Chapter 1 and Problems in hw1

Mingjia Huo

Problem 1.Part A. *Proof.* Define $A = \{s \in S \mid \Pr[X = s] \geq \Pr[Y = s]\}$, so $A \subseteq S$.

First, we have

$$1 = \sum_{s \in S} \Pr[X = s] = \sum_{s \in S} \Pr[Y = s],$$

$$\Rightarrow \sum_{s \in A} \Pr[X = s] + \sum_{s \in S \setminus A} \Pr[X = s] = \sum_{s \in A} \Pr[Y = s] + \sum_{s \in S \setminus A} \Pr[Y = s],$$

$$\Rightarrow \sum_{s \in A} (\Pr[X = s] - \Pr[Y = s]) = \sum_{s \in S \setminus A} (\Pr[Y = s] - \Pr[X = s]).$$

So

$$\begin{aligned}
&\frac{1}{2} \sum_{s \in S} \mid \Pr[X = s] - \Pr[Y = s] \mid \\
=&\frac{1}{2} \left(\sum_{s \in A} (\Pr[X = s] - \Pr[Y = s]) + \sum_{s \in S \setminus A} (\Pr[Y = s] - \Pr[X = s]) \right) \\
=&\sum_{s \in A} (\Pr[X = s] - \Pr[Y = s])
\end{aligned}$$

And

$$\max_{T \subseteq S} (\Pr[X \in T] - \Pr[Y \in T]) = \sum_{s \in A} (\Pr[X = s] - \Pr[Y = s]),$$

so

$$\Delta(X, Y) = \frac{1}{2} \sum_{s \in S} \mid \Pr[X = s] - \Pr[Y = s] \mid = \max_{T \subseteq S} (\Pr[X \in T] - \Pr[Y \in T]).$$

$\square$

Problem 1.Part B. *Proof.* Define $B = \{s \in S \mid D(s) = 1\}$, so $B \subseteq S$.

So we have

$$\begin{aligned}
\Pr[D(X) = 1] - \Pr[D(Y) = 1] &= \sum_{s \in B} \Pr[X = s] - \sum_{s \in B} \Pr[Y = s] \\
&= \Pr[X \in B] - \Pr[Y \in B] \\
&\leq \max_{T \subseteq S} (\Pr[X \in T] - \Pr[Y \in T]) \\
&= \Delta(X, Y)
\end{aligned}$$

$\square$

Problem 1.Part C. *Proof.* By the definition of D and X, we have

$$\Pr[D(X) = 1] = \sum_{s \in S} \Pr[X = s] p_s.$$

So

$$\Pr[D(X) = 1] - \Pr[D(Y) = 1] = \sum_{s \in S} (\Pr[X = s] - \Pr[Y = s]) p_s \tag{1}$$

First, we prove that the optimal D satisfies $p_s \in \{0, 1\}$ for every $s \in S$. If not, there is some $s_0 \in S$, such that $p_{s_0} = q \in (0, 1)$ for D.

1. If $\Pr[X = s_0] \geq \Pr[Y = s_0]$, adjust p_{s_0} to 1. Then by equation (1), the value is non-decreasing.

2. If $\Pr[X = s_0] < \Pr[Y = s_0]$, adjust p_{s_0} to 0. Then by equation (1), the value is non-decreasing.

So the optimal D satisfies $p_s \in \{0, 1\}$ for every $s \in S$.

With the proof of Part B, when D is deterministic, we have

$$\Pr[D(X) = 1] - \Pr[D(Y) = 1] = \Pr[X \in B] - \Pr[Y \in B]$$

Assume $\arg\max_{T \subseteq S} (\Pr[X \in T] - \Pr[Y \in T]) = T_{max}$. Then define

$$D_{max}(s) = \begin{cases} 1, & s \in T_{max} \\ 0, & s \notin T_{max} \end{cases}$$

Thus,

$$\begin{aligned} \Pr[D_{max}(X) = 1] - \Pr[D_{max}(Y) = 1] &= \Pr[X \in T_{max}] - \Pr[Y \in T_{max}] \\ &= \max_{T \subseteq S} (\Pr[X \in T] - \Pr[Y \in T]) \\ &= \Delta(X, Y) \end{aligned}$$

On the other hand, we have

$$\Pr[D(X) = 1] - \Pr[D(Y) = 1] \leq \Delta(X, Y).$$

So

$$\max_{D} (\Pr[D(X) = 1] - \Pr[D(Y) = 1]) = \Delta(X, Y)$$

□

Problem 2. Given $c \in \mathcal{C}$, we can get at most $\mid \mathcal{K} \mid$ plaintexts. That is, there are at least $(\mid \mathcal{M} \mid - \mid \mathcal{K} \mid)$ plaintexts $m \in \mathcal{M}$, s.t.

$$\Pr[\mathrm{Enc}_K(m) = c] = 0.$$

Define $\mathcal{M}_c = \{m \mid \Pr[\text{Enc}_K(m) = c] = 0\}$.

On the other hand,

$$\begin{aligned}\Delta(\text{Enc}_K(m_0), \text{Enc}_K(m_1)) &= \max_{C \in \mathcal{C}}(\Pr[\text{Enc}_K(m_0) \in C] - \Pr[\text{Enc}_K(m_1) \in C]) \\ &= \sum_{c \in \mathcal{C}} \max\{\Pr[\text{Enc}_K(m_0) = c] - \Pr[\text{Enc}_K(m_1) = c], 0\}\end{aligned}$$

Fix m_0, compute:

$$\begin{aligned}\sum_{m_1 \in \mathcal{M}} \Delta(\text{Enc}_K(m_0), \text{Enc}_K(m_1)) &= \sum_{m_1 \in \mathcal{M}} \sum_{c \in \mathcal{C}} \max\{\Pr[\text{Enc}_K(m_0) = c] - \Pr[\text{Enc}_K(m_1) = c], 0\} \\ &= \sum_{c \in \mathcal{C}} \sum_{m_1 \in \mathcal{M}} \max\{\Pr[\text{Enc}_K(m_0) = c] - \Pr[\text{Enc}_K(m_1) = c], 0\} \\ &\geq \sum_{c \in \mathcal{C}} \sum_{m_1 \in \mathcal{M}_c} \max\{\Pr[\text{Enc}_K(m_0) = c] - \Pr[\text{Enc}_K(m_1) = c], 0\} \\ &= \sum_{c \in \mathcal{C}} \sum_{m_1 \in \mathcal{M}_c} \Pr[\text{Enc}_K(m_0) = c] \\ &\geq \sum_{c \in \mathcal{C}} (\mid \mathcal{M} \mid - \mid \mathcal{K} \mid) \Pr[\text{Enc}_K(m_0) = c] \\ &= \mid \mathcal{M} \mid - \mid \mathcal{K} \mid .\end{aligned}$$

The last inequality is because that there are at least $(\mid \mathcal{M} \mid - \mid \mathcal{K} \mid)$ items in $\mathcal{M}_c$. Then with Pigeonhole principle, there exists m_1^*, such that

$$\Delta(\text{Enc}_K(m_0), \text{Enc}_K(m_1^*)) \geq \frac{\mid \mathcal{M} \mid - \mid \mathcal{K} \mid}{\mid \mathcal{M} \mid} = 1 - \frac{\mid \mathcal{K} \mid}{\mid \mathcal{M} \mid}.$$

Problem 3. First Part. A *secret-sharing scheme* is a 5-tuple $(\mathcal{M}, \mathcal{L}, \mathcal{R}, \text{Enc}, , \text{Dec})$ where $\mathcal{M}, \mathcal{L}$ and $\mathcal{R}$ are finite sets and where

- Enc: $\mathcal{M} \to \mathcal{L} \times \mathcal{R}$ is a randomized algorithm;
- Dec: $\mathcal{L} \times \mathcal{R} \to \mathcal{M}$ is a deterministic algorithm.

Moreover, we require that $\text{Dec}(\text{Enc}(m)) = m$ for every $m \in \mathcal{M}$.

Second Part. Let $\Pi = (\mathcal{M}, \mathcal{L}, \mathcal{R}, \text{Enc}, \text{Dec})$ be a secret-sharing scheme. Let $\mathcal{L}$ be the distribution over L induced by the first output of Enc, write down as $\text{Enc}.L$. So $\text{Enc}.L$ is a random variable of range $\mathcal{L}$. Similarly, $\text{Enc}.R$ is a random variable of range $\mathcal{R}$.

We say that Π is *perfectly secure* if for every two messages $m_0, m_1 \in \mathcal{M}$,

$$\Delta(\text{Enc}(m_0).L, \text{Enc}(m_1).L) = 0, \quad \Delta(\text{Enc}(m_0).R, \text{Enc}(m_1).R) = 0.$$

Problem 1.3. Assume the plaintext is in English. That is, the letter is from a to z. First, we define two functions:

- *to_number*: $\{a, b, \cdots, z\} \to \{0, 1, \cdots, 25\}$

- *to_letter*: $\{0, 1, \cdots, 25\} \rightarrow \{A, B, \cdots, Z\}$

Let $\Pi = (\mathcal{K}, \mathcal{M}, \mathcal{C}, \text{Gen}, \text{Enc}, \text{Dec})$ be a Vigenere cipher encryption, where $\mathcal{K}, \mathcal{M}$ and $\mathcal{C}$ are finite sets which represent key, plaintext and ciphertext space respectively. And

- Gen: $\{0, 1\}^* \rightarrow \mathcal{K}$ is a randomized algorithm which generate a key k with its length equal to the input string.

- Enc: $\mathcal{K} \times \mathcal{M} \rightarrow \mathcal{C}$ is a deterministic encryption algorithm. Assume the length of k is l. Assume $m = m_1 m_2 \cdots m_n (m_i \in \{a, b, \cdots, z\}), k = k_1 k_2 \cdots k_l (k_i \in \{a, b, \cdots, z\})$, then define ciphertext $c = c_1 c_2 \cdots c_n (c_i \in \{A, B, \cdots, Z\})$ as followed:

  $$\forall i \in \{1, 2, \cdots, n\}, c_i = \mathit{to\_letter}((\mathit{to\_number}(m_i) + \mathit{to\_number}(k_{i'})) \mod 26),$$

  where $i' \equiv i \mod l$.

- Dec: $\mathcal{K} \times \mathcal{C} \rightarrow \mathcal{M}$ is a deterministic decryption algorithm. To get plaintext m, do:

  $$\forall i \in \{1, 2, \cdots, n\}, m_i = \mathit{to\_number}^{-1}((\mathit{to\_letter}^{-1}(c_i) - \mathit{to\_number}(k_{i'})) \mod 26),$$

  where $i' \equiv i \mod l$.