Aix*Marseille université

Domaine Sciences et Technologies
Portail 1 René Descartes

Introduction à l'informatique : TD 3
Code UE : S14IN1I1

Année 2021–2022

Premiers algorithmes et cryptologie

La cryptologie (étym. science du secret) est un domaine à la frontière des mathématiques et de l'informatique. Elle se sépare en deux pans de même importance. Le premier consiste à transformer une information afin de la rendre secrète, autrement dit à la "crypter" ou "chiffrer". Il s'agit de la *cryptographie* (étym. écriture secrète). Le second consiste à analyser les informations cryptées et trouver des méthodes et techniques afin d'en dévoiler le sens caché. Il s'agit de la *cryptanalyse*. Dans ce TD, nous allons nous intéresser à quelques procédés historiques simples de cryptologie, ce qui permettra d'écrire quelques algorithmes simples.

Exercice 1 (Codage de César)

Historiquement, un procédé de cryptographie bien connu est le codage de César que Jules César utilisait dans ses correspondances. Le principe de chiffrement est simple. Étant donné un alphabet (ici, nous utiliserons l'alphabet latin) et un message, le message chiffré s'obtient en remplaçant chacune des lettres du message d'origine par une lettre à distance fixe toujours dans la même direction. Pour les dernières lettres, dans le cas d'une distance à droite, on reprend au début de l'alphabet. Il s'agit d'un chiffrement par décalage. À titre d'exemple, avec un décalage de 5, 'a' devient 'f', 'b' devient 'g', ..., 'y' devient 'd' et 'z' devient 'e'.

1. Soit le message "La vie est un long fleuve tranquille". Donnez ses représentations chiffrées selon le codage de César avec les clés 3 et -7.
2. En utilisant des phrases (pas du pseudo-code), donnez une description précise de l'algorithme de chiffrement utilisé pour le codage de César.
3. Une façon plus formelle de décrire l'algorithme est d'utiliser le pseudo-code vu en cours. Pour cela, on stocke les chaînes de caractères (le message à chiffrer et le texte chiffré par exemple) dans un tableau. Par exemple, le tableau de caractères

'a'	'l'	'g'	'o'	'r'	'i'	't'	'h'	'm'	'e'

   permet de représenter la chaîne de caractères « algorithme ». Dans le codage de César, on code les caractères ('a', 'b', ..., 'z') avec des entiers. Supposons ici que la lettre 'a' est codée par l'entier 0, la lettre 'b' par l'entier 1, etc. On se donne alors une fonction `code(c: caractère)` permettant de connaître le code d'un caractère `c` et une fonction `caractère(p: entier)` renvoyant le caractère dont le code est `p` (seulement s'il est entre 0 et 25). On a par exemple

   ```
   code('g') = 6        caractère(12) = 'm'        caractère(code('x')) = 'x'
   ```

   À l'aide de ces deux fonctions, on peut donc écrire un algorithme procédant au chiffrement d'un message :

   ```
   fonction chiffrement_Cesar(message: tableau de caractères, clé:
       entier):
       n := longueur(message)
       chiffré := tableau vide de longueur n
       Pour i de 0 à n−1 faire
           p := code(message[i])
           p_décalé := (p + clé) mod 26     # on décale puis on revient
                                            # dans l'intervalle [0,25]
           chiffré[i] := caractère(p_décalé)
       FinPour
       retourner(chiffré)
   ```

   Quelles en sont les entrées et sorties de cet algorithme ?

4. Exécuter l'algorithme précédent pour trouver ce que retourne l'appel `chiffrement_Cesar(['i','n','t','o','x'], 3)`. Pour exécuter la boucle `Pour`, on utilisera le tableau suivant :

i	message[i]	p	p_décalé	chiffré
avant la boucle				[, , , ,]
0	'i'	8	??	??
1	??	??	??	??
2	??	??	??	??
⋮				

5. L'algorithme précédent ne prend pas en charge les espaces auxquels on ne souhaite pas apporter de décalage. Ainsi, on aimerait que `chiffrement_Cesar(['u','n',' ','d','e','u','x'], 3)` renvoie le tableau `['x', 'q', ' ', 'g', 'h', 'x', 'a']`. Modifier l'algorithme précédent pour y parvenir.
6. Quelle est la complexité de l'algorithme de chiffrement de César ? On demande de compter le nombre d'opérations élémentaires effectuées par l'algorithme en fonction du nombre n de caractères d'un message donné. Après un calcul précis, essayer d'écrire ce nombre comme $\mathcal{O}(...)$.
7. Proposez un algorithme de déchiffrement, prenant en entrée le message chiffré et une clé et renvoyant le message décodé. À titre d'exemple, déchiffrez le message « kajex »sachant que la clé de chiffrement vaut 9.
8. Admettons que quelqu'un vous envoie un message chiffré en vous spécifiant qu'il s'agit d'un codage de César mais sans vous donner la clé. Est-il possible de le déchiffrer ? Si oui, comment et est-ce efficace ?

Exercice 2 (Codage spartiate)

Dans l'armée de Spartes (autour du Vème siècle avant J.-C.), les militaires étaient parfois amenés à se transmettre des messages chiffrés. Pour ce faire, ils utilisaient un bâton, appelé bâton de Plutarque (ou scytale). L'émetteur du message prenait une fine lanière de tissu ne pouvant contenir qu'une seule lettre dans sa largeur, l'enroulait en spirale autour d'un bâton, et écrivait son message ligne par ligne dans la longueur du bâton de façon à avoir toutes les lignes remplies sauf éventuellement la dernière ligne :

Une fois déroulée, la lanière contenait le message chiffré. Pour déchiffrer ce message, le récepteur devait posséder un bâton de même diamètre (ayant le même nombre de circonvolutions) que celui de l'émetteur. Ce type de codage est un chiffrement par transposition.

À titre d'exemple, considérons le message suivant : « La vie c'est comme une boîte de chocolats, on ne sait jamais sur quoi on va tomber. » (*Forrest Gump*, de Robert Zemeckis) Considérons un bâton dont le périmètre permet 6 circonvolutions (*nous dirons pour simplifier que la clé du chiffrement est* 6). Le message chiffré est alors « Lo␣␣moamdoan␣meni␣ve␣␣svi␣cn␣aeuhes␣␣no␣utcecsro'␣oa␣mebliqbsoatuetît␣or␣tsji.ce,a␣␣ »

1. Chiffrez la phrase « Les cons ça ose tout. » (*Les tontons flingueurs*, Michel Audiard) avec la clé 4 en vous aidant d'un ruban de papier que vous enroulerez autour d'un stylo par exemple. (C'est plus facile à faire à deux...)
2. Proposez une méthode décrivant ce procédé de chiffrement. En particulier, comment déterminer le nombre de colonnes à utiliser sur le bâton ? Illustrez votre méthode en chiffrant la phrase de la question précédente avec les clés 3 et 5.

3. Étant donné un message chiffré m de longueur n et une clé c, donnez un algorithme permettant de découvrir le message d'origine.
4. Déchiffrez les messages suivants :
   — "Ce'e'␣oceànos␣␣ntçln␣aeam␣siêq␣tmur." avec $c = 4$;
   — "Càq␣s'␣up␣emea ?so␣r␣titl␣␣␣ue␣" avec $c = 5$;
   — "L,␣snru␣s␣␣ekjutp.eeioè␣" avec $c = 6$.
5. À présent, vous recevez un message chiffré m sans la clé de chiffrement. Donnez un algorithme permettant de retrouver le message d'origine et évaluez-en l'efficacité en termes de temps.

Exercice 3 (Codage de Vigenère)

Revenons sur le codage de César. Le principal point faible de ce codage par décalage provient justement du type de décalage qui est identique quelle que soit la position de la lettre décalée dans le message d'origine.

1. Comment pourrait-on transformer le codage de César pour qu'il soit plus difficile à casser ?

L'objectif du codage de Vigenère est justement de remédier à ce défaut. Il a été décrit pour la première fois au XVIème siècle C'est également un chiffrement par décalage, mais la substitution est cette fois-ci poly-alphabétique. Cela signifie qu'une même lettre dans le message d'origine peut, selon sa position dans ce dernier, être remplacée par différentes lettres dans le message chiffré.

Plus précisément, ce nouveau chiffrement va utiliser une notion de clé différente des précédents. Ici, la clé va être une suite de caractères, qui prend généralement la forme d'un mot ou d'une phrase. Pour procéder au chiffrement, il faut parcourir le message d'origine lettre après lettre tout en parcourant circulairement les lettres de la clé pour effectuer la substitution. Bien sûr, plus la clé est longue et variée, plus le chiffrement est solide. La substitution à opérer est donnée par la table illustrée en Figure 1 page 4, appelée table de Vigenère, dans laquelle la première ligne correspond aux lettres du message d'origine et la première colonne à celles de la clé.

Ainsi, on remplace chaque lettre ℓ du message m d'origine par celle contenue dans la case (k, ℓ), avec k la lettre de la clé c correspondant à la position de ℓ dans m, modulo $|c|$.

À titre d'exemple, le message "Peu lui importe de quoi demain sera fait" associé à la clé "petit frère" sera chiffré de la manière suivante (en supprimant les accents) :

```
Message d'origine : Peu lui importe de quoi demain sera fait
Clé               : pet itf rerepet it frer epetit frer epet
                    ||ligne 't' colonne 'u' -> 'u' devient 'n'
                    |ligne 'e' colonne 'e' -> 'e' devient 'i'
                    ligne 'p' colonne 'p' -> 'P' devient 'E'
```

ce qui mène au message chiffré "Ein tnn zqgsgxx lx vlsz htqtqg xvvr jpmm".

2. Soit le message d'origine "Chacun voit sa voie de toi à moi". Chiffrez ce message avec les deux clés suivantes : "ntm" et "de personne je ne serai la cible". Qu'observez-vous ?
3. Étant donné un message d'origine m et une clé c, donnez un algorithme de chiffrement de m selon c.
4. Étant donné un message chiffré m et une clé c, donnez un algorithme de déchiffrement de m selon c. Continuez en déchiffrant le message "Diepe dorr n'owg naj k vrnnvr" avec la clé "keny arkana".
5. Expliquez les processus de chiffrement et déchiffrement plus formellement, c'est-à-dire mathématiquement.

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

FIGURE 1 – Table de Vigenère.