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:

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

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.* Let $M \in \mathcal{M}$ be a plaintext of length $x \in \mathbb{Z}$. Let $p_0, p_1, p_2, ..., p_{x-1}$ be positions of letters in plaintext M. Given key $w_1 \in \mathcal{K}$ of length $m \in \mathbb{Z}$ and key $w_2 \in \mathcal{K}$ of length $n \in \mathbb{Z}$, let $k_0, k_1, k_2, ..., k_{m-1}$ be positions of letters in key w_1, and $l_0, l_1, l_2, ..., l_{n-1}$ be positions of letters in key w_2. To encrypt plaintext p_i, we use a key k_j, where i is letter position from 0 to $x - 1$.

$$j \equiv i \pmod{m}$$

Let ciphertext C_i be ciphertext that corresponds to p_i

$$C_i \equiv p_i + k_j \pmod{26}$$

Where $k_j \equiv i \pmod{m}$.

Let ciphertext C_{2_i} be ciphertext that corresponds to C_i. Therefore, the second round of encryption, using the key w_2, results in following:

$$\begin{aligned}
C_{2_i} &= C_i + l_j && \pmod{26} \\
&= (p_i + k_j) + l_z && \pmod{26} \\
&= p_i + (k_j + l_z) && \pmod{26}
\end{aligned}$$

Where $z \equiv i \pmod{n}$ and $l_z \equiv i \pmod{n}$. Therefore that ensures that length of resulting key will be x. $\square$

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. Out of 94 printable characters, we only have 52 that correspond to letters of English alphabet. That means probability of getting a lowercase letter is $p(X) = 1 - \frac{1}{52}$. Therefore, the entropy of this key space is:

$$\begin{aligned} H(X) &= 8 \times (1 - \frac{1}{52}) \log_2 \frac{1}{1 - \frac{1}{52}} \\ &\approx 0.22 \end{aligned}$$

5. (a)

   (b)

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

1. Given conditional entropy

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

   Where $p(x|y) \neq 0$, and $p(M_i) = p(C_i) = \frac{1}{4}$ for $1 \leq i, j \leq 4$, the resulting $H(\mathcal{M}|\mathcal{C})$ is:

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

2.
3.