



Cryptographie

Math Bac Info: partie 3

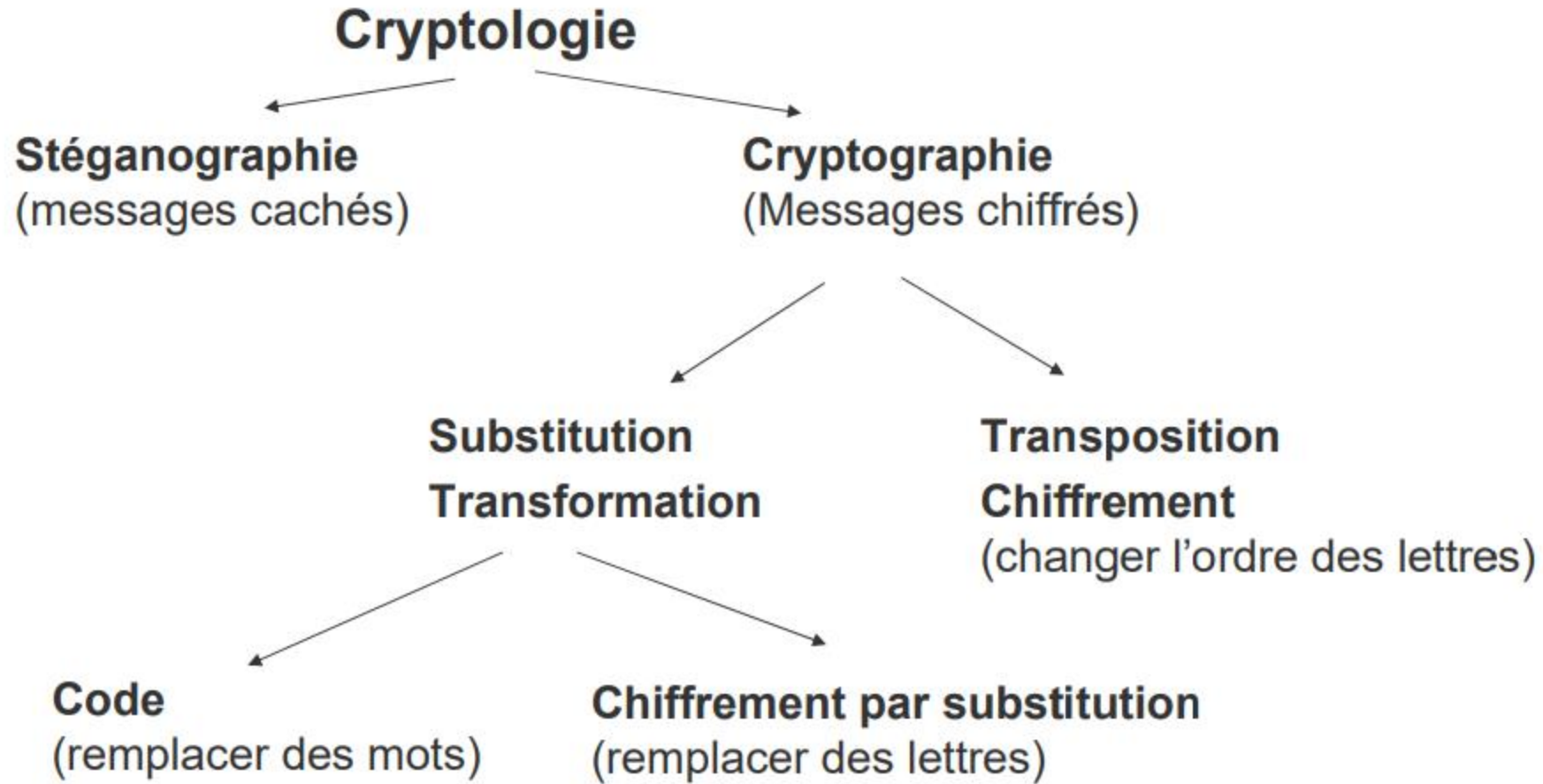
G.Barmarin
2023-2024

La **Cryptologie** est la science des messages secrets.

Elle se décompose en deux disciplines :

- la **Cryptographie**, art de transformer un message clair en un message inintelligible par celui qui ne possède pas la clé de déchiffrement. Cependant, on utilise souvent le mot cryptographie comme synonyme de cryptologie.
- la **Cryptanalyse**, art d'analyser un message chiffré afin de le décrypter quand on ne possède pas la clé de déchiffrement.

Cryptographie classique



Si un message est intercepté, il devrait ne pas être compris ou ne pas être déchiffré facilement.

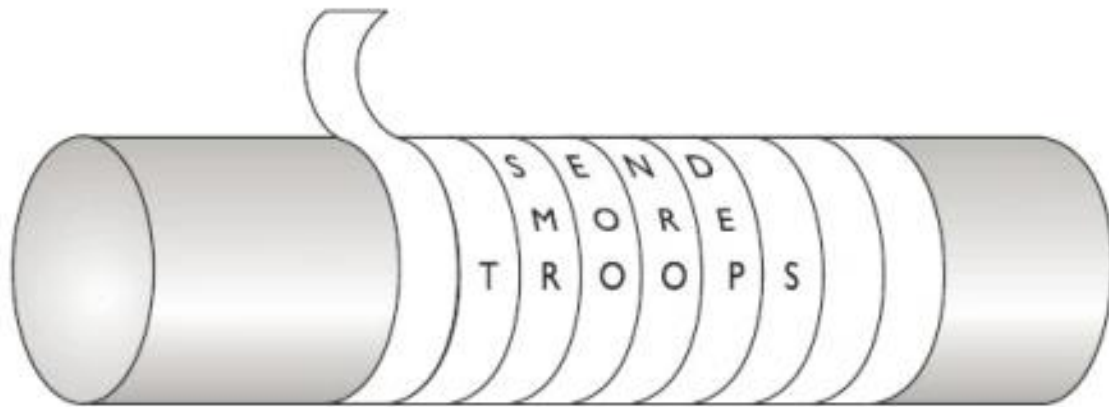
Les premières traces de la cryptographie remontent au XVIème siècle avant JC.

Au début il était question de chiffrement symétrique, c'est-à-dire, avec une clé permettant de chiffrer et déchiffrer un message.

Depuis le milieu des années 70, une méthode révolutionnaire de partage de messages est apparue avec des clés publiques et privées : le cryptage asymétrique. Actuellement, ces deux types de cryptage sont utilisés conjointement. Les algorithmes asymétriques pour transmettre des clés de chiffrement et les algorithmes symétriques afin de chiffrer les données à protéger.

Premières traces de cryptographie

Aux alentours du XVIème siècle avant J.-C, un potier en Irak avait gravé sur une table en argile sa recette en supprimant les consonnes et en modifiant l'orthographe des mots. Par la suite, entre le Xème et le VIIème siècle avant J.-C., les Grecs utilisaient des **scytales**, des sortes de bâtons en bois. Quand l'émetteur voulait communiquer, il enroulait une bande de cuir sur la scytale et y inscrivait le message (une lettre par bout de bande). Une fois la bande déroulée, les lettres n'étaient plus ordonnées et n'avaient donc plus aucun sens. Le seul moyen de pouvoir comprendre le message était d'enrouler la bande sur une scytale de même diamètre pour que les lettres puissent s'aligner correctement.



Inverser l'ordre des lettres d'un message

Premier truc un peu enfantin mais qui permet de nous échauffer sur la manipulation de chaînes de caractères : inverser l'ordre des lettres !

patate —> etatap

On prend chaque lettre du message une à une et on les ajoute à l'envers:

"p"

"ap"

"tap"

"atap"

"tatap"

"etatap"

En python:

```
message = "patate"  
inverse = ""  
for lettre in message :  
    inverse = lettre + inverse  
print(inverse)
```

Principe de Kerckhoffs

En 1883, le cryptologue hollandais Auguste Kerckhoffs publiait un essai intitulé « La cryptologie militaire ».

Cet essai présente une liste des 6 règles à respecter en cryptographie afin d'assurer un système confidentiel.

Bien que ce soient des règles anciennes (elles font par exemple référence à la correspondance par télégramme), elles sont toujours applicables de nos jours.

Principe de Kerckhoffs

1. Le système doit être matériellement, sinon mathématiquement, **indéchiffrable**;
2. Il faut qu'il **n'exige pas le secret**, et qu'il puisse sans inconvénient tomber entre les mains de l'ennemi;
3. La **clé** doit pouvoir en être communiquée et **retenue** sans le secours de notes écrites, et être **changée** ou modifiée au gré des correspondants;
4. Il faut qu'il soit applicable à la correspondance télégraphique ;
5. Il faut qu'il soit **portatif**, et que son maniement ou son fonctionnement n'exige pas le concours de plusieurs personnes;
6. Enfin, il est nécessaire, vu les circonstances qui en commandent l'application, que le système soit d'un **usage facile**, ne demandant ni tension d'esprit, ni la connaissance d'une longue série de règles à observer.

Il y a essentiellement deux types de cryptographie :

- La **cryptographie à clé secrète ou cryptographie symétrique**. C'est la plus ancienne.
- La **cryptographie à clé publique ou cryptographie asymétrique**. C'est la plus récente

On considère généralement que la cryptographie asymétrique est née en 1976 avec l'article de Diffie et Hellman : "New directions in cryptography".

Le chiffrement symétrique est beaucoup plus rapide que le chiffrement asymétrique mais a l'inconvénient de nécessiter le partage au préalable d'une clé secrète.

En pratique, on utilise d'abord un chiffrement asymétrique pour échanger la clé secrète et ensuite un chiffrement symétrique pour l'échange des données.

La clé publique permet l'encodage, mais pas le décodage. Il faut la clé privée pour effectuer le déchiffrement. Le tout est basé sur le produit de 2 nombres premiers

DSZQUPHSBQIJF TZNFSJRVF (*)



(*) Cryptographie symétrique

Fonctionnement de la cryptographie symétrique



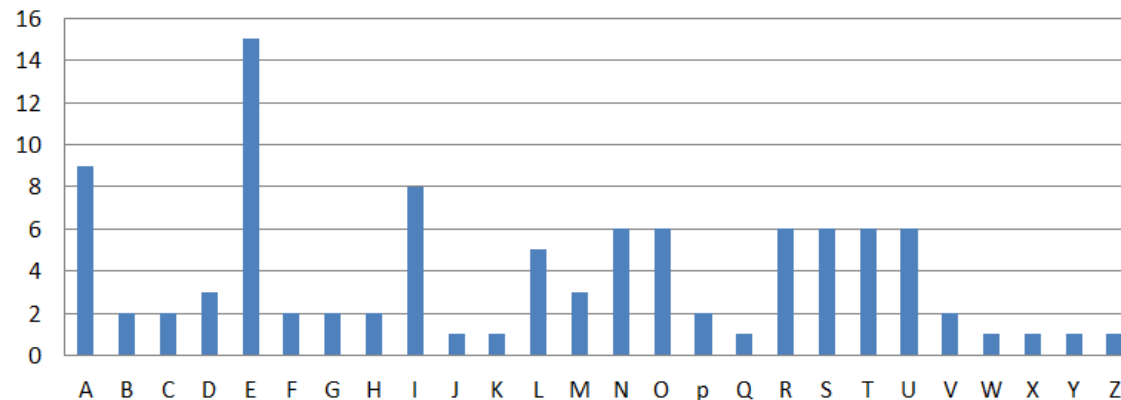
Faiblesse: L'émetteur doit transmettre la clé aux personnes à qui il désire transmettre le message s'il veut que son message puisse être lu.

Substitution monoalphabétique

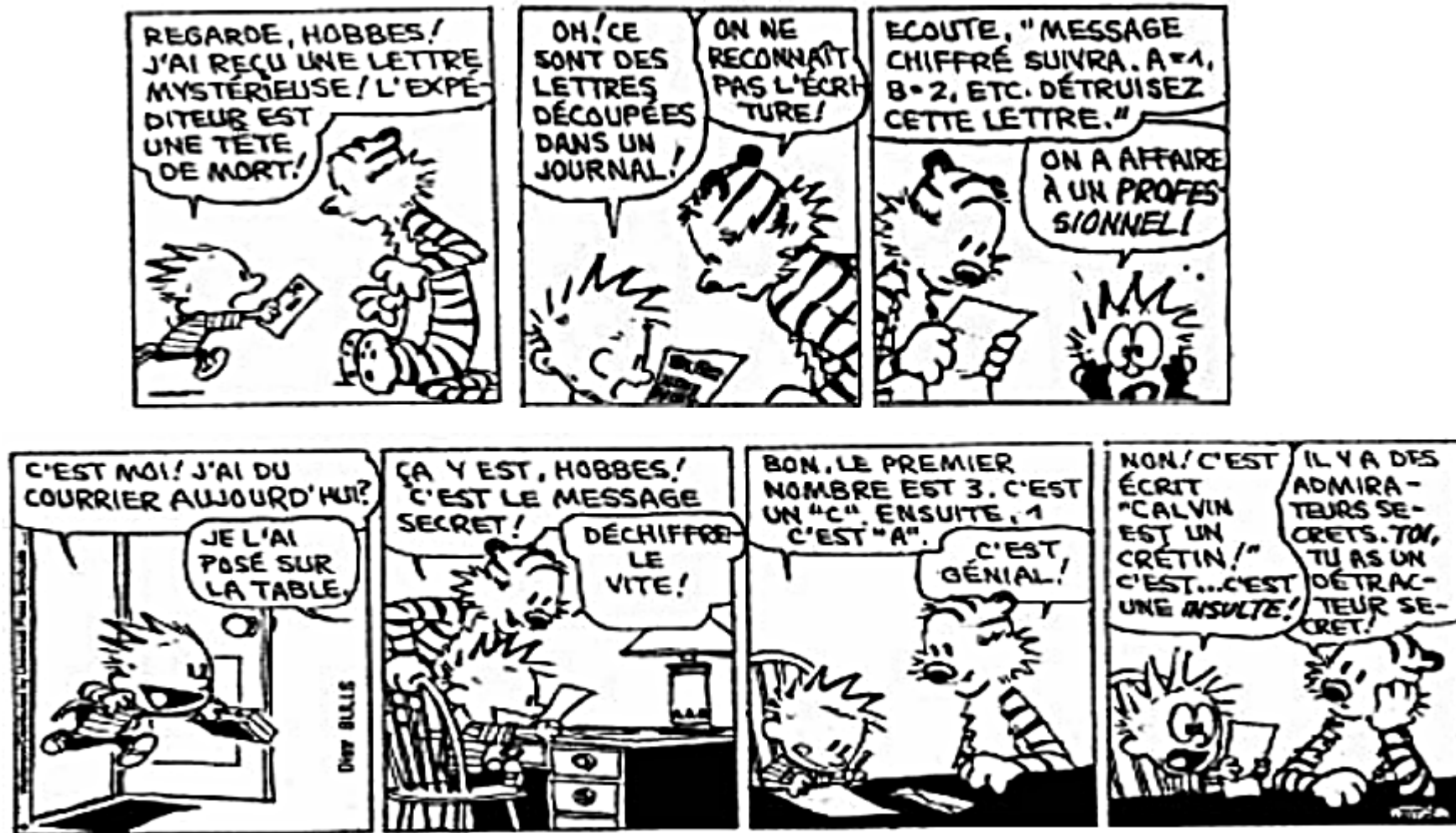
Dans les substitutions simples (qu'on appelle aussi monoalphabétiques), chaque lettre est remplacée par une autre lettre ou un autre symbole.

Dans cette catégorie, on peut citer **le chiffre de César**, **les alphabets désordonnés** ou encore **le chiffre affine**.

Toutes les substitutions simples sont **vulnérables à une analyse des fréquences d'apparition des lettres**.



Le message que déchiffre Calvin est aussi un exemple de substitution simple.



Chiffrement de César (1er siècle avant J.-C.)

Principe : très simple, il suffit de substituer chaque caractère du message d'origine par un autre dans l'alphabet, qui se trouve toujours à une distance fixe, cela revient à décaler les lettres de l'alphabet.

- Chiffrement : $C = E(p) = (p + k) \bmod 26$ (modulo: reste de la division euclidienne)
- Déchiffrement : $p = D(C) = (C - k) \bmod 26$

algorithme connu → cryptanalyse par force brute très simple → 25 clés possibles !

Pourquoi force brute ?

Algorithme connu 25 clés à essayer

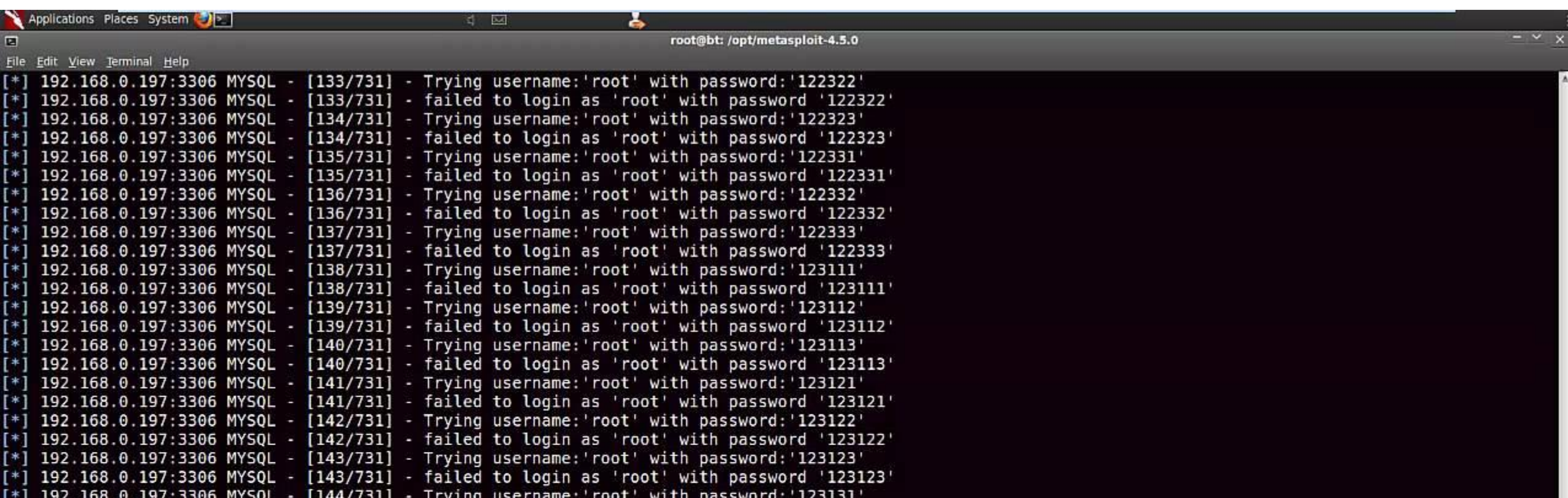
Langage initial connu

→ c'est la longueur de la clé qui rend cette attaque inutilisable



- une compression ou un **langage inconnu** rendent l'attaque plus difficile
- une permutation des 26 caractères alphabétiques → $26!$ clés ($> 4 \cdot 10^{26}$ clés).
(voir plus loin)

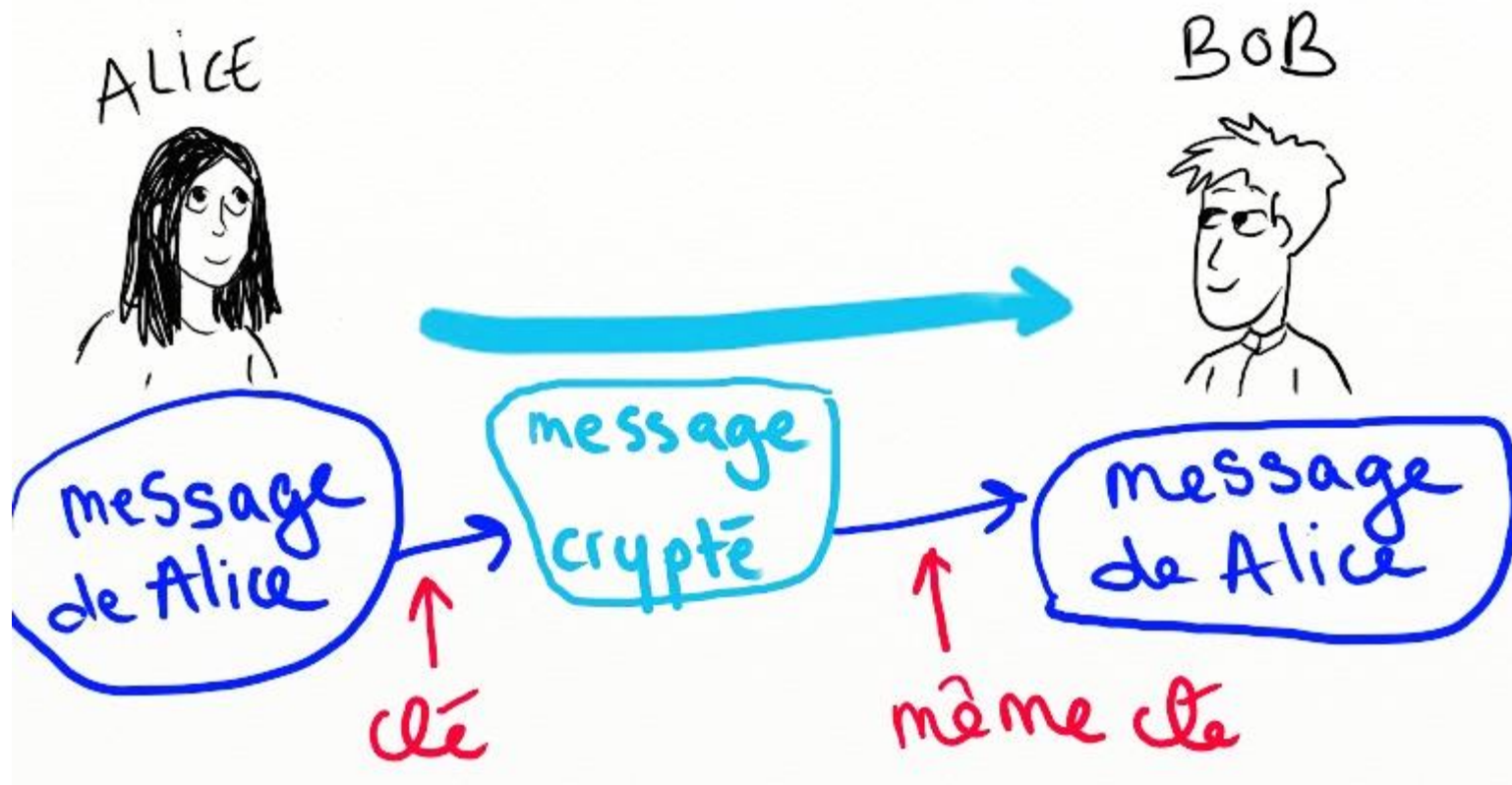
L'attaque par force brute est alors éliminée...



```
root@bt: /opt/metasploit-4.5.0
File Edit View Terminal Help
[*] 192.168.0.197:3306 MYSQL - [133/731] - Trying username:'root' with password:'122322'
[*] 192.168.0.197:3306 MYSQL - [133/731] - failed to login as 'root' with password '122322'
[*] 192.168.0.197:3306 MYSQL - [134/731] - Trying username:'root' with password:'122323'
[*] 192.168.0.197:3306 MYSQL - [134/731] - failed to login as 'root' with password '122323'
[*] 192.168.0.197:3306 MYSQL - [135/731] - Trying username:'root' with password:'122331'
[*] 192.168.0.197:3306 MYSQL - [135/731] - failed to login as 'root' with password '122331'
[*] 192.168.0.197:3306 MYSQL - [136/731] - Trying username:'root' with password:'122332'
[*] 192.168.0.197:3306 MYSQL - [136/731] - failed to login as 'root' with password '122332'
[*] 192.168.0.197:3306 MYSQL - [137/731] - Trying username:'root' with password:'122333'
[*] 192.168.0.197:3306 MYSQL - [137/731] - failed to login as 'root' with password '122333'
[*] 192.168.0.197:3306 MYSQL - [138/731] - Trying username:'root' with password:'123111'
[*] 192.168.0.197:3306 MYSQL - [138/731] - failed to login as 'root' with password '123111'
[*] 192.168.0.197:3306 MYSQL - [139/731] - Trying username:'root' with password:'123112'
[*] 192.168.0.197:3306 MYSQL - [139/731] - failed to login as 'root' with password '123112'
[*] 192.168.0.197:3306 MYSQL - [140/731] - Trying username:'root' with password:'123113'
[*] 192.168.0.197:3306 MYSQL - [140/731] - failed to login as 'root' with password '123113'
[*] 192.168.0.197:3306 MYSQL - [141/731] - Trying username:'root' with password:'123121'
[*] 192.168.0.197:3306 MYSQL - [141/731] - failed to login as 'root' with password '123121'
[*] 192.168.0.197:3306 MYSQL - [142/731] - Trying username:'root' with password:'123122'
[*] 192.168.0.197:3306 MYSQL - [142/731] - failed to login as 'root' with password '123122'
[*] 192.168.0.197:3306 MYSQL - [143/731] - Trying username:'root' with password:'123123'
[*] 192.168.0.197:3306 MYSQL - [143/731] - failed to login as 'root' with password '123123'
[*] 192.168.0.197:3306 MYSQL - [144/731] - Trying username:'root' with password:'123131'
```

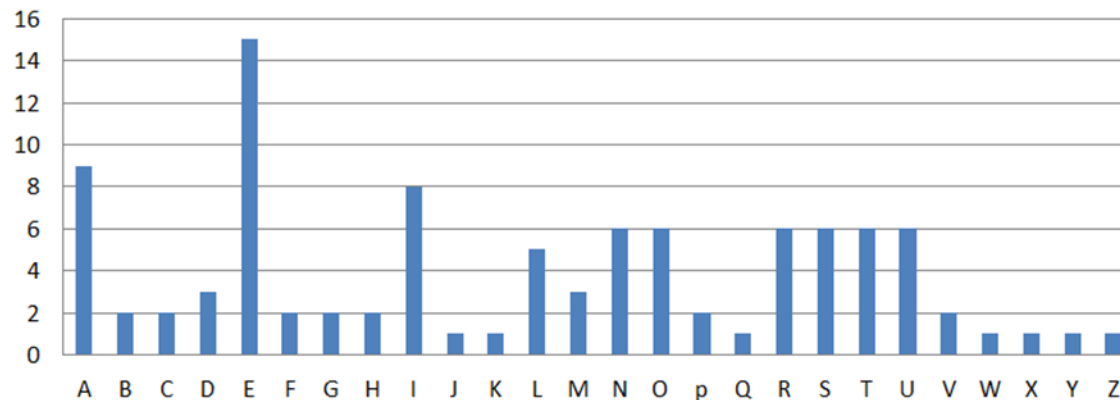
Le chiffrement symétrique consiste à utiliser une clé pour chiffrer un message et la même clé pour le déchiffrer.

Le code de César est donc un chiffrement symétrique .



Analyse de fréquence

- Si la langue de départ et la technique de chiffrement sont connus
→ **exploiter les régularités du langage.**
- **Analyse de la fréquence d'une lettre**
- Cette technique ne fonctionne bien que si le message chiffré est **suffisamment long pour avoir des moyennes significatives.**



Analyse de fréquence

QTJYCOQTQYVJYIOUOMPEGOJQIOYIUPQPFN

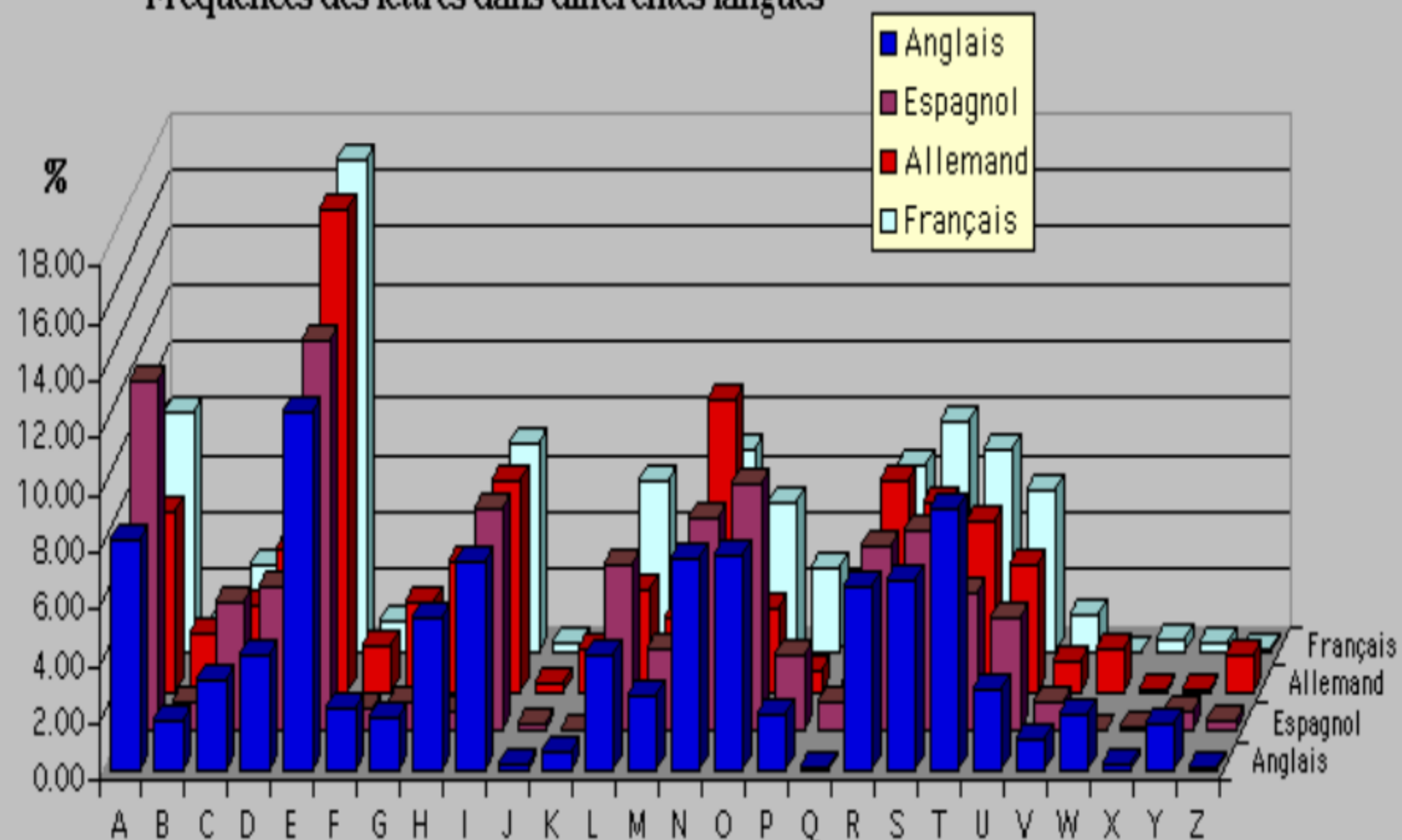
S E S S E E E ES

YOUOMGOBJOQSOYJGJQYEWAFOWOYYPHOSTUO

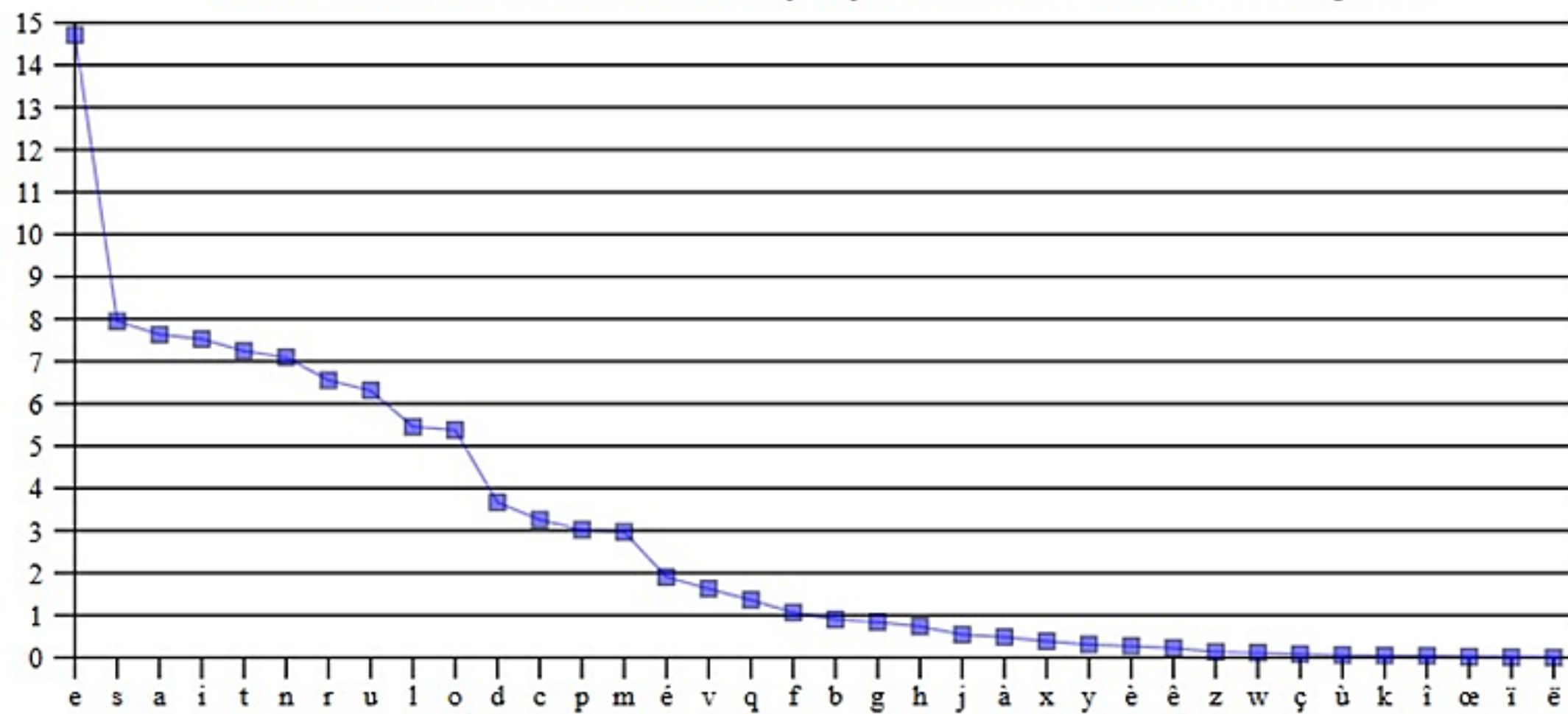
SE E E E ES S E SS E E

- Fréquences : O = 14, Y = 9, Q = 7, J = 6...
- Fréquences en français : E , A, S, I, T, ...
- Dans notre cas : O ?= E, Y ?= A ou Y ?= S, ...
- NOUSVENONSJUSTEDEFIREUNTESTDANALY
SEDEFREQUENCESURUNSIMPLEMESSAGECODE

Fréquences des lettres dans différentes langues



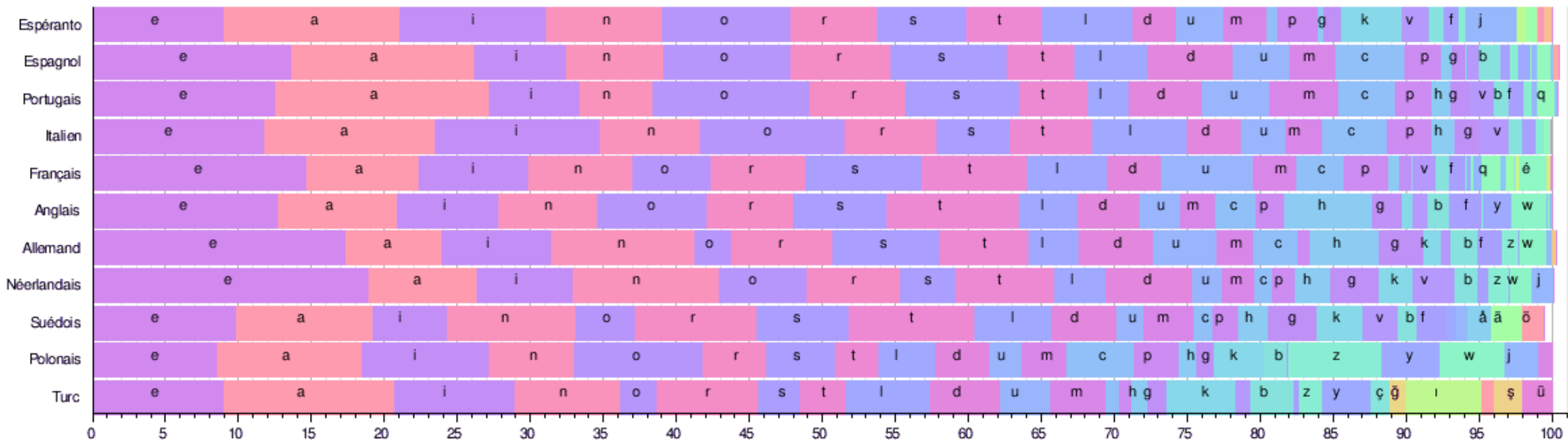
Distribution des lettres (%) dans un texte français



Letter ⇅	Anglais ⇅	Français ³ ⇅	Allemand ⁴ ⇅	Espagnol ⁵ ⇅	Portugais ⁶ ⇅	Espéranto ⁷ ⇅	Italien ⁸ ⇅	Turc ⁹ ⇅	Suédois ¹⁰ ⇅	Polonais ¹¹ ⇅	Néerlandais ¹² ⇅	Danois ¹³ ⇅	Islandais ¹⁴ ⇅	Finnois ¹⁵ ⇅	Tchèque ⇅
a	8.167%	7.636%	6.516%	11.525%	14.634%	12.117%	11.745%	11.920%	9.383%	8.910%	7.486%	6.025%	10.110%	12.217%	8.421%
b	1.492%	0.901%	1.886%	2.215%	1.043%	0.980%	0.927%	2.844%	1.535%	1.470%	1.584%	2.000%	1.043%	0.281%	0.822%
c	2.782%	3.260%	2.732%	4.019%	3.882%	0.776%	4.501%	0.963%	1.486%	3.960%	1.242%	0.565%	0	0.281%	0.740%
d	4.253%	3.669%	5.076%	5.010%	4.992%	3.044%	3.736%	4.706%	4.702%	3.250%	5.933%	5.858%	1.575%	1.043%	3.475%
e	12.702%	14.715%	16.396%	12.181%	12.570%	8.995%	11.792%	8.912%	10.149%	7.660%	18.91%	15.453%	6.418%	7.968%	7.562%
f	2.228%	1.066%	1.656%	0.692%	1.023%	1.037%	1.153%	0.461%	2.027%	0.300%	0.805%	2.406%	3.013%	0.194%	0.084%
g	2.015%	0.866%	3.009%	1.768%	1.303%	1.171%	1.644%	1.253%	2.862%	1.420%	3.403%	4.077%	4.241%	0.392%	0.092%
h	6.094%	0.737%	4.577%	0.703%	0.781%	0.384%	0.636%	1.212%	2.090%	1.080%	2.380%	1.621%	1.871%	1.851%	1.356%
i	6.966%	7.529%	6.550%	6.247%	6.186%	10.012%	10.143%	8.600%*	5.817%	8.210%	6.499%	6.000%	7.578%	10.817%	6.073%
j	0.153%	0.613%	0.268%	0.493%	0.397%	3.501%	0.011%	0.034%	0.614%	2.280%	1.46%	0.730%	1.144%	2.042%	1.433%
k	0.772%	0.074%	1.417%	0.011%	0.015%	4.163%	0.009%	4.683%	3.140%	3.510%	2.248%	3.395%	3.314%	4.973%	2.894%
l	4.025%	5.456%	3.437%	4.967%	2.779%	6.104%	6.510%	5.922%	5.275%	2.100%	3.568%	5.229%	4.532%	5.761%	3.802%
m	2.406%	2.968%	2.534%	3.157%	4.738%	2.994%	2.512%	3.752%	3.471%	2.800%	2.213%	3.237%	4.041%	3.202%	2.446%
n	6.749%	7.095%	9.776%	6.712%	4.446%	7.955%	6.883%	7.487%	8.542%	5.520%	10.032%	7.240%	7.711%	8.826%	6.468%
o	7.507%	5.796%	2.594%	8.683%	9.735%	8.779%	9.832%	2.476%	4.482%	7.750%	6.063%	4.636%	2.166%	5.614%	6.695%
p	1.929%	2.521%	0.670%	2.510%	2.523%	2.755%	3.056%	0.886%	1.839%	3.130%	1.57%	1.756%	0.789%	1.842%	1.906%
q	0.095%	1.362%	0.018%	0.877%	1.204%	0	0.505%	0	0.020%	0.140%	0.009%	0.007%	0	0.013%	0.001%
r	5.987%	6.693%	7.003%	6.871%	6.530%	5.914%	6.367%	6.722%	8.431%	4.690%	6.411%	8.956%	8.581%	2.872%	4.799%
s	6.327%	7.948%	7.270%	7.977%	6.805%	6.092%	4.981%	3.014%	6.590%	4.320%	3.73%	5.805%	5.630%	7.862%	5.212%
t	9.056%	7.244%	6.154%	4.632%	4.336%	5.276%	5.623%	3.314%	7.691%	3.980%	6.79%	6.862%	4.953%	8.750%	5.727%
u	2.758%	6.311%	4.166%	2.927%	3.639%	3.183%	3.011%	3.235%	1.919%	2.500%	1.99%	1.979%	4.562%	5.008%	2.160%
v	0.978%	1.838%	0.846%	1.138%	1.575%	1.904%	2.097%	0.959%	2.415%	0.040%	2.85%	2.332%	2.437%	2.250%	5.344%
w	2.360%	0.049%	1.921%	0.017%	0.037%	0	0.033%	0	0.142%	4.650%	1.52%	0.069%	0	0.094%	0.016%
x	0.150%	0.427%	0.034%	0.215%	0.253%	0	0.003%	0	0.159%	0.020%	0.036%	0.028%	0.046%	0.031%	0.027%
y	1.974%	0.128%	0.039%	1.008%	0.006%	0	0.020%	3.336%	0.708%	3.760%	0.035%	0.698%	0.900%	1.745%	1.043%
z	0.074%	0.326%	1.134%	0.467%	0.470%	0.494%	1.181%	1.500%	0.070%	5.640%	1.39%	0.034%	0	0.051%	1.599%
à	~0%	0.486%	0	0	0.072%	0	0.635%	0	0	0	0	0	0	0	0
â	~0%	0.051%	0	0	0.562%	0	~0%	~0%	0	0	0	0	0	0	0

Fréquence des caractères² sur le corpus de Wikipédia en français

Rang ↕	Caractère ↕	Nombre d'occurrences ↕	Pourcentage ↕	
1	e	115 024 205	12.10%	<div></div>
2	a	67 563 628	7.11%	<div></div>
3	i	62 672 992	6.59%	<div></div>
4	s	61 882 785	6.51%	<div></div>
5	n	60 728 196	6.39%	<div></div>
6	r	57 656 209	6.07%	<div></div>
7	t	56 267 109	5.92%	<div></div>
8	o	47 724 400	5.02%	<div></div>
9	l	47 171 247	4.96%	<div></div>
10	u	42 698 875	4.49%	<div></div>
11	d	34 914 685	3.67%	<div></div>
12	c	30 219 574	3.18%	<div></div>
13	m	24 894 034	2.62%	<div></div>
14	p	23 647 179	2.49%	<div></div>
15	é	18 451 937	1.94%	<div></div>
17	g	11 684 140	1.23%	<div></div>
18	b	10 817 171	1.14%	<div></div>
19	v	10 590 858	1.11%	<div></div>
20	h	10 583 562	1.11%	<div></div>
21	f	10 579 192	1.11%	<div></div>
22	q	6 140 307	0.65%	<div></div>
23	y	4 351 953	0.46%	<div></div>
24	x	3 588 990	0.38%	<div></div>
25	j	3 276 064	0.34%	<div></div>
26	è	2 969 466	0.31%	<div></div>
27	à	2 966 029	0.31%	<div></div>
28	k	2 747 547	0.29%	<div></div>
29	w	1 653 435	0.17%	<div></div>
30	z	1 433 913	0.15%	<div></div>



Fréquence d'apparition des lettres en français

- Le calcul de la fréquence des lettres dans une langue est difficile et soumis à interprétation. On compte la fréquence des lettres d'un texte arbitrairement long, mais un certain nombre de paramètres influencent les résultats :
- Le style narratif : s'il y a beaucoup de verbes à la 2e personne du pluriel (le vouvoiement, présent dans beaucoup de dialogues), il y aura significativement plus de « Z ».
- Le vocabulaire spécifique du document : si l'on parle de chemins de fer, il y aura beaucoup plus de « W » (wagon) ; si l'un des protagonistes se dénomme Loïs, le nombre d'« ï » s'en ressentira.
- Le type de document : des petites annonces en France présenteront souvent le symbole Euro (€), qui est absent de la plupart des autres documents.

Fréquence d'apparition des lettres en français (2)

- La langue d'origine du texte : les noms propres restant généralement les mêmes entre sa version originale et sa version traduite, certaines variations de fréquences de lettres rares en français peuvent se faire ressentir. Dans un texte d'un auteur anglais par exemple, les noms propres auront tendance à faire augmenter les fréquences de lettres relativement communes dans cette langue, telles que le H, le W ou le Y.
- L'époque à laquelle le texte a été rédigé : un texte français du dix-huitième siècle ne contiendra pas ou peu de W, car cette lettre était à cette époque beaucoup moins utilisée qu'aujourd'hui.

Fréquence d'apparition des lettres en français (3)

- Les paramètres techniques : on peut facilement calculer des statistiques sur des textes informatisés, mais souvent ceux-ci ne comportent pas de majuscules accentuées (car difficiles à entrer sur certains ordinateurs) et il arrive aux auteurs d'oublier des accents. La graphie de l'e-dans-l'o (œ) est impossible à représenter dans le codage latin-1 qui est souvent utilisé pour les textes en français. La présence de caractères non alphabétiques (symboles de ponctuation, chiffres, parenthèses et accolades, symboles mathématiques courants...) peut ou non être prise en compte ; la virgule, le point ou l'apostrophe sont par exemple plus fréquents que plus de la moitié des lettres.

Si ces paramètres ont un impact spectaculaire sur les symboles les moins fréquents (la fréquence du œ varie entre 0,002 % et 0,09 % pour trois textes pris au hasard), elle est également sensible même pour les lettres les plus fréquentes (l'ordre de fréquence des lettres A, S, I, T et N, qui sont les plus fréquentes à part E, fluctue d'un texte à l'autre).

Codage: Le chiffre affine

L'idée est d'utiliser comme fonction de chiffrage une fonction affine ($ax+b$) du type:

$$y = (ax + b) \bmod 26,$$

où a et b sont des constantes, et où x et y sont des nombres correspondant aux lettres de l'alphabet ($A=0, B=1, \dots$)

On peut remarquer que si $a=1$, alors on retrouve le chiffre de César et b est le décalage.

On remarquera aussi que si $b=0$, alors "a" est toujours chiffré "A".

Rappel fonction « modulo »

En informatique, l'opération modulo, ou opération mod, est une opération qui associe à deux entiers naturels le reste de la division euclidienne du premier par le second, le reste de la division de a par n ($n \neq 0$) est noté $a \bmod n$ ($a\%n$ en python).

Ainsi

$$9 \bmod 4 = 1, \quad \text{car } 9 = 2 \times 4 + 1 \quad \text{et} \quad 0 \leq 1 < 4$$

$$9 \bmod 3 = 0, \dots$$

En python:

```
Print (9%4)
```

```
1
```

```
Print (3*11+13) %26)
```

```
20
```

Le chiffre affine - fonctionnement

$$\text{Clé} = (k_1, k_2) \quad k_1, k_2 \in [0, 25] \quad \text{gcd}(k_1, 26) = 1$$

Transformation de chiffrement :

$$c_i = f(m_i) = (k_1 * m_i + k_2) \bmod 26$$

Transformation de déchiffrement :

$$m_i = f^{-1}(c_i) = k_1^{-1} * (c_i - k_2) \bmod 26$$

Exemple d'utilisation

Clé = $(k_1, k_2) = (3, 11)$ m=message, c=codage

Transformation de chiffrement :

$$c_i = f(m_i) = (3 * m_i + 11) \bmod 26$$

Transformation de déchiffrement :

$$k_1^{-1} = 3^{-1} \bmod 26 = 9 \text{ car } (3 * 9) \bmod 26 = 1$$

$$m_i = f^{-1}(c_i) = 9 * (c_i - 11) \bmod 26$$

Exemple :

NSA \rightarrow 13 – 18 – 0 \rightarrow 24 – 13 – 11 \rightarrow YNL

13 18 0

24 13 11

N=13

$$3 * 13 + 11 = 50$$

$$50 \bmod 26 = 24$$

24=Y

Donc N=Y

Y=24

$$9 * (24 - 11) = 117$$

$$117 \bmod 26 = 13$$

13=N

Donc Y=N

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

Le chiffre affine – cryptanalyse

1. Établir la fréquence relative de chaque lettre du texte chiffré
→ analyse de fréquence

Exemple:

GHUYI DEGRS YTGOR RYOVG EOHGA
HKEIA AOTDG SBINN TGKGR HENNI
RGSGH HGNYI ASI

G : 10 x

H : 6 x

Langue présumée: Anglais ?

Le chiffre affine – cryptanalyse

2. Sur base de l'analyse de fréquence, dériver les équations correspondantes:

Hypothèse : E et T sont les lettres les plus fréquentes en anglais

Equations correspondantes :

$$E \rightarrow G \quad f(E) = G$$

$$T \rightarrow H \quad f(T) = H$$

$$4 \rightarrow 6 \quad f(4) = 6$$

$$19 \rightarrow 7 \quad f(19) = 7$$

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

Le chiffre affine – cryptanalyse

3. Résoudre les équations pour k_1 et k_2 inconnus

$$f(4) = 6$$

$$f(19) = 7$$

$$4*k_1 + k_2 \equiv 6 \pmod{26}$$

$$19*k_1 + k_2 \equiv 7 \pmod{26}$$

$$15 k_1 \equiv 1 \pmod{26}$$

(soustraction des 2 équations ci-dessus)

$$k_1 = 7$$

$$4*7 + k_2 \equiv 6 \pmod{26}$$

$$28 + k_2 = 1*26 + 6 = 32$$

$$k_2 = 4$$

Contre-mesures

- Utiliser des homophones (faire des fote d'ortograf quoi ;-)
- remplacer une lettre non pas par un symbole unique, mais par un symbole choisi au hasard parmi plusieurs.
- Dans sa version la plus sophistiquée, on choisira un nombre des symboles proportionnel à la fréquence d'apparition de la lettre.
→ Renversement des fréquences.
- faire disparaître complètement les indications fournies par la fréquence
- Contrecarrer par les digrammes et les trigrammes

Digrammes (ou Bigrammes) et trigrammes

Les 20 bigrammes les plus fréquents

Bigrammes	ES	DE	LE	EN	RE	NT	ON	ER	TE	EL	AN	SE	ET	LA	AI	IT	ME	OU	EM	IE
Nombres	3318	2409	2366	2121	1885	1694	1646	1514	1484	1382	1378	1377	1307	1270	1255	1243	1099	1086	1056	1030

Les 20 trigrammes les plus fréquents

Trigrammes	ENT	LES	EDE	DES	QUE	AIT	LLE	SDE	ION	EME	ELA	RES	MEN	ESE	DEL	ANT	TIO	PAR	ESD	TDE
Nombres	900	801	630	609	607	542	509	508	477	472	437	432	425	416	404	397	383	360	351	350

Le chiffrement de Playfair (1854)

Chiffrement à lettre multiples (digrammes)

On dispose les 25 lettres de l'alphabet (W exclu car inutile, on utilise V à la place) dans une grille 5x5 (d'où l'exclusion d'une lettre!), ce qui donne la clef.
(La variante anglaise consiste à garder le W et à fusionner I et J.)

4 règles à appliquer

Déchiffrement : appliquer les règles dans l'autre sens.



Playfair - règles

Si les 2 lettres sont :

1. sur des « coins » → lettres chiffrées = les 2 autres coins.
2. sont sur la même ligne → prendre les 2 lettres qui les suivent immédiatement à leur droite
3. la même colonne → prendre les 2 lettres qui les suivent immédiatement en dessous
4. Identiques → insérer une nulle (usuellement le X) entre les deux pour éliminer ce doublon

B	Y	D	G	Z
J	S	F	U	P
L	A	R	K	X
C	O	I	V	E
Q	N	M	H	T

Règle 1

B	Y	D	G	Z
J	S	F	U	P
L	A	R	K	X
C	O	I	V	E
Q	N	M	H	T

Règle 2

B	Y	D	G	Z
J	S	F	U	P
L	A	R	K	X
C	O	I	V	E
Q	N	M	H	T

Règle 3

Pour former les grilles de chiffrement, on utilise un mot-clef secret pour créer un alphabet désordonné avec lequel on remplit la grille ligne par ligne. Les autres lettres de l'alphabet sont alors ajoutées dans l'ordre dans la grille pour la compléter

Playfair – cryptanalyse et critique

Si le cryptogramme est assez long → analyse de la fréquence des digrammes. Il faut ensuite essayer de reconstituer la grille de chiffrement.

Avantages :

- 26 lettres → $26 \times 26 = 676$ digrammes
- Analyse de fréquence difficile+...

Inconvénient :

- facile à casser car il conserve la structure du texte clair

Utilisé pendant les 2 guerres mondiales par les alliés

Chiffrement de Hill (1929)

Chiffrement

Les lettres sont d'abord remplacées par leur rang dans l'alphabet.

Les lettres P_k et $P_{k+1} \rightarrow C_k$ et C_{k+1}

$$\begin{pmatrix} C_k \\ C_{k+1} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} P_k \\ P_{k+1} \end{pmatrix} \pmod{26}$$

Chaque digramme clair (P_1 et P_2) sera chiffré (C_1 et C_2) selon :

$$C_1 \equiv aP_1 + bP_2 \pmod{26}$$

$$C_2 \equiv cP_1 + dP_2 \pmod{26}$$

Matrice de chiffrement

On ne peut pas prendre n'importe quoi comme matrice de chiffrement.

Ses composantes doivent tout d'abord être des nombres entiers positifs.

Il faut aussi qu'elle ait une matrice inverse dans \mathbb{Z}_{26} .

Le chiffre affine peut être vu comme la version unidimensionnelle du chiffrement de Hill.

Exemple de chiffrement

Alice prend comme clef de cryptage la matrice

$$\begin{pmatrix} 9 & 4 \\ 5 & 7 \end{pmatrix}$$

pour chiffrer le message « je vous aime »

Après avoir remplacé les lettres par leur rang dans l'alphabet (a=1, b=2, etc.), elle obtiendra:

$$C_1 \equiv 9 \cdot 10 + 4 \cdot 5 \pmod{26} = 110 \pmod{26} = 6$$

$$C_2 \equiv 5 \cdot 10 + 7 \cdot 5 \pmod{26} = 85 \pmod{26} = 7$$

Exemple de chiffrement

Elle fera de même avec les 3e et 4e lettres, 5e et 6e, etc.
Elle obtiendra finalement

Lettres	j	e	v	o	u	s	a	i	m	e
Rangs (P_k)	10	5	22	15	21	19	1	9	13	5
Rangs chiffrés (C_k)	6	7	24	7	5	4	19	16	7	22
Lettres chiffrées	F	G	X	G	E	D	S	P	G	V

Chiffre de Hill

Déchiffrement

Pour déchiffrer, le principe est le même que pour le chiffrement: on prend les lettres deux par deux, puis on les multiplie par une matrice

$$\begin{pmatrix} P_1 \\ P_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} \pmod{26}$$

Cette matrice doit être l'inverse de la matrice de chiffrement (modulo 26).
Ordinairement cet inverse est:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Exemple de déchiffrement

Pour déchiffrer le message d'Alice, Bob doit calculer :

$$\begin{pmatrix} 9 & 4 \\ 5 & 7 \end{pmatrix}^{-1} = \frac{1}{43} \begin{pmatrix} 7 & -4 \\ -5 & 9 \end{pmatrix} \pmod{26} = (43)^{-1} \begin{pmatrix} 7 & -4 \\ -5 & 9 \end{pmatrix} \pmod{26}$$

Comme $\gcd(43, 26) = 1$, $(43)^{-1}$ existe dans Z_{26} et $(43)^{-1} = 23$.

Bob a la matrice de déchiffrement :

$$\begin{pmatrix} 9 & 4 \\ 5 & 7 \end{pmatrix}^{-1} = 23 \begin{pmatrix} 7 & -4 \\ -5 & 9 \end{pmatrix} \pmod{26} = \begin{pmatrix} 161 & -92 \\ -115 & 207 \end{pmatrix} \pmod{26} = \begin{pmatrix} 5 & 12 \\ 15 & 25 \end{pmatrix} \pmod{26}$$

Si $u = 43^{-1} \pmod{26}$
Alors $43u = \mathbf{1} \pmod{26}$
Qui a pour solution
dans Z_{26} $u = 23$ car
 $43 \times 23 = 989$ et
 $989 \pmod{26} = \mathbf{1}$
Car $38 \times 26 = 988$
 $989 - 988 = \mathbf{1}$

Exemple de déchiffrement

Bob prend donc cette matrice pour déchiffrer le message "FGXGE DSPGV".

Après avoir remplacé les lettres par leur rang dans l'alphabet (A=1, B=2,etc.), il obtiendra:

$$P_1 \equiv 5 \cdot 6 + 12 \cdot 7 \pmod{26} = 114 \pmod{26} = 10$$

$$P_2 \equiv 15 \cdot 6 + 25 \cdot 7 \pmod{26} = 265 \pmod{26} = 5$$

Il fera de même avec les 3e et 4e lettres, 5e et 6e, etc. Il obtiendra finalement:

Lettres chiffrées	F	G	X	G	E	D	S	P	G	V
Rangs chiffrés (C_k)	6	7	24	7	5	4	19	16	7	22
Rangs (P_k)	10	5	22	15	21	19	1	9	13	5
Lettres	j	e	v	o	u	s	a	i	m	e

Substitutions polyalphabétique

Plutôt que d'utiliser un décalage fixe, on se base sur une clé qui va déterminer le décalage pour chaque caractère

Les substitutions polyalphabétiques (aussi appelées à double clef ou à alphabets multiples), utilisent plusieurs "alphabets", ce qui signifie qu'une même lettre peut être remplacée par plusieurs symboles.

L'exemple le plus fameux de chiffre polyalphabétique est sans doute le **chiffre de Vigenère**, qui résista aux cryptanalystes pendant trois siècles.

Chiffre de Vigenère (1568)

- Amélioration décisive du chiffre de César.
- Sa force réside dans l'utilisation non pas d'un, mais de 26 alphabets décalés pour chiffrer un message. (carré de Vigenère).
- Ce chiffre utilise une clef qui définit le décalage pour chaque lettre du message (A: décalage de 0 cran, B: 1 cran, C: 2 crans, ..., Z: 25 crans).

Chiffre de Vigenère

Exemple: chiffrer le texte "CHIFFRE DE VIGENERE" avec la clef "BACHELIER"
(cette clef est éventuellement répétée plusieurs fois pour être aussi longue que le texte clair)

Clair	C	H	I	F	F	R	E	D	E	V	I	G	E	N	E	R	E
Clef	B	A	C	H	E	L	I	E	R	B	A	C	H	E	L	I	E
Décalage	1	0	2	7	4	11	8	4	17	1	0	2	7	4	11	8	4
Chiffré	D	H	K	M	J	C	M	H	V	W	I	I	L	R	P	Z	I

La grande force du chiffre de Vigenère est que la même lettre sera chiffrée de différentes manières ce qui rend inutilisable l'analyse de fréquence classique.

Carré de Vignère

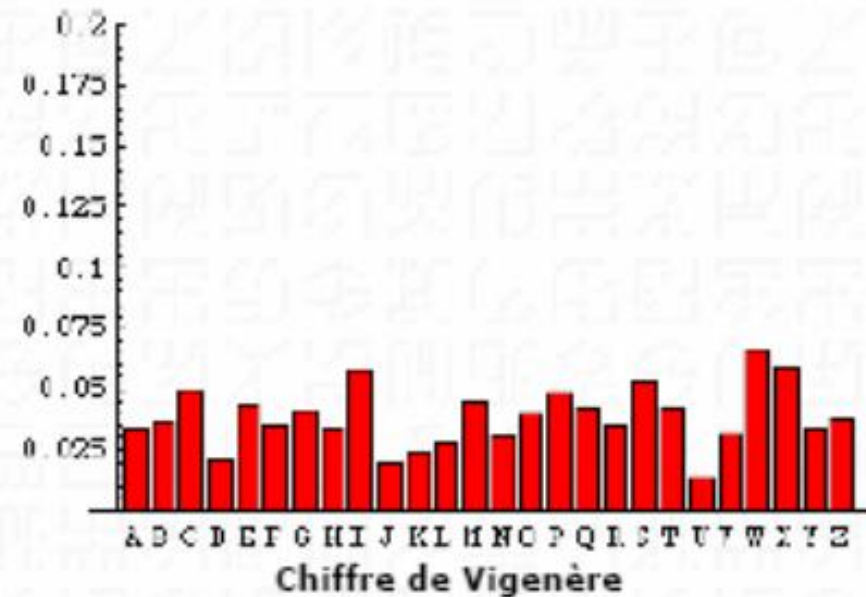
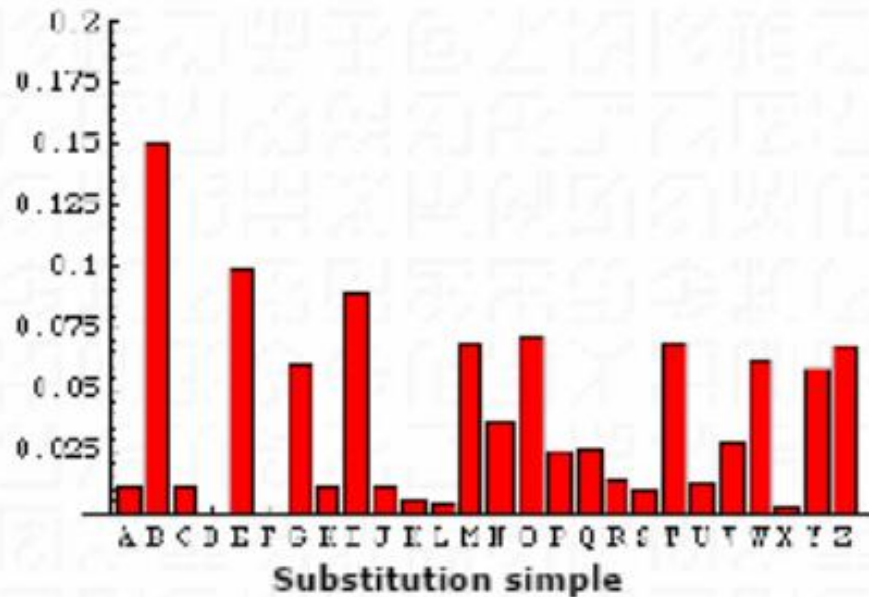
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	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

La lettre de la clef est dans la colonne la plus à gauche en rose, la lettre du message clair est dans la ligne tout en haut en vert. La lettre chiffrée est à l'intersection des deux en jaune.

Si la clef est aussi longue que le texte clair, et moyennant quelques précautions d'utilisation, le système est appelé...**masque jetable!**



Chiffre de Vigenère (1568)



fréquences des lettres d'une fable de la Fontaine (le chat, la belette et le lapin) chiffrée avec une substitution simple(gauche) et avec le chiffre de Vigenère (droite)

Carré de vigenère :

Chiffrement

$$c_i = f_i \bmod d(m_i) = m_i + k_i \bmod d \bmod 26$$

Déchiffrement

$$m_i = f_i^{-1} \bmod d(m_i) = m_i - k_i \bmod d \bmod 26$$

Clés

$$k_1, k_2, \dots, k_{d-1}$$

Nombre de clés : pour une longueur de clé d : 26^d

Cryptanalyse – Méthode Kasiski

1. chercher des séquences de lettres qui apparaissent plus d'une fois dans le texte:
 - soit la même séquence de lettres du texte clair a été cryptée avec la même partie de la clef
 - soit deux suites de lettres différentes dans le texte clair auraient (possibilité faible) par pure coïncidence engendré la même suite dans le texte chiffré.

Le 1er cas = le plus probable → le nombre de facteurs de la clef

2. méthode de fréquence de distribution des lettres cryptées → les lettres du texte clair.

En prenant par exemple la clef KILO, la lettre E peut être chiffrée en O, M, P ou S selon que K, I, L ou O sont utilisés pour la chiffrer. Ainsi le mot thé peut être chiffré en DPP, BSS, EVO ou HRM.

K	I	L	O	K	I	L	O	K	I	L	O	K	I	L	O	K	I	L	O	K				
t	h	e	r	u	s	s	e	t	h	e	j	a	s	m	i	n	t	h	e	c	h	i	n	e
D	P	P	F	E	A	D	S	D	P	P	X	K	A	X	W	X	B	S	S	M	P	T	B	O

Dans l'exemple ci-dessus, le mot "thé" est chiffré "DPP" 2 fois et "BSS" 1 fois.

Des répétitions de cette sorte offrent la prise nécessaire pour attaquer Vigenère et trouver la clef si elle est courte et le texte long.

Le premier pas consiste à deviner la longueur de la clef.

On cherche pour cela des séquences de plusieurs lettres consécutives (par exemple 3 ou plus) apparaissant plusieurs fois.

Ce renseignement est capital. Si, par exemple, la longueur de la clef est 3, cela signifie que les caractères de rang 1, 4, 7, 10, ..., $3k+1$, sont simplement décalés à la manière du chiffre de César. On peut donc appliquer maintenant l'analyse de fréquence à ces caractères et trouver la première lettre de la clef.

Pour la deuxième lettre de la clef, on analysera les fréquences des caractères de rang $3k+2$ et pour la dernière lettre les fréquences des caractères de rang $3k$.

Cryptanalyse – test de Friedman

Le test de Friedman (aussi appelé test kappa) a pour premier objectif de déterminer si un texte a été chiffré avec un chiffre monoalphabétique ou polyalphabétique.

Comme second bénéfice, il suggère la longueur du mot-clef si le chiffre est polyalphabétique.

Pour réaliser cela, le test de Friedman s'appuie sur une métrique appelée Indice de Coïncidence (IC), qui est la probabilité que deux lettres choisies aléatoirement dans un texte soient identiques.

Cryptanalyse – test de Friedman

L'Indice de Coïncidence (IC) est la probabilité que deux lettres choisies aléatoirement dans un texte soient identiques.

Soient:

n le nombre de lettres dans le texte

n_1 = nombre de A, n_2 = nombre de B ,..., n_{26} = nombre de Z

La probabilité de tirer deux A parmi les n lettres du texte est:

$$P(2 \text{ fois } A) = \frac{C_2^{n_1}}{C_2^n} = \frac{\frac{n_1(n_1 - 1)}{2}}{\frac{n(n - 1)}{2}} = \frac{n_1(n_1 - 1)}{n(n - 1)}$$

Cryptanalyse – test de Friedman

La probabilité de tirer 2 lettres identiques :

$$IC = \sum_{i=1}^{26} \frac{n_i(n_i - 1)}{n(n - 1)}$$

Exemples d'indices calculés sur des textes contemporains dans différentes langues:

Langue	allemand	anglais	espagnol	esperanto	français	italien	norvégien	suédois
IC	0.072	0.065	0.074	0.069	0.074	0.075	0.073	0.071

Cryptanalyse – test de Friedman

Remarques importantes:

Pour un langage de 26 lettres où chaque lettre a la même fréquence soit $1/26$, $IC = 0.038$

Pour tout chiffre monoalphabétique, la distribution des fréquences est invariante, donc l'IC sera le même que pour le texte clair.

Donc, si on applique ce test à un texte chiffré avec un chiffre monoalphabétique, on devrait trouver IC égal environ à 0.074 (en français). Si IC est beaucoup plus petit (p. ex. 0.050), le chiffre est probablement polyalphabétique.

Trouver la longueur de la clé avec l'IC

Soit le message suivant, chiffré avec Vigenère (369 lettres):

PERTQ UDCDJ XESCW MPNLV MIQDI ZTQFV XAKLR PICCP QSHZY
DNCPW EAJWS ZGCLM QNRDE OHCGE ZTQZY HELEW AUQFR OICWH
QMYRR UFGBY QSEPV NEQCS EEQWE EAGDS ZDCWE OHYDW QERLM
FTCCQ UNCPP QSKPY FEQOI OHGPR EERWI EFSDM XSYGE UELEH
USNLV GPMFV EIVXS USJPW HIEYS NLCDW MCRTZ MICYX MNMFZ
QASLZ QCJPY DSTTK ZEPZR ECMYW OICYG UESIU GIRCE UTYTI
ZTJPW HIEYI ETYYH USOFI XESCW HOGDM ZSNLV QSQPY JSCAV
QSQLM QNRLP QSRLM XLCCG AMKPG QLYLY DAGEH GERCI RAGEI
ZNMGI YBPP

On va considérer les sous-chaînes obtenues en prenant les lettres à intervalle donné:

Intervalle de 1: PERTQ UDCDJ XESCW MPNLV . . . (texte original)

Intervalle de 2: PRQDD XSWPL . . . et ETUCJ ECMNV . . .

Intervalle de 3: PTDJS MLIIQ . . . , EQCXC PVQZF . . . et RUDEW

Trouver la longueur de la clé avec l'IC

On calcule ensuite les IC pour toutes ces sous-chaînes:

Intervalle	Indice de coïncidence
1	0.0456107
2	0.0476954, 0.0443098
3	0.044249, 0.0494469, 0.0426771
4	0.0465839, 0.0453894, 0.0449116, 0.0425227
5	0.0799704, 0.0925583, 0.0836727, 0.0795282, 0.0684932
6	0.0512956, 0.0407192, 0.0371585, 0.0382514, 0.0661202, 0.0431694

On remarque que quand l'intervalle est de 5, l'IC correspond plus ou moins avec l'IC caractéristique du français (en tout cas, c'est cette ligne qui s'approche le plus de 0.074, les autres lignes étant plutôt proches de 0.038). La longueur de la clef utilisée est donc probablement 5. Pour découvrir la clef elle-même, on peut ensuite procéder comme le faisait Kasiski

Trouver la longueur de la clé avec l'IC

Si un message en français de longueur n et d'indice de coïncidence IC est chiffré avec un carré de Vigenère, alors r , la longueur du mot-clef composé de lettres distinctes, est donné par la formule:

$$r \approx \frac{(0.074 - 0.038) n}{(n - 1)IC - 0.038 n + 0.074} \approx \frac{0.036 n}{(n - 1)IC - 0.038 n + 0.074}$$

En appliquant cette formule au texte précédent, on trouve $r = 4.69$, ce qui confirme ce que l'on avait trouvé ci-dessus.

Chiffre de Vernam (One-Time Pad)

Masque jetable = chiffre de Vigenère avec comme caractéristique que la clef de chiffrement a la même longueur que le message clair

Exemple :

Clair	M	A	S	Q	U	E	J	E	T	A	B	L	E
Clef	X	C	A	A	T	E	L	P	R	V	G	Z	C
Décalage	23	2	0	0	19	4	11	15	17	21	6	25	2
Chiffré	J	C	S	Q	N	I	U	T	K	V	H	K	G

Chiffre de Vernam (One-Time Pad) ou Méthode du masque jetable

Il faut :

1. choisir une clef aussi longue que le texte à chiffrer,
2. utiliser une clef formée d'une suite de **caractères aléatoires**,
3. protéger votre clé
4. ne jamais réutiliser une clé (d'où le nom de masque jetable)
5. écrire des textes clairs ne contenant que les lettres (sans ponctuation et sans espaces).

Le système du masque jetable, avec les précautions indiquées ci-dessus, est absolument inviolable si l'on ne connaît pas la clef.

Il est ou a été couramment utilisé par les États.

En effet, ceux-ci peuvent communiquer les clefs à leurs ambassades de manière sûre via la valise diplomatique .

Cet algorithme qui est théoriquement incassable est cependant difficile à mettre en œuvre en pratique.

Néanmoins, la ligne téléphonique entre le Kremlin et la Maison Blanche a été, à un moment, protégée par un tel système.

Difficultés

Le problème de ce système est de communiquer les clefs de chiffrage ou de trouver un algorithme de génération de clef commun aux deux partenaires :

1. La création de grandes quantités de clefs aléatoires :
n'importe quel système fortement utilisé pourrait exiger des millions de caractères aléatoires de façon régulière.
2. La distribution des clés :
une clé de longueur égale est nécessaire pour l'expéditeur et pour le récepteur. Cela nécessite une bonne organisation.

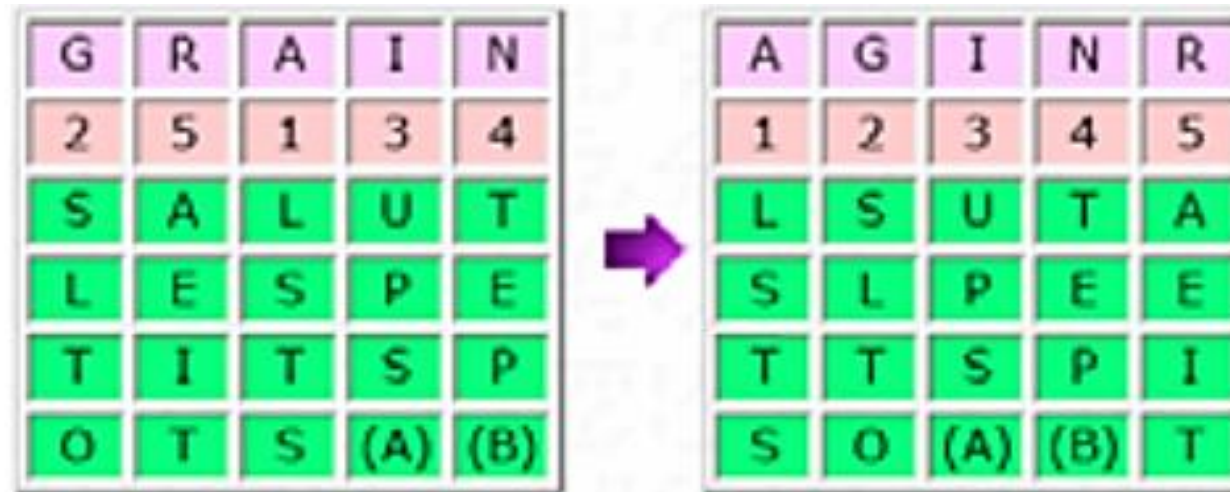
Codage par Transposition

- Consiste à changer l'ordre des lettres
- Pour de très courts messages : la méthode est peu sûre car il y a peu de variantes. Par exemple un mot de trois lettres ne peut être modifié que de 6 ($=3!$) façons différentes. Ainsi col ne peut se transformer qu'en col, clo, ocl, olc, lco, loc
- Lorsque le nombre de lettres croît : impossible de retrouver le texte original sans connaître le procédé de brouillage. Par exemple, une phrase de 35 lettres peut être disposée de $35! = 1040$ manières
- Nécessite un procédé rigoureux convenu auparavant entre les parties. Une transposition au hasard des lettres semblerait offrir un très haut niveau de sécurité, mais il y a un inconvénient: l'expéditeur et l'envoyeur doivent s'être préalablement entendus sur la méthode.

Transposition – exemple

Une transposition rectangulaire consiste à écrire le message dans une grille rectangulaire, puis à arranger les colonnes de cette grille selon un mot de passe donné (le rang des lettres dans l'alphabet donne l'agencement des colonnes).

Dans l'exemple ci- dessous, on a choisi comme clef GRAIN pour chiffrer le message SALUT LES PETITS POTS. En remplissant la grille, on constate qu'il reste deux cases vides, que l'on peut remplir avec des nulles ou pas.



Machines à rotor (WWII) cryptage par substitution (polyalphabétique)

L'entre-deux-guerres voit le début de la mécanisation de la cryptographie.

Des outils mécaniques, comme les cylindres chiffants, sont mis à disposition des opérateurs, et des machines électromécaniques, sont mises au point.

Ces machines fonctionnent sur le principe des rotors et des contacts électriques, afin de réaliser des formes de substitution polyalphabétique dont la clé a une longueur gigantesque de l'ordre de centaines de millions de lettres, au lieu de quelques dizaines dans les méthodes artisanales, comme le chiffre de Vigenère.

Enigma est la machine à chiffrer et déchiffrer qu'utilisèrent les armées allemandes du début des années trente jusqu'à la fin de Seconde Guerre Mondiale. Elle automatise le chiffrement par substitution. Cette machine ressemble à une machine à écrire. Quand on presse sur une touche, deux choses se passent:

Premièrement, une lettre s'allume sur un panneau lumineux: c'est la lettre chiffrée.

Deuxièmement, un mécanisme fait tourner le rotor de droite d'un cran;

toutes les 26 frappes, le deuxième rotor tourne d'un cran, toutes les 676 frappes (26 au carré), c'est le troisième rotor qui tourne d'un cran. Certaines Enigmas avaient 3 rotors, celles de la Kriegsmarine en avaient 4 ou 5.

Machines à rotor (WWII)

Ces rotors tournants modifient les connexions électriques dans la machine, ce qui fait que la touche "A" allumera peut-être le "B" la première fois, mais le "X" la deuxième, le "E" la troisième, etc.

Un "tableau de connexions" et un "réflecteur" complique encore le système.

Le côté génial de cette machine est que même si elle tombe entre les mains ennemies, sa sécurité n'est pas compromise. En effet, c'est le nombre faramineux de réglages de la machine qui fait sa force et les réglages changeaient évidemment chaque jour. On peut en effet changer l'ordre de rotors, leur orientation initiale et les branchement du tableau de connexions. Par exemple, on pouvait spécifier la clef du jour ainsi:

- Position des rotors : 2 - 3 - 1
- Orientations des rotors : 2 - 23 - 5
- Branchements des connexions : A/L - P/R - T/D - B/W - K/F - O/Y
- Indicateurs : B - W - E

Ainsi, connaître le fonctionnement de la machine n'aide (presque) pas à décrypter les messages qu'elle produit. Tout le problème est de retrouver le bon réglage. C'est dans ce but qu'ont été produites les bombes de Turing

La machine Enigma



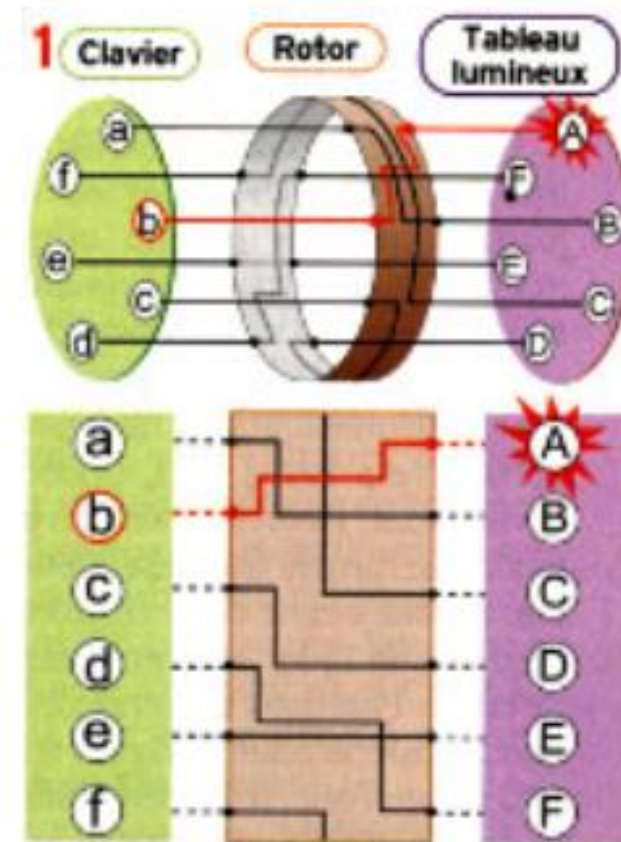
La machine Enigma – Principe (1)

Si on frappe la lettre b sur le clavier, un courant électrique est envoyé dans le rotor, suit la câblage interne, puis ressort à droite pour allumer la lettre A sur le tableau lumineux.

Autre principe de base: chaque fois qu'une lettre est tapée au clavier, le rotor tourne d'un cran.

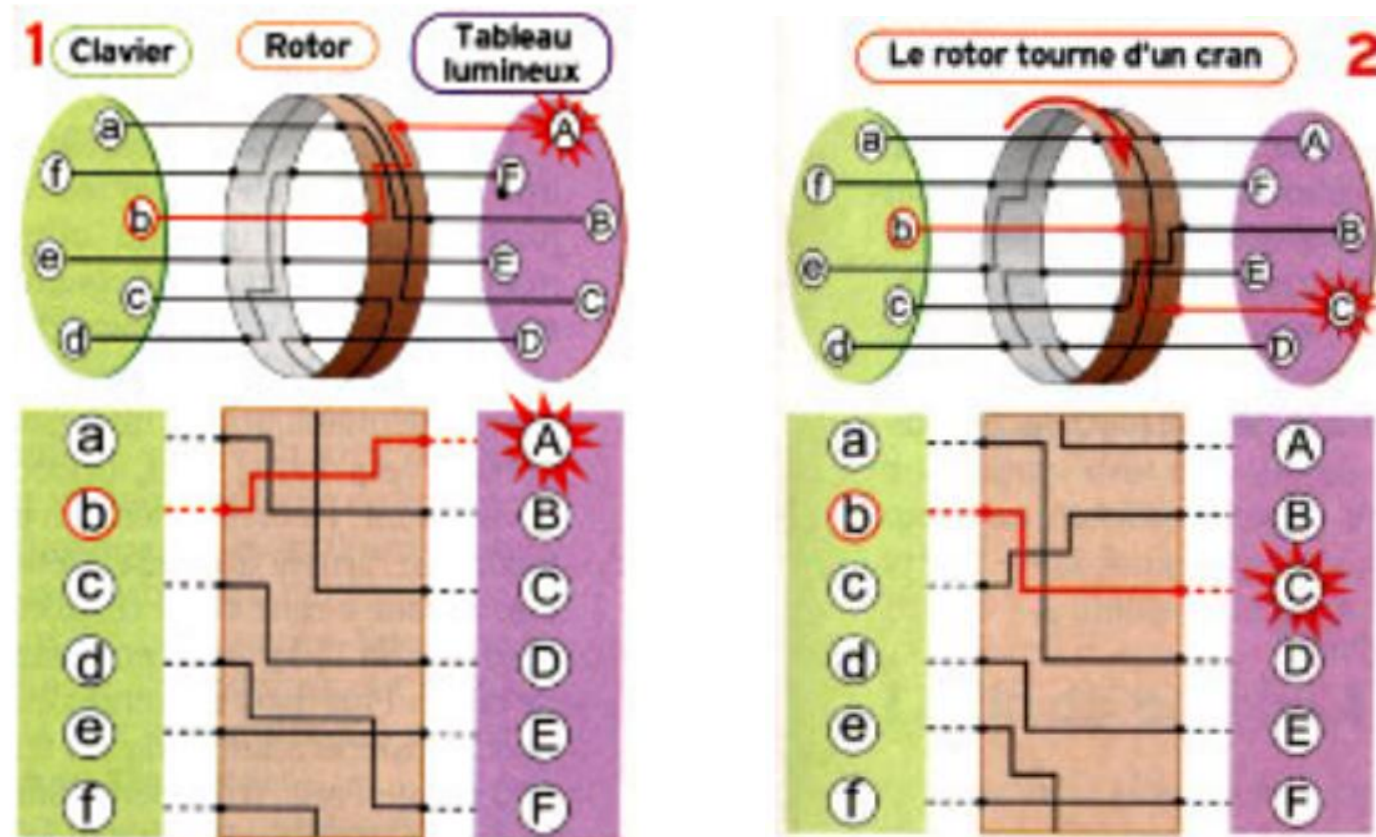
Ainsi, b devient A la première fois, mais b devient C la deuxième fois, puis b devient E, etc.

Le principe de base des machines Enigma conçues par Scherbius repose sur l'utilisation de rotors qui transforment l'alphabet clair (noté en minuscules) en alphabet chiffré (en majuscules). Pour l'illustrer, nous nous limiterons à un alphabet de six lettres. Voici la représentation de l'un de ces fameux rotors, ainsi que le schéma équivalent qui permet de mieux suivre l'opération.



La machine Enigma – Principe (2)

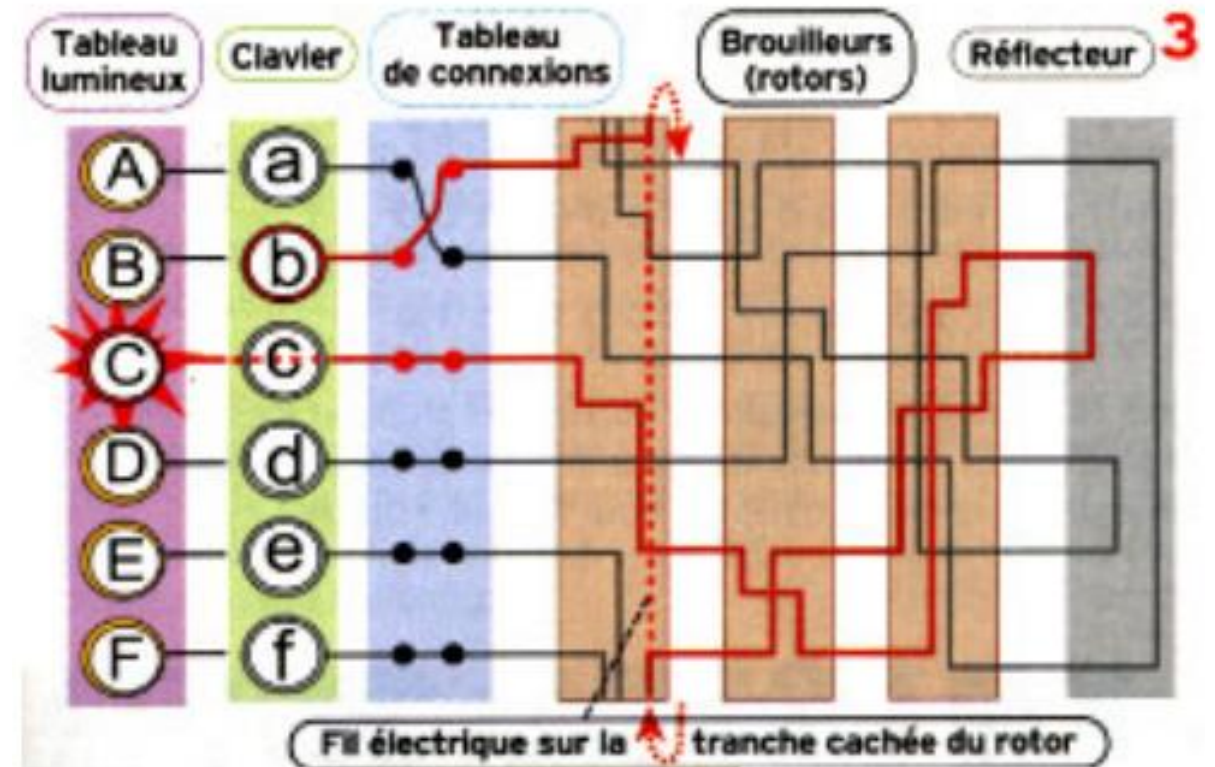
Dans notre exemple le mot bac est chiffré ADD (et non ABD si le rotor était resté immobile).



La machine Enigma – Principe (3)

Pour augmenter le nombre de combinaisons possibles et déjouer les tentatives des cryptanalystes, Scherbius a associé plusieurs dispositifs:

- Un tableau de connexions → brouiller les pistes en reliant 2 lettres du clavier entre elles
- trois brouilleurs associés multiplient le nombre de combinaisons.
- Un réflecteur qui renvoie le courant dans le dispositif jusqu'au panneau lumineux où la lettre cryptée s'affiche. Le réflecteur complexifie encore le chiffrement du caractère car il refait passer le courant une deuxième fois par les rotors selon le même principe qu'à l'aller.



Le tableau de connexions permet de brouiller les pistes en reliant 2 lettres du clavier entre elles (ici a et b). Ainsi quand on tape b le courant prend en fait le circuit prévu pour a et vice-versa. 10 câbles reliaient 20 lettres entre elles (et donc 6 lettres ne changeaient pas)

Les 3 brouilleurs associés multiplient le nombre de combinaisons.

Bien que le câblage ne soit pas linéaire (le courant passant par la première position, ne va pas forcément ressortir par la première position), si aucun autre mécanisme n'est utilisé, le cryptage serait monoalphabétique.

C'est pourquoi on utilise des rotors, ainsi, à chaque frappe de clavier, le rotor de droite va se décaler d'une position. Une fois qu'il a tourné 26 fois (pour les 26 lettres de l'alphabet) et a donc effectué un tour complet, le deuxième rotor va se décaler d'un cran. Ensuite, une fois que le deuxième rotor a été décalé de 26 positions, c'est au tour du troisième de se décaler d'une position. Le mécanisme est comparable à celui d'un compteur kilométrique d'une voiture, qui comptabilise les centaines de mètres, les kilomètres, les centaines et les milliers de kilomètres

Quant au réflecteur, il renvoie le courant dans le dispositif jusqu'au panneau lumineux où la lettre cryptée s'affiche. Son rôle n'est pas d'augmenter le nombre de combinaisons possibles, mais de faciliter considérablement la tâche du destinataire. En effet, si b devient C dans notre exemple (en rouge), on a aussi c devient B. Et c'est valable pour toutes les paires de lettres claire/cryptée. Conséquence: si le mot « efface » est chiffré ACBFEB par l'émetteur, il suffira à l'opérateur qui reçoit le message crypté de taper acbfeb sur son clavier pour voir les lettres E, F, F, A, C, E s'allumer.

Seule condition: les deux opérateurs distants doivent avoir réglé leur machine Enigma de la même façon. En effet, les 3 rotors peuvent être choisis parmi 5, leur position de montage interchangeables, et leur position initiale peut ne pas être à 1. Le tableau de connexions doit aussi être configuré de la même façon sur chaque machine.

La machine Enigma – Principe (4)

Au final, on a:

$26 \times 26 \times 26 = 17576$ combinaisons liées à l'orientation de chacun des trois brouilleurs (Les rotors ont 26 contacts électriques: un pour chaque lettre de l'alphabet sur chaque face du rotor),

6 combinaisons possibles liées à l'ordre dans lequel sont disposés les brouilleurs, si on utilise 3 brouilleurs qui peuvent être montés dans n'importe quel ordre ($3! = 3 \times 2 \times 1 = 6$). Avec 3 brouilleurs choisis parmi 5 différents, on monte à 60 ($5 \times 4 \times 3$)

$\pm 10^{11}$ branchements possibles quand on relie les six paires de lettres dans le tableau de connexions.

Les machines Enigma peuvent donc chiffrer un texte selon $17576 \times 6 \times 100391791500 = 10^{16}$ combinaisons différentes!

La machine Enigma comportait une faille, qui au premier abord semblait ne pas avoir d'importance. En effet, une lettre ne peut pas être chiffrée par elle-même.

Si un opérateur tapait la lettre A, n'importe quel autre caractère aurait pu ressortir, sauf le A.

C'est donc de cette façon qu'Alan Turing a pu déchiffrer des messages.

Les opérateurs allemands envoyaient un rapport météorologique journalier à 6h du matin, qui comportait en en-tête « Wetterbericht » (qui signifie en allemand bulletin météorologique) et signaient le message par « Heil Hitler ».

En partant de l'idée qu'une lettre ne peut pas être remplacée par elle-même, Turing et son équipe tentaient de placer ces termes dans un message intercepté, à un endroit où aucune lettre ne correspondait.

Exemple de message intercepté comparé à « Wetterbericht »

Message chiffré	A	E	T	V	K	L	Z	U	D	F	X	Y	A	A	L	M	J	N
Mot	W	E	T	T	E	R	B	E	R	I	C	H	T					

Exemple de message intercepté qui n'a aucun caractère en commun

Message chiffré	A	E	T	V	K	L	Z	U	D	F	X	Y	A	A	L	M	J	N
Mot				W	E	T	T	E	R	B	E	R	I	C	H	T		

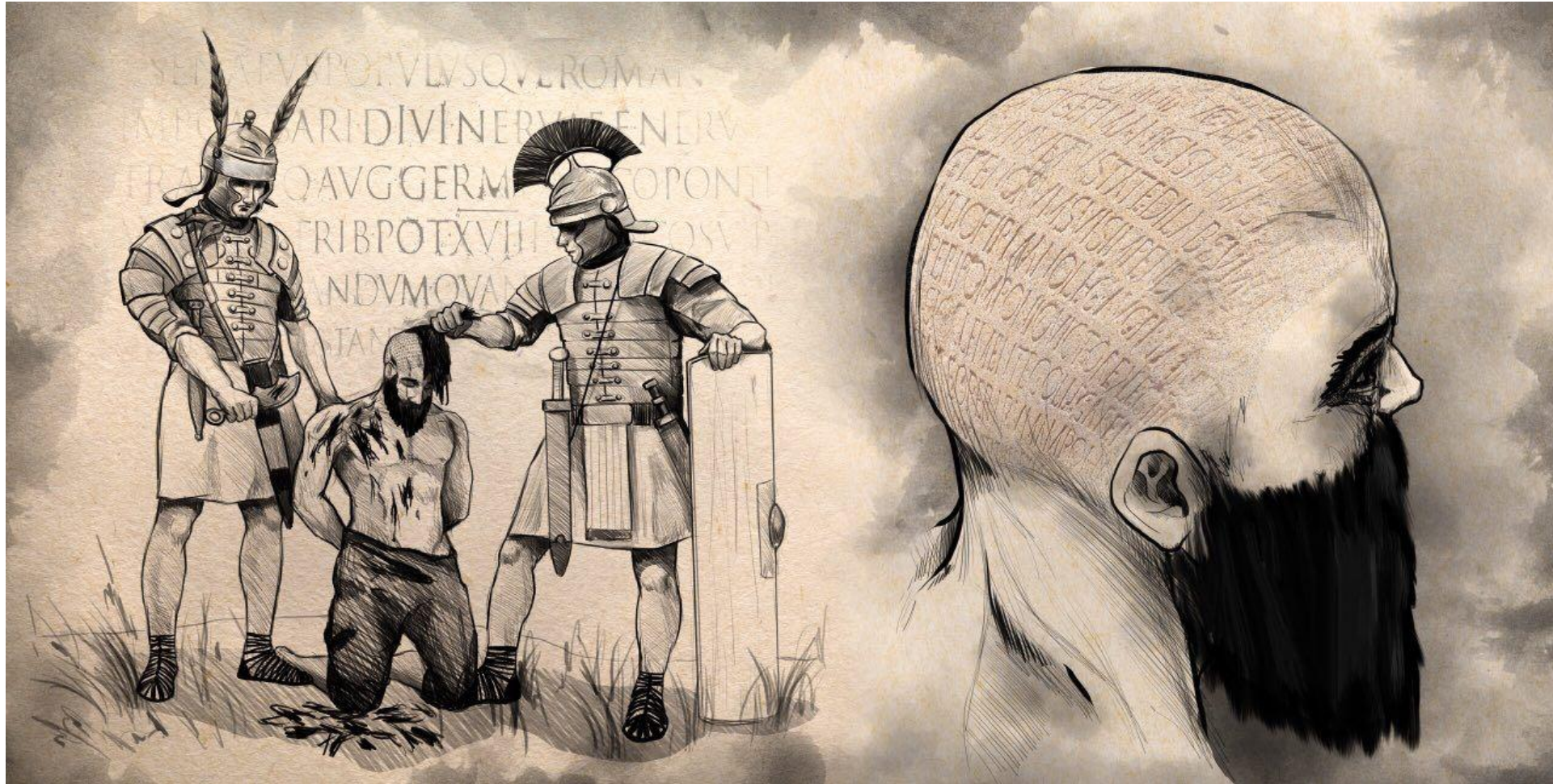
Grâce à cette découverte, un message qui aurait pris plusieurs millions d'années à être déchiffré a pu l'être en moins de vingt minutes

STEGANOGRAPHIE

La stéganographie (en grec «l'écriture couverte») cache les messages dans un support anodin:

- Encre invisible
- Gammes de musique
- Lettres de Georges Sand
- Images
- ...

Les premiers emplois attestés de la stéganographie se lisent chez Hérodote vers le Ve siècle avant Jésus-Christ: un certain Histiée, voulant prendre contact avec le tyran Aristagoras de Milet, choisit un esclave dévoué, lui rasa la tête, et y inscrivit le message à transmettre. Il attendit que ses cheveux repoussent pour l'envoyer à Aristagoras avec l'instruction de se faire raser le crâne.

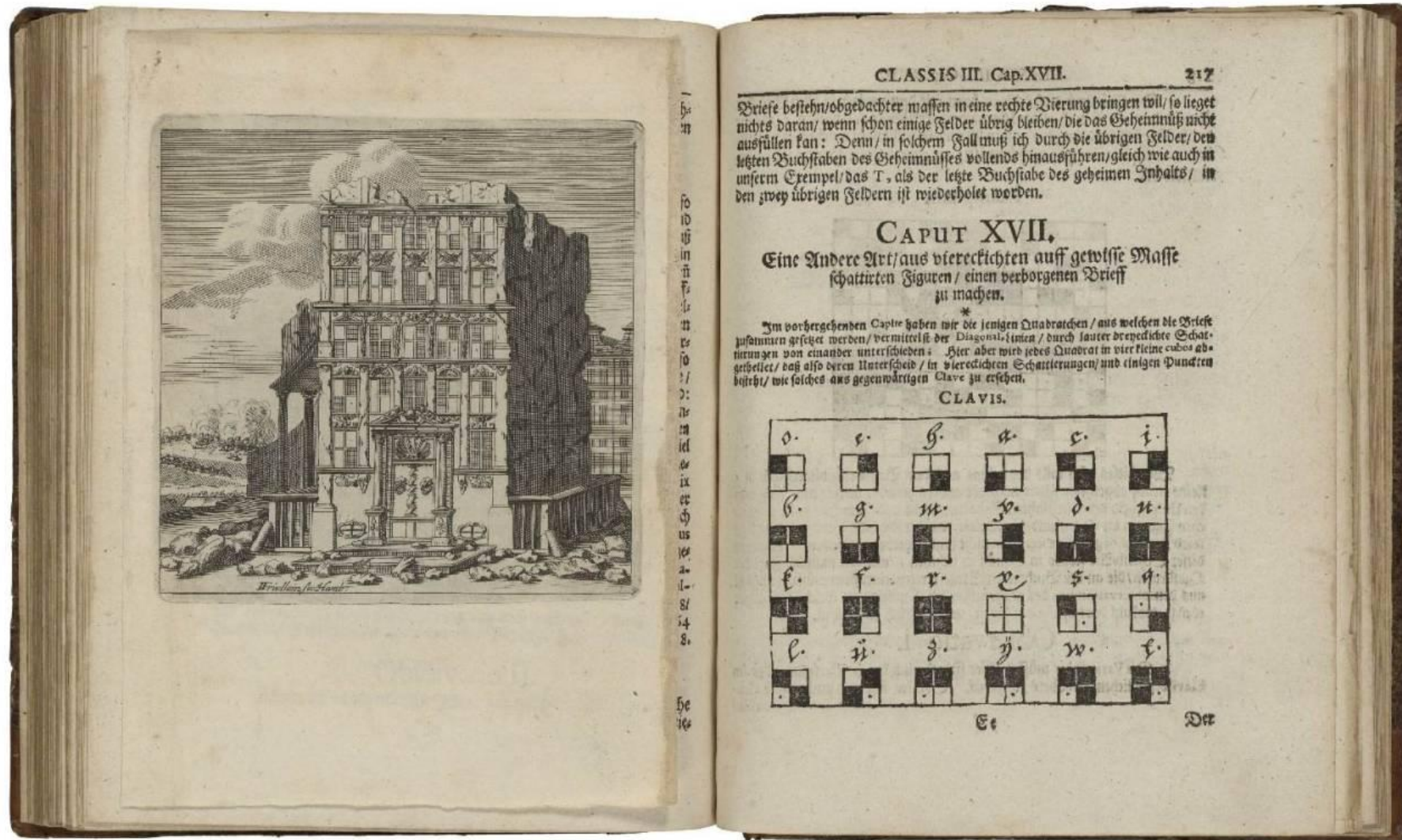


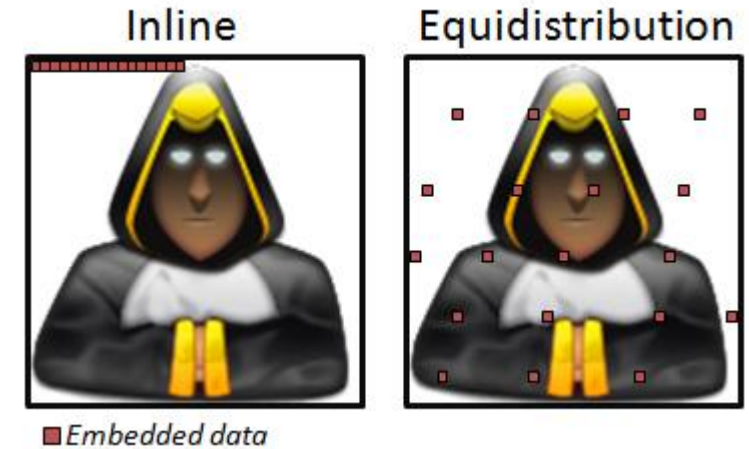
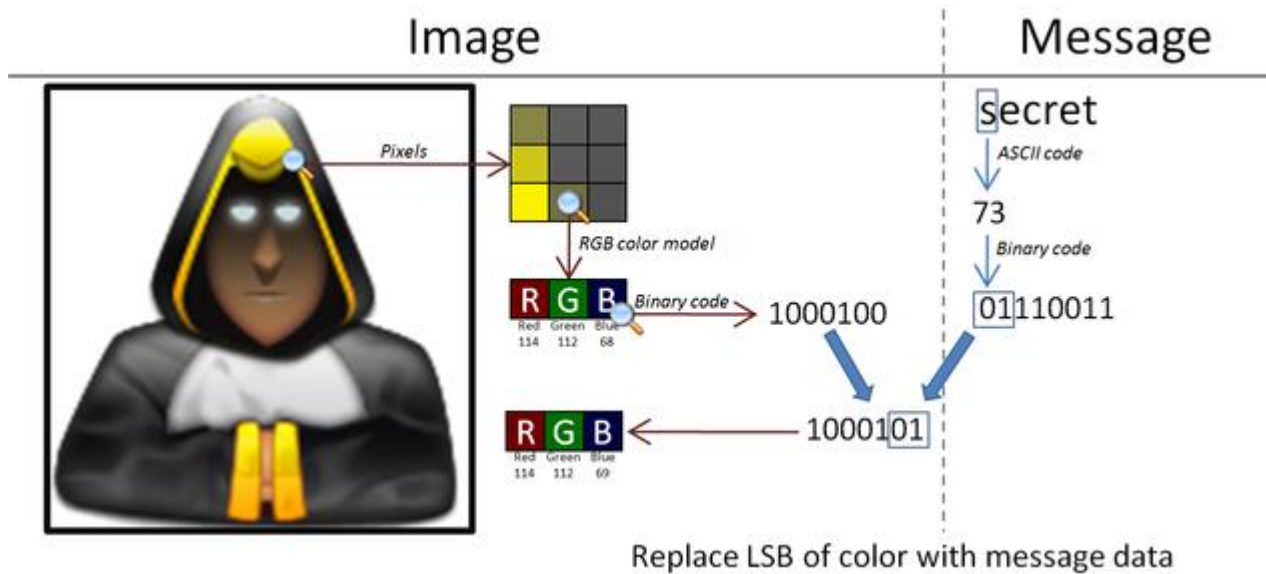
Toujours d'après Hérodote, pour informer les Spartiates de l'attaque imminente des Perses, un certain Démarate utilisa un élégant stratagème: il prit des tablettes, en racla la cire et grava sur le bois le message secret, puis il recouvrit les tablettes de cire. De cette façon, les tablettes, apparemment vierges, n'attirèrent pas l'attention.

En Chine ancienne, on écrivait les messages sur un fin ruban de soie dont on faisait une petite boule en l'englobant dans de la cire. Le messenger avalait ensuite cette boule, vous devinez la suite...

Au XVI^e siècle, le scientifique italien Giovanni Porta découvrit comment cacher un message dans un oeuf dur: il suffit d'écrire sur la coquille avec une encre contenant une once d'alun pour une pinte de vinaigre; la solution pénètre la coquille et dépose sur la surface du blanc d'oeuf le message que l'on lira aisément après avoir épluché l'œuf!

Dans son livre *Cryptographia oder Geheime Schrifft-münd-und Würrkliche Correspondendentz* (1684), Johannes Balthasar Friderici montre un dessin apparemment anodin (page de gauche) mais il contient un message secret. Ce dernier est codé par les fenêtres de l'immeuble : WIR HABEN KEIN PULVER MEHR (nous n'avons plus de poudre) avec la correspondance visible sur la page de droite :





En modifiant uniquement les bits de poids faible, la couleur du pixel ne change pas beaucoup.
Attention: cela ne fonctionne qu'avec des systèmes de codage de l'image sans compression ou avec compression sans perte d'information



Source:

<https://k-lfa.info/tools-stegano/>

SilentEye utilise la technique du bit de poids faible (LSB) pour cacher les informations dans la représentation RVB d'un pixel. Cette méthode est utilisable uniquement pour une image de type BMP, étant donné que ce format est non compressé et ne souffre donc pas de la problématique de perte de données.

<https://achorein.github.io/silenteYE/embedding/?i1s2>

Et il y en a bien d'autres pour crypter ou tester si un fichier contient un message:

- LSB-Steganographie : <https://github.com/RobinDavid/LSB-Steganography>
- StegoVeritas: Outil disposant de nombreuses fonctionnalités : vérifier les métadonnées, créer de nombreuses images transformées, Brute forces LSB, ... <https://github.com/bannsec/stegoVeritas>
- Zsteg : Permet de détecter le type de stéganographie du fichier (LSB openstego, Camouflage ...) <https://github.com/zed-0xff/zsteg>
- Stegdetect: Effectue des tests statistiques pour déterminer si un outil stego a été utilisé



Mieux vaut
prendre une image
sans à plat de
couleurs!



L'image de Mortimer et de Blake a été mélangée à l'image du message exactement de même taille.

Chaque pixel de la photo de départ comme de l'image à cacher contient 3 octets de couleur rouge-vert et bleu, chacun codé de 0 à 255 soit en hexadécimal de 00 à FF.

L'astuce que Mortimer a utilisée consiste à garder les 4 bits forts de chaque octet de couleur pour l'image qui va masquer l'image du message et de remplacer les 4 bits faibles par les 4 bits forts du pixel équivalent de l'image à masquer.

Voici un exemple : si le pixel $[i,j]$ de la photo de Blake et Mortimer a pour valeur décimale 180, soit la valeur B4 en hexadécimal, on ne retient que les bits de poids forts donc B et on met les autres à 0, soit au final B0.

Pour l'image à masquer, on prend le même pixel $[i,j]$ dont on ne retient également que les bits forts, qu'on ajoute du côté des bits faibles dans l'image codée. Ainsi, si la valeur décimale du pixel $[i,j]$ de l'image à cacher est 172, sa valeur hexa est AC, on retient les bits forts soit A, c'est à dire 0A.

Le codage de l'octet du pixel $[i,j]$ de l'image codée (mélange de la photo et du message donc) sera B0 + 0A, c'est à dire BA. Et voilà, il ne restait plus à Mortimer qu'à appliquer cette règle à tous les pixels (225x225) ... et à vous d'appliquer le système inverse pour obtenir l'image du message :

- Ouvrez le fichier secret nommé secret.bmp par exemple avec PIL et parcourez 1 à 1 tous les pixels
- Le même traitement a été appliqué aux 3 couleurs (RGB) pour garder les mêmes tonalités dans l'image.
- Astuce 1 : Le reste de la division modulo 16 d'un octet fournit la valeur des 4 bits de poids faible.
- Astuce 2 : Multiplier une valeur par 16 revient à la faire glisser de 4 rangs vers la gauche en binaire...
- Une fois calculée votre matrice finale 225x225 pixels correspondant à l'image cachée, sauver-là dans le format de votre choix et affichez-là à l'écran pour voir le message !

Exercice de stéganographie Blake & Mortimer

Encryptage

```
from PIL.Image import *

g1=open("image2.bmp") #image encore visible
g2=open("ima_decache01.bmp") #image cachée
(xmax,ymax)=g1.size

im=new('RGB',(xmax,ymax),(255,255,255))

for j in range(ymax):
    for i in range(xmax):
        c1=Image.getpixel(g1,(i,j))
        c2=Image.getpixel(g2,(i,j))
        r=16*int(c1[0]/16)+int(c2[0]/16)
        v=16*int(c1[1]/16)+int(c2[1]/16)
        b=16*int(c1[2]/16)+int(c2[2]/16)
        Image.putpixel(im,(i,j),(r,v,b))

im.save("ima_cache.jpg","JPEG")
im.save("ima_cache.png","PNG")
im.save("ima_cache.bmp","BMP")
im.save("ima_cache.gif","GIF")
```



Décryptage

```
from PIL.Image import *

g1=open("ima_cache04.bmp") #image codée
(xmax,ymax)=g1.size

im=new('RGB',(xmax,ymax),(255,255,255))

for j in range(ymax):
    for i in range(xmax):
        c1=Image.getpixel(g1,(i,j))
        r=16*(c1[0]%16)
        v=16*(c1[1]%16)
        b=16*(c1[2]%16)
        Image.putpixel(im,(i,j),(r,v,b))

im.save("ima_decache01.png","PNG")
```

la question 5
ne fait pas partie
de l'examen, vous ne
devez pas la faire.
Sa solution se trouve
sur Teams sous le
nom 5.py

En général on utilise plus une méthode basée sur le code aASCII de la lettre et plus douce par rapport au codage des couleurs (seulement 1 ou 2 LSB utilisés):

Pixel d'origine:

	R (rouge)	G (green)	B (blue)	Couleur
Pixel 1	17 -> 00010001	77 -> 01001101	217 -> 110110001	
Pixel 2	230 -> 11100110	226 -> 11100010	9 -> 00001001	
Pixel 3	128 -> 10000000	9 -> 00001001	230 -> 11100110	

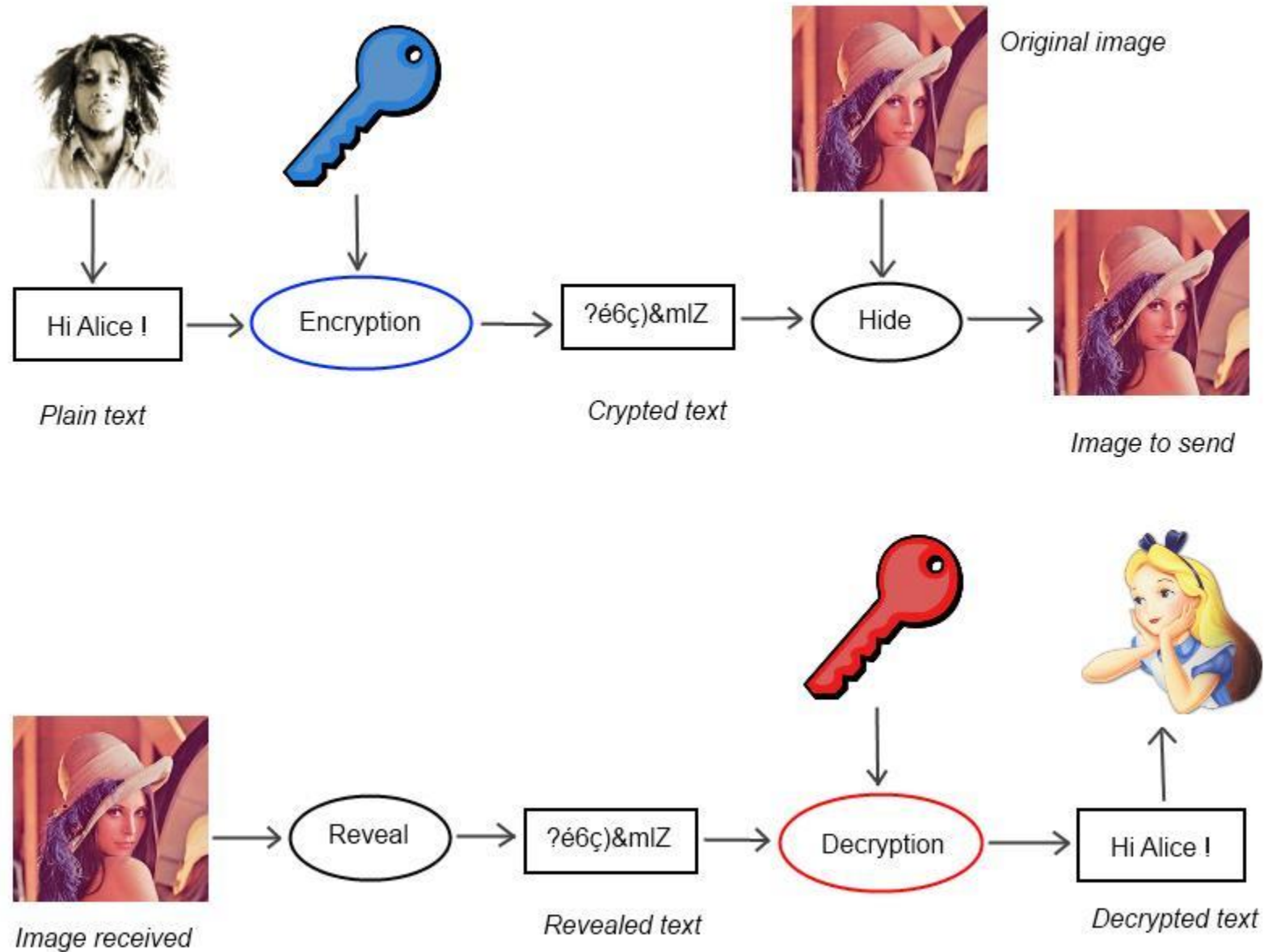
Si l'on modifie seulement la valeur du bit de poids faible de chaque composante (le bit le plus à droite de l'octet), l'incidence sur le code couleur sera mineure. En utilisant le codage ASCII les lettres majuscules sont codées entre 65 et 90. La lettre A correspond au code 65 soit 01000001. Modifions les bits de poids faible de l'exemple précédent pour dissimuler un A dans l'image :

Pixels encodés:

	R (rouge)	G (green)	B (blue)	Couleur
Pixel 1	17 -> 00010000	77 -> 01001101	217 -> 110110000	
Pixel 2	230 -> 11100110	226 -> 11100010	9 -> 00001000	
Pixel 3	128 -> 10000000	9 -> 00001001	230 -> 11100110	

Les modifications de couleur sont invisibles à l'œil nu mais une personne avertie sera capable de récupérer l'information. Ce type de technique peut être employé pour la diffusion de malwares (LokiBot par exemple) et permet de tromper la vigilance des antivirus.

Et on peut combiner la stéganographie avec la cryptographie...



On peut aussi cacher un message texte ou image dans un fichier audio banal...

Audiostego (hideme)

Outil permettant de cacher/trouver de la donnée cachée dans un fichier WAV/MP3

<https://github.com/danielcardenas/AudioStego>

spectrology

Encoder une image dans un spectrogramme audio

<https://github.com/solusipse/spectrology>

L'encre magique

Le plus connu des procédés de stéganographie est sans doute l'utilisation d'encres sympathiques, mentionnée par Pline l'Ancien dès le 1er siècle avant J.-C. On écrit, au milieu des textes écrits à l'encre, un message à l'aide de jus de citron, de lait, de certains produits chimiques, ou même d'urine! Il est invisible à l'oeil, mais une simple flamme, ou un bain dans un réactif chimique, révèle le message. L'exemple suivant a été réalisé à l'aide de lait :



L'historien de la Grèce Antique Enée le Tacticien imagina d'envoyer un message secret en piquant de **minuscules trous sous certaines lettres d'un texte anodin**. La succession de ces lettres fournit le texte secret.

Deux mille ans plus tard, les épistoliers anglais employèrent la même méthode, non pour assurer le secret à leurs envois, mais pour éviter de payer des taxes excessives. En effet, avant la réforme du service postal, dans les années 1850, envoyer une lettre coûtait environ un shilling tous les cent miles, ce qui était hors de portée de la plupart des gens, mais les journaux ne payaient pas de taxe. Grâce aux piqûres d'épingles, les Anglais malins pouvaient envoyer leurs messages gratuitement.

Ce procédé a été aussi utilisé par les Allemands pendant la première guerre mondiale. Au cours de la seconde guerre mondiale, ils améliorèrent le procédé en **cochant les lettres de journaux avec de l'encre sympathique**.

Un texte apparemment innocent peut aussi révéler un message important. Voici un exemple d'un tel message, envoyé par un espion allemand pendant la seconde guerre mondiale:

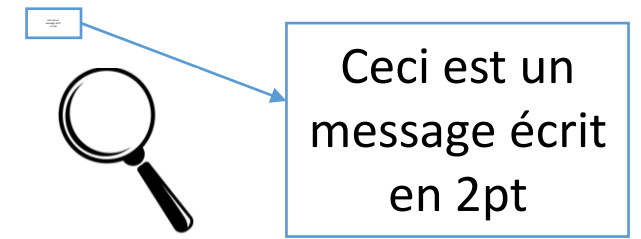
« *Apparently neutral's protest is thoroughly discounted and ignored. Isman hard it. Blockade issue affects pretext for embargo on byproducts, ejecting suets and vegetable oils.* »

(Apparemment la protestation des pays neutres est totalement ignorée. Isman frappe fort. L'issue du blocus donne des prétextes pour un embargo sur certains produits, mis à part graisses animales et huiles végétales.)

Si l'on prend la deuxième lettre de chaque mot, le message suivant émerge:

Pershing sails from NY June 1. ([Le Pershing part de New-York le 1er juin](#))

Les espions allemands de la deuxième guerre mondiale utilisaient aussi des micropoints pour faire voyager discrètement leurs informations. C'est une photographie de la taille d'un point de ponctuation, qu'il suffit d'agrandir pour voir apparaître clairement le message (c'est une sorte de microfilm). Ce micropoint pouvait être inséré dans une lettre anodine, parfois sous un timbre, etc.



Ressources

<http://nomis80.org/cryptographie/cryptographie.html>

<http://www.01adfm.com/win-xp/hacking/Hacking08.htm>

<http://www.01adfm.com/win-xp/hacking/Hacking07.htm>

<http://perso.clubinternet.fr/guidovdi/codes/lapagecryptologie.htm>

http://www.protechnix.com/information/crypto/pages/vernam_base.html

<http://www.chez.com/nopb/crypto2.html#transposition>

<http://jf.morreeuw.free.fr/vigenere/vigenere.html>

http://www.pro-technix.com/information/crypto/pages/vernam_base.html

<https://ecolepratique.com/steganographie-introduction-a-la-dissimulation-de-donnees/>

<https://fr.wikipedia.org/wiki/St%C3%A9ganographie>

Et pour terminer, voici les messages
cachés dans les échanges entre
Georges Sand et Musset...

Lettre d'Alfred de Musset à Georges Sand:

Quand je mets à vos pieds un éternel hommage,
Voulez-vous qu'un instant je change de visage ?
Vous avez capturé les sentiments d'un coeur
Que pour vous adorer forma le créateur.
Je vous chéris, amour, et ma plume en délire
Couche sur le papier ce que je n'ose dire.
Avec soin de mes vers lisez les premiers mots,
Vous saurez quel remède apporter à mes maux.

Et la réponse de Georges Sand à Alfred de Musset:

Cette insigne faveur que votre coeur réclame
Nuit à ma renommée et répugne à mon âme.

Lettre d'Alfred de Musset à Georges Sand:

Quand je mets à vos pieds un éternel hommage,
Voulez-vous qu'un instant je change de visage ?
Vous avez capturé les sentiments d'un coeur
Que pour vous adorer forma le créateur.
Je vous chéris, amour, et ma plume en délire
Couche sur le papier ce que je n'ose dire.
Avec soin **de mes vers lisez les premiers mots**,
Vous saurez quel remède apporter à mes maux.

Et la réponse de Georges Sand à Alfred de Musset:

Cette insigne faveur que votre coeur réclame
Nuit à ma renommée et répugne à mon âme.