TP 10

Cryptographie

Ce sujet est librement inspiré de l'épreuve écrite d'informatique posée à l'Ecole Polytechnique en 2008 dans la filière MP.

On cherche à crypter (c'est-à-dire à coder) un message donné en Français, composé de caractères (les 26 lettres de l'alphabet) en minuscules, non accentués et sans aucun espace entre les mots. Le message est donné sous la forme d'une liste `L` de lettres minuscules. Pour illustrer ce qui suit, nous utiliserons le message *cestsupermaple* représenté par la liste `L` suivante de longueur 14:

```
> L:=[c,e,s,t,s,u,p,e,r,m,a,p,l,e];
```

L:=[c,e,s,t,s,u,p,e,r,m,a,p,l,e]

Comme il est plus facile de travailler sur des nombres que sur des lettres, nous commencerons par transformer les messages en nombres.

Remarque. Pour éviter toute confusion entre les caractères de l'alphabet et les variables locales, on utilisera des lettres majuscules pour nommer les variables locales.

1 Chiffrage et déchiffrage d'un message

On numérote les lettres dans l'ordre lexicographique (en commençant à compter à partir de 0): la lettre a est remplacée par 0, b par 1, et ainsi de suite jusqu'à z remplacée par le nombre 25. Un message `L` sera donc transformé en une liste de nombres `C` :

Liste `L`	c	e	s	t	s	u	p	e	r	m	a	p	l	e
Liste `C`	2	4	18	19	18	20	15	4	17	12	0	15	11	4
Indices dans `L` et `C`	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Cette opération s'appelle le *chiffrage*. En utilisant la correspondance réciproque (remplacer le nombre $\in \{0, \ldots, 25\}$ par la lettre corrspondante), on peut effectuer l'opération inverse: transformer une liste de nombres `C` en un message `L`. C'est l'opération de *déchiffrage*.

Exercice 1

1. Ecrire une procédure `chiffrage(L)` qui, étant donné un message `L` (représenté sous forme d'une liste), retourne le chiffrage `C`.
2. Ecrire une procédure `dechiffrage(C)` qui, étant donné une liste de nombres $\in \{0, \ldots, 25\}$ `C`, retourne le déchiffrage `L` de `C`.

2 Le codage de César

Le codage de César est le plus rudimentaire que l'on puisse imaginer. Il a été utiliser par Jules César (et même auparavant...) pour certaines de ses correspondances. Le principe est de décaler les lettres de l'alphabet vers la droite de une ou plusieurs positions. Par exemple, en décalant les lettres d'une position, le caractère a se transforme en b, le b en c, ..., et le z en a. Le message *cestsupermaple* est donc codé par $dftutvqfsnbqmf$.

2.1 Codage et décodage d'un message connaissant la clé

Exercice 2

1. Ecrire une procédure `codagecesar(L,d)` prenant en entrée un message `L` et un décalage d et qui retourne le message `L` décalé de d lettres (on utilisera les procédures `chiffrage` et `dechiffrage`).
2. Ecrire une procédure `decodagecesar(L,d)` prenant en entée un message `L` et un décalage d et qui réalise le décodage du message `L` (on utilisera ici aussi les procédures `chiffrage` et `dechiffrage`).

2.2 "Casser" un code de César

Pour réaliser le décodage d'un message codé par le codage de César, il faut connaître la valeur du décalage, appelée *clé* du codage. Il existe des méthodes pour retrouver cette clé et ainsi pouvoir déchiffrer les messages cryptés: on dit alors qu'on a *cassé* le code. Une manière de la déterminer est d'essayer de deviner cette valeur en observant la fréquence d'apparition de chaque lettre de l'alphabet dans le message crypté. En effet, la lettre la plus fréquente dans un texte suffisamment long en français est le e.

Exercice 3

1. Ecrire une procédure `frequences(C)` qui prend en entrée une liste `C` représentant le chiffrage du message codé et qui retourne la liste `L` de longueur 26 telle que `L[i]` est égale au nombre d'apparitions du nombre $i-1$ dans la liste chiffrée `C`, pour $1 \leq i \leq 26$.
2. Ecrire une procédure `clecodage(C)` qui prend en entrée une liste `C` représentant le chiffrage du message codé et qui retourne la valeur de la clé, c'est-à-dire la valeur du décalage d utilisée pour coder (on utilisera la procédure `frequences`).
3. Ecrire une procédure `decodagecesarsanscle(C)` qui prend en argument un message codé `C` et qui retourne le message d'origine (on utilisera les procédures `chiffrage`, `clecodage` et `decodagecesar`).

3 Le codage de Vigenère

Au XVI^e siècle, Blaise de Vigenère a modernisé le codage de César très peu résistant de la manière suivante. Au lieu de décaler toutes les lettres du texte de la même manière, on utilise un texte clé qui donne une suite de décalages. Prenons par exemple la clé *concours*. Pour coder un texte, on code la première lettre en utilisant le décalage qui envoie le a sur le c (la première lettre de la clé). Pour la deuxième lettre, on prend le décalage qui envoie le a sur le o (la seconde lettre de la clé) et ainsi de suite. Pour la huitième lettre, on utilise le décalage de a sur s, puis, pour la neuvième, on reprend la clé à partir de la première lettre. Sur l'exemple `ecolepolytechnique` avec la clé `concours`, on obtient: (la première ligne donne le texte, la seconde le texte crypté et la troisième la lettre de la clé utilisée pour le décalage)

e	c	o	l	e	p	o	l	y	t	e	c	h	n	i	q	u	e
g	q	b	n	s	j	f	d	a	h	r	e	v	h	z	i	w	s
c	o	n	c	o	u	r	s	c	o	n	c	o	u	r	s	c	o

Exercice 4

1. Ecrire une procédure `codagevigenere(L,cle)` prenant comme arguments le message `L` à coder et la clé choisie sous la forme d'une liste `cle` et retournant le message codé (on utilisera les procédures `chiffrage` et `dechiffrage`).
2. Ecrire une procédure `decodagevigenere(L,cle)` prenant comme arguments le message `L` à décoder et la clé sous la forme d'une liste `cle` et retournant le message décodé (on utilisera ici aussi les procédures `chiffrage` et `dechiffrage`).

Ce code très performant n'a été cassé que trois siècles plus tard (la difficulté est de trouver la longueur de la clé). On pourra consulter l'épreuve écrite d'informatique posée au concours d'admission 2008 de l'Ecole Polytechnique en MP où il est présenté une méthode pour décrypter le chiffre de Vigenère, inventée par Babagge et Kasiski.