Foundation of Cryptography, Fall 2011 — Iftach Haitner

Problem set 1. Exercises 1-3

November 9, 2011 — Due: Nov 20

- Send your solutions in a PDF format to foc.exc@gmail.com.
- Solution for each exercise should be emailed *separately*, title: Exe # (e.g., 3), Id (Israel id) (write the same details in the body of the attached file).
- Please *don't* write your name in the email/attached file.
- Write clearly and shortly using sub claims if needed. The emphasize in most questions is on the proofs (no much point is writing a "solution" w/o proving its correctness)
- For Latex users, a solution example can be found in the course web site.
- In case you work in (small) groups, please write the id list of your partners in the solution file. I stress that each student should write his solution by *himself* (join forces is only allowed in the "thinking phase")
- The notation we use appear in the first lecture (www.cs.tau.ac.il/∼iftachh/Courses/FOC/Fall11/Introduction_admin.pdf), section "Notation"

Exe 1 (basic probability): Let P and Q be distributions over a finite set $\mathcal{U}$.

a. (2 points) Prove that $\mathrm{SD}(P,Q) = \max_{\mathcal{S}\subseteq\mathcal{U}}(P(\mathcal{S}) - Q(\mathcal{S}))$ (recall that $\mathrm{SD}(P,Q) := \frac{1}{2}\sum_{u\in\mathcal{U}}|P(u) - Q(u)|$).

b. (3 points) Prove that $\mathrm{SD}(P^2,Q^2) \leq 2 \cdot \mathrm{SD}(P,Q)$ (see "Notation" in the first class slides for the definition of P^2, Q^2).

Let $\mathcal{Q} = \{Q_n\}_{n\in\mathbb{N}}$, $\mathcal{P} = \{P_n\}_{n\in\mathbb{N}}$ and $\mathcal{R} = \{R_n\}_{n\in\mathbb{N}}$ be distribution ensembles.

c. (2 points) Given that $\mathcal{Q} \overset{\mathrm{c}}{\approx} \mathcal{P}$ (i.e., $\mathcal{Q}$ is computationally indistinguishable from $\mathcal{P}$) and $\mathcal{P} \overset{\mathrm{c}}{\approx} \mathcal{R}$, prove that $\mathcal{Q} \overset{\mathrm{c}}{\approx} \mathcal{R}$.

d. (3 points) Give an example for ensemble $\mathcal{Q}$ and $\mathcal{P}$ such that: (1) $\mathrm{Supp}(Q_n) = \mathrm{Supp}(P_n)$ for every $n \in \mathbb{N}$, and (2) $\mathrm{SD}(Q_n, P_n) = 1 - \mathrm{neg}(n)$ (i.e., for every $p \in \mathrm{poly}$, exists $n' \in \mathbb{N}$ with $\mathrm{SD}(Q_n, P_n) > 1 - \frac{1}{p(n)}$ for every $n > n'$)

Exe 2 (one way functions):

a.(3 points) Give a simple example of a function f such that :

- Given y, it is easy to find $x \in f^{-1}(y)$ if such exists.
- Given $f(x)$, it is hard to find x: for any algorithm A, it holds that $\Pr[A(1^n, f(U_n)) = U_n] = \text{neg}(n)$.
  Is you function a one-way functions?

b.(4 points) Prove that OWFs cannot have polynomially-bounded cycles.

Namely, for every function $f\colon \{0,1\}^* \mapsto \{0,1\}^*$ and $x \in \{0,1\}^*$, define $\text{cyc}_f(x)$ as the smallest *positive* integer i with $f^i(x) = x$, where $f^i(x) = f(f^{i-1}(x))$ and $f^0(x) = x$. Prove that if f is (even weakly) one-way, then for every $p \in \text{poly}$ and large enough n (i.e., for every $n > n'$), it holds that $\text{E}[\text{cyc}_f(U_n)] > p(n)$ (i.e., the expected value of $\text{cyc}_f(x)$ over $x \leftarrow \{0,1\}^n$ is large).

Hint: Recall Markov inequality: for any non-negative random variable X and a real number $r > 0$,

$$\Pr[X \geq r \cdot \text{E}(X)] \leq \frac{1}{r}.$$

b.(3 points) Let f be an efficiently computable function such that $\text{E}[\text{cyc}_f(U_n)]$ is not polynomially-bounded. Is it necessary that f is (weakly) one-way? What if $\min_{x \in \{0,1\}^n}(\text{cyc}_f(x))$ is not polynomially-bounded?

Exe 3 (efficient length preserving one way functions):

a.(3 points)

Definition 1 (pairwise independent hash functions)**.** *A function family $\mathcal{H}$ from $\{0,1\}^n$ to $\{0,1\}^m$ (i.e., each $h \in \mathcal{H}$ is from $\{0,1\}^n$ to $\{0,1\}^m$) is* pairwise independent*, if for every $x \neq x' \in \{0,1\}^n$ and $y, y' \in \{0,1\}^m$, it holds that $\Pr_{h \leftarrow \mathcal{H}}[h(x) = y \wedge h(x') = y')] = 2^{-2m}$.*

Let $\mathcal{H}$ be the function family from $\{0,1\}^n$ to $\{0,1\}^m$, naturally defined by the set $\left\{(A, \bar{b}) \in \mathcal{F}_2^{m \times n} \times \mathcal{F}_2^m\right\}$. That is, the members of $\mathcal{H}$ are all pairs of Boolean matrices of dimension $m \times n$ and Boolean vectors of dimension m, and $(A, b)(x) := Ax + b \mod 2$. Prove that $\mathcal{H}$ is a pairwise independent family.

b.(7 points).

Definition 2 (efficient function family)**.** *An ensemble of function families $\mathcal{F} = \{\mathcal{F}_n\}_{n \in \mathbb{N}}$ is* efficient*, if the following hold:*

Samplable. *$\mathcal{F}$ is samplable in polynomial-time: there exists a* PPT *that given 1^n, outputs (the description of) a uniform element in $\mathcal{F}_n$.*

Efficient. *There exists a polynomial-time algorithm that given $x \in \{0,1\}^n$ and (a description of) $f \in \mathcal{F}_n$, outputs $f(x)$.*

Let $f \colon \{0,1\}^n \mapsto \{0,1\}^{\ell(n)}$ be a OWF, and let $\mathcal{H} = \{\mathcal{H}_n\}$ be an efficient ensemble of function families, where each $\mathcal{H}_n$ is a pairwise independent family form $\{0,1\}^{\ell(n)}$ to $\{0,1\}^{2n}$.

Prove that $g \colon \{0,1\}^{2n} \times \mathcal{H}_n \mapsto \{0,1\}^{2n} \times \mathcal{H}_n$,[1] defined as $g(x,h) = (h(f(x_{1,\ldots,n})), h)$, is a length-preserving OWF.

Hint: Given an "inverter" A for g, consider the following inverter for f:

Algorithm 3. *(M)*

Input: 1^n, $y \in \{0,1\}^{\ell(n)}$.

Operation:

(a) Choose $h \leftarrow \mathcal{H}_n$.

(b) Set $(x,h) = A(1^n, h(y), h)$.

(c) Output $x_{1,\ldots,n}$.

[1] I.e., here we view $\mathcal{H}_n$ as a set of strings – the function descriptions.