MTAT.07.003 Cryptology II
Spring 2008 / Homework 2

1. Recall that a game is a two-party protocol between the challenger $\mathcal{G}$ and an adversary $\mathcal{A}$ and that the output of the game $\mathcal{G}^{\mathcal{A}}$ is always determined by the challenger. Prove the following claims:

   (a) Any hypothesis testing scenario $\mathcal{H}$ can be formalised as a game $\mathcal{G}$ such that $\Pr[\mathcal{A} = b|\mathcal{H}] = \Pr[\mathcal{G}^{\mathcal{A}} = b]$ for all adversaries $\mathcal{A}$.

   (b) For two simple hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, there is a game $\mathcal{G}$ such that

   $$\mathsf{cd}^t_\star(\mathcal{H}_0, \mathcal{H}_1) = 2 \cdot \max_{\mathcal{A}\text{ is }t\text{-time}} \left|\Pr[\mathcal{G}^{\mathcal{A}} = 1] - \tfrac{1}{2}\right| \enspace .$$

   (c) The computational distance between games defined as

   $$\mathsf{cd}_\star(\mathcal{G}_0, \mathcal{G}_1) = \max_{\mathcal{A}\text{ is }t\text{-time}} |\Pr[\mathcal{G}_0^{\mathcal{A}} = 1] - \Pr[\mathcal{G}_1^{\mathcal{A}} = 1]|$$

   is pseudo-metric, i.e. it satisfies following constraints:

   $$\mathsf{cd}^t_\star(\mathcal{G}_0, \mathcal{G}_1) = \mathsf{cd}^t_\star(\mathcal{G}_1, \mathcal{G}_0) \enspace ,$$
   $$\mathsf{cd}^t_\star(\mathcal{G}_0, \mathcal{G}_2) \leq \mathsf{cd}^t_\star(\mathcal{G}_0, \mathcal{G}_1) + \mathsf{cd}^t_\star(\mathcal{G}_1, \mathcal{G}_2) \enspace .$$

   When is the computational distance a proper metric, i.e.,

   $$\mathsf{cd}^t_\star(\mathcal{G}_0, \mathcal{G}_1) \neq 0 \qquad \Leftrightarrow \qquad \mathsf{sd}_\star(\mathcal{G}_0, \mathcal{G}_1) = 0 \enspace ?$$

   (d) Usually, the statistical distance $\mathsf{sd}_\star(\mathcal{G}_0, \mathcal{G}_1)$ is defined as a limiting value $\mathsf{sd}_\star(\mathcal{G}_0, \mathcal{G}_1) = \lim_{t\to\infty} \mathsf{cd}^t_\star(\mathcal{G}_0, \mathcal{G}_1)$. Give an alternative interpretation in terms of output distributions.

2. Let $\mathcal{A}$ be a t-time distinguisher and let $\alpha(\mathcal{A}) = \Pr[\mathcal{A} = 1|\mathcal{H}_0]$ and $\beta(\mathcal{A}) = \Pr[\mathcal{A} = 0|\mathcal{H}_1]$ be the ratios of false negatives and false positives. Show that for any c there exists a $t + \mathrm{O}(1)$-time adversary $\mathcal{B}$ such that

   $$\alpha(\mathcal{B}) = (1 - c) \cdot \alpha(\mathcal{A}) \qquad \text{and} \qquad \beta(\mathcal{B}) = c + (1 - c) \cdot \beta(\mathcal{A}) \enspace .$$

   Are there any practical settings where such trade-offs are economically justified? Give some real world examples.

   Hint: What happens if you first throw a fair coin and run $\mathcal{A}$ only if you get tail and otherwise output 0?

3. Let $\mathcal{X}_0$ and $\mathcal{X}_1$ efficiently samplable distributions that are (t, ε)-indistinguishable. Show that distributions $\mathcal{X}_0$ and $\mathcal{X}_1$ remain computationally indistinguishable even if the adversary can gets n samples.

(a) First estimate computational distances between following games

```
$\mathcal{G}_{00}^{\mathcal{A}}$
    $x_0 \leftarrow \mathcal{X}_0$
    $x_1 \leftarrow \mathcal{X}_0$
    return $\mathcal{A}(x_0, x_1)$
```

```
$\mathcal{G}_{01}^{\mathcal{A}}$
    $x_0 \leftarrow \mathcal{X}_0$
    $x_1 \leftarrow \mathcal{X}_1$
    return $\mathcal{A}(x_0, x_1)$
```

```
$\mathcal{G}_{11}^{\mathcal{A}}$
    $x_0 \leftarrow \mathcal{X}_1$
    $x_1 \leftarrow \mathcal{X}_1$
    return $\mathcal{A}(x_0, x_1)$
```

Hint: What happens if you feed a sample $x_0 \leftarrow \mathcal{X}_0$ together an unknown sample $x_1 \leftarrow \mathcal{X}_i$ to $\mathcal{A}$ and use the reply to guess i.

(b) Generalise the argumentation to the case, where the adversary $\mathcal{A}$ gets n samples from a distribution $\mathcal{X}_i$. That is, define the corresponding sequence of games $\mathcal{G}_{00\ldots0}, \ldots, \mathcal{G}_{11\ldots1}$.

(c) Why do we need to assume that distributions $\mathcal{X}_0$ and $\mathcal{X}_1$ are efficiently samplable?

4. Consider the following game, where an adversary $\mathcal{A}$ gets three values x_1, x_2 and x_3. Two of them are sampled from the efficiently samplable distribution $\mathcal{X}_0$ and one of them is sampled from the efficiently samplable distribution $\mathcal{X}_1$. The adversary wins the game if it correctly determines which sample is taken from $\mathcal{X}_1$.

   (a) Find an upper bound to the success probability if distributions $\mathcal{X}_0$ and $\mathcal{X}_1$ are (t, ε)-indistinguishable.

   (b) How does the bound on the success change if we modify the game in the following manner. First, the adversary can first make its initial guess i_0. Then the challenger reveals $j \neq i_0$ such that x_j was sampled from $\mathcal{X}_0$ and then the adversary can output its final guess i_1.

   Hint: How well the adversary can perform if the challenger gives no samples to the adversary? How can you still simulate the game to the adversary who expects these samples?

5. Let $\mathcal{G}_s$ denote an interactive hypothesis testing game, where the challenger $\mathcal{G}$ acts according to a fixed program that depends on a private variable s. Prove that if $\mathsf{cd}_\star^t(\mathcal{G}_{s_i}, \mathcal{G}_{s_j}) \leq \varepsilon$ for all $s_i, s_j \in \mathrm{supp}(\mathcal{S}_0)$, then for any function $g : \mathcal{S}_0 \rightarrow \{0,1\}^*$ and for any $t - \mathrm{O}(t)$-time adversary $\mathcal{A}$:

$$\mathsf{Adv}^{\mathsf{sem}}_{\mathcal{G}_\star,g}(\mathcal{A}) = \Pr\left[s \leftarrow \mathcal{S}_0 : \mathcal{G}^{\mathcal{A}} = g(s)\right] - \max_s \Pr\left[g(s)\right] \leq \varepsilon \enspace .$$

   Hint: Look at the standard IND⇒SEM proof and repeat the same proof steps in the interactive setting.

   (a) Show that the guessing advantage $\mathsf{Adv}^{\mathsf{sem}}_{\mathcal{G}_\star,g}(\mathcal{A})$ is maximised by a deterministic function $g : \mathcal{S}_0 \rightarrow \mathbb{Z}$. Show also that a t-time adversary $\mathcal{A}$ can predict only values $0, \ldots, t$.

   (b) Use the classical sampling idiom to rewrite $\mathcal{G}_\star$ without changing its meaning. Prove formally that the sampling idiom does not change the distribution of s.

(c) Prove formally that indistinguishability of games $\mathcal{G}_{s_i}$ and $\mathcal{G}_{s_j}$ assures also indistinguishability of following games:

$$\begin{array}{lcl} \overline{\mathcal{G}}_i^{\mathcal{A}} & & \overline{\mathcal{G}}_j^{\mathcal{A}} \\ \left[\begin{array}{l} s \leftarrow \mathcal{S}_i \\ \mathsf{return}\ \mathcal{G}_s^{\mathcal{A}} \end{array}\right. & \quad\text{and}\quad & \left[\begin{array}{l} s \leftarrow \mathcal{S}_j \\ \mathsf{return}\ \mathcal{G}_s^{\mathcal{A}} \end{array}\right. \end{array}$$

where $\mathcal{S}_i$, $\mathcal{S}_j$ are the distributions introduced by the sampling idiom.

(d) Use the indistinguishability of games $\overline{\mathcal{G}}_i$ and $\overline{\mathcal{G}}_j$ to prove the final claim. Essentially, repeat the proof of guessing between many hypotheses.