CPSC 418 / MATH 318 — Introduction to Cryptography
ASSIGNMENT 1

Name: Artem Golovin
Student ID: 30018900

Problem 1 — Superencipherment for substitution ciphers, 12 marks

1. (a) *Proof.* Encryption using Shift cipher is given by $E_K(M) \equiv (M + K) \mod 26$. Given $\mathcal{M} = \mathcal{C} = \mathcal{K} = \mathbb{Z}_{26}$, $K_1, K_2 \in \mathcal{K}$ and $M \in \mathcal{M}$:
   Let $C_1 \in \mathcal{C}$ be a ciphertext that results from encrypting plaintext M with a key K_1:
   i. $E_{K_1}(M) \equiv (M + K_1) \mod 26$
   ii. $C_1 \equiv E_{K_1}$
   iii. $C_1 \equiv (M + K_1) \mod 26$
   iv. Let $C_2 = E_{K_2}(C_1)$, where $E_{K_2}(C_1) \equiv (C_1 + K_2) \mod 26$
   v. Therefore, by substituting C_1,

$$\begin{aligned} C_2 &= C_1 + K_2 && \pmod{26} \\ &= M + (K_1 + K_2) && \pmod{26} \\ &= M + K_3 && \pmod{26} \end{aligned}$$

   Where $K_3 \in \mathcal{K}$ and $K_3 = K_1 + K_2$. Therefore, resulting key of multiple encipherment is K_3.

   $\square$

   (b) *Proof.* Based on previous proof, superencipherment using shift cipher can be defined as follows

$$E_{K_i}(M) = (M + \sum_{\substack{k_i \in \\ i=1}}^{n} k_i) \pmod{26}$$

   Where $M \in \mathcal{M}$, $K_i \in \mathcal{K}$, $n \in \mathbb{Z}$, and $i \geq 1, i \in \mathbb{Z}$.
   i. Base Case: let $n = 1$,

$$\begin{aligned} C &= M + \sum_{\substack{K_i \in \mathcal{K} \\ i=1}}^{1} K_i && \pmod{26} \\ &= M + K_1 && \pmod{26} \end{aligned} \tag{1}$$

   ii. Induction hypothesis: Assume $n = m$, where $m \in \mathbb{Z}$. Therefore:

$$C = M + \sum_{\substack{K_i \in \mathcal{K} \\ i=1}}^{m} K_i \pmod{26}$$

Let:

$$K^{'} = \sum_{\substack{K_i \in \mathcal{K} \\ i=1}}^{m} K_i \tag{2}$$
$$C = M + K^{'} \pmod{26}$$

Still results in a shift cipher, according to definition.

iii. Inductive case: According to induction hypothesis, we can show that $m + 1$ holds true as well:

$$\begin{aligned}
C &= (M + \sum_{\substack{K_i \in \mathcal{K} \\ i=1}}^{m+1} K_i) && \pmod{26} \\
&= (M + (\sum_{\substack{K_i \in \mathcal{K} \\ i=1}}^{m} K_i + \sum_{\substack{K_i \in \mathcal{K} \\ i=1}}^{1} K_i)) && \pmod{26} \\
&= (M + (\sum_{\substack{K_i \in \mathcal{K} \\ i=1}}^{m} K_i + K_1)) && \pmod{26} \\
&= (M + (K^{'} + K_1)) && \pmod{26} \\
&= (M + K^{"}) && \pmod{26}
\end{aligned}$$

where K_1 is our base case (1) and $K^{'}$ is our induction hypothesis (2). Since $K_1 \in \mathcal{K}$ and $K^{'} \in \mathcal{K}$, therefore $K^{"} = K_1 + K^{'}, K^{"} \in \mathcal{K}$. Hence, the key of multiple encipherment is sum of all given keys.

□

2. *Proof.* Assume there exist two leys w_1 of length m and w_2 of length n.

Since we know lengths of given keys, to find the length of new keyword w, find the least common multiple of m and n, that is $x = LCM(m, n)$, where x is the length of the key w.

The key is obtained by adding each letters of given keys: $b_c = k_i + l_j \pmod{26}$, where b_c is a letter of key w, that is $b_c \in \{b_0, b_1, ..., b_{x-1}\}$, $k_i \in \{k_1, k_2, \ldots, k_m\}$, and $l_j \in \{l_1, l_2, \ldots, l_n\}$. The position of key is obtained by doing following:

$$i = c \mod m$$
$$j = c \mod n$$

Where i, j, and c are positions of letters in w_1, w_2, and w. That ensures that key wraps around and matches the size of w. Therefore, the length of key w is $LCM(m, n)$ and key is obtained by adding each letter of w_1 and w_2. □

Problem 2 — Key size versus password size, 21 marks

1. 256^8
2. (a) 94^8

   (b) $\frac{94^8}{256^8} \times 100 \approx 0.033\%$
3. Assuming that all characters are chosen equally likely, then $p(X_i) = \frac{1}{94}$. Therefore, entropy of the key space will be:

$$\begin{aligned} H(X) &= 8 \times \frac{1}{94} \log_2 94 \\ &\approx 0.56 \end{aligned}$$

4.

$$\begin{aligned} H(X) &= 8 \times \frac{1}{26} \log_2 \frac{1}{26} \\ &\approx 1.45 \end{aligned}$$

5. Given $H(X) = 128$, let $n \in \mathbb{Z}$ be a password length.

   (a) $p(X) = \frac{1}{94}$

$$\begin{aligned} n \times \frac{1}{94} \log_2(94) &= 128 \\ n &= \frac{128 \times 94}{\log_2(94)} \\ n &\approx 1836 \end{aligned}$$

   (b) $p(X) = \frac{1}{26}$

$$\begin{aligned} n \times \frac{1}{26} \log_2(26) &= 128 \\ n &= \frac{128 \times 26}{\log_2(26)} \\ n &\approx 709 \end{aligned}$$

Problem 3 — Equiprobability maximizes entropy for two outcomes, 12 marks

1.
$$
\begin{aligned}
H(X) &= \frac{1}{4}\log_2 4 + \frac{3}{4}\log_2 \frac{4}{3} \\
&\approx 0.811
\end{aligned}
$$

2. $H(X)$ is maximized if and only if all outcomes are equally likely. For any n, $H(X) = \log_2 n$ is maximal if and only if $p(X_i) = \frac{1}{n}$ for $1 \leq i \leq n$.

   Proof. Let's consider function $H(X) = p\log_2(\frac{1}{p}) + (1-p)\log_2(\frac{1}{(1-p)}) = -p\log_2 p - (1-p)\log_2(1-p)$ as a function of p:

$$
\begin{aligned}
H(p) &= p\log_2(\frac{1}{p}) + (1-p)\log_2(\frac{1}{(1-p)}) \\
&= -p\log_2 p - (1-p)\log_2(1-p)
\end{aligned}
$$

   By taking the derivative of $H(p)$, we can determine maximum of the function

$$
\begin{aligned}
H'(p) &= (-p\log_2 p)' - ((1-p)\log_2(1-p)) \\
&= -\frac{\ln(p)+1}{\ln(2)} + \frac{\ln(1-p)+1}{\ln(2)} \\
&= \frac{\ln(1-p) - \ln(p)}{\ln(2)} \\
&= \log_2(1-p) - \log_2(p)
\end{aligned}
$$

   Maximum of $H'(p) = \log_2(1-p) - \log_2(p)$:

$$
\begin{aligned}
\log_2(1-p) - \log_2(p) &= 0 \\
\log_2(\frac{(1-p)}{p}) &= 0 \\
1 &= \frac{(1-p)}{p} \\
p &= 1-p \\
p &= \frac{1}{2}
\end{aligned}
$$

   $\square$

3. Maximal value of $H(X)$, given $p = \frac{1}{2}$:

$$
\begin{aligned}
H(X) &= -\frac{1}{2}\log_2(\frac{1}{2}) - (1-\frac{1}{2})\log_2(1-\frac{1}{2}) \\
&= \frac{1}{2} + \frac{1}{2} \\
&= 1
\end{aligned}
$$

Problem 4 — Conditional entropy, 12 marks

Given conditional entropy

$$H(X|Y) = \sum_{y\in Y}\sum_{x\in X} p(x,y)\log_2(\frac{1}{p(x|y)})$$

1. Before computing $H(\mathcal{M}|\mathcal{C})$, let's compute $p(M_i|C_j)$

$$\begin{array}{llll} p(M_1|C_1) = \frac{1}{2} & p(M_1|C_2) = 0 & p(M_1|C_3) = 0 & p(M_1|C_4) = \frac{1}{2} \\ p(M_2|C_1) = \frac{1}{2} & p(M_2|C_2) = 0 & p(M_2|C_3) = \frac{1}{2} & p(M_2|C_4) = 0 \\ p(M_3|C_1) = 0 & p(M_3|C_2) = \frac{1}{2} & p(M_3|C_3) = \frac{1}{2} & p(M_3|C_4) = 0 \\ p(M_4|C_1) = 0 & p(M_4|C_2) = \frac{1}{2} & p(M_4|C_3) = 0 & p(M_4|C_4) = \frac{1}{2} \end{array}$$

   $H(\mathcal{M}|\mathcal{C})$ results in:

$$H(\mathcal{M}|\mathcal{C}) = 4 \times \frac{1}{4} \times (8 \times \frac{1}{2} \times \log_2(2)) = \frac{8}{2} = 4$$

2. Assuming that the system provides perfect secrecy $p(M|C) = p(M)$ and $p(M) > 0$ for all $M \in \mathcal{M}$. Given that $H(\mathcal{M}|\mathcal{C}) = \sum_{C\in\mathcal{C}}\sum_{M\in\mathcal{M}} p(M,C)\log_2(\frac{1}{p(M|C)})$ and $H(\mathcal{M}) = \sum_{M\in\mathcal{M}} p(M)\log_2(\frac{1}{p(M)})$, show that $H(\mathcal{M}|\mathcal{C}) = H(\mathcal{M})$. Assume that $H(\mathcal{M}|\mathcal{C}) = H(\mathcal{M})$:

   Proof.

$$H(\mathcal{M}|\mathcal{C}) = H(\mathcal{M})$$

$$\sum_{C\in\mathcal{C}} p(C) \sum_{M\in\mathcal{M}} p(M|C)\log_2(\frac{1}{p(M|C)}) = H(\mathcal{M})$$

   Given that the system provides perfect secrecy, we get

$$\sum_{C\in\mathcal{C}} p(C) \sum_{M\in\mathcal{M}} p(M)\log_2(\frac{1}{p(M)}) = H(\mathcal{M})$$

$$\sum_{C\in\mathcal{C}} p(C)H(\mathcal{M}) = H(\mathcal{M})$$

   Dividing both sides by $H(\mathcal{M})$, we get

$$\sum_{C\in\mathcal{C}} p(C) = 1$$

   Therefore, as shown above $\sum_{C\in\mathcal{C}} p(C) = 1$. Using that, we can now show that $H(\mathcal{M}|\mathcal{C}) = H(\mathcal{M})$:

$$\begin{aligned} H(\mathcal{M}|\mathcal{C}) &= \sum_{C\in\mathcal{C}} p(C) \sum_{M\in\mathcal{M}} p(M|C)\log_2(\frac{1}{p(M|C)}) \\ &= \sum_{M\in\mathcal{M}} p(M|C)\log_2(\frac{1}{p(M|C)}) \\ &= H(\mathcal{M}) \end{aligned}$$

$\square$

3. *Proof.* Assume the system provides perfect secrecy. Therefore, the following consition should be met $p(M|C) = p(M)$ for $M \in \mathcal{M}$ and $C \in \mathcal{C}$. Let's calculate $p(M_1|C_1)$, if the system provides perfect secrecy, $p(M_1|C_1) = p(M_1) = \frac{1}{4}$

$$p(M_1|C_1) = p(M_1) + p(M_2) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

Therefore, the system does not provide perfect secrecy. $\square$