Cryptographie

Introduction

Le principe du codage est de transformer/coder des données de manière à pouvoir ensuite les retrouver/décoder. Le but du codage est la confidentialité d'un message de manière à ce qu'il ne soit compréhensible/décodable que par son destinataire.

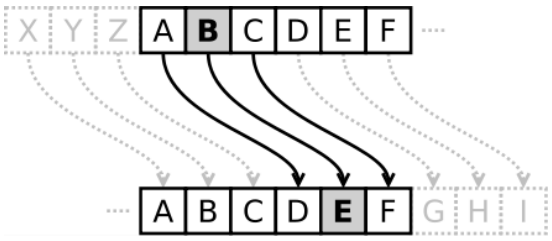
1. Le codage de César

En cryptographie, le **chiffrement par décalage**, aussi connu comme le **chiffre de César** ou le **code de César**, est une méthode de chiffrement très simple utilisée par Jules César dans ses correspondances secrètes (ce qui explique le nom « chiffre de César »).

Codage :

Le texte chiffré s'obtient en remplaçant chaque lettre du texte clair original par une lettre à distance fixe, toujours du même côté, dans l'ordre de l'alphabet. Pour les dernières lettres (dans le cas d'un décalage à droite), on reprend au début.

Par exemple avec un décalage de 3 vers la droite, A est remplacé par D, B devient E, et ainsi jusqu'à W qui devient Z, puis X devient A etc. Il s'agit d'une permutation circulaire de l'alphabet.



La longueur du décalage, 3 dans l'exemple évoqué, constitue la **clé du chiffrement** qu'il suffit de transmettre au destinataire (s'il sait déjà qu'il s'agit d'un chiffrement de César) pour que celui-ci puisse déchiffrer le message.

Dans le cas de l'alphabet latin, le chiffre de César n'a que 26 clés possibles (y compris la clé nulle, qui ne modifie pas le texte).

Exemple :

Soit le message suivant :

M = "LE TRESOR EST CACHE DANS LE CHATEAU DE: SUITE AU PROCHAIN MESSAGE!"

Alors le codage de César de M est :

MC = "OH WUHVRU HVW FDFKH GDQV OH FKDWHDX GH: VXLWH DX SURFKDLQ PHVVDJH!"

Décodage

Pour déchiffrer, si on connaît la clé, on peut facilement décoder, on fait tout simplement l'inverse.

Il suffit de décaler de -3 : C devient A, etc.

Exercice 1.

Q1. Ecrire une fonction **cesar(m,d)** qui décale d'un entier **d** les caractères d'un mot **m** écrit avec l'alphabet "ABCDEFGHIJKLMNOPQRSTUVWXYZ". Laisser inchangés les caractères hors de l'alphabet (espaces, ponctuation, minuscules etc.).

Q2. Ecrire une fonction **decesar(m,d)** qui décode le mot **m** avec la clé **d**.

Attaque du code de César

Une attaque consiste, à partir d'un message codé, à déterminer la clé et le message *décodé*. On dit aussi *casser le code*, et on parle de *cryptanalyse*.

Un codage de César est facile à casser : en français la lettre la plus fréquente est le E, il suffit donc de déterminer la lettre la plus fréquente dans le message codé pour connaître avec une forte probabilité la clé du code.

Exemple :

Dans le message codé "SL AYLZVY LZA JHJOL KHUZ SL JOHALHB KL: ZBPAL HB WYVJOHPU TLZZHNL!"

La lettre la plus fréquente est le L, qui a pour rang 12 dans l'alphabet, et E a pour rang 5, donc la clé est

12 - 5 = 7.

On vérifie que le décodage donne le message "LE TRESOR ...".

Q3. Ecrire une fonction **cassecesar(m)** qui décode un mot **m** codé avec un code de César dont on ignore la clé. Tester avec ces mots : OVRA, NIRP IBHF, IREPVATRGBEVK NHENVG CRHG-RGER TNTAR!

2. Le code de Vigenère

Principe

Toujours décaler les lettres, mais au lieu de décaler toutes les lettres du texte de la même manière, on utilise un texte clé qui donne une suite de décalages de manière variable et périodique.

Codage :

La clé est une chaîne de caractères, pour coder, on décale le $i^{ème}$ caractère du message du rang dans l'alphabet de la $i^{ème}$ lettre de la clé. Quand on arrive au bout de la clé on repart du début.

Exemple :

Le message à coder est « CETTE PHRASE NE VEUT RIEN DIRE ».

Associer à chacune des 26 lettres des entiers de 0 à 25.

A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Choisir une clé de codage constitué de **k** nombres entiers (n_1, n_2, …, n_k) compris entre 0 et 25.

Par exemple, la clé « DBFC » constitué de 4 nombres entiers (3, 1, 5, 2).

Découper la phrase à coder en bloc de 4 lettres « CETT EPHR ASEN EVEU TRIE NDIR E »

Remarque : Les espaces sont purement indicatifs. Dans la première phrase, l'espace sépare les mots, alors que dans la deuxième, elle sépare les blocs.

Le chiffrement consiste maintenant à effectuer un chiffrement de type *César* (décalage), mais qui dépend de la position de chaque lettre dans le bloc :

- Un décalage de n_1 positions pour la $1^{ère}$ lettre de chaque bloc.
- Un décalage de n_2 positions pour la $2^{ème}$ lettre de chaque bloc.
- …
- Un décalage de n_k positions pour la $k^{ème}$ (dernière) lettre de chaque bloc.

Exemple :
Chiffrons le bloc CETT avec la clé (3, 1, 5, 2)
↳ Un décalage de 3 lettres pour C donne F.
↳ Un décalage de 1 lettre pour E donne F.
↳ Un décalage de 5 lettres pour T donne Y.
↳ Un décalage de 2 lettres pour T donne V.

⇨ CETT devient FFYV

Remarque :
Dans le bloc initial «CETT » ; les deux lettres 'T' ne sont pas cryptées par la même lettre.
Dans le bloc crypté « FFYV » ; les deux lettres 'F' ne cryptent pas la même lettre.

Enfin, Pour crypter le message entier, il suffit ensuite de répéter l'opération pour tous les blocs.

Décodage
Pour décoder, le principe est le même avec des décalages négatifs.

Exercice 2.
Dans la suite, l'alphabet utilisé est "ABCDEFGHIJKLMNOPQRSTUVWXYZ".
Q1. Ecrire une fonction **rang(c)** qui rend le **rang ≥ 0** d'un caractère **c** dans l'alphabet. Ecrire la fonction réciproque **lettre(r)**.
Q2. Ecrire une fonction **vigenere(m,cle)** qui code un mot **m** avec la clé **cle** selon la méthode de *Vigenère*.
Q3. Ecrire une fonction **devigenere(m,cle)** qui décode le mot **m** avec la clé **cle**.