
Homework 2 – Due Tuesday, October 2 in Class

Reminders: Collaboration is permitted on homeworks, but you must (a) list your collaborators on the problem set and (b) **understand and write your solutions yourself**. If you worked entirely alone, write “Collaborators: none”.

Clarity is an important component of the grade on your problem sets. Poorly explained solutions will not receive full marks, even if they are correct.

1 Negligible Functions

Recall that $f : \mathbb{N} \rightarrow \mathbb{R}^+$ is negligible if for all constants $c > 0$, we have $f(n) \in O(\frac{1}{n^c})$. Equivalently, a function f is negligible if and only if for every constant $c > 0$, there exists an integer $n_0(c)$ such that for all $n > n_0(c)$, we have $f(n) < \frac{1}{n^c}$.

1. Write each of the following functions in the form $2^{g(n)}$, and indicate whether the original function is negligible in n :

(a) 0.99^n

(b) $\frac{1}{2^{100 \cdot n}}$

(c) $\frac{1}{n \log(n)}$

(d) $\frac{1}{n!}$

(e) $n^{-\frac{1}{\ln(n)}}$

(f) $n^{-\sqrt{n}}$

2. Pick one of the functions $f(n)$ above and *prove* that it is negligible by explaining how to compute $n_0(c)$ such that for all $n > n_0(c)$, $f(n) < \frac{1}{n^c}$.
3. Suppose that $\epsilon(n)$ and $\delta(n)$ are negligible functions. For each of the following statements, give either a proof or a counterexample.

(a) $\epsilon(n) + \delta(n)$ is negligible.

(b) $\frac{1}{\log(\frac{1}{\epsilon(n)})}$ is negligible

4. Suppose you are told that a particular one-time, fixed length encryption scheme is asymptotically secure (as defined in class). You decide that a distinguishing advantage of 2^{-80} is a sufficient level of security for a system you are currently building. What extra information do you need in order to decide what security parameter, n , to use?

2 Indistinguishability (KL Exercise 3.5 in disguise)

Recall that two ensembles of distributions $\{X_n\}$ and $\{Y_n\}$ are *indistinguishable in polynomial time* if for all PPT algorithms D , we have $|\Pr(D(1^n, X_n) = 1) - \Pr(D(1^n, Y_n) = 1)| \leq \text{negl}(n)$.

For this problem, consider an alternative version of the definition of indistinguishability, based on the following mental experiment, which takes as input a security parameter n (in unary) and an algorithm A :

- Guessing-Game($1^n, A$):
- Select $b \leftarrow \{0, 1\}$.
 - If $b = 0$, compute $b' \leftarrow A(1^n, X_n)$.
 - If $b = 1$, compute $b' \leftarrow A(1^n, Y_n)$.
 - Output “win” if $b' = b$, and “lose” otherwise.

Definition 1 *Two ensembles of probability distributions $\{X_n\}, \{Y_n\}$ are guessing-game indistinguishable if for every PPT algorithm A , the probability that Guessing-Game($1^n, A$) outputs “win” is bounded above by $\frac{1}{2} + \text{negl}(n)$.*

Prove that guessing game indistinguishability is equivalent to the version of polynomial-time indistinguishability defined in class.

3 Hybrid Arguments

You are given two pseudorandom generators G_0, G_1 with expansion functions $\ell_0(n), \ell_1(n)$, and you wish to create a new pseudorandom generator G' by concatenating the outputs of G_0 and G_1 . Thus G' takes a seed s of even length n and splits it into two pieces s_0 and s_1 of length $n/2$. The output is $G'(s_0, s_1) = G_0(s_0) \| G_1(s_1)$, where “ $\|$ ” denotes concatenation of binary strings.

In this problem, you will prove that G' is indeed a pseudorandom generator, using a hybrid argument.

1. What is the expansion function $\ell'(n)$ of G' as a function of ℓ_0 and ℓ_1 ?
2. The first step is to define three probability distributions $\text{Exp}_0, \text{Exp}_1, \text{Exp}_2$ where Exp_0 is a random string of length $\ell'(n)$ (denoted $U_{\ell'(n)}$), Exp_2 is the output of G' on a random seed $s = s_0, s_1$ of total length n , and Exp_1 is a distribution “between” the Exp_0 and Exp_2 .
Suggest an intermediate distribution Exp_1 (there is more than one natural choice).
3. Using your suggestion for Exp_1 , prove that any algorithm D' which can distinguish $G'(s)$ from $U_{\ell'(n)}$ with advantage $\epsilon(n)$ can be used either (i) to construct a PPT algorithm D_0 that distinguishes $G_0(s)$ from random with advantage at least $\epsilon(2n)/2$, or (ii) to construct a PPT algorithm D_1 that distinguishes $G_1(s)$ from random with probability advantage at least $\epsilon(2n)/2$. Conclude that G' is a secure p.r.g. as long as both G_0 and G_1 are secure.
4. Does the construction produce a p.r.g. if G_0 and G_1 are the same?
5. Does the construction always produce a p.r.g. if we use the same seed for both generators, that is, if we set $G'(s) = G_0(s) \| G_1(s)$?

4 Encryption

1. Given a probability distribution M on $\{0, 1\}^n$, define the *predictability* of M to be the probability mass of the most likely element in the distribution, that is:

$$\text{pred}_M = \max_{m \in \{0, 1\}^n} \Pr(M = m).$$

Suppose $(\text{Gen}, \text{Enc}, \text{Dec})$ is a one-time, fixed-length encryption scheme which has indistinguishable ciphertexts (KL Def. 3.9). Prove that for any ensemble of message distributions $\{M_n\}$, and any PPT adversary A , the probability that A can guess the message given a ciphertext is approximately bounded by the predictability of the message distribution, that is, for any ensemble of probability distributions $\{M_n\}$:

$$\Pr_{m \leftarrow M_n} (A(\text{Enc}_K(m)) = m) \leq \text{pred}_{M_n} + \text{negl}(n).$$

2. In order to provide authentication in addition to encryption, I.N. Security Services Inc. has taken a secure encryption scheme and modified it as follows. They start with a hash function $h : \{0, 1\}^* \rightarrow \{0, 1\}^{160}$ (for example, the popular hash function “SHA-1”). The new encryption scheme Enc' appends the hash of the message to the ciphertext:

$$\text{Enc}'_k(m) = \text{Enc}_k(m) \| h(m).$$

The key generation algorithm remains the same and the decryption function is basically unchanged (it ignores the hash value and decrypts the first part of the ciphertext).

- (a) Does the resulting scheme possess indistinguishable ciphertexts (Def. 3.9)? (Note: In this problem, we do not discuss what sort of authenticity this provides. That will come later in the course.)
- (b) Can you suggest a simple modification that retains the hash information but makes the new scheme secure?

5 Extra Credit

Consider the following two restrictions on the adversary \mathcal{A} in the “eavesdropping indistinguishability experiment” on page 63 of KL:

1. In stage 1, $\mathcal{A}(1^n)$ has to produce messages $m_0, m_1 \in \{0, 1, 2, \dots, 2^n - 1\}$ whose binary encodings differ in a single bit position, that is the Hamming distance between the encodings of m_0 and m_1 is at most 1.
2. In stage 1, $\mathcal{A}(1^n)$ has to produce messages $m_0, m_1 \in \{1, 2, 3, \dots, 2^n - 1\}$ which differ by at most 1 as integers, that is $|m_0 - m_1| = 1$.

For each of the restrictions above, decide whether or not the resulting definition of security is equivalent to KL Definition 3.9. If so, give a proof. If not, give a counter example (i.e. modify an encryption scheme so that it satisfies one definition but not the other).