TD 2: Pseudo Random Generators (corrected version)

Exercise 1. *PRG implies smCPA with hybrid argument*

Let $\mathsf{G} : \{0,1\}^n \to \{0,1\}^m$ be a function, with $m > n$.

1. Recall the definition of a PRG from the lecture.

☞ $\mathsf{G} : \{0,1\}^n \to \{0,1\}^m$ is a PRG if there exists no ppt $\mathcal{A} : \{0,1\}^m \to \{0,1\}$ that distinguish with non-negligible probability between $\mathcal{U}(\{0,1\}^m)$ and $G(\mathcal{U}(\{0,1\}^n))$.

Let $\mathsf{Enc} : \{0,1\}^n \times \{0,1\}^m \to \{0,1\}^m$ defined by $\mathsf{Enc}(k, m) = \mathsf{G}(k) \oplus m$.

2. Give the associated decryption algorithm.

☞ Enc = Dec

3. Recall the smCPA security notion from the lecture.

☞ Two experiments, Exp_b for $b \in \{0,1\}$ are defined as follows:

1. The challenger $\mathcal{C}$ chooses k uniformly.
2. The adversary $\mathcal{A}$ chooses m_0, m_1 distinct of identical bitlength.
3. The challenger $\mathcal{C}$ returns $\mathsf{Enc}(k, m_b)$.
4. The adversary $\mathcal{A}$ outputs a guess b'.

This is summed up in the following sketch:

$\mathcal{C}$	$\mathcal{A}$
$k \hookleftarrow U(K)$	Choose and send $(m_0, m_1) \in P' := \{(m, m') \in P^2, m \neq m' \wedge \lvert m\rvert = \lvert m'\rvert\}$
Send $\mathsf{Enc}(k, m_b)$	
	Output $b' \in \{0,1\}$.

The advantage of $\mathcal{A}$ is defined as $\mathsf{Adv}(\mathcal{A}) := \left|\Pr\left(\mathcal{A} \xrightarrow{\mathsf{Exp}_0} 1\right) - \Pr\left(\mathcal{A} \xrightarrow{\mathsf{Exp}_1} 1\right)\right|$. Then $(\mathsf{Enc}, \mathsf{Dec})$ is said smCPA-secure if no efficient adversary has non-negligible advantage.

Let $m_1, m_2 \in \{0,1\}^m$ be arbitrary messages.

4. What is the statistical distance between the distributions $\mathcal{U}_1 = m_1 \oplus \mathcal{U}(\{0,1\}^m)$ and $\mathcal{U}_2 = m_2 \oplus \mathcal{U}(\{0,1\}^m)$?

☞ They are the same distributions, so 0.

5. Prove that if G is a PRG, then $(\mathsf{Enc}, \mathsf{Dec})$ is smCPA-secure using a hybrid argument.

☞ TODO

(Bonus) We just proved that G PRG $\Rightarrow$ $(\mathsf{Enc}, \mathsf{Dec})$ smCPA-secure. We are going to prove $(\mathsf{Enc}, \mathsf{Dec})$ not smCPA-secure $\Rightarrow$ G not PRG.

6. Let $\mathcal{A}$ be an distinguisher between two games G_0 and G_1. We say that $\mathcal{A}$ wins if it output 0 (resp 1) during the game G_0 (resp G_1). Show that

$$\mathsf{Adv}_{\mathcal{A}}(G_0, G_1) = 2 \cdot \left| \Pr_{b \sim \mathcal{U}(\{0,1\})}(\mathcal{A} \text{ wins in } G_b) - \frac{1}{2} \right|$$

☞

$$\Pr_{b \sim \mathcal{U}(\{0,1\})}(\mathcal{A} \text{ wins in } G_b) = \frac{1}{2} \cdot \Pr_{G_0}(\mathcal{A} \to 0) + \frac{1}{2} \cdot \Pr_{G_1}(\mathcal{A} \to 1) = \frac{1}{2}\left(\Pr_{G_0}(\mathcal{A} \to 0) + 1 - \Pr_{G_1}(\mathcal{A} \to 0)\right)$$

Hence the result.

7. Assume that $\mathcal{A}$ is an adversary with non-negligible advantage ε against the smCPA-security of $(\mathsf{Enc}, \mathsf{Dec})$. Construct an explicit distinguisher between $\mathcal{U}(\{0,1\}^m)$ and $G(\mathcal{U}(\{0,1\}^n))$ and compute its advantage.

☞ We define the $\mathcal{A}'$ to be the following:

1. Get k from the distribution $G = G(\mathcal{U}(\{0,1\}^n))$ or $\mathcal{U}(\{0,1\}^m)$.
2. Get m_1, m_2 from $\mathcal{A}$.
3. Sample b from $\mathcal{U}(\{0,1\})$.
4. Send $k \oplus m_b$ to $\mathcal{A}$ and get the output b'.
5. If $b = b'$, output "G" else output "$\mathcal{U}$".

The advantage of $\mathcal{A}'$ is $|\Pr_{k \sim G}(\mathcal{A}' \to G) - \Pr_{k \sim \mathcal{U}}(\mathcal{A}' \to G)|$.

Assume $k \sim \mathcal{U}$ and define Y_0 the game played when $b = 0$ and Y_1 the game played when $b = 1$. Since $k \sim \mathcal{U}$, we have that $m_b \oplus k$ is independent from m_b, hence $\Pr_{m_0,k}(\mathcal{A}(m_0 \oplus k) \to 1) = \Pr_{m_1,k}(\mathcal{A}(m_1 \oplus k) \to 1)$ and hence the advantage of $\mathcal{A}$ between Y_0 and Y_1 is 0. By the previous question we have $\Pr_{b \sim \mathcal{U}(\{0,1\}),k}(\mathcal{A} \text{ wins when given } m_b \oplus k) = 1/2$.

Assume $k \sim G$ and define Y_0' the game played when $b = 0$ and Y_1' the game played when $b = 1$. We have $\Pr_{k \sim G}(\mathcal{A}' \to G) = \Pr_b(\mathcal{A} \text{ wins } Y_b')$.

Finaly, $\mathsf{Adv}_{\mathcal{A}'} = |\Pr_{k \sim G}(\mathcal{A}' \to G) - \Pr_{k \sim \mathcal{U}}(\mathcal{A}' \to G)| = |\Pr_b(\mathcal{A} \text{ wins } Y_b') - 1/2| = \varepsilon/2$.

Exercise 2. *smCPA does not imply PRG*

Let $(\mathsf{Enc}, \mathsf{Dec})$ be an encryption scheme over $K \times P \times \{0,1\}^n$.

1. In this question, we assume that $(\mathsf{Enc}, \mathsf{Dec})$ is smCPA-secure. Prove that there exists a smCPA-secure encryption scheme $(\mathsf{Enc}', \mathsf{Dec}')$ such that $G : k \mapsto \mathsf{Enc}'(k, 0)$ is not a secure PRG. *Hint: try to concatenate constant bits to every ciphertext.*

☞ Define $\mathsf{Enc}' : (k, m) \mapsto 1^\ell || \mathsf{Enc}(k, m)$. The decryption algorithm Dec' ignores the first ℓ bits and calls Dec on the remaining ones. We have two things to prove:

- The pair $(\mathsf{Enc}', \mathsf{Dec}')$ is a smCPA-secure encryption scheme.
- $G : k \mapsto 1^\ell || \mathsf{Enc}(k, 0)$ is not a secure PRG.

We start with the first claim. If we assume by contradiction that there exists an efficient adversary $\mathcal{A}$ that breaks the smCPA-security of $(\mathsf{Enc}', \mathsf{Dec}')$, we build $\mathcal{A}'$ against the smCPA-security of $(\mathsf{Enc}, \mathsf{Dec})$ the following way. It starts by calling $\mathcal{A}$. When $\mathcal{A}$ outputs two messages m_0, m_1, $\mathcal{A}'$ outputs the same messages to the challenger. When the challenger outputs a ciphertext c, $\mathcal{A}'$ sends to $\mathcal{A}$ the ciphertext $1^\ell || c$. When $\mathcal{A}$ outputs a bit b', $\mathcal{A}'$ outputs the same. This is summed up in the following sketch:

$\mathcal{C}$	$\mathcal{A}'$	$\mathcal{A}$
$k \hookleftarrow U(K)$		
	Call $\mathcal{A}$	
		Choose and send $(m_0, m_1) \in P'$
	Send the same messages (m_0, m_1)	
Send $c := \mathsf{Enc}(k, m_b)$		
	Compute and send to $\mathcal{A}$: $c' := 1^\ell || c$	
		Output b'
	Output b'	

In these games, the view of $\mathcal{A}$ is the same as in the previous question. This means that it behaves the same way as in the Exp_b games for the encryption scheme $(\mathsf{Enc}', \mathsf{Dec})$. By definition of the advantage, $\mathsf{Adv}(\mathcal{A}') = \mathsf{Adv}(\mathcal{A})$. Thus, this breaks the security of $(\mathsf{Enc}, \mathsf{Dec})$.

We move on to prove the second claim by exhibiting an efficient distinguisher $\mathcal{B}$. It does the following: upon receiving a sample from either $G(U(K))$ or the uniform distribution, it outputs 1 if the first ℓ bits are 1 and 0 otherwise. Its advantage is $1 - \frac{1}{2^\ell}$. It is non-negligible as soon as $\ell \geq 1$.

Exercise 3. *Enlarge your PRG*

Let $G : \{0,1\}^k \to \{0,1\}^{k+1}$ be a secure pseudo-random generator.

1. Let $\ell < k + 1$ and define $G_\ell : \{0,1\}^k \to \{0,1\}^\ell$ such that $G_\ell(x) = [G(x)]_{1\ldots\ell}$, where this denotes the first ℓ bits of $G(x)$. Prove that G_ℓ satisfies the security notion of a PRG[1].

☞ Assume that there exists some ℓ for which G_ℓ is not a secure PRG and let $\mathcal{A}$ be a distinguisher between $G_\ell(U(\{0,1\}^k))$ and $U(\{0,1\}^\ell)$ with non-negligible advantage. We build a distinguisher $\mathcal{A}'$ between distributions $G(U(\{0,1\}^k))$ and $U(\{0,1\}^{k+1})$ that does the following. Upon receiving a sample y, it keeps only the ℓ first bits of it and runs $\mathcal{A}([y]_{1\ldots\ell})$. It outputs the bit b' that was returned.

[1] It is however NOT a PRG as its input size is less than its output size.

If y follows the uniform distribution over $\{0,1\}^{k+1}$ then $[y]_{1\dots\ell}$ follows the uniform distribution over $\{0,1\}^{\ell}$. If y follows the distribution $G(U(\{0,1\}^k))$ then $[y]_{1\dots\ell}$ follows the distribution $G_\ell(U(\{0,1\}^k))$. We see that $\mathcal{A}$ is always called upon samples from the distributions it distinguishes from with non-negligible advantage. As $\mathcal{A}'$ outputs the same answer as $\mathcal{A}$ it holds that $\mathsf{Adv}(\mathcal{A}') = \mathsf{Adv}(\mathcal{A})$. This contradicts the security of G.

2. Consider $G^{(1)} : \{0,1\}^k \to \{0,1\}^{k+2}$ defined as follows. On input $x \in \{0,1\}^k$, algorithm $G^{(1)}$ first evaluates $G(x)$ and obtains $(x^{(1)}, y^{(1)}) \in \{0,1\}^k \times \{0,1\}$ such that $G(x) = x^{(1)} \parallel y^{(1)}$. It then evaluates G on $x^{(1)}$ and eventually returns $G(x^{(1)}) \parallel y^{(1)}$. Show that if G is a secure PRG, then so is $G^{(1)}$.

☞ The intuition of the proof is the following: as G is a secure PRG, we can replace the output of $G(x)$ with the uniform distribution, and no adversary will notice. Then we compare the distribution $G(U(\{0,1\}^k))||U(\{0,1\})$, which is the distribution of $G^{(1)}(U(\{0,1\}^k))$ with this replacement, with $U(\{0,1\}^{k+2})$ and the security of the PRG once again saves us.

We study three distributions:

- D_0: the PRG distribution $G^{(1)}(U(\{0,1\}^k))$.
- D_1: the hybrid distribution $G(U(\{0,1\}^k))||U(\{0,1\})$.
- D_2: the uniform distribution $U(\{0,1\}^{k+2})$.

This proof is based on the hybrid argument: we will prove that no efficient distinguisher can distinguish with non-negligible advantage between D_0 and D_1, and between D_1 and D_2. This will prove that D_0 and D_2 cannot be distinguished.

Step 1: Assume that there exists an efficient distinguisher $\mathcal{A}$ between D_0 and D_1 with non-negligible advantage. We build $\mathcal{A}'$ a distinguisher between $G(U(\{0,1\}^k))$ and $U(\{0,1\}^{k+1})$ the following way. Upon receiving $x = (x_0, x_1) \in \{0,1\}^k \times \{0,1\}$, algorithm $\mathcal{A}'$ computes $x' := G(x_0)||x_1$) and sends it to $\mathcal{A}$. Note that x' follows exactly the distribution D_0 or D_1 depending on whether x is sampled with the PRG or uniformly. Then $\mathcal{A}'$ outputs exactly the same bit as $\mathcal{A}$. Its advantage is exactly the advantage of $\mathcal{A}$, which contradicts the security of G.

Step 2: Assume that there exists an efficient distinguisher $\mathcal{B}$ between D_1 and D_2 with non-negligible advantage. We build $\mathcal{B}'$ a distinguisher between $G(U(\{0,1\}^k))$ and $U(\{0,1\}^{k+1})$ that does the following. Upon receiving a sample x from either of these distributions, it flips a coin and get a uniform bit y. It calls $\mathcal{B}$ on $x||y$ and answers the same bit as $\mathcal{B}$. Note that $x||y$ is exactly distributed as D_1 or D_2 depending on whether x is sampled with the PRG or uniformly. Then $\mathcal{B}'$ has non-negligible advantage (equal to the advantage of $\mathcal{B}$).

Finally, algorithm $G^{(1)}$ is a secure PRG, as the advantage of any distinguisher for D_0 and D_2 is at most the sum of the advantage of any distinguisher for D_0 and D_1, and D_1 and D_2.

3. **(a)** Let $n \geq 1$. Propose a construction of a PRG $G^{(n)} : \{0,1\}^k \to \{0,1\}^{k+n+1}$ based on G. Show that if G is a secure PRG, then so is $G^{(n)}$.

☞ We iterate the previous construction, i.e., assuming that $G^{(i)}$ exists and is a secure PRG, we build $G^{(i+1)}$ that does the following. Upon receiving a key $x \in \{0,1\}^k$, run $G^{(i)}(x) =: (x_0, x_1) \in \{0,1\}^k \times \{0,1\}^i$. Return $G(x_0)||x_1$.

Note that the security proof is exactly the same as before except that **Step 1** now relies on the security of $G^{(i)}$ instead of the security of G. Remark: it is possible to rely only on the security of G, by using more hybrid distributions and more steps in the previous proof.

(b) Evaluate the cost of your construction.

☞ One evaluation of $G^{(i)}$ costs $i+1$ times the complexity of G. Let ε denote the advantage of the best adversary against the security game of G, then the security loss is as follows. The advantage of the best adversary against $G^{(i)}$ is at most $(i+1)\varepsilon$.

4. In this question only, we assume that $G : \{0,1\}^k \to \{0,1\}^{2k}$ is a secure PRG. Adapt the previous questions to build a secure PRG $G' : \{0,1\}^k \to \{0,1\}^{2^n \cdot k}$ for any $n \geq 1$. Evaluate the cost of your construction and compare it with the previous one.

☞ Let $G^{(1)} := G$. Assume that G is ε-secure. We build $G^{(i+1)}$ by induction, by assuming the existence of a secure PRG $G^{(i)} : \{0,1\}^k \to \{0,1\}^{2^i k}$ that calls at most $2^i - 1$ times G and is $i\varepsilon$-secure. On input $x \in \{0,1\}^k$, the PRG $G^{(i+1)}$ computes $G(x) =: (x_0, x_1) \in \{0,1\}^k \times \{0,1\}^k$ and outputs $G^{(i)}(x_0)||G^{(i)}(x_1)$. Then it calls G at most $2(2^i - 1) + 1 = 2^{i+1} - 1$ times.

We prove the security of the PRG $G^{(i+1)}$ by using a hybrid argument with two hybrid distributions. Namely

- D_0 is the distribution $G^{(i+1)}(U(\{0,1\}^k))$.
- D_1 is the distribution $G^{(i)}(U(\{0,1\}^k))||G^{(i)}(U(\{0,1\}^k))$.
- D_2 is the distribution $G^{(i)}(U(\{0,1\}^k))||U(\{0,1\}^{2^i k})$.
- D_3 is the distribution $U(\{0,1\}^{2^{i+1} k})$.

Under the assumption that G is secure, D_0 and D_1 are indistinguishable. The distributions D_1 and D_2 are indistinguishable under the security assumption of $G^{(i)}$, and this is also the case for D_2 and D_3. This proves the security of $G^{(i+1)}$.

To be more precise, the advantage of any distinguisher against $G^{(i+1)}$ is at most $(2^{i+1} - 1)\varepsilon$, assuming that the advantage of any distinguisher against G is at most ε.

When compared to the previous question, we gain a factor k, at the cost of having a PRG with output size 2^{2k} instead of 2^{k+1}.

An arbitrary-length PRG is a function $G^\star$ taking as inputs $x \in \{0,1\}^n$ and $\ell \geq 1$ in unary, and returning an element of $\{0,1\}^\ell$. It is said to be secure if for all ℓ polynomially bounded with respect to n, the distributions $G^\star(U(\{0,1\}^n), 1^\ell)$ and $U(\{0,1\})^\ell$ are computationally indistinguishable.

5. Let $n \geq 1$. Propose a construction of an arbitrary-length PRG G^* based on G. Show that if G is a secure PRG, then so is G^*.

   ☞ Goldreich-Goldwasser-Micali:

   We construct a pseudo-random generator $G_\ell : \{0,1\}^n \to \{0,1\}^\ell$ for any $\ell > 0$:

   - if $\ell \leq n$, let $G_\ell : x \mapsto [G(x)]_{1\ldots\ell}$ (recall the first question);
   - if $\ell \geq n+1$, let $G_\ell : x \mapsto G(G^{(\ell-1)}(x)]_{1\ldots n}) \parallel [G^{(\ell-1)}(x)]_{n+1\ldots\ell+1}$.

   From the previous questions, and since $G = G^{(n)}$ is secure, we know by induction that all G_ℓ are secure.

   Finally, $G^\star : (x, 1^\ell) \mapsto G_\ell(x)$ is a secure arbitrary-length PRG.

To be continued...