CS6111 (2025): Assignment 1

Instructor: John Augustine

Due: Aug 24, 2025 (11.59 PM IST).

- Use LaTeX to typeset your answer. You must submit both the latex file as well as the PDF file. A turnitin submission link will be provided.
- You can use theorems/lemmas from the textbook (Katz and Lindell, Edition 3 only) without reproving them. However, you should mention the theorem/lemma number from the textbook and paraphrase the statement of the theorem/lemma to make your solution self-contained.
- The deadline is Aug 24, 2025 (11:59 PM Indian Standard Time).
- Please follow the academic honesty policy for assignments. If you are stuck and need a hint, please post a question on moodle.
- The assignment is 30 marks with equal distribution. It will be later scaled down to 5 marks.

1. (5 marks) In this exercise, we study conditions under which the shift, mono-alphabetic substitution, and Vigenère ciphers are perfectly secret:
   (a) Prove that if only a single character is encrypted, then the shift cipher is perfectly secret.

Solution: Given $m \in \{0, 1, 2, \ldots, 25\}$ the algorithm is as follows

$$k \leftarrow \{0, 1, 2, \ldots, 25\}$$

$$Enc_k(m) = (m + k) \ \ mod\ 26$$

$$Dec_k(c) = (c - k) \ \ mod\ 26$$

By bayes theorem,

$$Pr[M = m|C = c] = \frac{Pr[M = m] * Pr[C = c|M = m]}{\sum_{m' \in \mathcal{M}} Pr[M = m'] * Pr[C = c|M = m']}$$

Clearly from the alogrithm, since key k is chosen uniformly

$$Pr[C = c|M = m] = Pr[C = m \oplus k|M = m] = Pr[K = k] = \frac{1}{26}$$

$$\implies Pr[M = m|C = c] = \frac{Pr[M = m] * \frac{1}{26}}{\sum_{m' \in \mathcal{M}} Pr[M = m'] * \frac{1}{26}} = \frac{Pr[M = m]}{\sum_{m' \in \mathcal{M}} Pr[M = m']}$$

By law of total probability,

$$\sum_{m' \in \mathcal{M}} Pr[M = m'] = 1$$

$$\implies Pr[M = m|C = c] = Pr[M = m]$$

Hence perfect secrect holds for the single character encrypted through the shift cipher.

(b) What is the largest plaintext space $\mathcal{M}$ you can find for which the mono-alphabetic substitution cipher provides perfect secrecy? (Note: $\mathcal{M}$ need not contain only valid English words.)

Solution: The substitution cipher is a bijection map from $\{0, 1, 2, \ldots, 25\}$ to $\{0, 1, 2, \ldots, 25\}$. Hence there are 26! different possibilities for the bijection.
The largest message space $\mathcal{M}$ that can be encrypted with perfect secrecy is of length 26 with non-repeated characters in the original message.
Consider $m \in \mathcal{M}$ of length 26 with non-repeated characters,

$$m = (m_1, m_2, \ldots, m_{26})$$

$$k \leftarrow \{0, 1, \ldots, 26!\}$$

$$Enc_k(m) = c = (c_1, c_2, \ldots, c_{26})$$

where k represents which bijection function is chosen for encryption.

Bayes theorem implies,

$$Pr[M = m|C = c] = \frac{Pr[M = m] * Pr[C = c|M = m]}{\sum_{m' \in \mathcal{M}} Pr[M = m'] * Pr[C = c|M = m']}$$

$$Pr[C = c|M = m] = Pr[K = k] = \frac{1}{26!}$$

$$\implies Pr[M = m|C = c] = \frac{Pr[M = m] * \frac{1}{26!}}{\sum_{m' \in \mathcal{M}} Pr[M = m'] * \frac{1}{26!}} = \frac{Pr[M = m]}{\sum_{m' \in \mathcal{M}} Pr[M = m']}$$

By law of total probability,

$$\implies Pr[M = m|C = c] = Pr[M = m]$$

Hence, perfect secrecy holds for message of length 26 or lower with non-repeated characters . But for length more than 26 characters there will be certain characters that are repeated which will have the same ciphertext character due to the bijective mapping which reveals some information about the message encrypted. Hence, The size of message space is 26!.

(c) Show how to use the Vigenère cipher to encrypt any word of length t so that perfect secrecy is obtained (note: you can choose the length of the key). Prove your answer.

Solution: Using the vigenere shift cipher we can encrypt any message of length t with perfect secrecy by choosing a key of the same length t. This is nothing but a one time pad now.

$$\implies k \leftarrow \mathcal{K} \qquad \textit{such that} \qquad |\mathcal{K}| = \frac{1}{26^t}$$

For such an encryption, we cannot perform frequency analysis on the streams of ciphers.

$$Enc_k(m_1, m_2, \ldots, m_t) = (m_1 \oplus k_1, m_2 \oplus k_2, \ldots, m_t \oplus k_t) = c$$

Bayes theorem implies,

$$Pr[M = m|C = c] = \frac{Pr[M = m] * Pr[C = c|M = m]}{\sum_{m' \in \mathcal{M}} Pr[M = m'] * Pr[C = c|M = m']}$$

Since key k is chosen uniformly at random,

$$Pr[C = c|M = m] = Pr[C = m \oplus k|M = m] = Pr[K = k] = \frac{1}{26^t}$$

$$\implies Pr[M = m|C = c] = \frac{Pr[M = m] * \frac{1}{26^t}}{\sum_{m' \in \mathcal{M}} Pr[M = m'] * \frac{1}{26^t}} = \frac{Pr[M = m]}{\sum_{m' \in \mathcal{M}} Pr[M = m']}$$

By law of total probability,

$$\implies Pr[M = m|C = c] = Pr[M = m]$$

Hence by choosing a key length of the same size as message a Vigenere cipher encryption is perfectly secret.

2. (5 marks) Let G be a pseudorandom generator with a polynomial expansion factor $\ell(n)$. In each of the following cases, say whether the defined G' is necessarily a pseudorandom generator. If yes, give a proof; if not, show a counterexample.

   (a) Define $G'(s) = G\left(s_1 \cdots s_{\lfloor log(n) \rfloor}\right)$, where $s = s_1 \cdots s_n$.

> **Solution:** Given G(s) is a pseudorandom generator,
>
> $$\implies |Pr[D(G(s)) = 1] - Pr[D(r) = 1]| \leq \ negl(n)$$
>
> Let us assume G'(s) : $\{0,1\}^{\lfloor \log n \rfloor} \rightarrow \{0,1\}^{l(\lfloor logn \rfloor)}$
> The number of possible seeds is
>
> $$2^{\lfloor \log n \rfloor} \leq 2^{\log n} = n$$
>
> Hence a polynomial time adversary can check all the n possible cases for the seeds and distinguish the generator G'(.) from a perfect random generator. Hence G'(.) is not a Pseudo Random Generator.

   (b) Define $G'(s) = G\left(x_1 \cdots x_n\right)$, where $x = G(s)$ and $s = s_1 \cdots s_n$.

> **Solution:** Let us assume that there exists a Distinguisher A that Distinguishes $G^{'}(.)$ from a PRT.
>
> $$\implies |Pr[A(G'(x)) = 1] - Pr[A(r) = 1]| \geq negl(n)$$
>
> where x is the truncated n bits of the result G(s).
> Now construct a distinguisher D based on reduction using A,
>
> 1. $s \leftarrow \{0,1\}^n$
> 2. Query G with s, x = G(s)
> 3. Query G with trucated n bits of x
> 4. $A \leftarrow G^{'}(x) = G(G(s)_{1:n})$
>
> $$G(s)_{1:n} \sim U_n$$
>
>   This is because when a pseudorandom string is truncated we get a pseudorandom string.
> 5. $D \leftarrow$ Output of A
>
> Since A distinguishes, $G^{'}(.)$, by reduction D distinguishes G(.).
>
> $$\implies G(.) \textit{ is not pseudorandom}$$

This contradicts the fact that G(.) is a PRG, hence our assumption is wrong.

$$\implies \textit{distinguisher A cannot exist}$$

Therefore, G'(.) is pseudorandom.

3. (5 marks) Consider the following modification to the one-time pad. Let the message m be of length n.

   Gen: Take two distinct integers i and j uniformly at random from $0 \leq i, j \leq 2^{\frac{n}{2}}$, output $k = 2^{\frac{n}{2}}i + j$.

   Enc: For a given message m and key k, output $m \oplus k$.

   Is this scheme perfectly secure?

Solution:

For this we use Theorem 2.10,

If (Gen, Enc, Dec) is a perfectly secret encryption scheme with message space $\mathcal{M}$ and key space $\mathcal{K}$, then $|\mathcal{K}| > |\mathcal{M}|$

Proof. Let $|\mathcal{K}| < |\mathcal{M}|$ and $c \in \mathcal{C}$ be a ciphertext with non zero probability. If $\mathcal{M}(\lfloor)$ is the set of all messages that can encrypt to cipher text c using some key k ,

$$\mathcal{M}(c) = \{m | m = Dec_k(c) \ \ \textit{for some key } k\}$$

Clerly,

$$|\mathcal{M}(c)| \leq |\mathcal{K}|$$

If

$$|\mathcal{K}| < |\mathcal{M}|$$

Then, there exists $m' \in \mathcal{M}(c)$ *and* $m' \notin \mathcal{M}$

$$Pr[M = m'|C = c] = 0 \neq Pr[M = m']$$

Hence, for perfect encryption $|\mathcal{K}| > |\mathcal{M}|$.
Now consider the construction of key k,

$$k = 2^{\frac{n}{2}}i + j$$

It is clearly a concatenation of distinct integers i and j (each $\frac{n}{2}$ bits)chosen such that $0 \leq i, j \leq 2^{\frac{n}{2}}$.

Since i, j are distinct,

$$|\mathcal{K}| < 2^n = |\mathcal{M}|$$

Hence the scheme is not perfectly secure.

4. (5 marks) In certain practical scenarios, we might be open to compromising with the "perfect" security and allow a little relaxation bound ϵ to it. We say that the scheme Π with security parameter n and message space $\mathcal{M} = \{0,1\}^n$ is almost perfect if

$$\Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1\right] \leq \frac{1}{2} + \varepsilon.$$

for some non-zero ϵ.

Consider the one-time pad with key space $\mathcal{K} \subset \{0,1\}^n$. What would be the minimum size of $\mathcal{K}$ as a function of the size of the message space and ϵ that will assure almost perfect security?

Solution: Let $b \in \{0,1\}$ be a random bit chosen,

$$k \leftarrow \mathcal{K}$$

$$c \leftarrow Enc_k(m_b) = m_b \oplus k$$

$$\Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1\right] = \frac{1}{2} * \Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1 \mid b = 0\right] + \frac{1}{2} * \Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1 \mid b = 1\right]$$

$$\Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1\right] = \frac{1}{2} * \Pr\left[A \textit{ outputs } 0 \mid b = 0\right] + \frac{1}{2} * \Pr\left[A \textit{ outputs } 1 \mid b = 1\right]$$

Say the Adversary is choosing two messages m_0, m_1 so as to maximize its gains,

Let, C_{m_0} *and* C_{m_1} denote the set of all cipher texts corresponding to a message m_0, m_1 respectively, we try to minimize the overlap between the two spaces.

$$\Pr\left[A \textit{ outputs } 0 \mid b = 0\right] = \sum_{c \in C_{m_0}} Pr[C = c] * Pr[A \textit{ outputs } 0 | C = c]$$

Since every message is mapped to $|K|$ keys,

$$Pr[C = c] = \frac{1}{|K|}$$

$$\Pr\left[A \textit{ outputs } 0 \mid b = 0\right] = \frac{1}{|K|} * \sum_{c \in C_{m_0}} Pr[A \textit{ outputs } 0 | C = c]$$

Let $M(c)$ be the set of messages that can be encrypted to a cipher text c using some key k,

$$Pr[A \textit{ outputs } 0 | C = c] = 1 * Pr[m_1 \notin M(c)] + \frac{1}{2} * Pr[m_1 \in M(c)]$$

If m_1 is not in M(c) then with a probability 1 Adversary will be able to predict the random bit b, if not its a random guess on b.

$$Pr[m_1 \notin M(c)] = Pr[c \in C_{m_0} \; and \; c \notin C_{m_1}]$$

A message m is mapped to $|K|$ through different keys in the key space,

$$|C_{m_0}| = |C_{m_1}| = |K|$$

$$|C_{m_0} \cap C_{m_1}| = (|C_{m_0}| + |C_{m_1}|) - |C_{m_0} \cup C_{m_1}| \geq 2|K| - 2^n$$
$$|C_{m_0}/C_{m_1}| = |C_{m_0}| - |C_{m_0} \cap C_{m_1}| \leq 2^n - |K|$$

$$Pr[m_1 \notin M(c)] = \frac{|C_{m_0}/C_{m_1}|}{|C_{m_0}|} \leq \frac{2^n - |K|}{|K|}$$

Hence,

$$Pr[A \; outputs \; 0|C = c] = 1 * Pr[m_1 \notin M(c)] + \frac{1}{2} * Pr[m_1 \in M(c)] = \frac{1}{2} + \frac{1}{2} Pr[m_1 \notin M(c)]$$

$$\implies Pr[A \; outputs \; 0|C = c] \leq \frac{1}{2} + \frac{1}{2} * \frac{2^n - |K|}{|K|}$$

Since, $|C_{m_0}| = |K|$

$$\implies \Pr\left[A \; outputs \; 0 \; |b = 0\right] leq \frac{1}{|K|} * \sum_{c \in C_{m_0}} \frac{1}{2} + \frac{1}{2} * \frac{2^n - |K|}{|K|} \leq \frac{1}{2} + \frac{1}{2} * \frac{2^n - |K|}{|K|}$$

By symmetry argument,

$$implies \Pr\left[A \; outputs \; 1 \; |b = 1\right] \leq \frac{1}{2} + \frac{1}{2} * \frac{2^n - |K|}{|K|}$$

$$\implies \Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} \leq 1\right] = \frac{1}{2} + \frac{1}{2} * \frac{2^n - |K|}{|K|}$$

By comparing with the bound on the question,

$$\frac{1}{2} * \frac{2^l - |K|}{|K|} \leq \epsilon$$

$$|K|(2\epsilon + 1) \geq 2^l \implies |K| \geq \frac{2^n}{2\epsilon + 1}$$

5. (5 marks) Let $G_1, G_2 : \{0,1\}^n \to \{0,1\}^\ell$ be two functions. Define a new generator

$$G : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^\ell : G(s_1, s_2) = G_1(s_1) \oplus G_2(s_2).$$

Suppose G is not a PRG. Then, prove that neither G_1 nor G_2 are PRGs.

Solution:

If $G(.)$ is a not a PRG, there exists a distinguisher D such that

$$|Pr[D(G(s_1, s_2)) = 1] - Pr[D(r) = 1]| \geq \ negl(n)$$

in other words,

$$G(s_1, s_2) \not\sim U_l$$

where, U_l is a uniform distribution on set of all random strings. If G_1 is pseudorandom,

$$G_1(s_1) \sim U_l$$

Then ,

$$G_1(s_1) \oplus G_2(s_2) \sim U_l$$

This contradicts the fact that there exists a Distinguisher D to distinguish $G(s_1, s_2)$.

Simmilarly, if G_2 is pseudorandom,

$$G_2(s_2) \sim U_l$$

Then ,

$$G_1(s_1) \oplus G_2(s_2) \sim U_l$$

This contradicts the fact that there exists a Distinguisher D to distinguish $G(s_1, s_2)$.

Hence, if $G(s_1, s_2)$ is not PRG, then both $G_1(s_1)$ and $G_2(s_2)$ are not PRG.

6. (5 marks) Consider the following definition of perfect secrecy for the encryption of two messages. An encryption scheme $(\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ over a message space $\mathcal{M}$ is perfectly-secret for two messages if for all distributions over $\mathcal{M}$, all $m, m' \in \mathcal{M}$, and all $c, c' \in \mathcal{C}$ with $\Pr[C = c \wedge C' = c'] > 0$:

$$\Pr[M = m \wedge M' = m' \mid C = c \wedge C' = c'] = \Pr[M = m \wedge M' = m'],$$

where m and m' are sampled independently from the same distribution over $\mathcal{M}$. Prove that no encryption scheme satisfies this definition. (*Hint:* Take $m \neq m'$ but $c = c'$.)

Solution: Let $m \neq m'$ and $c = c'$,

$$Pr[M = m \wedge M' = m'] = 1 - \frac{1}{|\mathcal{M}|}$$

For the same key k, if two messages $m \neq m'$ have the same cipher texts, then the correctness property of encryption cannot hold, hence

$$Pr[M = m \wedge M' = m' | C = c \wedge C = c'] = 0$$

Clearly, $LHS \neq RHS$, therefore no encryption scheme satisfies,

$$\Pr[M = m \wedge M' = m' \mid C = c \wedge C' = c'] = \Pr[M = m \wedge M' = m']$$

Acknowledgment

I would like to thank Shri Prathaa M (EE22B144), Pranay (EE22B132) for useful discussion on solving the problems. Of course, you must replace the names appropriately.