the uniform distribution over $\{0,1\}^{\ell(k)}$ are computationally indistinguishable? Is $G'(s) \equiv h(G(s))$ a pseudorandom generator?

**2)** Is it always the case that $\{s \leftarrow \{0,1\}^k : G(h(s))\}$ and the uniform distribution over $\{0,1\}^{\ell(k)}$ are computationally indistinguishable? Is $G'(s) \equiv G(h(s))$ a pseudorandom generator?

**3)** If you know that $h$ is polynomial-time computable, do your answers to (1) and (2) change?

## Problem 3

Let $G_1$ and $G_2$ be pseudorandom generators with respective expansion factors $\ell_1$ and $\ell_2$. For each of the candidates below, justify whether the function is a pseudorandom generator or not.

**A:** $G_A(x) = \mathsf{reverse}(G_1(x))$, where $\mathsf{reverse}(\cdot)$ reverses the bits of its argument.

**B:** $G_B(x) = G_1(x)||G_2(x)$.

**C:** $G_C(x||y) = G_1(x)||G_2(y)$, where $|x| = |y|$ or $|x| = |y| + 1$.

**D:** $G_D(x) = G_2(G_1(x))$.

**E:** $G_E(x) = G_1(x) \oplus \big(x||0^{\ell_1(|x|)-|x|}\big)$.

## Problem 4

Let $\mathcal{F} = \{F_s \colon \{0,1\}^k \to \{0,1\}^k\}_{s \in \{0,1\}^k}$ be a pseudorandom function. For each of the candidates below, justify whether the function is a pseudorandom function or not.

1. $G_s(x) = F_s(x) || F_s(\bar{x})$.
2. $G_s(x) = F_{0^k}(x) || F_s(x)$.
3. $G_s(x) = F_{s_1}(x) || F_{s_2}(x)$, where $s_1 \equiv F_s(0^k)$ and $s_2 \equiv F_s(1^k)$.
4. $G_s(x) = F_x(s)$.
5. $G_s(x) = F_s(x) \oplus s$.
6. $G_{s_1,s_2}(x) = (F_{s_1}(x) \oplus s_2) || F_{s_2}(x)$ (where $|s_1| = |s_2| = k$; consider only even-length seeds for $G$).
7. $G_s(x) = F_{F_s(x)}(x)$.

## Problem 5

In this problem we consider two other ways of modeling what it means to be a pseudorandom function family, and investigate how these new definitions compare to the one we discussed.

1. In the definition of a PRF, we allow for an adversary to *adaptively* query its oracle in order to distinguish whether the oracle is truly random or pseudorandom. Suppose we now consider *non-adaptive* tests: an adversary provides a list of testing points, then receives the values of the oracle at each of those testing points, and finally makes a decision (without consulting the oracle again).

   **Definition 1 (Non-Adaptive Pseudo-Random Function Families)** *A function family is* non-adaptively *pseudorandom if a random member of the family is indistinguishable from a random function, under all polynomial-time non-adaptive tests.*

   Is the above definition of PRF strictly stronger, strictly weaker, equivalent, or incomparable to our original adaptive notion? Prove your answer.

2. Now we consider a different kind of test in which we see if not being able to *predict* an output of a function is equivalent to the function seeming random. A *predictor* is allowed to adaptively query the oracle on several points, and then outputs a pair $(x, y)$. The predictor succeeds in the test if: (1) it has not already queried the oracle on point $x$, and (2) the value of the oracle at $x$ is equal to $y$.

   **Definition 2 (Unpredictable Function Family)** *A function family is* unpredictable *if no polynomial-time oracle machine can succeed in the prediction experiment with non-negligible advantage over random guessing.*

   Is the above definition of PRF strictly stronger, strictly weaker, equivalent, or incomparable to our original adaptive notion? Prove your answer.