La cryptographie

[Année 2016- 2017]

Noms et Prénoms des élèves, niveaux : CLEMENT Agathe, CHAUVEAU Alice, GORIN Antoine, PIEROTTI David, LOUCHARD Théo, GUERGUER Thimothée, AUBRY Paul, CHEBOUN Jasmine, FRANCHINI Baptiste 1ère S

Établissements : Lycée Marguerite de Navarre et Alain Fournier BOURGES

Enseignant-e-s : PELLETIER Guillaume, HERMINIER Nathalie, CRECHET Olivier, DUFOUR Emmanuelle, PENNETIER Antonin

Chercheur : NGUYEN Benjamin (INSA Centre Val de Loire)

[**Présentation du sujet**]

On cherche à déchiffrer deux messages secrets en sachant simplement qu'il s'agit de textes en français.

1) Déchiffrer les textes.

2) Proposer un programme ou un algorithme pour déchiffrer automatiquement tout type de texte chiffré de ces manières.

Jcq qylejmrq jmleq Bcq tgmjmlq Bc j ysrmklc Zjcqqclr kml amcsp B slc jylescsp Kmlmrmlc Rmsr qsddmaylr Cr zjckc osylb Qmllc j fcspc Hc kc qmstgclq Bcq hmspq ylagclq Cr hc njcspc Cr hc k cl tygq Ys tclr kystygq Osg k cknmprc Bcay bcjy Nypcgj y jy Dcsgjjc kmprc

d1

Tqxw qk wwwd Xueoqurm viphb jz Sfkqi Jz uwb kxahfh Vrmm cg mr gk m fp vszbpmwxp Xn xdyv nxhveq rnukqxvx gwznc wm csxdv Nbyijb jz nvkw wttum u rpges Auj bhomo p cm vppw nj jlunecq Ysh cragm ywkq ket odmsy ymbdzzf tpsv s yuxa Qymdju tyj yvcb Vp bbbi tv fhm wnxq oatuh Hjy lbnbyqyg guxskxn h llce tk oexyl Droyzr zp dlam mjjkc k hfwui Qkz rxece f sc lffm dz zbkdusg

d2

[**Annonce des conjectures et résultats obtenus**]

Pour déchiffrer des textes, nous nous sommes intéressés aux méthodes de chiffrement les plus simples et connues comme le chiffrement de César et de substitution. Nous avons essayé aussi des méthodes plus complexes comme le chiffrement de Vigenère ou affine. Nous avons choisi ces méthodes de chiffrement par déduction car elles étaient les plus simples et les plus connues parmi toutes les méthodes qui existent.

[**Texte de l'article**]

Veiller à la numérotation des paragraphes .
Les dessins et images doivent être légendés et lisibles .
Si possible, fournir les images dans des fichiers séparés.

1) L'historique du cryptage

Au cours des siècles, plusieurs méthodes de cryptage ont été utilisées. La technique du rouleau assyrien qui consiste à faire un anagramme en écrivant un message sur une feuille enroulée autour d'un rouleau d'un certain diamètre et pour lire le message, on enroule la feuille autour d'un rouleau de même diamètre.Il existe aussi le décalage de César ou encore le code de Vigenère que nous verrons dans cette présentation mais également la machine Enigma qui fut utilisée durant la seconde guerre mondiale. D'autres méthodes de cryptage plus robustes existent.

En jaune le papier sur lequel est écrit le message

En bleu un baton dont le diametre est connu de l'emetteur et du destinataire

Lorsque l'on deroule le papier les lettres sont dans le desordre: le message est crypté

Il suffit d'enrouler le papier autour d'un baton de meme diametre pour décrypter le message.

©http://wakaziva.pagesperso-orange.fr/crypto/2.htm

2) Le chiffre de César

César utilisait une méthode de chiffrement qui porte aujourd'hui son nom.

Le chiffre de César est la méthode de cryptographie la plus ancienne communément admise par l'histoire. Il consiste en une substitution mono-alphabétique. Chaque lettre est remplacée, c'est à dire substituée, par une seule et même lettre tout au long du texte, d'où le terme mono-alphabétique, selon un certain décalage dans l'alphabet ou de façon arbitraire. César avait coutume d'utiliser un décalage de 3 lettres : A devient D, B devient E, C devient F, etc. Il écrivait donc son message normalement, puis remplaçait chaque lettre par celle qui lui correspondait :

CLAIR 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

Ici, il y a un décalage de trois lettres.

CODE 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

exemple : d'après cette méthode, "VIVE LES MATHS" devient donc "YLYH OHV PDWKV" !

Nous avons supposé que le texte (d1) (voir ci-dessus) était chiffré avec cette méthode donc nous l'avons utilisée

pour le déchiffrer.

3) Le chiffrement par substitution

Le chiffrement par substitution est un chiffrement utilisé depuis bien longtemps, le chiffrement de César en fait partie. Cette méthode consiste à remplacer une lettre du texte, ou du message, à coder par une autre : exemple A par F. Sans clé particulière. (sans décalage de lettre définie, comme le code de César).

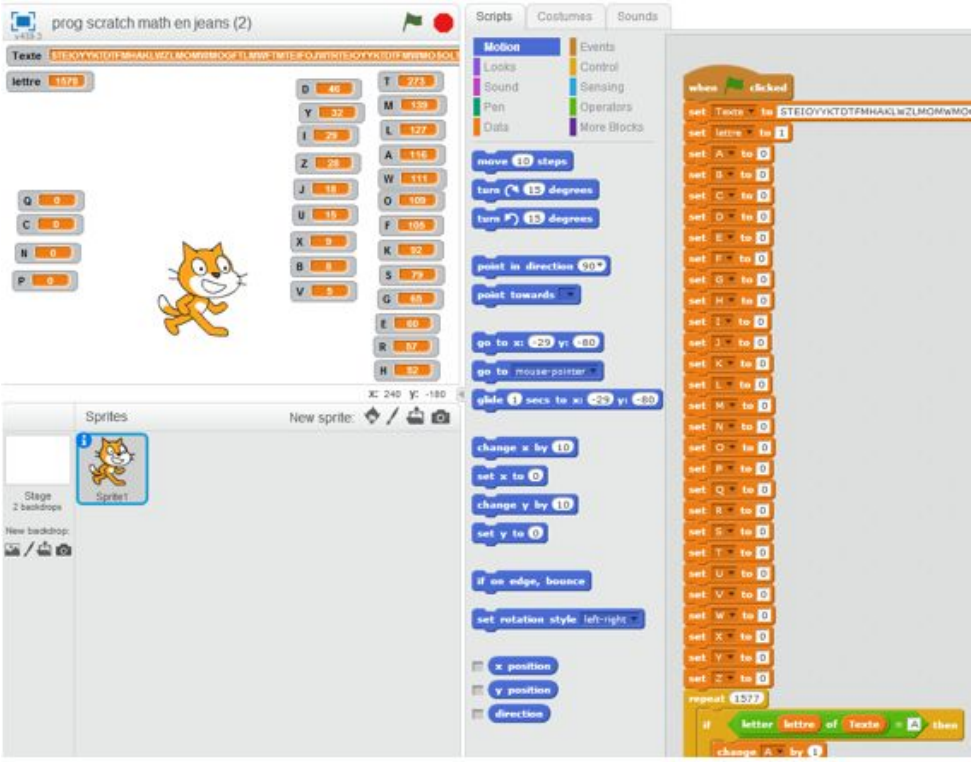
Le chiffrement par substitution est un chiffrement mono-alphabétique, dans lequel une lettre ne peut être remplacée que par une seule autre lettre, et non pas par plusieurs comme dans un chiffrement polyalphabétique. Un exemple ci-dessous de chiffrement mono-alphabétique :

S T E I O Y Y K T D T F M

L E C H I F F R E M E N T

4) Le programme

Pour pouvoir avoir une idée de la fréquence des lettres d'un texte crypté, nous avons réalisé un programme qui permettrait, en entrant le texte, de non seulement compter le nombre de lettres mais aussi de compter chacune des lettres séparément.

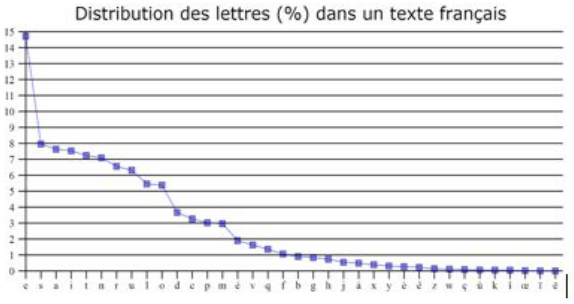


Après avoir eu les résultats de chaque lettre, nous avons recherché les fréquences de chaque lettre dans la langue française.

Grâce à ce résultat, nous avons déterminé des associations.

Par exemple, dans le texte codé T est associé à E dans le texte clair et M à S.

Mais cette méthode reste très limitée car étant donné le nombre restreint de lettres dans notre texte, certaines fréquences dans le tableau ci-dessous étaient proches les unes des autres, et nous n'avons pas pu différencier certaines lettres.



Aussi nous en avons conclu que plus le texte est long, plus la fréquence se rapprochera de la fréquence théorique et plus cette méthode est efficace.

5) Le chiffre de Vigenère

Le chiffre de Vigenère est un système de chiffrement polyalphabétique. C'est un chiffrement par substitution, mais une même lettre du message clair peut être remplacée par des lettres différentes suivant sa position dans le texte. Contrairement à un système de chiffrement mono-alphabétique comme le chiffre de César.

Cette méthode résiste ainsi à l'analyse de fréquences, ce qui est un avantage décisif sur les chiffrements mono-alphabétiques. Cependant le chiffre de Vigenère a été cassé par le major prussien Friedrich Kasiski qui a publié sa méthode en 1863. Il n'offre plus aucune sécurité depuis cette époque.

Pour utiliser le tableau de Vigenère, nous devons prendre des lettres clés pour pouvoir traduire un texte.

Table de Vigenère.

Clé Utilisée \ Lettre en clair	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
26 lettres chiffrées																										
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

Exemple :

Nous avons choisi les lettres : H, E et R
Puis il faut prendre le texte à traduire et reporter chaque lettre en alternant entre les colonnes H, E et R.
CONGRES donne une fois traduit : JSENVVZ
On se place sur la ligne correspondant à la première lettre « C » du mot CONGRES, puis on se déplace jusqu'à la colonne H et nous obtenons la lettre « J ».
Ensuite on doit prendre la lettre suivante « O » et on la projette dans la colonne E pour obtenir un « S ».
Nous devons répéter cette étape pour toutes les lettres du texte à coder en alternant avec les colonnes H, E et R.
Nous pouvons aussi l'utiliser avec toutes les lettres de l'alphabet en faisant : +1, +2, +3 ....
Puis il suffit de décaler de +1, +2, +3 .... chaque lettre du texte à traduire.
Mais le chiffrement de Vigenère est aujourd'hui possible à décrypter mais cette méthode est longue et fastidieuse pour déchiffrer un message.

6) Le codage affine

Le codage affine consiste à remplacer chaque lettre de l'alphabet A, B...Z par son rang entre 0 et 25 (A=0 jusqu'à Z=25) grâce à une fonction affine. On choisit deux nombres entiers a et b compris entre 0 et 25. On nomme x le rang de la lettre et r(x) le reste de la division euclidienne de y=ax+b par 26. R(x) est alors le rang codé de la lettre. Chaque lettre est toujours codée par la même lettre ce qui signifie que c'est un chiffrement par substitution mono-alphabétique.

Lettre	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
Rang	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

Nous avons essayé cette méthode de déchiffrement pour l'un de nos textes. Nous avons pris plusieurs possibilités pour a et b. Après plusieurs essais, nous nous sommes rendus compte que si le **"a"** choisi au départ

n’est pas premier avec 26 plusieurs lettres peuvent être codées par la même lettre donc il faut choisir une valeur de **“a”** qui n’est pas dans la table de 26 et donc qui est première avec 26.

7) Les difficultés

Nous avons tout d'abord tenté de décrypter les messages à l'aide des méthodes classiques de substitution mono-alphabétique telles que le chiffrement affine. Cette dernière n'étant pas celle utilisée, nous nous sommes dirigés vers l'analyse fréquentielle . Cette analyse ne permettant pas de décoder les messages, , nous avons dû essayer une autre méthode. Par la suite,après de nombreux essais avec différentes clés, nous avons trouvé que la méthode de César était celle utilisée dans le texte d1. Nous avons donc supposé que la méthode était probablement la même utilisée pour le texte d2. Seulement, la présence de groupes de lettres surprenants tels que « wwwd » a causé le rejet de l'hypothèse puisque dans la langue française un mot commençant par trois lettres identiques n'existe pas. Nous avons alors pu en déduire qu'une lettre ne code pas une autre lettre spécifique dans l'alphabet et donc que la substitution mono-alphabétique n'est pas le moyen de cryptage employée dans le texte d2.
Nous avons finalement supposé que la méthode à adopter était le décryptage par la substitution polyalphabétique.

8) Les résultats

Pour le premier texte nous avons cherché à décrypter le texte avec des méthodes simples, en effet comme c'était le premier texte nous avons pensé qu'il devait être codé d'une manière classique. Après quelques essais nous avons enfin trouvé la clef, le texte était crypté à l’aide du chiffre de César**.**

Clair	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
Chiffré	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

Ici le décalage est de 3, c'est la clef de César « classique ». Pour notre texte le décalage était de 2. Avant de trouver cette clef nous avons essayé +1 et -1, la solution paraît simple mais lorsque nous avons vu le texte pour la première nous avons été très surpris. Pour le seconde texte les choses se sont compliquées.

Clair	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
Chiffré	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

Ainsi ce texte d1 est décryptable en appliquant une clef de +2

Jcq qylejmrq jmleq
Bcq tgmjmlq
Bc j ysrmklc
Zjcqqclr kml amcsp
B slc jylescsp
Kmlmrmlc

Les sanglots longs
Des violons
De l'automne
Blessent mon cœur
D'une langueur
Monotone.

Pour commencer le décryptage du texte d2 nous avons d'abord essayé la clef de César, sans résultat. Au bout de plusieurs tentatives et de nombreuses heures de réflexion nous sommes arrivés à trouver la clef. Certains groupes de lettres suspects nous ont mis sur la piste :

- **WWWD**
- **BBBI**

Ne connaissant aucun mot dans la langue française comportant trois lettres identiques à la suite, nous en avons déduit que le chiffrement était **polyalphabétique c'est à dire qu'une lettre peut être cryptée par plusieurs lettres, ici toutes.**

Pour le second texte la clef était : -1 ;-2 ;-3...-26

Après plusieurs essais, nous avons trouvé par hasard cette clef.
Lorsqu'on arrive à la 26^{e} lettre du texte la clef est relancée, c'est à dire que l'on redémarre à -1, ainsi de suite jusqu'à ce que tout le texte soit décrypté. Ici le texte était encore un poème : « Le pont Mirabeau » de Guillaume Apollinaire.

[Conclusion]

Le premier texte à décrypter correspondait à « Chanson d'Automne » de Paul Verlaine, le second plus compliqué était « Le pont Mirabeau » de Guillaume Apollinaire.

Pour conclure, après avoir étudié quelques types de cryptographies comme le chiffrement par substitution ou la méthode de chiffrement affine et après avoir trouvé leur fonctionnement, nous avons tenté de l'appliquer à nos textes que nous devions déchiffrer. Aujourd'hui encore il y a des chiffrements qui résistent aux spécialistes comme le chiffrement quantique et même le chiffrement RSA actuellement utilisé.

Ce fut une bonne expérience qui nous a permis de voir les mathématiques sous un autre angle que ce soit les séances, les réunions ou le congrès à Paris.