Université Claude Bernard Lyon 1
Institut de Science financière et d'Assurances (ISFA)

50 avenue Tony Garnier
69007 Lyon, France

Master 1 informatique
Université Claude Bernard Lyon 1

Cryptologie : TP n° 1

Le but de ce TP est de cryptanalyser des systèmes de chiffrements historiques, qui succombent à des attaques par analyse de fréquences.

Pour cela, vous allez devoir programmer les algorithmes de chiffrement et de déchiffrement, et des outils de cryptanalyse. Ces derniers devront être les plus interactifs et les plus visuels possibles.

Vous trouverez en annexe des tableaux sur les fréquences des lettres (ou de groupe de lettres) en français qui pourront vous être utiles.

L'énoncé du TD est disponible à cette adresse :
`http://perso.ens-lyon.fr/fabien.laguillaumie/teaching/M1_TP_1_crypto.pdf`.

Vous allez devoir opérer comme des cryptanalystes pour retrouver des messages interceptés chiffrés. En particulier, vous devez récupérer les fichiers dans sous format traitable, traiter des chaînes de caractères, compter des occurrences de lettres, etc...

Exercice 1 (Chiffrement de César)**.** Décrypter les textes suivants qui ont été obtenus en appliquant le chiffrement de César sur un texte en langue française et dans lequel les espaces ont été supprimées :

1. `vcfgrwqwfsbhfsntowbsobgfsbhfsnqvsnjcigsghqsoixcifrviwtshseicwb`
   `sgojsnjcigdogeisjcigoihfsgofhwgobgjcigbsrsjsnqwfqizsfrobgzsgfi`
   `sgzsgxcifgcijfopzsgeiojsqzsggwubsgrsjchfsdfctsggwcbdofzseiszsg`
   `hhcbashwsf`

2. `hcihszouoizssghrwjwgsssbhfcwgdofhwsgrcbhzibssghvopwhssdofzsgps`
   `zusgzoihfsdofzsgoeiwhowbgzohfcwgwsasdofqsileiwrobgzsifzobuisgs`
   `bcaasbhqszhsgshrobgzobchfsuoizcwg`

Exercice 2 (Chiffrement par transposition - scytale)**.** Décryptez le texte suivant qui a été obtenu en appliquant un chiffrement par transposition à l'aide d'une scytale.

`lelnracsrunanatuvllerrmcnjeeaeseiaetanctsagsgeemftqdnnentrarau`
`eneciliianeredofaesntdneenignpcdaishdcaoaeenede`

Exercice 3 (Chiffrement affine)**.** Le chiffrement affine est un système de chiffrement par subsitution mono-alphabétique, dont la clé est un couple d'entier $(a, b) \in (\mathbb{Z}/26\mathbb{Z})^\star \times \mathbb{Z}/26\mathbb{Z}$. Une lettre du texte clair $i \in \{0, \ldots, 25\}$ est alors remplacée dans le texte chiffré par la lettre $a \cdot i + b \pmod{26}$. L'application $x \mapsto ax + b \pmod{26}$ est une permutation car a est inversible modulo dans $\mathbb{Z}/26\mathbb{Z}$ (par définition).

Décrypter les textes suivants qui ont été obtenus en appliquant un chiffrement affine sur un texte en langue française et dans lequel les espaces ont été supprimées :

1. `ntjmpumgxpqtstgqpgtxpnchumtputgfsftgthnngxnchumwxootrtumhpyctg`
   `ktjqtjchfooxujqhgztumxpotjxotfoqtohrxumhzutwftgtopfmntjmpuatmf`
   `mshodpfrxpjjtqtghbxuj`

2. `spaxhnnvjupkytpcpuppycxklppygpkpycxkyejkpapzzphktvkkjppyrjpsxh`
   `zppyrjhzxhzppyrjhkpnytwxrjpavhnkhyvjyxaxhyoxzpzpkhyvjyxaxhyjkp`
   `xjycppyzxhzppyzptvzgcpke`

Exercice 4 (Chiffrement par substitution poly-alphabétique)**.** Le chiffrement de Vigénère est un système de substitution poly-alphabétique élaboré par B. de Vigenère en 1586. Ce procédé de chiffrement repose sur l'utilisation périodique de plusieurs alphabets de subsitution déterminés par la clé (en général, un mot). Pour pouvoir chiffrer un texte clair, à chaque caractère est associée une lettre de la clé pour effectuer le décalage correspondant comme dans le chiffrement de César.

La principale difficulté pour attaquer ce chiffrement est de trouver la longueur de la clé. Pour cela, on peut utiliser une méthode connue sous le nom de test de Kasiski : elle repose sur le fait que si deux groupes de lettres (polygrammes) du chiffré sont égaux, alors il s'agit probablement du même polygramme dans le texte clair chiffré avec la même partie de la clé. La taille de l'intervalle qui sépare ces deux polygrammes identiques dans le chiffré sera donc, dans la majorité des cas, un multiple de la longueur de la clé.

Decryptez les textes suivants obtenus par un chiffrement de Vigenère.

1. `zbpuevpuqsdlzgllksousvpasfpddggaqwptdgptzweemqzrdjtddefekeferd`
   `prrcyndgluaowcnbptzzzrbvpssfpashpncotemhaeqrferdlrlwwertlussfi`
   `kgoeuswotfdgqsyasrlnrzppdhtticfrciwurhcezrpmhtpuwiyenamrdbzyzw`
   `elzucamrptzqseqcfgdrfrhrpatsepzgfnaffisbpvblisrplzgnemswaqoxp!d`
   `seehbeeksdptdttqsdddgxurwnidbdddplncsd`

2. `iefomntuohenwfwsjbsfftpgsnmhzsbbizaomosiuxycqaelrwsklqzekjvwsi`
   `vijmhuvasmvwjewlzgubzlavclhgmuhwhakookakkgmrelgeefvwjelksedtyh`
   `sgghbamiyweeljcemxsohlnzujagkshakawwdxzcmvkhuwswlqwtmlshojbsgu`
   `elgsumlijsmlbsixuhsdbysdaolfatxzofstszwryhwjenuhgukwzmshbagigz`
   `zgnzhzsbtzhalelosmlasjdttqzeswwwrklfguzl`

Exercice 5 (Chiffrement par subsitution mono-alphabétique)**.** Décryptez le texte suivant.

```
v ubcfb osu ymoqsuu n cxqfj dqmfnu ub vjcfqu juz amqjmruz zmsscfusb bquflu auoquz hfszbms zwfba
ju wusbms qusbqu ncsz ju vmo z uddmqvcfb n uxfbuq ju xusb wcoxcfz fj eczzc qcefnuwusb jc emqbu
xfbquu no ijmv nuz wcfzmsz nu jc xfvbmfqu ecz czzul qcefnuwusb vueusncsb emoq uweuvauq kou z
usrmoddqu us wuwu buwez kou jof os bmoqifjjms nu emozzfuqu ub nu zciju

ju acjj zusbcfb ju vamo vofb ub ju xfuog bcefz c j osu nu zuz ugbquwfbuz osu cddfvau nu vmojuoq
bqme xczbu emoq vu nuejmfuwusb fsbuqfuoq ubcfb vjmouu co woq ujju quequzusbcfb zfwejuwusb os
usmqwu xfzcru jcqru nu ejoz n os wubqu ju xfzcru n os amwwu n usxfqms kocqcsbu vfsk csz c j
uecfzzu wmozbcvau smfqu cog bqcfbz cvvusbouz ub iucog

hfszbms zu nfqfruc xuqz j uzvcjfuq fj ubcfb fsobfju n uzzcpuq nu equsnqu j czvuszuoq wuwu cox
wufjjuoquz uemkouz fj dmsvbfmsscfb qcquwusb cvboujjuwusb n cfjjuoqz ju vmoqcsb ujuvbqfkou ubcfb
```

vmoeu ncsz jc ymoqsuu v ubcfb osu nuz wuzoquz n uvmsmwfu eqfzuz us xou nu jc zuwcfsu nu jc acfsu zms ceecqbuwusb ubcfb co zuebfuwu hfszbms kof cxcfb bqusbu suod csz ub zmoddqcfb n os ojvuqu xcqfkouog co nuzzoz nu jc vauxfjju nqmfbu wmsbcfb jusbuwusb fj z cqqubc ejozfuoqz dmfz us vauwfs emoq zu quemzuq c vackou ecjfuq zoq osu cddfvau vmjjuu co woq dcvu c jc vcru nu j czvuszuoq j usmqwu xfzcru xmoz dfgcfb no qurcqn v ubcfb os nu vuz emqbqcfbz cqqcsruz nu bujju zmqbu kou juz puog zuwijusb zofxqu vujof kof eczzu osu jurusnu zmoz ju emqbqcfb nfzcfb ifr iqmbauq xmoz qurcqnu

c j fsbuqfuoq nu j ceecqbuwusb nu hfszbms osu xmfg zovquu dcfzcfb usbusnqu osu zuqfu nu smwiquz kof cxcfusb bqcfb c jc eqmnovbfms nu jc dmsbu jc xmfg eqmxuscfb n osu ejckou nu wubcj mijmsrou wfqmfq buqsu usvczbqu ncsz ju woq nu nqmfbu hfszbms bmoqsc os imobms ub jc xmfg nfwfsoc nu xmjowu wcfz juz wmbz ubcfusb usvmqu nfzbfsvbz ju zms nu j ceecqufj no bujuvqcs vmwwu ms nfzcfb emoxcfb ubqu czzmoqnf wcfz fj s p cxcfb covos wmpus nu j ubufsnqu vmwejubuwusb hfszbms zu nfqfruc xuqz jc dusubqu fj ubcfb nu zbcboqu dquju ejobmb eubfbu ub zc wcfrquoq ubcfb zmojfrsuu ecq jc vmwifscfzms ijuou osfdmqwu no ecqbf fj cxcfb juz vauxuog bquz ijmsnz ju xfzcru scboqujjuwusb zcsrofs jc euco noqvfu ecq ju zcxms rqmzzfuq juz jcwuz nu qczmfq uwmozzuuz ub ju dqmfn nu j afxuq kof xuscfb nu equsnqu dfs

A Fréquences des lettres en français

Rang	Lettre	Nombre	Pourcentage
3.	a	117110	7,636
19.	b	13822	0,901
12.	c	50003	3,260
11.	d	56269	3,669
1.	e	225947	14,715
18.	f	16351	1,066
20.	g	13288	0,866
21.	h	11298	0,737
4.	i	115465	7,529
22.	j	8351	0,545
32.	k	745	0,049
9.	l	83668	5,456
14.	m	45521	2,968
6.	n	108812	7,095
10.	o	82762	5,378
13.	p	46335	3,021
17.	q	20889	1,362
7.	r	100500	6,553
2.	s	121895	7,948
5.	t	111103	7,244
8.	u	96785	6,311
16.	v	24975	1,628
29.	w	1747	0,114
24.	x	5928	0,387
25.	y	4725	0,308
28.	z	2093	0,136
23.	à	7449	0,486
30.	ç	1306	0,085
26.	è	4160	0,271
15	é	29206	1,904
27.	ê	3445	0,225
36.	ë	7	0,000
33.	î	695	0,045
35.	ï	84	0,006
31.	ù	890	0,058
34.	œ	283	0,018

Bigrammes les plus fréquents (en nombre d'apparition sur 10000 lettres).

es 305	te 163	ou 118	ec 100	eu 89	ep 82
le 246	se 155	ai 117	ti 98	ur 88	nd 80
en 242	et 143	em 113	ce 98	co 87	ns 79
de 215	el 141	it 112	ed 96	ar 86	pa 78
re 209	qu 134	me 104	ie 94	tr 86	us 76
nt 197	an 30	is 103	ra 92	ue 85	sa 75
on 164	ne 124	la 101	in 90	ta 85	ss 73
er 163					

Les bigrammes constitués par deux consonnes les plus fréquents sont :

nt	tr	ns	st
197	86	79	61

Ceux constitués de deux voyelles sont :

ou	ai	ie	eu	ue	ui	au	oi	io
118	117	94	89	85	68	64	52	49

Ceux formés de deux lettres identiques les plus fréquents sont

ss	ee	ll	tt	nn	mm	rr	pp	ff	cc	aa	uu	ii
73	66	29	24	20	17	16	10	8	3	3	2	1