1 Statistical distance of $\mathcal{D}'_0$ and $\mathcal{D}'_1$

Here are some facts about statistical distance:

Fact 1.1. For any three distributions $\mathcal{D}_0, \mathcal{D}_1, \mathcal{D}_2$,

$$\mathsf{SD}(\mathcal{D}_0, \mathcal{D}_2) \le \mathsf{SD}(\mathcal{D}_0, \mathcal{D}_1) + \mathsf{SD}(\mathcal{D}_1, \mathcal{D}_2).$$

Fact 1.2. For any two distributions $\mathcal{D}_0, \mathcal{D}_1$, $\mathsf{SD}(\mathcal{D}_0, \mathcal{D}_1) = \epsilon$ if and only if there exists an adversary $\mathcal{A}$ (not necessarily polynomial time) that can win the following game with probability $1/2 + \epsilon/2$:

- Challenger picks a bit $b \leftarrow \{0, 1\}$, samples $u \leftarrow \mathcal{D}_b$ and sends u to $\mathcal{A}$.
- $\mathcal{A}$ sends its guess b' and wins if $b = b'$.

We will use the above facts to prove the following theorem.

Theorem 1.3. Suppose $\mathsf{SD}(\mathcal{D}_0, \mathcal{D}_1) \le \epsilon$. Then $\mathsf{SD}(\mathcal{D}'_0, \mathcal{D}'_1) \le t \cdot \epsilon$.

Proof. The proof follows via a sequence of hybrid distributions $\mathcal{H}_0, \ldots, \mathcal{H}_t$, where $\mathcal{H}_0 \equiv \mathcal{D}'_0$, $\mathcal{H}_t \equiv \mathcal{D}'_1$, and $\mathcal{H}_i$ is defined as follows:

$\mathcal{H}_i$ = {output (t-i) samples chosen independently from $\mathcal{D}_0$ and output i samples chosen independently from $\mathcal{D}_1$ }

Claim 1.4. For any $i \le t$, $\mathsf{SD}(\mathcal{H}_{i-1}, \mathcal{H}_i) \le \epsilon$.

Proof. Let Ω be sample space for $\mathcal{D}_0$ and $\mathcal{D}_1$.
Let Ω ={ $a_1, a_2, ..., a_{n-1}, a_n$}.

Let define a new sample space Ω^t ={ output (t-i) samples chosen independently from $\mathcal{D}_0$ and output i samples chosen independently from $\mathcal{D}_1$ for i $\in$ [0,t] }

$\mathsf{SD}(\mathcal{H}_{i-1}, \mathcal{H}_i) = \frac{1}{2}(\sum_{z\in\Omega^t} |\Pr_{x\leftarrow\mathcal{H}_{i-1}}[x = z] - \Pr_{x\leftarrow\mathcal{H}_i}[x = z]|). --- (1)$

Let z = $(z_1, z_2, .., z_t) \in \Omega^t$ where each $z_i \in \Omega$.

$\mathsf{SD}(\mathcal{H}_{i-1}, \mathcal{H}_i) = \frac{1}{2}(\sum_{z\in\Omega^t} |\Pr_{x\leftarrow\mathcal{H}_{i-1}}[x = (z_1, z_2, .., z_t)] - \Pr_{x\leftarrow\mathcal{H}_i}[x = (z_1, z_2, .., z_t)] |). --- (2)$

As we know that samples is chosen independently .So we will use probability of independent event formula.

Pr[x=$(z_1, z_2, .., z_t)$]=Pr[x_1=z_1]Pr[x_2=z_2]...Pr[x_t=z_t].

$\mathsf{SD}(\mathcal{H}_{i-1}, \mathcal{H}_i)$

$=\frac{1}{2}(\sum_{z\in\Omega^t} |\Pr_{x_1\leftarrow\mathcal{D}_0}[x_1 = z_1]\Pr_{x_2\leftarrow\mathcal{D}_0}[x_2 = z_2]....\Pr_{x_{t-i+1}\leftarrow\mathcal{D}_0}[x_{t-i+1} = z_{t-i+1}]\Pr_{x_{t-i+2}\leftarrow\mathcal{D}_1}[x_{t-i+2} = z_{t-i+2}]...\Pr_{x_t\leftarrow\mathcal{D}_1}[x_t = z_t] - \Pr_{x_1\leftarrow\mathcal{D}_0}[x_1 = z_1]\Pr_{x_2\leftarrow\mathcal{D}_0}[x_2 = z_2]....\Pr_{x_{t-i}\leftarrow\mathcal{D}_0}[x_{t-i} = z_{t-i}]\Pr_{x_{t-i+1}\leftarrow\mathcal{D}_1}[x_{t-i+1} = z_{t-i+1}]...\Pr_{x_t\leftarrow\mathcal{D}_1}[x_t = z_t] |).$

$=1\frac{}{2}. \{\sum_{z\in\Omega^t} |\Pr_{x_1\leftarrow\mathcal{D}_0}[x_1 = z_1]\Pr_{x_2\leftarrow\mathcal{D}_0}[x_2 = z_2]....\Pr_{x_{t-i}\leftarrow\mathcal{D}_0}[x_{t-i} = z_{t-i}](\Pr_{x_{t-i+1}\leftarrow\mathcal{D}_0}[x_{t-i+1} = z_{t-i+1}] - \Pr_{x_{t-i+1}\leftarrow\mathcal{D}_1}[x_{t-i+1} = z_{t-i+1}])\Pr_{x_{t-i+2}\leftarrow\mathcal{D}_1}[x_{t-i+2} = z_{t-i+2}]...\Pr_{x_t\leftarrow\mathcal{D}_1}[x_t = z_t] |\}.$ —(3)

Let simplify above expression :

Let define $p_{0,i}$ be the probability that a_i is chosen from $\mathcal{D}_0$ during sampling. Similarly $p_{1,i}$ be the probability that a_i is chosen from $\mathcal{D}_1$ during sampling.

Then according to total probability theorem :

$$\sum_{i=1}^{n} p_{0,i} = 1 \qquad --(4)$$

$$\sum_{i=1}^{n} p_{1,i} = 1 \qquad --(5)$$

As we know that $z_i \in \Omega$, hence z_i can take "n" different values.

Let fixed z_{t-i-1} say to a_1 and let other z_j vary and take values from Ω for j $\in$ [1,n]- {t-i+1}.

Note that for each $z_k \in \Omega$ where k $\in$ [1,t-i] :

$\sum_{z_k \in \Omega, k \in [1,t-i]} | \Pr_{x_1 \leftarrow \mathcal{D}_0}[\mathrm{x}_1 = z_1] \Pr_{x_2 \leftarrow \mathcal{D}_0}[\mathrm{x}_2 = z_2] .... \Pr_{x_{t-i} \leftarrow \mathcal{D}_0}[\mathrm{x}_{t-i} = z_{t-i}] \, | = (\, p_{0,1} + p_{0,2} + .. + p_{0,n} \,)^{t-i} = 1.$
(from equation 4) $--$ (6)

Similarly for each $z_k \in \Omega$ where k $\in$ [t-i+2,t] :

$\sum_{z_k \in \Omega, k \in [t-i+2,t]} | \Pr_{x_{t-i+2} \leftarrow \mathcal{D}_1}[\mathrm{x}_{t-i+2} = z_{t-i+2}] ... \Pr_{x_t \leftarrow \mathcal{D}_1}[\mathrm{x}_t = z_t] \, | = (\, p_{1,1} + p_{1,2} + .. + p_{1,n} \,)^{i-1} = 1.$
(from equation 5) $--$ (7)

Now from Equation 6 and 7 we can re-write the Equation 3 where value of z_{t-i+1} is fixed to a_1.

$\mathsf{SD}(\mathcal{H}_{i-1}, \mathcal{H}_i)$

$= \frac{1}{2}. \, \{ \sum_{z \in \Omega^t} | \Pr_{x_1 \leftarrow \mathcal{D}_0}[\mathrm{x}_1 = z_1] \Pr_{x_2 \leftarrow \mathcal{D}_0}[\mathrm{x}_2 = z_2] .... \Pr_{x_{t-i} \leftarrow \mathcal{D}_0}[\mathrm{x}_{t-i} = z_{t-i}] (\Pr_{x_{t-i+1} \leftarrow \mathcal{D}_0}[\mathrm{x}_{t-i+1} = a_1]$ - $\Pr_{x_{t-i+1} \leftarrow \mathcal{D}_1}[\mathrm{x}_{t-i+1} = a_1]) \Pr_{x_{t-i+2} \leftarrow \mathcal{D}_1}[\mathrm{x}_{t-i+2} = z_{t-i+2}] ... \Pr_{x_t \leftarrow \mathcal{D}_1}[\mathrm{x}_t = z_t] \, | \} + \frac{1}{2}. \, \{ \sum_{\substack{z_{t-i+1} \neq a_1 \\ z \in \Omega^t}} | \Pr_{x_1 \leftarrow \mathcal{D}_0}[\mathrm{x}_1 = z_1] \Pr_{x_2 \leftarrow \mathcal{D}_0}[\mathrm{x}_2 = z_2] .... \Pr_{x_{t-i} \leftarrow \mathcal{D}_0}[\mathrm{x}_{t-i} = z_{t-i}] (\Pr_{x_{t-i+1} \leftarrow \mathcal{D}_0}[\mathrm{x}_{t-i+1} = a_1]$ - $\Pr_{x_{t-i+1} \leftarrow \mathcal{D}_1}[\mathrm{x}_{t-i+1} = a_1]) \Pr_{x_{t-i+2} \leftarrow \mathcal{D}_1}[\mathrm{x}_{t-i+2} = z_{t-i+2}] ... \Pr_{x_t \leftarrow \mathcal{D}_1}[\mathrm{x}_t = z_t] \, | \}$

$= \frac{1}{2}. \, | \, (\, p_{0,1} + p_{0,2} + .. + p_{0,n} \,)^{t-i} (\Pr_{x_{t-i+1} \leftarrow \mathcal{D}_0}[\mathrm{x}_{t-i+1} = a_1]$ - $\Pr_{x_{t-i+1} \leftarrow \mathcal{D}_1}[\mathrm{x}_{t-i+1} = a_1]) \, (\, p_{1,1} + p_{1,2} + .. + p_{1,n} \,)^{i-1} | \, + \, \frac{1}{2}. \, \{ \sum_{\substack{z_{t-i+1} \neq a_1 \\ z \in \Omega^t}} | \Pr_{x_1 \leftarrow \mathcal{D}_0}[\mathrm{x}_1 = z_1] \Pr_{x_2 \leftarrow \mathcal{D}_0}[\mathrm{x}_2 = z_2] .... \Pr_{x_{t-i} \leftarrow \mathcal{D}_0}[\mathrm{x}_{t-i} = z_{t-i}] (\Pr_{x_{t-i+1} \leftarrow \mathcal{D}_0}[\mathrm{x}_{t-i+1} = a_1]$ - $\Pr_{x_{t-i+1} \leftarrow \mathcal{D}_1}[\mathrm{x}_{t-i+1} = a_1]) \Pr_{x_{t-i+2} \leftarrow \mathcal{D}_1}[\mathrm{x}_{t-i+2} = z_{t-i+2}] ... \Pr_{x_t \leftarrow \mathcal{D}_1}[\mathrm{x}_t = z_t] \, |$

$= \frac{1}{2}. \, \{ \sum_{z \in \Omega^t} |1.(\Pr_{x_{t-i+1} \leftarrow \mathcal{D}_0}[\mathrm{x}_{t-i+1} = a_1]$ - $\Pr_{x_{t-i+1} \leftarrow \mathcal{D}_1}[\mathrm{x}_{t-i+1} = a_1]).1 \, | \} \, + \, \frac{1}{2}. \, \{ \sum_{\substack{z_{t-i+1} \neq a_1 \\ z \in \Omega^t}} | \Pr_{x_1 \leftarrow \mathcal{D}_0}[\mathrm{x}_1 = z_1] \Pr_{x_2 \leftarrow \mathcal{D}_0}[\mathrm{x}_2 = z_2] .... \Pr_{x_{t-i} \leftarrow \mathcal{D}_0}[\mathrm{x}_{t-i} = z_{t-i}] (\Pr_{x_{t-i+1} \leftarrow \mathcal{D}_0}[\mathrm{x}_{t-i+1} = a_1]$ - $\Pr_{x_{t-i+1} \leftarrow \mathcal{D}_1}[\mathrm{x}_{t-i+1} = a_1]) \Pr_{x_{t-i+2} \leftarrow \mathcal{D}_1}[\mathrm{x}_{t-i+2} = z_{t-i+2}] ... \Pr_{x_t \leftarrow \mathcal{D}_1}[\mathrm{x}_t = z_t] \, |$

Similarly we can fix remaining values of $z_{t-i+1} \in \{a_2, a_3, ..., a_n \}$ and we will ge following expression :

$\frac{1}{2}$. $\{\sum_{j\in\Omega}|(\underset{x_{t-i+1}\leftarrow\mathcal{D}_0}{Pr}[\mathrm{x}_{t-i+1}=\mathrm{j}]$ - $\underset{x_{t-i+1}\leftarrow\mathcal{D}_1}{Pr}[\mathrm{x}_{t-i+1}=\mathrm{j}]).\ |\} = \mathsf{SD}(\mathcal{D}_0,\mathcal{D}_1)$.

Hence we got $\mathsf{SD}(\mathcal{H}_{i-1},\mathcal{H}_i)$ =SD$(\mathrm{D}_0,\mathcal{D}_1)\leq\epsilon$ for i $\in$ [1,t]. –(8)

Let use Triangle Inequality given as Fact 1.1 to derive final result.

$\mathsf{SD}(\mathcal{D}_0',\mathcal{D}_1')\leq\mathsf{SD}(\mathcal{D}_0',\mathcal{H}_1)+,\mathsf{SD}(\mathcal{H}_1,\mathcal{D}_1',)$ –(9)

Again we can expand $\mathsf{SD}(\mathcal{H}_1,\mathcal{D}_1',)$ using Triangle Inequality as :

$\mathsf{SD}(\mathcal{H}_1,\mathcal{D}_1',)\leq\mathsf{SD}(\mathcal{H}_1,\mathcal{H}_2)+\mathsf{SD}(\mathcal{H}_2,\mathcal{D}_1',)$ –(10)

Similarly we will keep applying triangle equality and the final expression for Equation 8 would be :

$\mathsf{SD}(\mathcal{D}_0',\mathcal{D}_1')\leq\mathsf{SD}(\mathcal{D}_0',\mathcal{H}_1)+\sum_{i=1}^{i=(t-2)}\mathsf{SD}(\mathcal{H}_i,\mathcal{H}_{i+1})$ + SD$(\mathrm{H}_{t-1},\mathcal{D}_1')=\sum_{i=0}^{i=(t-1)}\mathsf{SD}(\mathcal{H}_i,\mathcal{H}_{i+1})--(11)$
(As $\mathcal{H}_0\equiv\mathcal{D}_0'$, $\mathcal{H}_t\equiv\mathcal{D}_1'$)

Now we will apply Equation (7) Result into Equation(10) and we will finally get :

$\mathsf{SD}(\mathcal{D}_0',\mathcal{D}_1')\leq\sum_{i=0}^{i=(t-1)}\mathsf{SD}(\mathcal{H}_i,\mathcal{H}_{i+1})\leq\mathrm{t}\epsilon\quad\Rightarrow\quad\mathsf{SD}(\mathcal{D}_0',\mathcal{D}_1')\leq\mathrm{t}\epsilon$. (Q.E.D)

□

□

2 Weak PRPs

Theorem 2.1. Assuming F is a secure PRF, the construction described in the assignment is a weak PRP.

Proof. We will prove that the construction is a weak PRP via a sequence of hybrid worlds. We first present the hybrid worlds below, then show that they are indistinguishable.

World 0: In this world, the challenger uses two PRF keys k_1, k_2. For every query, the challenger picks (x_i, y_i) uniformly at random, and sends the output of the PRP construction.

- The challenger chooses two uniformly random PRF keys k_1, k_2.
- For the i^{th} query, the challenger chooses uniformly randomly x_i, y_i and then computes $v_i = y_i \oplus F(x_i, k_1)$. It then sends (x_i, y_i) together with $(v_i, x_i \oplus F(v_i, k_2))$.
- Adversary sends b'.

Hybrid 1:

- The challenger chooses $\underline{\text{a uniformly random function } f \leftarrow \mathsf{Func}[\mathcal{X}, \mathcal{X}]}$ and PRF key k_2.
- For the i^{th} query, the challenger chooses uniformly random x_i, y_i and then computes $\underline{v_i = y_i \oplus f(x_i)}$. It then sends (x_i, y_i) together with $(v_i, x_i \oplus F(v_i, k_2))$.
- Adversary sends b'.

Hybrid 2:

- The challenger chooses $\underline{\text{two uniformly random functions } f_1, f_2.}$
- For the i^{th} query, the challenger chooses uniformly randomly x_i, y_i and then computes $v_i = y_i \oplus f_1(x_i)$. It then sends (x_i, y_i) together with $\underline{(v_i, x_i \oplus f_2(v_i))}$.
- Adversary sends b'.

World 1:

- The challenger chooses $\underline{\text{ uniformly random permutation } P \leftarrow \mathsf{Perm}[\mathcal{X}^2].}$
- For the i^{th} query, the challenger chooses uniformly randomly x_i, y_i and then sends (x_i, y_i) together with $\underline{P(x_i, y_i)}$.
- Adversary sends b'.

We will now prove that the hybrids are indistinguishable.

Claim 2.2. Suppose there exists a p.p.t adversary $\mathcal{A}$ such that $p_0 - p_{\text{hyb},1} = \epsilon$. Then there exists a **p.p.t** reduction algorithm $\mathcal{B}$ that breaks the PRF security of F with probability $1/2 + \epsilon/2$.

Proof. We construct the reduction algorithm $\mathcal{B}$ as follows:

- $\mathcal{B}$ sends x_i to the PRF challenger. The PRF challenger sends $ct = F(x_i, k_1)$ if $b = 0$ and $ct = f(x_i)$ if $b = 1$.
- $\mathcal{B}$ chooses y_i, k_2 u.a.r from $\{0,1\}^n$. Sets $v_i = y_i \oplus ct$ and sends $(v_i, x_i \oplus F(v_i, k_2))$ to $\mathcal{A}$.

- After polynomially many queries, the adversary $\mathcal{A}$ outputs b' sends to $\mathcal{B}$ and $\mathcal{B}$ forwards it to the PRG challenger.

  If $b = 0$, then the adversary $\mathcal{A}$ is in the **World 0** and if $b = 1$, then the adversary $\mathcal{A}$ is in **Hybrid 1**. Clearly, $Pr(b' = 0|b = 0)$ is same as p_0 and $Pr(b' = 0|b = 1)$ is same as $p_{hyb,1}$.

□

Claim 2.3. Suppose there exists a p.p.t adversary $\mathcal{A}$ such that $p_{\text{hyb},1} - p_{\text{hyb},2} = \epsilon$. Then there exists a **p.p.t** reduction algorithm $\mathcal{B}$ that breaks the PRF security of F with probability $1/2 + \epsilon/2$.

Proof. This proof is very similar to the proof of previous claim. □

Claim 2.4. Hybrid 2 is indistinguishable from world 1.

Proof. To prove indistinguishability between Hybrid 2 and World 1, we go through the following sequence of hybrids :

Hybrid 2.1 ≡ Hybrid 2

- The challenger chooses $\underline{\text{two uniformly random functions } f_1, f_2.}$
- For the i^{th} query, the challenger chooses uniformly randomly x_i, y_i and then computes $v_i = y_i \oplus f_1(x_i)$. It then sends (x_i, y_i) together with $\underline{(v_i, x_i \oplus f_2(v_i))}$.
- Adversary sends b'.

Hybrid 2.2:

- The challenger chooses a uniformly random function f and a random number r.
- For the i^{th} query, the challenger chooses uniformly randomly x_i, y_i and then computes $v_i = y_i \oplus r$. It then sends (x_i, y_i) together with $\underline{(v_i, x_i \oplus f(v_i))}$.
- Adversary sends b'.

Hybrid 2.3

- The challenger chooses a uniformly random number r_1 and a random number r_2.
- For the i^{th} query, the challenger chooses uniformly randomly x_i, y_i and then computes $v_i = y_i \oplus r_1$. It then sends (x_i, y_i) together with $\underline{(v_i = y_i \oplus r_1, x_i \oplus r_2)}$.
- Adversary sends b'.

Hybrid 2.4

- The challenger chooses a uniformly random function $f : \{0,1\}^{2n} \to \{0,1\}^{2n}$.
- For the i^{th} query, the challenger chooses uniformly randomly x_i, y_i and then sends $f(x_i||y_i)$.
- Adversary sends b'.

Hybrid 2.5 ≡ World 2:

- The challenger chooses a uniformly random permutation $P : \{0,1\}^{2n} \to \{0,1\}^{2n}$.
- For the i^{th} query, the challenger chooses uniformly randomly x_i, y_i and then sends $P(x, y)$.
- Adversary sends b'.

Analysis

- Indistinguishability between Hybrid 2.1 and 2.2 follows from the property of random function.
- Similiarly, inditinguishability between Hybrid 2.2 and 2.3 follows from the property of random function.
- Indistinguishability between Hybrid 2.3 and 2.4 follows from the fact that XOR of a string with some random string outputs a random string, now concatenating two random strings of size n gives another random string of size 2n and then using the property of random function, the two are equivalent.
- Indistinguishability between Hybrid 2.4 and 2.5 follows from birthday bound.

This proves that Hybrid 2 and World 1 are indistinguishable.

$\square$

$\square$

3 Composing PRGs and PRFs

3.1

Theorem 3.1. Assuming F is a secure PRF and G is a secure PRG, F' is a secure PRF.

Proof. We will prove this theorem via a sequence of hybrid experiments, where world-0 (= hybrid-0) corresponds to the challenger choosing a PRF key, and world-1 (= final hybrid) corresponds to the challenger choosing a uniformly random function.

Description of hybrids: We describe the hybrids in the following way:

World 0: In this world, the challenger uses PRG G and PRF F.

- In i^{th} query, the adversary $\mathcal{A}$ queries for x_i.
- Challenger outputs $G(F(x_i, k))$.
- Adversary $\mathcal{A}$ outputs b'.

Hybrid 1: In this world, the challenger uses PRG G and random function f.

- In i^{th} query, the adversary $\mathcal{A}$ queries for x_i.
- Challenger outputs $G(f(x_i))$.
- Adversary $\mathcal{A}$ outputs b'.

Hybrid 2: In this world, the challenger uses PRG G and samples a random number r for each query. For repeated queries, it uses the same previously sampled random number.

- In i^{th} query, the adversary $\mathcal{A}$ queries for x_i.
- Challenger samples a random number r outputs $G(r)$.
- Adversary $\mathcal{A}$ outputs b'.

Hybrid 3: In this world, the challenger samples a random number r for each query. For repeated queries, it uses the same previously sampled random number.

- The adversary $\mathcal{A}$ queries for x.
- Challenger samples a random number r and outputs r.
- Adversary $\mathcal{A}$ outputs b'.

World 1: In this world, the challenger uses a random function f for each query x and outputs $f(x)$.

- The adversary $\mathcal{A}$ queries for x.
- Challenger outputs $f(x)$.
- Adversary $\mathcal{A}$ outputs b'.

Next, we show that the consecutive hybrids are computationally indistinguishable.

Analysis: We show the indistinguishability between hybrids in the following way:

- Indistinguishability of World 0 and Hybrid 1 follows from the fact that F is a secure PRF.
- Indistinguishability of Hybrid 1 and Hybrid 2 follows from the property of the random function. (Note that Hybrid 1 and Hybrid 2 are equivalent in the way that despite using a random function f in Hybrid 2, we sample the value $f(x)$ on fly for query x).
- Indistinguishability of Hybrid 2 and Hybrid 3 follows from the fact that G is a secure PRG and is secure even for polynomial number of queries (done in class).
- Indistinguishability of Hybrid 3 and World 1 follows from the property of the random function. (Note that Hybrid 3 and World 1 are equivalent in the way that despite using a random function f in Hybrid 3, we sample the value $f(x)$ for query x on fly).

This proves that World 0 and World 1 are indistinguishable. □

3.2

3.2.1 Construction of $\mathcal{G}'$

We define $\mathcal{G}'(x||y) = \mathcal{G}(x)$.

Claim 3.2. Suppose there exists a p.p.t adversary $\mathcal{A}$ that breaks the security of the PRG $\mathcal{G}'$ then there exists a p.p.t adversary $\mathcal{B}$ that breaks the PRG security of $\mathcal{G}$.

Proof. We construct the adversary $\mathcal{B}$ as follows:

- Challenger of PRG G outputs u_b to adversary $\mathcal{B}$.
- $\mathcal{B}$ sends u_b to $\mathcal{A}$.
- $\mathcal{A}$ outputs b' to $\mathcal{B}$ and $\mathcal{B}$ forwards the same to challenger of PRG G.

Let adversary $\mathcal{A}$ have the probability of outputting 0 in World 0 and World 1 wrt PRG $\mathcal{G}'$ as p_0 and p_1 respectively. Clearly, in the World 0 wrt PRG $\mathcal{G}$, $\Pr(b' = 0) = p_0$ (adversary $\mathcal{A}$ receives $\mathcal{G}(x)$ for some x) and similiarly in the World 1, $\Pr(b' = 0) = p_1$ (adversary $\mathcal{A}$ receives a random string of length n as expected). Clearly, if $p_0 - p_1$ is non-negligible, so is the difference $\Pr(b' = 0|b = 0)$ - $\Pr(b' = 0|b = 1)$ in case of PRG $\mathcal{G}$. Thus, $\mathcal{B}$ breaks the PRG security of $\mathcal{G}$. □

Claim 3.3. F' is not a secure PRF.

Proof. We construct an adversary $\mathcal{A}$ that breaks the PRF security of F' as follows:

- The adversary $\mathcal{A}$ sends polynomially many queries of the form $0^n||x$ where x is any random string of size n.
- Clearly, the adversary $\mathcal{A}$ receives $F(G(0^n), k)$ as a result of each query.
- Clearly, adversary $\mathcal{A}$ can distinguish between a random function and F'.

This proves that F' is not secure. □

4 CBC mode

Theorem 4.1. Assuming F is a secure PRP, and $|\mathcal{X}|$ is super-polynomial in the security parameter, the CBC mode of encryption satisfies $\mathsf{No\text{-}Query\text{-}Semantic\text{-}Security}$.

Proof. As discussed in class (Lecture 11, Section 2), this proof goes through a sequence of hybrids.

World 0:

- Adversary $\mathcal{A}$ sends two messages m_0,m_1 s.t $|m_0| = |m_1| = n \cdot \ell$. Let $m_b = (m_{b,1} \mid\mid \ \ldots \ \mid\mid m_{b,\ell})$.
- Challenger chooses PRP key $k \leftarrow \mathcal{K}$. It computes $\mathsf{ct}_1 = F(m_{0,1}, k)$. For all $i > 1$, it computes $\mathsf{ct}_i = F(m_{0,i} \oplus \mathsf{ct}_{i-1}, k)$.

  Finally, it sends $(\mathsf{ct}_1, \ldots, \mathsf{ct}_\ell)$ to $\mathcal{A}$.
- Adversary sends b'

Hybrid 1:

- Adversary $\mathcal{A}$ sends two messages m_0,m_1 s.t $|m_0| = |m_1| = n \cdot \ell$. Let $m_b = (m_{b,1} \mid\mid \ \ldots \ \mid\mid m_{b,\ell})$.
- Challenger chooses $\underline{f \leftarrow \mathsf{Perm}[\mathcal{X}]}$. It computes $\underline{\mathsf{ct}_1 = f(m_{0,1})}$. For all $i > 1$, it computes $\underline{\mathsf{ct}_i = f(m_{0,i} \oplus \mathsf{ct}_{i-1})}$.

  Finally, it sends $(\mathsf{ct}_1, \ldots, \mathsf{ct}_\ell)$ to $\mathcal{A}$.
- Adversary sends b'

Hybrid 2:

- Adversary $\mathcal{A}$ sends two messages m_0,m_1 s.t $|m_0| = |m_1| = n \cdot \ell$. Let $m_b = (m_{b,1} \mid\mid \ \ldots \ \mid\mid m_{b,\ell})$.
- Challenger chooses $f \leftarrow \mathsf{Perm}[\mathcal{X}]$. It computes $\underline{\mathsf{ct}_1 = f(m_{1,1})}$. For all $i > 1$, it computes $\underline{\mathsf{ct}_i = f(m_{1,i} \oplus \mathsf{ct}_{i-1})}$.

  Finally, it sends $(\mathsf{ct}_1, \ldots, \mathsf{ct}_\ell)$ to $\mathcal{A}$.
- Adversary sends b'

World 1:

1. A sends two messages m_0,m_1 s.t $|m_0| = |m_1| = n \cdot \ell$.
2. Challenger chooses $\underline{\text{PRP key } k \leftarrow \mathcal{K}}$ and computes $\underline{\mathsf{ct}_1 = F(m_{1,1}, k)}$. For all $i > 1$, it computes $\underline{\mathsf{ct}_i = F(m_{1,i} \oplus \mathsf{ct}_{i-1}, k)}$.

   Finally, it sends $(\mathsf{ct}_1, \ldots, \mathsf{ct}_\ell)$ to $\mathcal{A}$.
3. Adversary sends b'.

Let $p_0, p_1, p_{\text{hyb},1}$ and $p_{\text{hyb},2}$ denote the probability of adversary $\mathcal{A}$ outputting 0 in world-0, world-1, hybrid-1 and hybrid-2 respectively.

Claim 4.2. Assuming F is a secure PRP, $p_0 \approx p_{\text{hyb},1}$.

Proof. This follows from the PRP security — for a uniformly random PRP key, $F(\cdot, k)$ is indistinguishable from a uniformly random permutation. $\square$

Claim 4.3. For any adversary $\mathcal{A}$, $p_{\text{hyb},1} - p_{\text{hyb},2} \leq \mu(n)$.

Proof. Let construct following Hybrids to proof the claim.

Hybrid 1.1:

- Adversary $\mathcal{A}$ sends two messages m_0,m_1 s.t $|m_0| = |m_1| = n \cdot \ell$. Let $m_0 = (m_{0,1} \mid\mid \ \dots \ \mid\mid m_{0,\ell})$.
- Challenger chooses a random function $F_0 \leftarrow \mathsf{Func}[\mathcal{X}, \mathcal{X}]$. It computes $\underline{\mathsf{ct}_1 = F_0(m_{0,1})}$. For all $i > 1$, it computes $\underline{\mathsf{ct}_i = F_0(m_{0,i} \oplus \mathsf{ct}_{i-1})}$.

  Finally, it sends $(\mathsf{ct}_1, \dots, \mathsf{ct}_\ell)$ to $\mathcal{A}$.
- Adversary sends b'

Hybrid 1.2:

- Adversary $\mathcal{A}$ sends two messages m_0,m_1 s.t $|m_0| = |m_1| = n \cdot \ell$. Let $m_0 = (m_{0,1} \mid\mid \ \dots \ \mid\mid m_{0,\ell})$.
- Challenger maintains a table T in which it stores the random strings chosen by the challenger corresponding to the input provided by the adversary.
- Challenger first check the table T for the entry m_0, if present it uses the random string $y = (y_1 || y_2 ... || y_l)$ used previously. If m_0 is not present in the table, then the challenger chooses uniformly random strings $y_i \leftarrow \mathcal{X}$ for each i. For all i, it computes $\underline{\mathsf{ct}_i = y_i \oplus m_{0,i}}$. Then, it stores $y = (y_1, y_2, .., y_l)$ corresponding to m_0.

  Finally, it sends $(\mathsf{ct}_1, \dots, \mathsf{ct}_\ell)$ to $\mathcal{A}$.
- Adversary sends b'

Hybrid 1.3:

- Adversary $\mathcal{A}$ sends two messages m_0,m_1 s.t $|m_0| = |m_1| = n \cdot \ell$. Let $m_0 = (m_{0,1} \mid\mid \ \dots \ \mid\mid m_{0,\ell})$.
- Challenger chooses uniformly random strings $y_i \leftarrow \mathcal{X}$ for each i. For all i, it computes $\underline{\mathsf{ct}_i = y_i \oplus m_{0,i}}$.

  Finally, it sends $(\mathsf{ct}_1, \dots, \mathsf{ct}_\ell)$ to $\mathcal{A}$.
- Adversary sends b'

Hybrid 1.4:

- Adversary $\mathcal{A}$ sends two messages m_0,m_1 s.t $|m_0| = |m_1| = n \cdot \ell$. Let $m_1 = (m_{1,1} \mid\mid \ \dots \ \mid\mid m_{1,\ell})$.
- Challenger chooses uniformly random strings $y_i \leftarrow \mathcal{X}$ for each i. For all i, it computes $\underline{\mathsf{ct}_i = y_i \oplus m_{1,i}}$.

  Finally, it sends $(\mathsf{ct}_1, \dots, \mathsf{ct}_\ell)$ to $\mathcal{A}$.
- Adversary sends b'

Hybrid 1.5:

- Adversary $\mathcal{A}$ sends two messages m_0,m_1 s.t $|m_0| = |m_1| = n \cdot \ell$. Let $m_1 = (m_{1,1} \mid\mid \ \dots \ \mid\mid m_{1,\ell})$.
- Challenger maintains a table T in which it stores the random strings chosen by the challenger corresponding to the input provided by the adversary.
- Challenger first check the table T for the entry m_1, if present it uses the random string $y = (y_1 || y_2 ... || y_l)$ used previously. If m_1 is not present in the table, then the challenger chooses uniformly random strings $y_i \leftarrow \mathcal{X}$ for each i. For all i, it computes $\underline{\mathsf{ct}_i = y_i \oplus m_{1,i}}$. Then, it stores $y = (y_1, y_2, .., y_l)$ corresponding to m_1.

  Finally, it sends $(\mathsf{ct}_1, \dots, \mathsf{ct}_\ell)$ to $\mathcal{A}$.

- Adversary sends b'

Hybrid 1.6:

- Adversary $\mathcal{A}$ sends two messages m_0,m_1 s.t $|m_0| = |m_1| = n \cdot \ell$. Let $m_b = (m_{b,1} \mid\mid \ \ldots \ \mid\mid m_{b,\ell})$.
- Challenger chooses a random function $F_1 \leftarrow \mathsf{Func}[\mathcal{X}, \mathcal{X}]$. It computes $\underline{\mathsf{ct}_1 = F_1(m_{1,1})}$. For all $i > 1$, it computes $\underline{\mathsf{ct}_i = F_1(m_{1,i} \oplus \mathsf{ct}_{i-1})}$.

  Finally, it sends $(\mathsf{ct}_1, \ldots, \mathsf{ct}_\ell)$ to $\mathcal{A}$.
- Adversary sends b'

Probability Analysis

- $p_{\text{hyb},1} \approx p_{\text{hyb},1.1}$ due to Birthday Bound.
- $p_{\text{hyb},1.1} = p_{\text{hyb},1.2}$, The two experiments are identical in the adversary's view. In one case, it receives the outputs of a random function on distinct inputs, while in the other case, it creates a random function on fly by mapping m to some y.
- $p_{\text{hyb},1.2} = p_{\text{hyb},1.3} \leq q^2/|\mathcal{X}|$, The only difference in the two hybrids happens if some m is sampled twice. Using the birthday bound, we know that this happens with probability at most $q^2/|\mathcal{X}|$.
- $p_{\text{hyb},1.3} = p_{\text{hyb},1.4}$, using security of Shannon's OTP.
- $p_{\text{hyb},1.4} = p_{\text{hyb},1.5} \leq q^2/|\mathcal{X}|$, The only difference in the two hybrids happens if some m is sampled twice. Using the birthday bound, we know that this happens with probability at most $q^2/|\mathcal{X}|$.
- $p_{\text{hyb},1.5} = p_{\text{hyb},1.6}$, The two experiments are identical in the adversary's view. In one case, it receives the outputs of a random function on distinct inputs, while in the other case, it creates a random function on fly by mapping m to some y.
- $p_{\text{hyb},1.6} \approx p_{\text{hyb},2}$ due to Birthday Bound.

From above Probability Analysis , we can conclude that $p_{\text{hyb},1} \approx p_{\text{hyb},2}$.

Hence, For any adversary $\mathcal{A}$, $p_{\text{hyb},1} - p_{\text{hyb},2} \leq \mu(n)$.

$\square$

Claim 4.4. Assuming F is a secure PRP, $p_{\text{hyb},2} \approx p_1$.

Proof. This proof is similar to the proof of Claim 4.2. $\square$

Putting together the above claims, it follows that the CBC mode of encryption satisfies $\mathsf{No\text{-}Query\text{-}Semantic\text{-}Security}$.

$\square$

5 OWFs

5.1

Let $f : \{0,1\}^n \to \{0,1\}^m$ be a OWF. Let $g : \{0,1\}^n \to \{0,1\}^{m-1}$, where $g(x)$ is computed by evaluating $f(x)$ and then removing the first bit from the output (i.e. if $f(x) = y_1, y_2, ..., y_m$ then $g(x) = y_2, ..., y_m$).

We begin by assuming that g is not an OWF. i.e. given $g(x)$ for some $x \in \{0,1\}^n$, there exists an adversary $\mathcal{A}$ that can output x' in polynomial time such that $g(x') = g(x)$.

Theorem 5.1. We show that if g is not an OWF then, f also can't be an OWF. We will be using the adversary $\mathcal{A}$ to show that f can't be an OWF.

Proof. Given $f(x) = f(x)_1 f(x)_2 .. f(x)_m$ ($f(x)_i$ is the i^{th} bit of $f(x)$), if the adversary $\mathcal{A}$ is given input $f(x)_2 .. f(x)_m$, then $\mathcal{A}$ outputs x' such that $g(x') = f(x)_2 .. f(x)_m$ for which either $f(x') = 0 || f(x)_2 .. f(x)_m$ or $f(x') = 1 || f(x)_2 .. f(x)_m$.

This means that the adversary $\mathcal{A}$ is able to output a valid value x' ($f(x') = f(x)$) in one of the above two cases depending whether $f(x)_1$ is 0 or 1. $--(1)$

Thus, we construct our adversary $\mathcal{B}$ for the OWF f as follows:

- The adversary $\mathcal{B}$ receives $f(x) = f(x)_1 f(x)_2 .. f(x)_m$. The adversary $\mathcal{B}$ forwards $f(x)_2 f(x)_3 .. f(x)_m$ to adversary $\mathcal{A}$.
- $\mathcal{A}$ sends x' to adversary $\mathcal{B}$ and $\mathcal{B}$ checks whether $f(x') = f(x)$.

From (1), we get that the adversary $\mathcal{B}$ will be able to output x' such that $f(x') = f(x)$ for 50% of the cases. This shows that $f(x)$ is not an OWF which is a contradiction. Thus our assumption was wrong that g is not an OWF.

This proves that g is also an OWF.

$\square$