TD 1 : Play with definitions

Exercise 1. [*Perfect security*]

Let (E, D) be a cipher over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$. Recall the definition of "perfect security" that was given in class. We are going to see that perfect security guarantees that the ciphertext reveals nothing about the message. Now consider a random experiment in which **k** and **m** are random variables, such that:

- **k** is uniformly distributed over $\mathcal{K}$
- **m** is distributed over $\mathcal{M}$, and
- **k** and **m** are independent

Define the random variable $c = E(k, m)$. Prove that:

- if (E, D) is perfectly secure, then **c** and **m** are independent;
- conversely, if **c** and **m** are independent, and each message in $\mathcal{M}$ occurs with nonzero probability, then (E, D) is perfectly secure.

Exercise 2. [*Variable length OTP is not secure*]

A *variable length one-time pad* is a cipher (E, D), where the keys are bit strings of some fixed length L, while messages and ciphertexts are variable length bit strings, of length at most L. Thus, (E, D) is defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$, where

$$\mathcal{K} := \{0,1\}^L \text{ and } \mathcal{M} := \mathcal{C} = \{0,1\}^{\leq L}$$

for some parameter L. Here, $\{0,1\}^{\leq L}$ denotes the set of all bit strings of length at most L (including the empty string). For a key $k \in \{0,1\}^L$ and a message $m \in \{0,1\}^{\leq L}$ of length l, the encryption function is defined as follows:

$$E(k, m) := k[0 \, .. \, l-1] \oplus m$$

Provide a counter-example showing that the variable length OTP is not secure.

Exercise 3. [*Distinguishability*]

We consider two distributions P_0 and P_1 over $\{0,1\}^L$.

1. Recall the definitions that were given in class for the notions of *distinguisher* and the advantage of a distinguisher. We say that P_0 and P_1 are ϵ-indistinguishable if for all distinguishers, the advantage is at most ϵ. Show that if P_0 and P_1 are 0-indistinguishable, then $P_0 = P_1$.

We are now going to give other slightly different definitions of ϵ-indistinguishability. The first one is based on the statistical distance.

$$\Delta(P_0, P_1) = \frac{1}{2} \sum_{a \in \{0,1\}^L} |P_0(a) - P_1(a)|.$$

2. Show that Δ satisfies the usual properties of a distance.

It will be useful in what follows to introduce random variables: let X have distribution P_0 and Y have distribution P_1. We will write $\Delta(X, Y)$ for $\Delta(P_0, P_1)$.

3. Show that for any function f we have, $\Delta(f(X), f(Y)) \leq \Delta(X, Y)$.

4. Show that $\Delta(X,Y) = \max_{T \subseteq \{0,1\}^L} |\Pr[X \in T] - \Pr[Y \in T]|$
5. Show that P_0 and P_1 are ϵ-indistinguishable if and only if $\Delta(X,Y) \leq \epsilon$.

Now, we consider a third definition of ϵ-indistinguishability. For this consider the following game.

$\mathcal{C}$	$\mathcal{A}$
sample $b \hookleftarrow U(0,1)$	
sample $x \hookleftarrow P_b$	
send x to $\mathcal{A}$	
	compute a bit b'
	send b' to $\mathcal{C}$
If $b = b'$, say "Win", else say "Lose".	

6. Show that there is a strategy for $\mathcal{A}$ such that the winning probability is $\frac{1}{2} + \frac{1}{2}\Delta(P_0, P_1)$. Moreover, show that for any strategy $\mathcal{A}$, the winning probability is at most $\frac{1}{2} + \frac{1}{2}\Delta(P_0, P_1)$. As such we could also define ϵ-indistinguishability of P_0 and P_1 by saying that the winning probability for this game is at most $\frac{1}{2} + \frac{1}{2}\epsilon$.

In cryptography, we will restrict the adversary $\mathcal{A}$ to be efficient. The distributions P_0 and P_1 are said to be ϵ-*computationally*-indistinguishable if all *efficient* distinguishers $\mathcal{A}$ have an advantage of at most ϵ. Note that we could equivalently define it by requiring that any adversary in the game defined above has a winning probability of at most $\frac{1}{2} + \frac{1}{2}\epsilon$.

7. Under reasonable assumptions, there exists functions $G : \{0,1\}^l \to \{0,1\}^{2l}$, such that $G(U(\{0,1\}^l))$ and $U(\{0,1\}^{2l})$ are ϵ-computationally indistinguishable for $\epsilon \leq \frac{1}{10}$ (in fact, we have ϵ that is smaller than any inverse polynomial in l). Show that there can be a large gap between computational indistinguishability and indistinguishability. More precisely, show that for large enough l, there is a distinguisher that has an advantage gets close to 1.

Exercise 4. [*More on encryption scheme*]

1. ***(Multiplicative OTP)*** We may also define a "multiplication mod p" variation of the one-time pad. This is a cipher (E, D), defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$, where $\mathcal{K} := \mathcal{M} := \mathcal{C} := \{1, \ldots, p-1\}$, where p is a prime. Encryption and decryption are defined as follows:

$$E(k,m) := k \times m \mod p \quad \text{and} \quad D(k,c) := k^{-1} \times c \mod p$$

   Here, k^{-1} denotes the multiplicative inverse of k modulo p. Verify the correctness property for this cipher and prove that it is perfectly secure.

2. ***(A good substitution cipher)*** Consider a variant of the substitution cipher (E, D) where every symbol of the message is encrypted using an independent permutation. That is, let $\mathcal{M} = \mathcal{C} = \Sigma^L$ for some a finite alphabet of symbols Σ and some L. Let the key space be $\mathcal{K} = S^L$ where S is the set of all permutations on Σ. The encryption algorithm $E(k,m)$ is defined as

$$E(k,m) := \big(k[0](m[0]), k[1](m[1]), \ldots, k[L-1](m[L-1])\big)$$

   Show that (E, D) is perfectly secure.

3. ***(Chain encryption)*** Let (E, D) be a perfectly secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ where $\mathcal{K} = \mathcal{M}$. Let (E', D') be a cipher where encryption is defined as

$$E'\big((k_1,k_2), m\big) := E\big((k_1,k_2), E(k_2,m)\big)$$

   Show that E' is perfectly secure.