Problem 5.1.

Within this homework, we consider that the Vigenère cipher is designed for 26 characters, where each letter is assigned its position in the alphabet: $A \to 0, B \to 1, \ldots, Z \to 25$. The encryption is done by adding the (repeated) key with the plaintext and then taking modulo 26. Assume that you have intercepted the following ciphertext which was encrypted either using monoalphabetic substitution or using the Vigenère cipher:
`EBGRYCXGBGHITURSYNEAVCGBGRYV`
Also, you have managed to find out 4 letters of the plaintext message:
`T*****************U**I**I***`

1. Can you tell if the message was encrypted with the Vigenère cipher or by means of monoalphabetic substitution?

Solution:

If monoalphabetic substitution was used, the letter `I` in the plaintext should always be replaced by the same letter in the ciphertext. Since letter `I` is encrypted as `C` the first time and as `G` the second time, the encryption scheme cannot be monoalphabetic substitution. Hence, we can conclude that the plaintext was encrypted with the Vigenère cipher.

Relevant slides : 276 - 277, 283

2. Using the previous question, can you find the key and the plaintext?
Hint: the key and plaintext consist of English words.

Solution:

We first "subtract" the known plaintext letters from the encrypted message to find parts of the key:

```
   E B G R Y C X G B G H I T U R S Y N E A V C G B G R Y V
⊖  T * * * * * * * * * * * * * * * * * U * * I * * I * * *
   L * * * * * * * * * * * * * * * * * K * * U * * Y * * *
```

Here, the operation $\ominus$ is defined as subtraction modulo 26 by assigning each letter its position in the alphabet $(A \to 0, B \to 1, \ldots)$.
The next step is to determine the length of the key. The key length cannot divide 18, 21 or 24, otherwise the initial `L` should be repeated on position 19, 22 and 25 respectively. Thus, the key cannot be of length $1, 2, 3, 4, 6, 7, 8, 9, 12, 18, 21, 24$. Knowing this, we start by looking for a valid key of length 5:

```
   E B G R Y C X G B G H I T U R S Y N E A V C G B G R Y V
⊖  T H * H A R D * R I W O * K T H E * U C K I * R I G E *
   L U * K Y L U * K Y L U * K Y L U * K Y L U * K Y L U *
```

Knowing that the key is an English word, it is easy to guess that the key is `LUCKY` and the message is `THE HARDER I WORK THE LUCKIER I GET`.

Relevant slides : 276 - 277, 283

Problem 5.2.

Let n be a positive integer, and M a uniformly distributed message of length $2n$ over the alphabet $\mathcal{A} = \{\mathtt{a}, \mathtt{b}, \dots, \mathtt{z}\}$. Given a key K, let $V_K(M)$ be the Vigenère encryption of M with key K.

1. You have a key K of length n, uniformly distributed in $\mathcal{A}^n$.

   (a) Does the encryption $V_K(M)$ provide perfect secrecy?

   Solution:

   We have $H(M) = \log_2(|\mathcal{A}^{2n}|) = 2n\log_2(26)$, while $H(K) = \log_2(|\mathcal{A}^n|) = n\log_2(26)$. So $H(K) < H(M)$, there is no perfect secrecy.

   Relevant slides : 286 - 290, 292 - 293

   (b) Is your answer to the previous question still true if M is not uniformly distributed?

   Solution:

   No: this method can provide perfect secrecy if the distribution of M is not uniform in $\mathcal{A}^{2n}$. A trivial example is if M takes the value $m_{\mathtt{a}} = \mathtt{aa}\dots\mathtt{a}$ with probability 1. Then, $M = m_a$ with probability 1 independently of the value of $V_K(M)$, so M and $V_K(M)$ are independent.
   A less trivial example would be to take M in the set of messages of the form $m||m \in \mathcal{A}^{2n}$ for $m \in \mathcal{A}^n$. We do not provide a full proof for this case (it is similar to question 2 (b)), but you can already notice that K now contains "enough" entropy, as $H(M) = n\log_2(26) = H(K)$.

   Relevant slides : 286 - 290

   (c) Is there any way to encrypt M with a key of length n taken from $\mathcal{A}^n$ that provides perfect secrecy?

   Solution:

   For any distribution of a key K' in $\mathcal{A}^n$, we have $H(K') \leq H(K)$ (uniform distribution maximizes the entropy). So irrespective of the encryption algorithm used, we always have $H(K') < H(M)$, so it is impossible to achieve perfect secrecy with a key of length n.

   Relevant slides : 292 - 293

2. You have two keys K_1 and K_2 each of length n chosen uniformly and independently in $\mathcal{A}^n$.

   (a) Does the double encryption $V_{K_2}(V_{K_1}(M))$ provide perfect secrecy?

   Solution:

   Double encryption using the Vigenère scheme does not increase the security at all. Indeed, let $K_3 = V_{K_2}(K_1)$ (i.e., for any i, the ith letter of K_3 is the sum modulo 26 of the ith letters of K_1 and K_2). Then, $V_{K_2}(V_{K_1}(M)) = V_{K_3}(M)$: the double encryption is equivalent to a simple encryption with K_3. So it does not provide perfect secrecy, since K_3 is of length n.

(b) Let $K_1||K_2$ denote the concatenation of the two keys. Does $V_{K_1||K_2}(M)$ provide perfect secrecy? Does the answer require M to be uniformly distributed in $\mathcal{A}^{2n}$?

Solution:

The key $K_3 = K_1||K_2$ is uniformly distributed in $\mathcal{A}^{2n}$, and for any $m \in \mathcal{A}^{2n}$, the map

$$V_\bullet(m) : \mathcal{A}^{2n} \longrightarrow \mathcal{A}^{2n} : k \longmapsto V_k(m)$$

is a bijection. Therefore, for any $m \in \mathcal{A}^{2n}$, $V_{K_3}(m)$ is also uniformly distributed in $\mathcal{A}^{2n}$. Then, for any $m, c \in \mathcal{A}^{2n}$, we have

$$p(V_{K_3}(M) = c|M = m) = p(V_{K_3}(m) = c) = \frac{1}{26^{2n}}$$

which does not depend on $M = m$, so M and $V_{K_3}(M)$ are independent, and the scheme provides perfect secrecy. We did not use the distribution of M: it remains true even if M is not uniformly distributed.

Relevant slides : 286 - 290, 292 - 293

3. Now fix $n = 4$. A crime lord learned about the betrayal of three of his men, Matt, Axel and Kyle. He decides that some will be killed, and some will simply be sent a warning, by receiving the kiss of death (as a caution). He sends the orders to his hitman: messages of the form $M = A||B$ where $A \in \{\texttt{kiss}, \texttt{kill}\}$ and $B \in \{\texttt{matt}, \texttt{axel}, \texttt{kyle}\}$.

(a) Suppose the messages are encrypted as $V_K(M)$ for keys K of length 4 (different for each message). Decrypt the two ciphertexts

$$C_1 = \texttt{iolgielz}, \text{ and } C_2 = \texttt{gikdiale}.$$

Solution:

For $i = 1, 2$, let $M_i = A_i||B_i$ be the clear message with A_i and B_i of length 4. Let K_i be the key of length 4 used to encrypt M_i. Then,

$$C_i = V_{K_i}(M_i) = V_{K_i}(A_i)||V_{K_i}(B_i).$$

Observe that $V_{K_1}(A_1)$ and $V_{K_1}(B_1)$ start with the same letter ($\texttt{i}$), so A_1 and B_1 also start with the same letter. Since A_1 must start with a $\texttt{k}$, we deduce that B_1 starts with a $\texttt{k}$, so $B_1 = \texttt{kyle}$. The 3rd letter of B_1 is an $\texttt{l}$, and it coincides with the 3rd letter of A_1, so $A_1 = \texttt{kill}$. Therefore $M_1 = \texttt{killkyle}$. The second letter of A_2 must be an $\texttt{i}$, and the second letter of $V_{K_2}(A_2)$ is an $\texttt{i}$, so the second letter of K_2 must be an $\texttt{a}$. Therefore the second letter of B_2 is exactly the second letter of $V_{K_2}(B_2)$, namely an $\texttt{a}$. So $B_2 = \texttt{matt}$. Let k_3 be the third letter of K_2. By looking at the third letter of B_2, we get $V_{k_3}(\texttt{t}) = \texttt{l}$, so $k_3 = \texttt{s}$ (because $V_\texttt{s}(\texttt{t}) = \texttt{l}$). Let a_3 be the third letter of A_2. We have

$$\texttt{k} = V_{k_3}(a_3) = V_\texttt{s}(a_3),$$

and we deduce that $a_3 = \texttt{s}$ (because $V_\texttt{s}(\texttt{s}) = \texttt{k}$). So $A_2 = \texttt{kiss}$, and $M_2 = \texttt{kissmatt}$.

(b) Suppose the key K is of length 8, but the same is used to encrypt all the messages. You intercept

$$C_1 = \texttt{xwcxlzkj}, \text{ and } C_2 = \texttt{xwjenbsy}.$$

Who will the hitman kill, and who will receive a warning?

Solution:

Let M_1 and M_2 be the clear messages of C_1 and C_2. Write $K = k_1k_2 \dots k_8$. Suppose that M_1 starts with `kiss`. Then, $V_{k_3}(\texttt{s}) = \texttt{c}$, so $k_3 = \texttt{k}$. But then, $V_{k_3}(\texttt{s}) = \texttt{c}$ and $V_{k_3}(\texttt{l}) = \texttt{v}$, none of which coincides with the third letter of C_2, a contradiction. So M_1 starts with `kill`, and as a direct consequence, M_2 starts with `kiss`.

Observe that the 5th letter of C_1 and C_2 are at distance 2 (`n` is two letters after `l`), and since they are both encrypted with k_5, the 5th letter of M_1 and M_2 must also be at distance 2. The only possibility is that M_1 ends with `kyle` and M_2 ends with `matt` (`m` is two letters after `k`). Kyle will be killed, and Matt will receive the warning.