Cryptographie

Cours 1

Maîtriser les concepts
et
algorithmes cryptographiques

Jérémy Briffaut
STI 2A

Plan

I. Histoire, définition et objectifs de la cryptographie
- Concepts et algorithmes de permutation et de substitution

II. Chiffrement Symétrique
- DES, 3DES, AES, IDEA

III. Chiffrement Asymétrique
- RSA, ElGamal

IV. Signature, Hachage et Scellement

V. Echange de clés
- Algorithme Deffie-Hellman

VI. Hachage : MD5, SHA-1, SHA-2

VII. Code d'Authentification & MAC

Objectifs de ce cours

- Maîtriser les concepts et algorithmes cryptographiques
- Introduire les bases de la cryptographie
- Comprendre les principes de bases

Services à assurer sur un hôte

- ***Disponibilité*** : garantie de la continuité du service.
- ***Intégrité*** : garantie que l'information n'est pas altérée.
- ***Confidentialité*** : garantie que de l'information n'est pas divulguée à des tiers non autorisés (frauduleusement ou non)

Services à assurer sur le réseau

- ***authentification*** : garantie de l'origine des données
- ***Intégrité*** : garantie que l'information n'est pas altérée.
- ***Confidentialité*** : garantie que de l'information n'est pas divulguée à des tiers non autorisés
- ***Disponibilité*** : garantie que l'information est disponible (dénie de service)
- ***Non répudiation*** :
  - Ensemble de moyens et techniques permettant de prouver la participation d'une entité dans un échange de données

Plan

I. Histoire, définition et objectifs de la cryptographie
   I. Définition
   II. Transposition, Substitution
   III. Cryptographie moderne

- Terminologie
  - **Cryptographie** : Science mathématique permettant d'effectuer des opérations sur un texte intelligible afin d'assurer une ou plusieurs propriétés de la sécurité de l'information
  - **Cryptanalyse** : la Science permettant d'étudier les systèmes cryptographiques en vue de les tester ou de les casser
  - **Cryptologie** = cryptographie + cryptanalyse

Assurer la confidentialité

- **Stéganographie** : écriture couverte

  - Information non-chiffrée

    Connaissance de l'existence de l'information

    =

    Connaissance de l'information

- Exemples :
  - Message *couvert :*
    - Tablette couverte de cire
    - Crane du messager
  - Message *invisible*
    - Encre invisible (Pline – 1er siècle avant JC)
  - Message *illicible*
    - Micro-film sous la forme d'un point

- **Cryptographie** : écriture cachée/brouillée

  - Information chiffrée

Connaissance de l'existence de l'information
≠
Connaissance de l'information

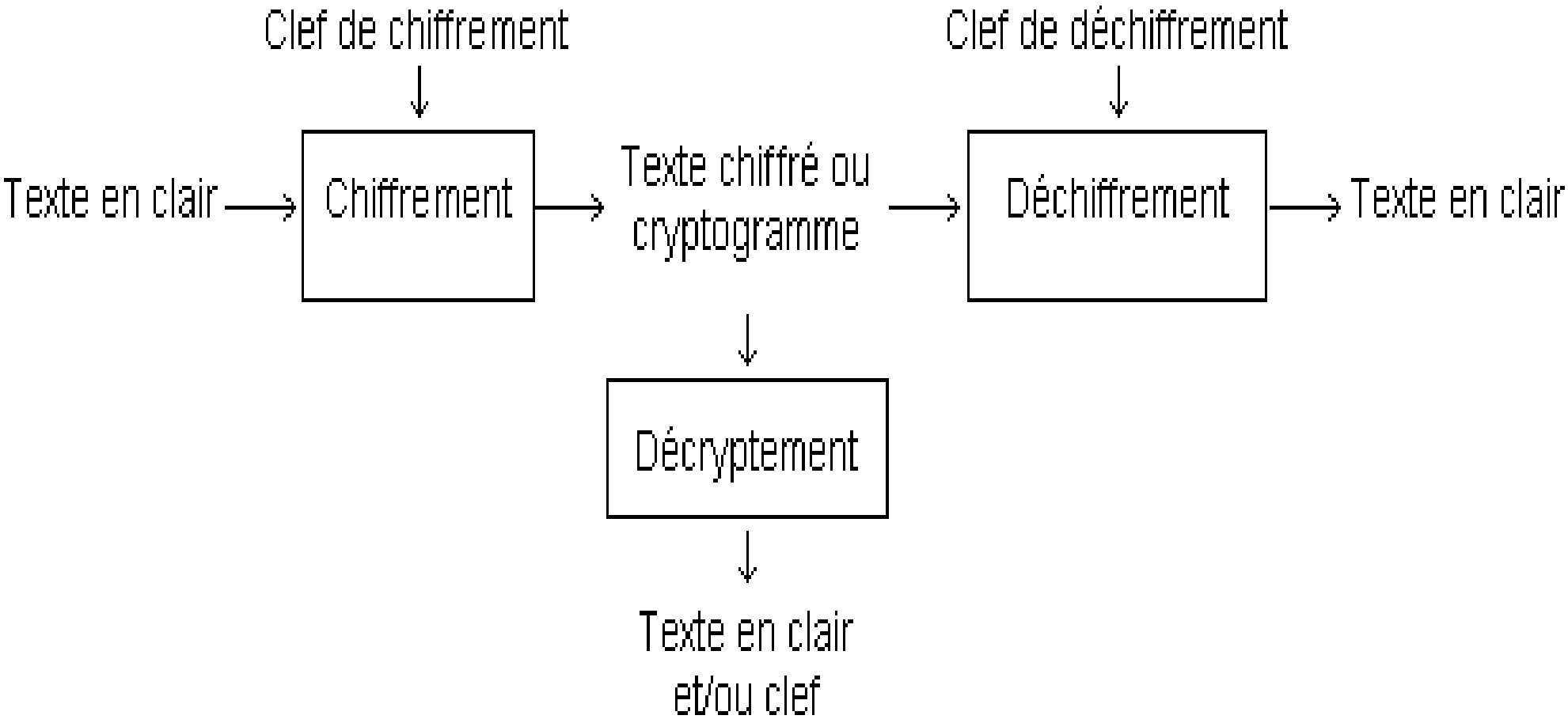
- Le Chiffrement
  - Ces algorithmes assurent la transformation d'un message en **clair** ("*plaintext*") en un message **brouillé** ("*ciphertext*")
  - Il existe deux grandes familles d'algorithmes
    - → algorithmes **symétriques**
      - imposent au système qui crypte de savoir décrypter
    - → algorithmes **asymétriques**
      - ne permettent pas au système qui crypte de décrypter
  - Pour les deux cas les algorithmes de chiffrement sont commutatifs.

Confusion et Diffusion

- Confusion :
  - Aucune propriété statistique ne peut être déduite du message chiffré.

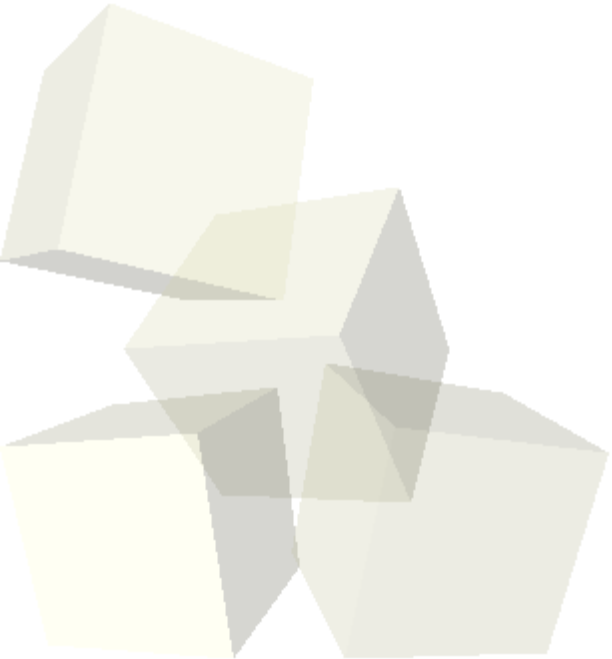
- Diffusion :
  - Toute modification du message en clair se traduit par une modification complète du chiffré.

- Chiffrement, déchiffrement et décryptement :

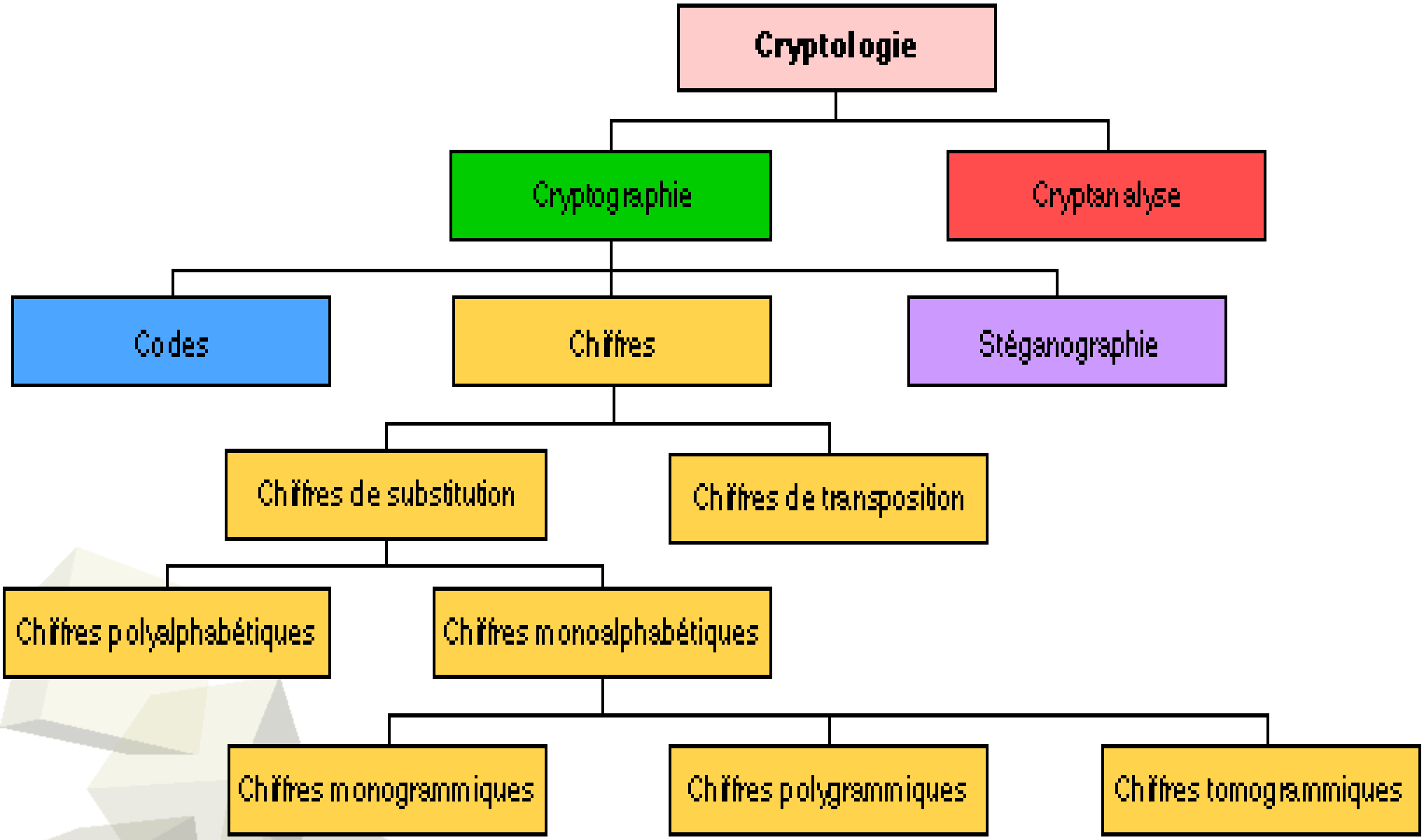


Plan

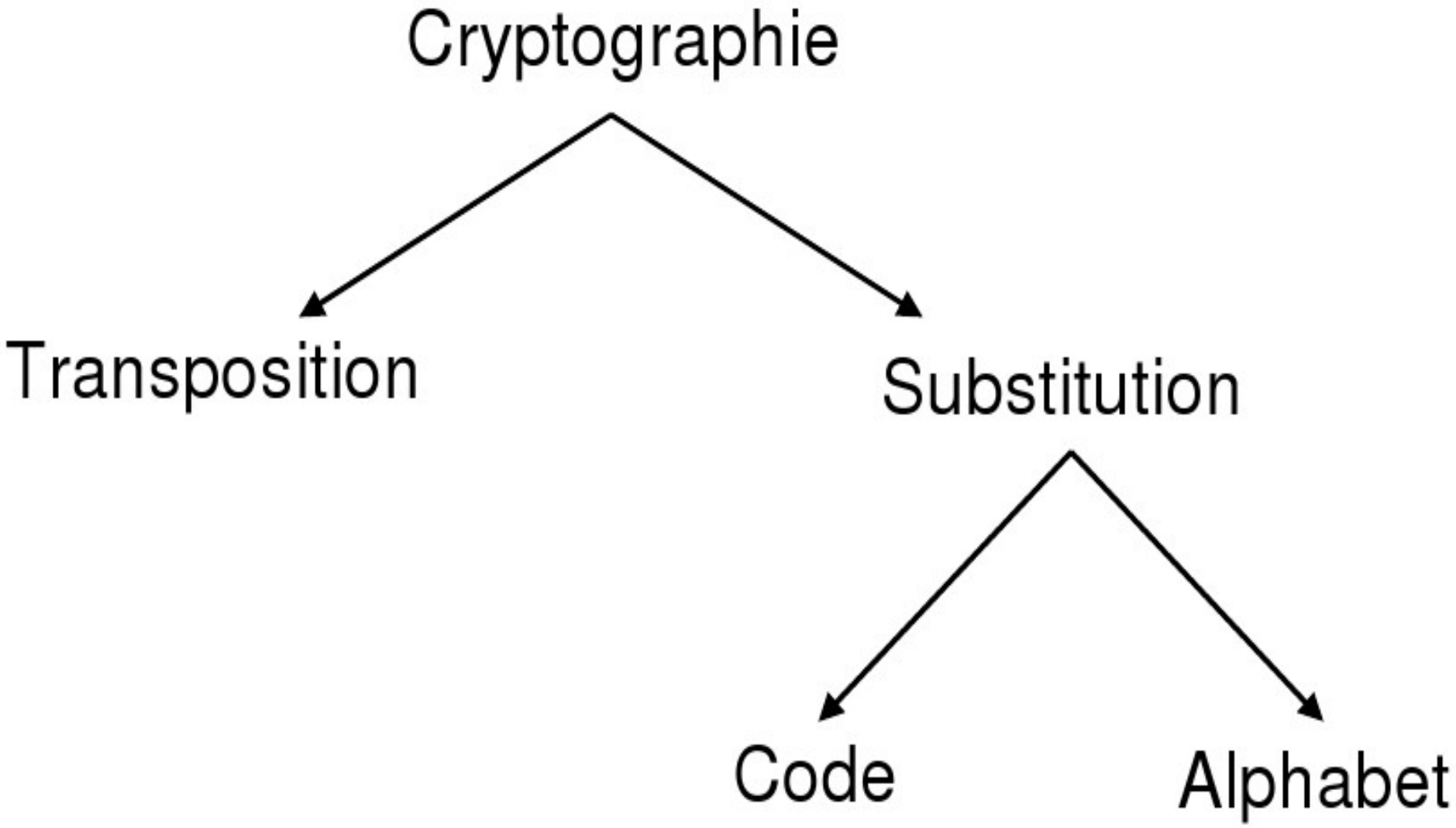
I. Histoire, définition et objectifs de la cryptographie
   I. Définition
   II. Transposition, Substitution
   III. Cryptographie moderne



Taxonomie


Cryptologie
Cryptographie
Cryptanalyse
Codes
Chiffres
Stéganographie
Chiffres de substitution
Chiffres de transposition
Chiffres polyalphabétiques
Chiffres monoalphabétiques
Chiffres monogrammiques
Chiffres polygrammiques
Chiffres tomogrammiques

Cryptographie ancienne



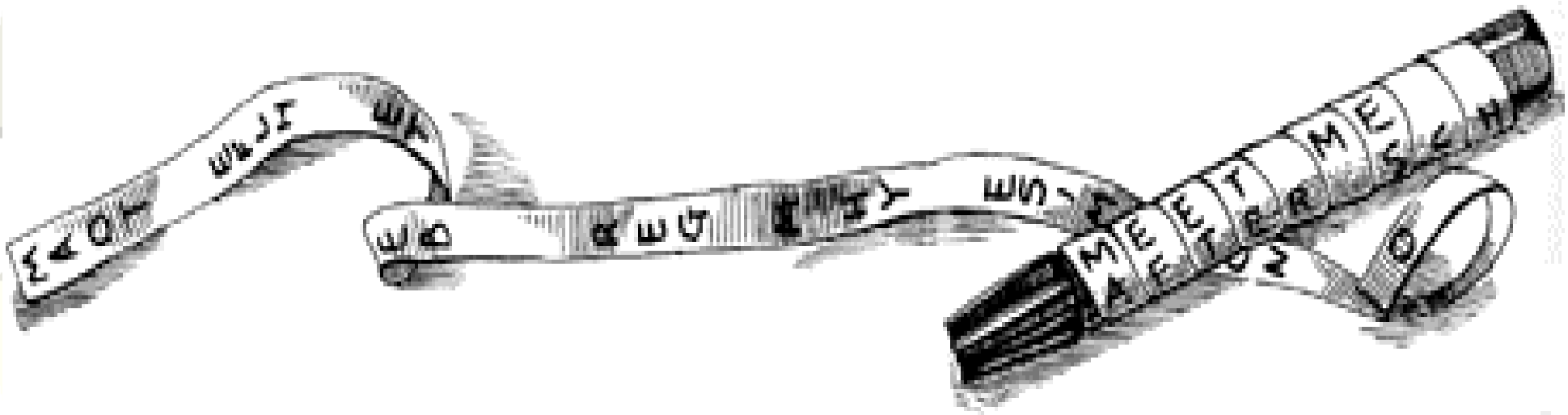
Transposition

- Chiffrement type anagramme.
  - Niveau de sécurité théorique :
    - ➔ Message de 35 lettres : 35! chiffrés possibles.

- Problèmes :
  - Confusion sur la syntaxe mais ...
  - ... chaque lettre conserve sa valeur.
  - Clé de chiffrement «complexe».

Exemple de transposition

- **La *scytale* spartiate (5 siècle av. JC) :**
  - premier dispositif de cryptographie militaire connu
  - un bâton de bois autour duquel est entourée une bande de cuir
  - L'expéditeur
    - écrit son message sur toute la longueur de la scytale
    - déroule ensuite la bande
      - apparaît alors couverte d'une suite de lettres sans signification
  - Le messager
    - emportera la bande de cuir, l'utilisant comme ceinture, les lettres tournées vers l'intérieur.
  - Le destinataire
    - enroulera alors cette bande sur son bâton (de même diamètre) pour lire le message clair.



Exemple de transposition

Rail Fence

- se traduit littéralement "palissade"
- connaît son heure de gloire aux débuts de la cryptographie

- Exemple
  - le message VIENS ME REJOINDRE A CINQ HEURES.
  - Le Rail Fence à deux niveaux dispose les lettres en «zig zag»

V E S E E O N R A I Q E R S
I N M R J I D E C N H U E

  - Nous obtenons alors :

VESEE ONRAI QERSI NMRJI DECNH UE

  - à trois niveaux:

V S E N A Q R
I N M R J I D E C N H U E
E E O R I E S

  - Nous obtiendrons alors :

VSENA QRINM RJIDE CNHUE EEORI ES.

Substitution

- Chiffrement en changeant d'alphabet.
  - Kama Sutra : *mlecchita-vikalpà* ou art de l'écriture secrète (4ème siècle av JC).
- Niveau de sécurité *théorique* :
  - Alphabet à 26 lettres : 26! alphabets possibles.
- Problèmes :
  - Confusion sur l'alphabet mais ...
  - *... chaque lettre conserve sa place d'origine.*
- Exemples :
  - substitutions simples (monoalphabétiques)
    - chiffre Pig Pen, le carré de Polybe, le chiffre Atbash, le chiffre de César, les alphabets désordonnés, le chiffre affine ...
  - substitutions polyalphabétiques (à double clef ou à alphabets multiples)
    - le chiffre de Vigenère, le chiffre de Gronsfeld, le cylindre de Jefferson, la machine Enigma ...
  - substitutions polygrammiques (polygraphiques)
  - des substitutions tomogrammiques (par fractions de lettres)

Exemple de substitution

- **Le chiffre de César**
  - consiste simplement à décaler les lettres de l'alphabet de quelques crans vers la droite ou la gauche.
  - Substitution monoalphabétiques

- Exemple
  - décalons les lettres de 3 rangs vers la gauche, comme le faisait Jules César (d'où le nom de ce chiffre):

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

  - Ainsi, le message

    Ave Caesar morituri te salutant

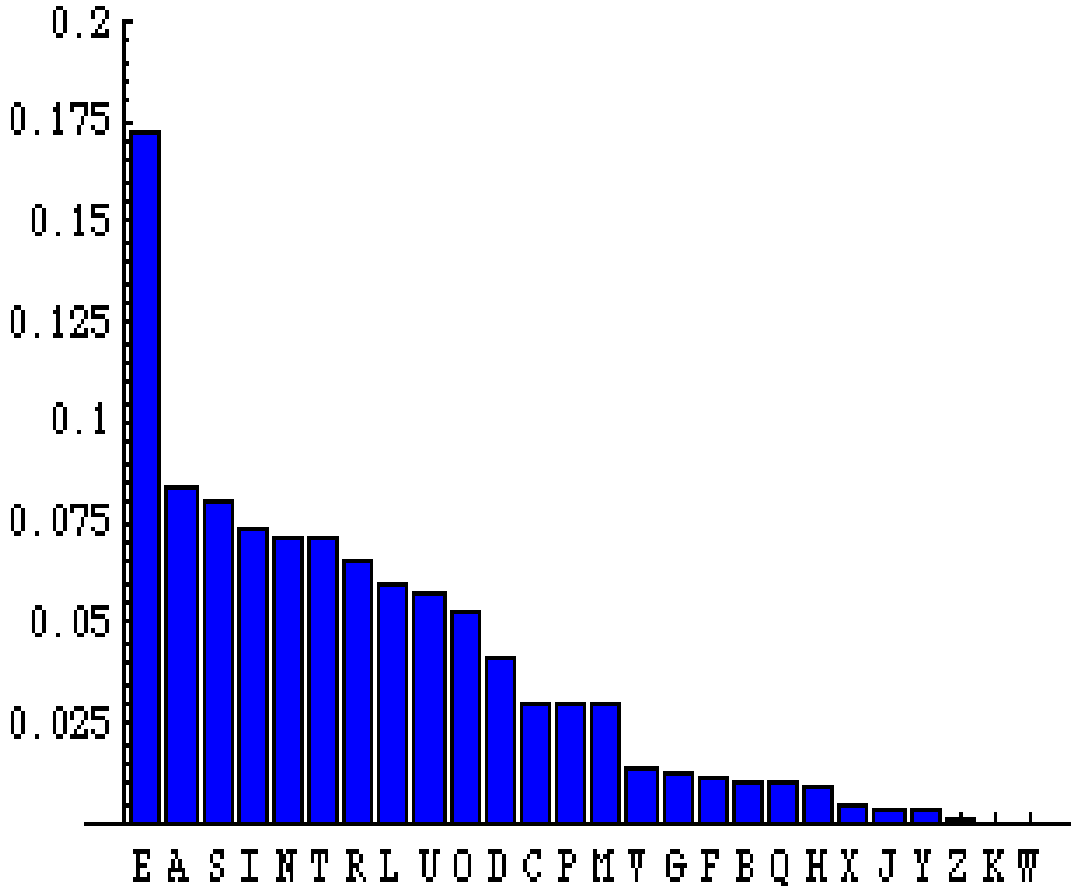
  - devient

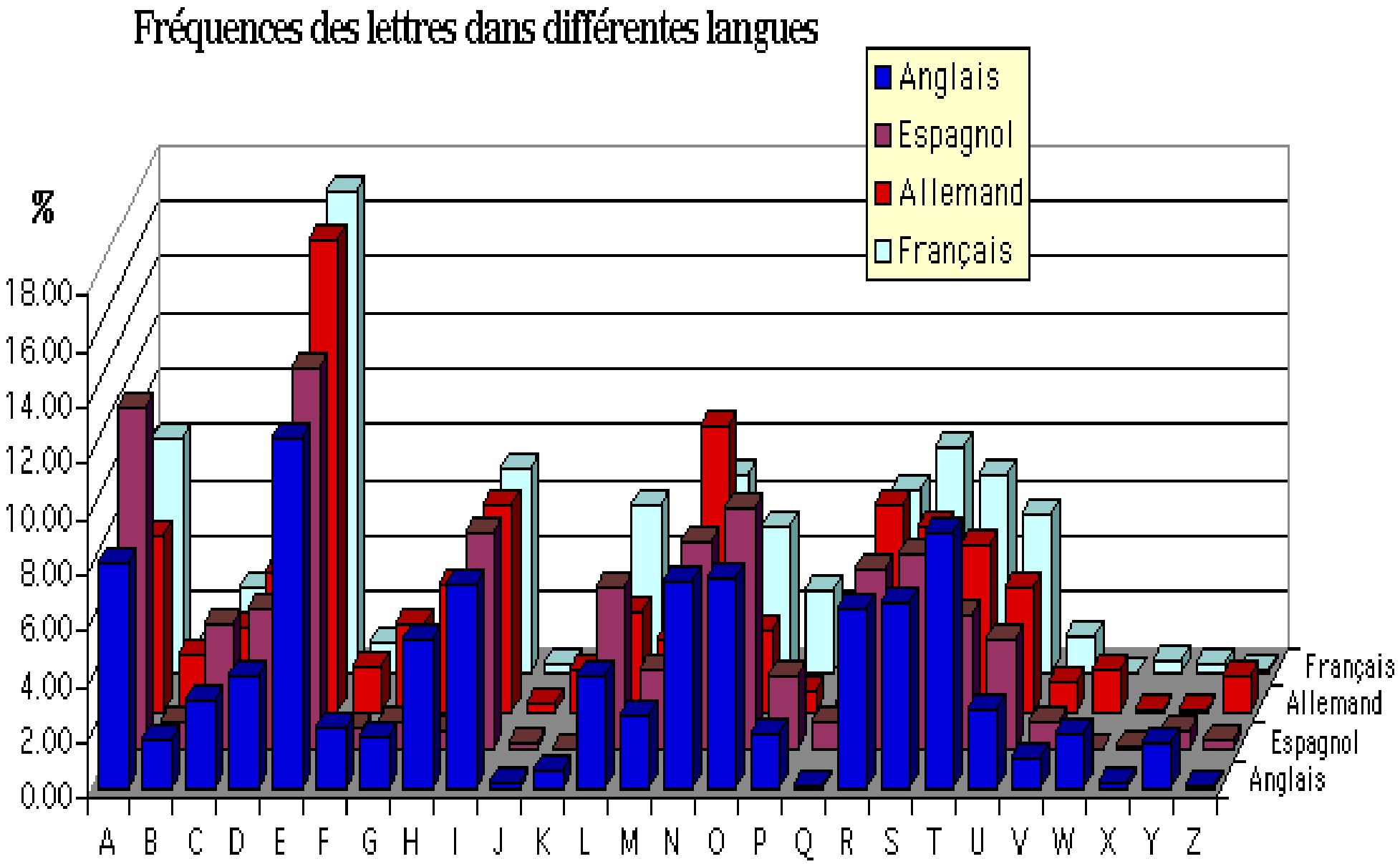
    DYHFD HVDUP RULWX ULWHV DOXWD QW

Cryptanalyse de la substitution monoalphabétique

- **Principe (Al-Kindi - 9ème siècle) :**
  - analyse des fréquences
  - ne fonctionne bien que si le cryptogramme est suffisamment long pour avoir des moyennes significatives.

Lettre	Fréquence	Lettre	Fréquence
A	8.40 %	N	7.13 %
B	1.06 %	O	5.26 %
C	3.03 %	P	3.01 %
D	4.18 %	Q	0.99 %
E	17.26 %	R	6.55 %
F	1.12 %	S	8.08 %
G	1.27 %	T	7.07 %
H	0.92 %	U	5.74 %
I	7.34 %	V	1.32 %
J	0.31 %	W	0.04 %
K	0.05 %	X	0.45 %
L	6.01 %	Y	0.30 %
M	2.96 %	Z	0.12 %




Fréquences des lettres dans différentes langues
Anglais
Espagnol
Allemand
Français
%
18.00
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
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
Français
Allemand
Espagnol
Anglais

Substitutions polyalphabétiques

- Utilisent plusieurs "alphabets"
  - ce qui signifie qu'une même lettre peut être remplacée par plusieurs symboles
  - Exemples :
    - chiffre de Vigenère
      - qui résista aux cryptanalystes pendant trois siècles
    - Des exemples plus récents s'inspirant de ce chiffre :
      - le chiffre de Beaufort
      - le chiffre de Gronsfeld
      - le cylindre de Jefferson
      - la machine Enigma
- La substitution homophonique
  - consiste à remplacer chaque lettre par un nombre de symboles proportionnel à sa fréquence d'apparition est une sous-catégorie.

Chiffre de Vigenère

- Chiffre de Vigenère
  - 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.
  - On peut résumer ces décalages avec un 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).

- La grande force du chiffre de Vigenère est que la même lettre sera chiffrée de différentes manières.
  - Par exemple le E du texte clair suivant a été chiffré successivement M V L P I, ce qui rend inutilisable l'analyse des fréquences classique.

Chiffre de Vigenère

- Exemple :
  - chiffrons 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

- Définition de la clé de chiffrement :
  - *Mot-clé* identifiant les alphabets à utiliser.

Carré de Vigenère

	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

Substitution polyalphabétique

- Confusion et Diffusion ?
  - *Idem substitution monoalphabétique*
- Confusion :
  - Confusion sur l'alphabet mais …
  - … analyse fréquentielle des lettres.
- Diffusion :
  - Pas du tout assurée.

Cryptanalyse de la substitution polyalphabétique

- C. Babbage (19ème siècle)
  - Principe en deux étapes :
    - Trouver la longueur du mot-clé.
    - Analyse fréquencielle sur chacun des alphabets.

- Longueur du mot clé :

h i v e r h i v e r h i v e r h i v e r
K E Y K E Y K E Y K E Y K E Y K E Y K E
R M T O V F S Z C B L G F I P R M T O V

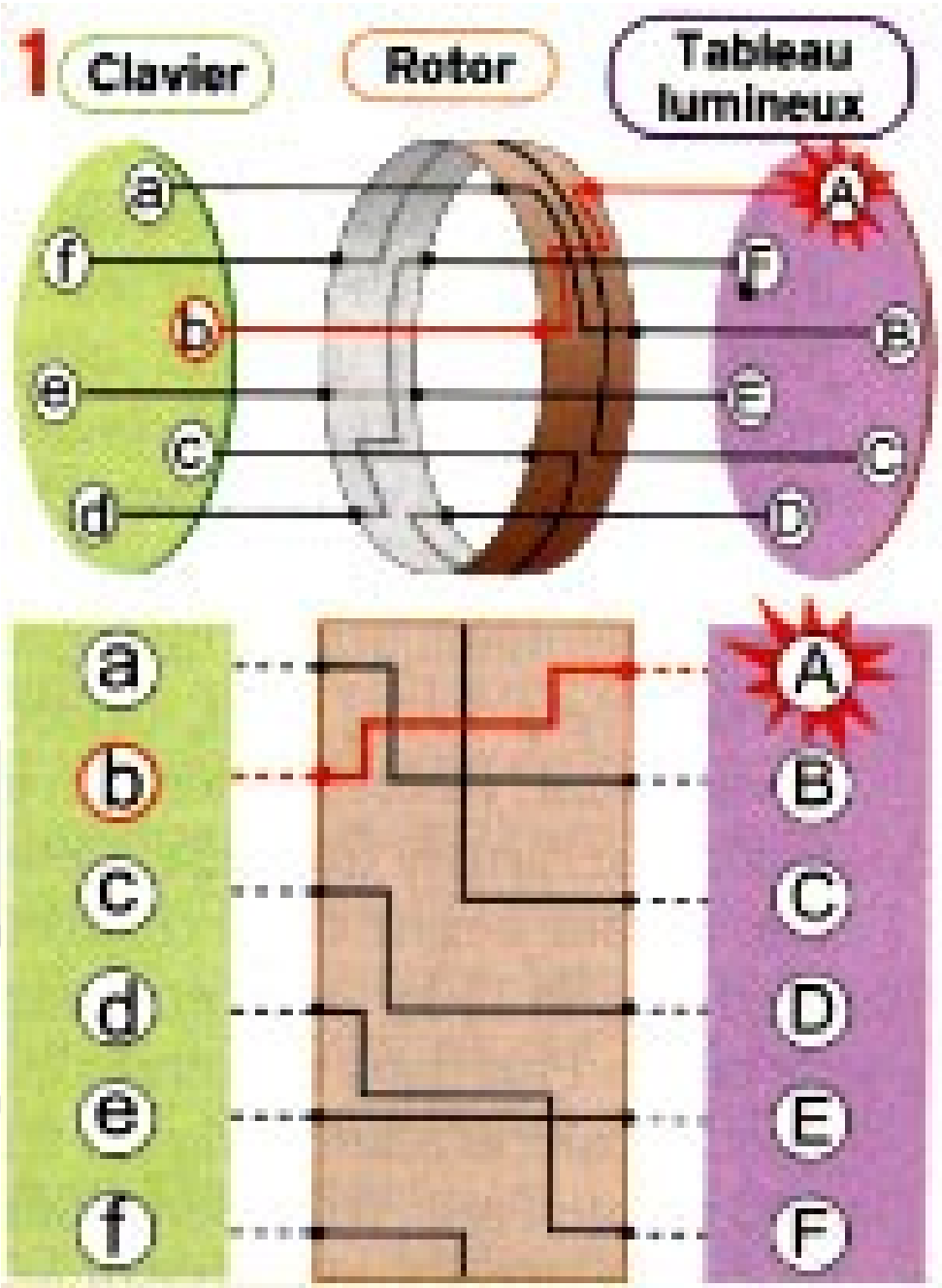
Extension à la substitution polyalphabétique

- Faiblesse de la substitution :
  - Taille du mot-clé : un digramme peut être chiffré plusieurs fois *de la même manière*.
- Idées :
  - *Utilisation de plus d'alphabets de chiffrement.*
  - Choisir des mot-clés plus grand.

Enigma

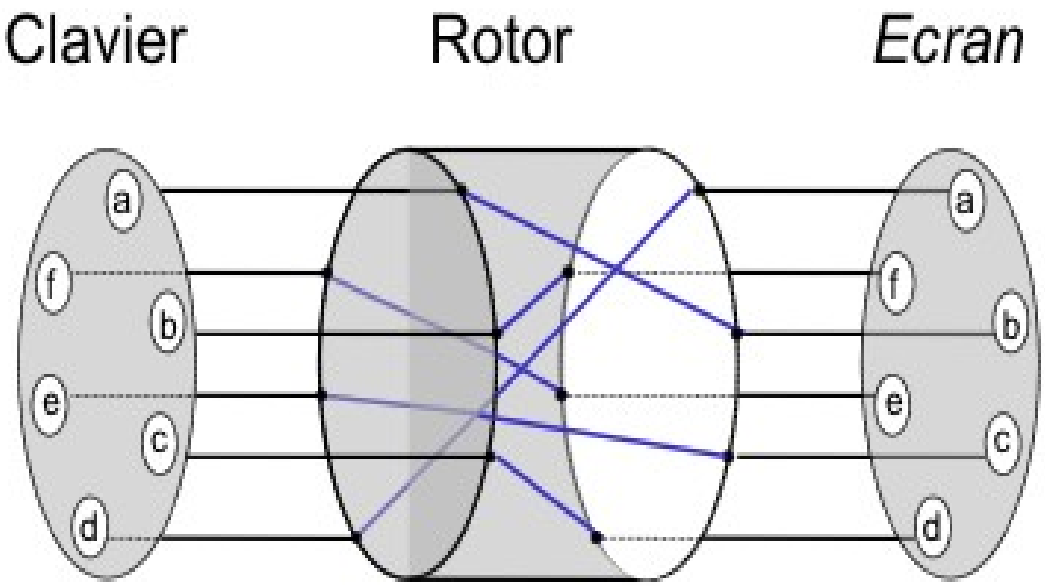
- La machine à chiffrer et déchiffrer qu'utilisèrent les armées allemandes du début des années trente jusqu'à la fin de Seconde Guerre Mondiale.
- Automatise le chiffrement par substitution.
- Principes de base :
  - Substitution polyalphabétique
- Techniques utilisées :
  - Rotors = substitutions poly-alphabétiques.
  - Connector = substitution

Enigma - Rotor

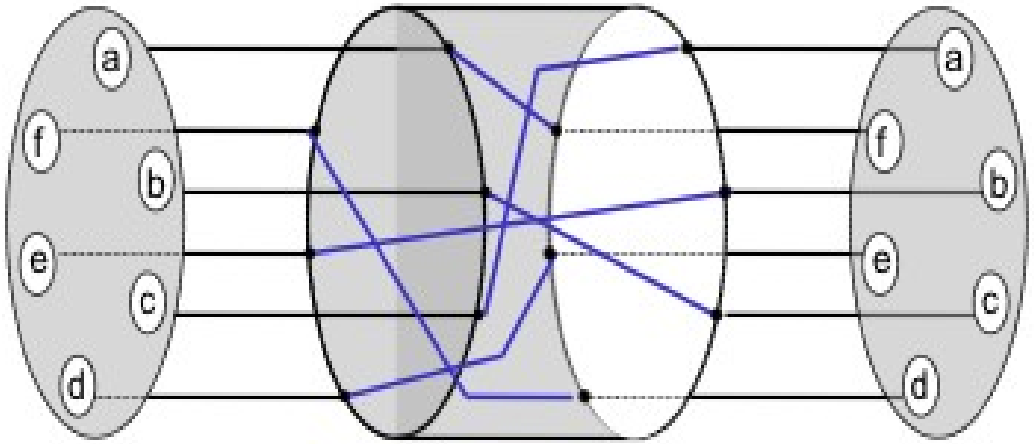


- Substitution Polyalphabétique
  - Si on frappe la lettre b sur le clavier, un courant électrique est envoyé dans le rotor, suit laecâblage interne, puis ressort à droite pour allumer la lettre A sur le tableau lumineux. B est donc chiffré en A *(B-> A)*.

Substitution Polyalphabétique

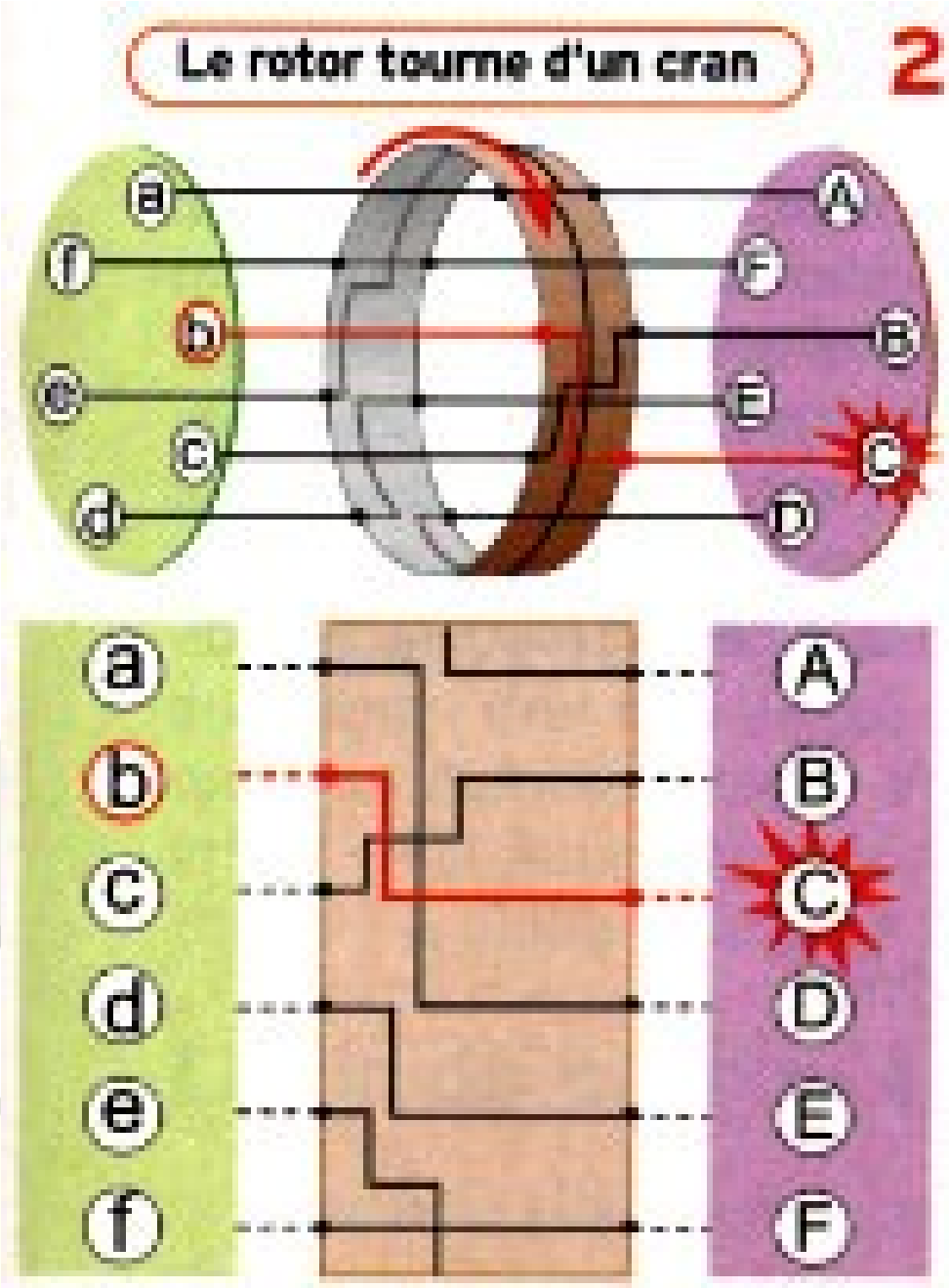


a	b	c	d	e	f
B	F	D	A	C	E



a	b	c	d	e	f
F	C	A	E	B	D

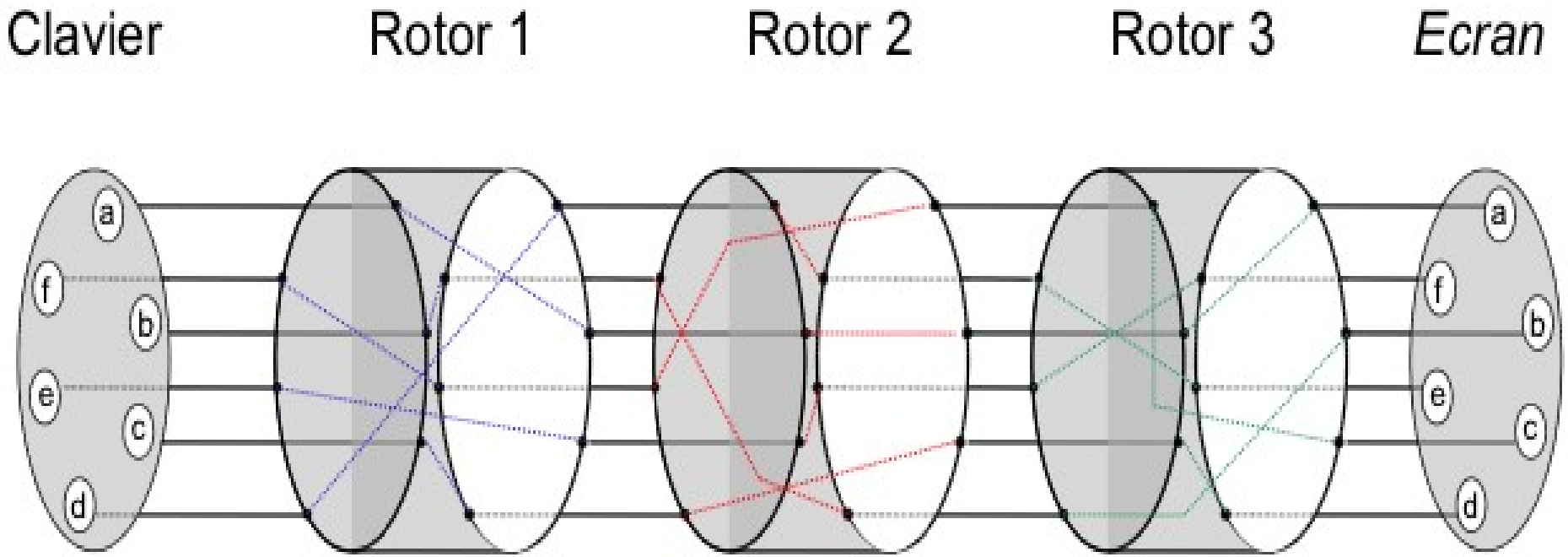
Enigma - Rotor



- 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 mot BAC est chiffré ADD (et non ABD si le rotor était resté immobile).

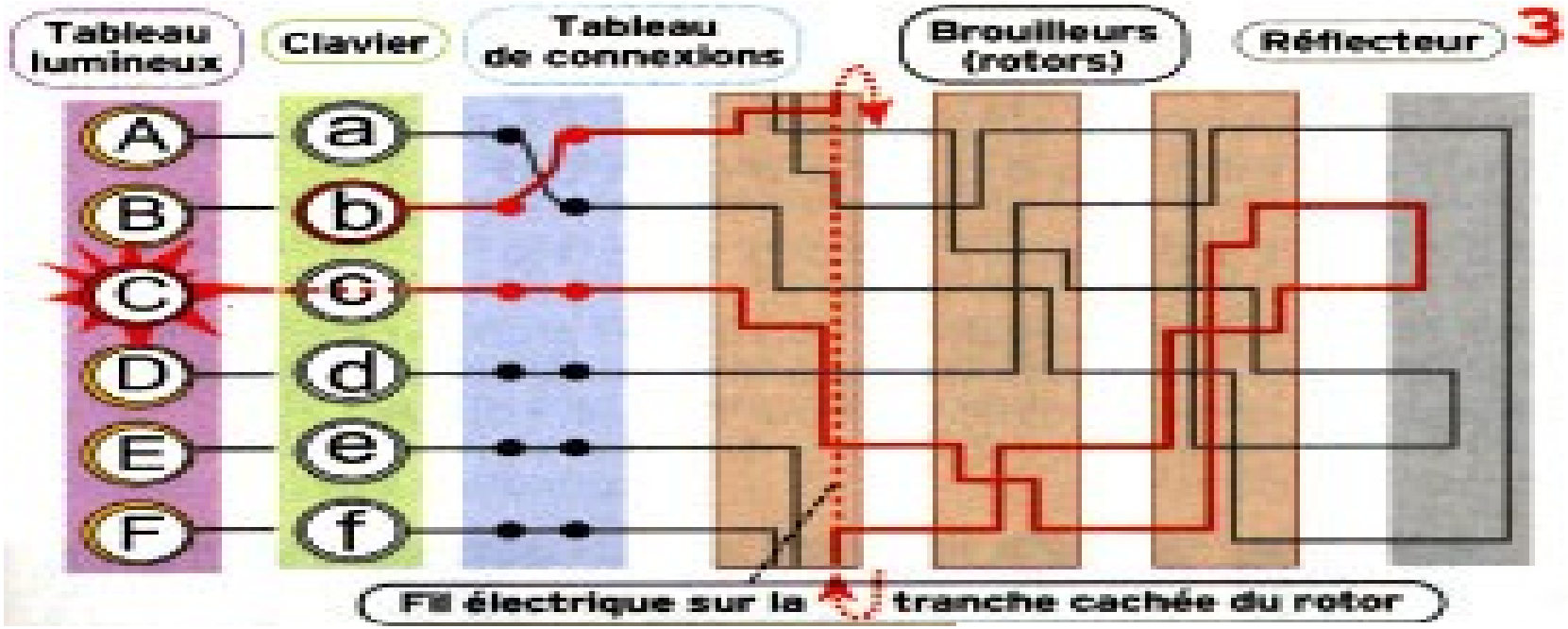
Substitution Polyalphabétique



Complexité de la substitution :

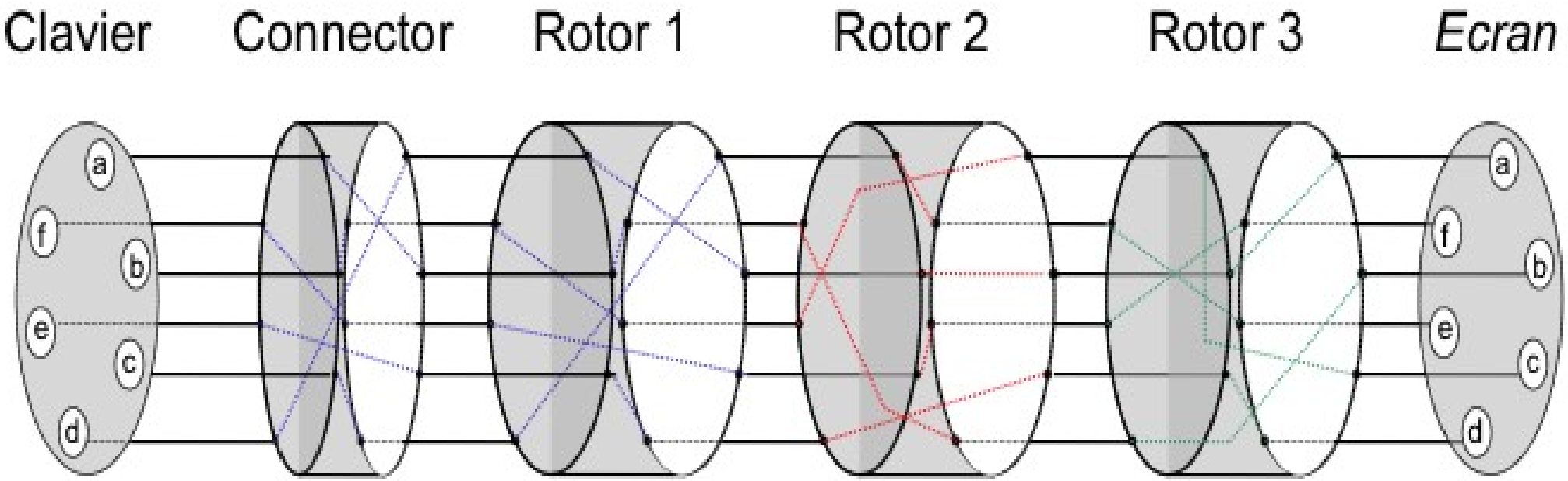
- 26 x 26 x 26 = 17 576 alphabets de chiffrement.

Enigma



- Le tableau de connexions permet de brouiller les pistes en reliant deux lettres du clavier entre elles.
  - Ainsi, quand on tape B, le courant prend en fait le circuit prévu pour A.
- Les trois brouilleurs associés multiplient ainsi le nombre de combinaisons.
- Le deuxième et le troisième avancent respectivement d'un cran quand le premier et le deuxième ont fait un tour complet.
- 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.

Substitutions élémentaires



Complexité de la substitution :

- 6 connexions possibles : 100 391 791 500 branchements possibles.

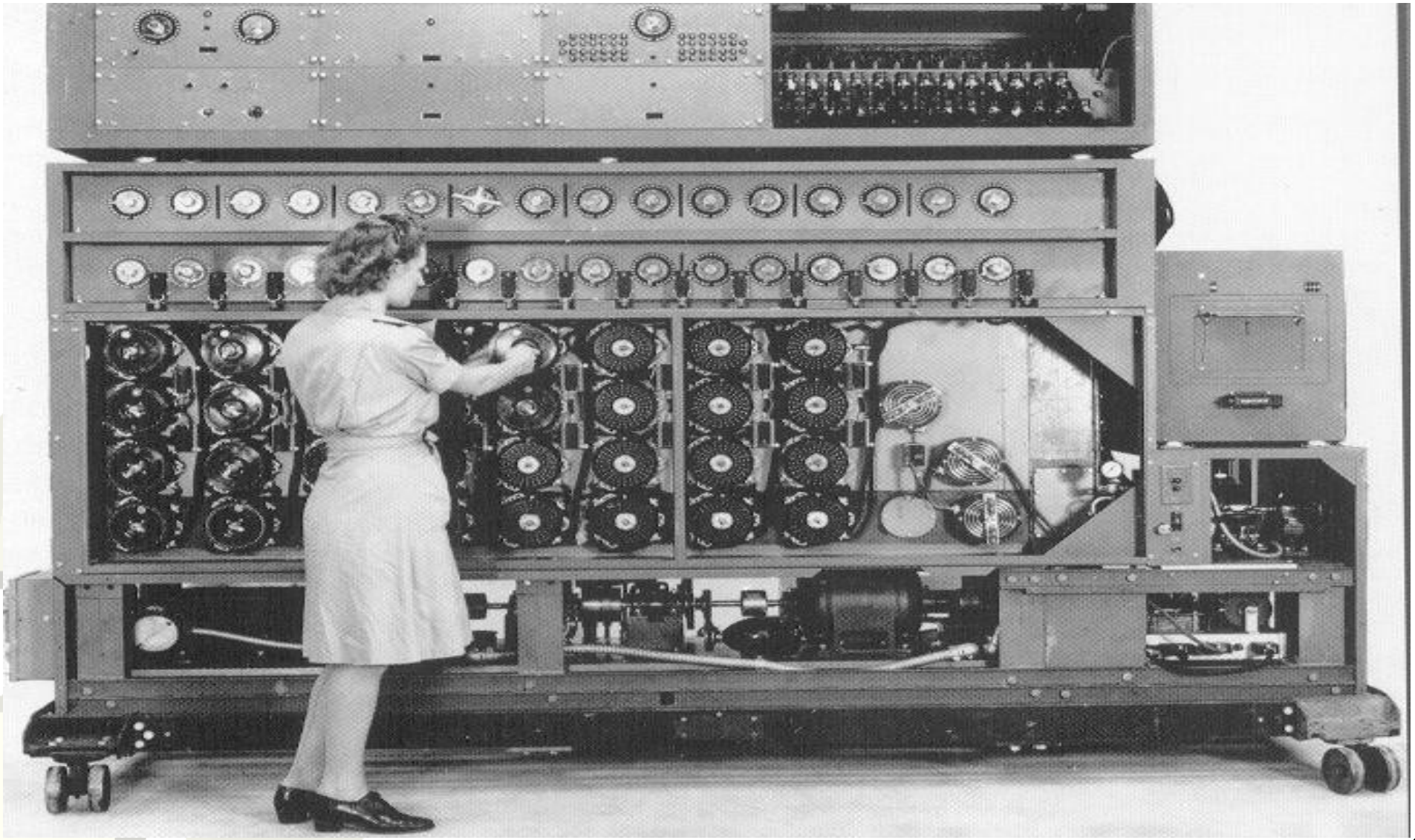
Enigma – Algorithme et clé

- Algorithme :
  - Substitutions des rotors.

- Clé de chiffrement :
  - Disposition des rotors.
  - Orientation initiale des rotors.
  - Connexions entre lettres de l’alphabet.

Cryptanalyse de Enigma

- La Bombe de Turing



Cryptanalyse de Enigma

- Principe :
  - Les bombes ont été construites pour retrouver le réglage de la machine Enigma.
  - L'idée était de deviner certains mots du message et de voir si l'on pouvait faire correspondre une partie du cryptogramme avec ce mot probable (crib en anglais).
    - Par exemple, le Allemands envoyaient souvent des prévisions météorologiques chiffrées avec Enigma; on pouvait donc essayer les mots "nuages", "pluie", etc.
- Résultat de la *Bombe* de Turing.
  - Performance :
    - Clé trouvée en 1 heure.
  - Limite de la *Bombe* :
    - Utilisation de plus de 5 rotors.
    - Pas de structure dans message.
- Décisif dans la victoire des alliers.

Extension à la substitution polyalphabétique

- Faiblesse de la substitution :
  - Taille du mot-clé : un digramme peut être chiffré plusieurs fois *de la même manière*.

- Idées :
  - Utilisation de plus d'alphabets de chiffrement.
  - *Choisir des mot-clés plus grand.*

Enigma - Complexité

- Au final, si l'on revient aux machines Enigma équipées pour 26 lettres, on a:
  - 26 x 26 x 26 = 17'576 combinaisons liées à l'orientation des chacun des trois brouilleurs,
  - 6 combinaisons possibles liées à l'ordre dans lequel sont disposés les brouilleurs,
  - 100'391'791'500 branchements possibles quand on relie les six paires de lettres dans le tableau de connexions.

- Les machines Enigma peuvent donc chiffrer un texte selon 17'576 x 6 x 100'391'791'500 = 10'000'000'000'000'000 combinaisons différentes!

Chiffrement parfait ???

- Longueur du mot-clé = longueur du message :
  - Garantie *a priori* un niveau de sécurité maximal mais …
- Cryptanalyse possible si :
  - Réutilisation du mot-clé.
  - Mot-clé *trivial*.

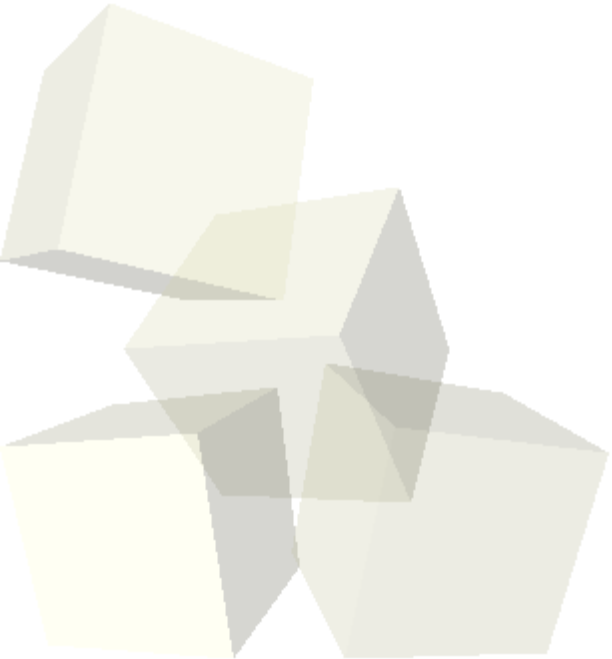
- Le chiffrement idéal : *One-time-pad*.
  - Longueur du mot-clé = longueur du message.
  - Mot-clé choisi *aléatoirement*.
  - Mot-clé *jamais* réutilisé.

Sécurité *mathématiquement* prouvée !

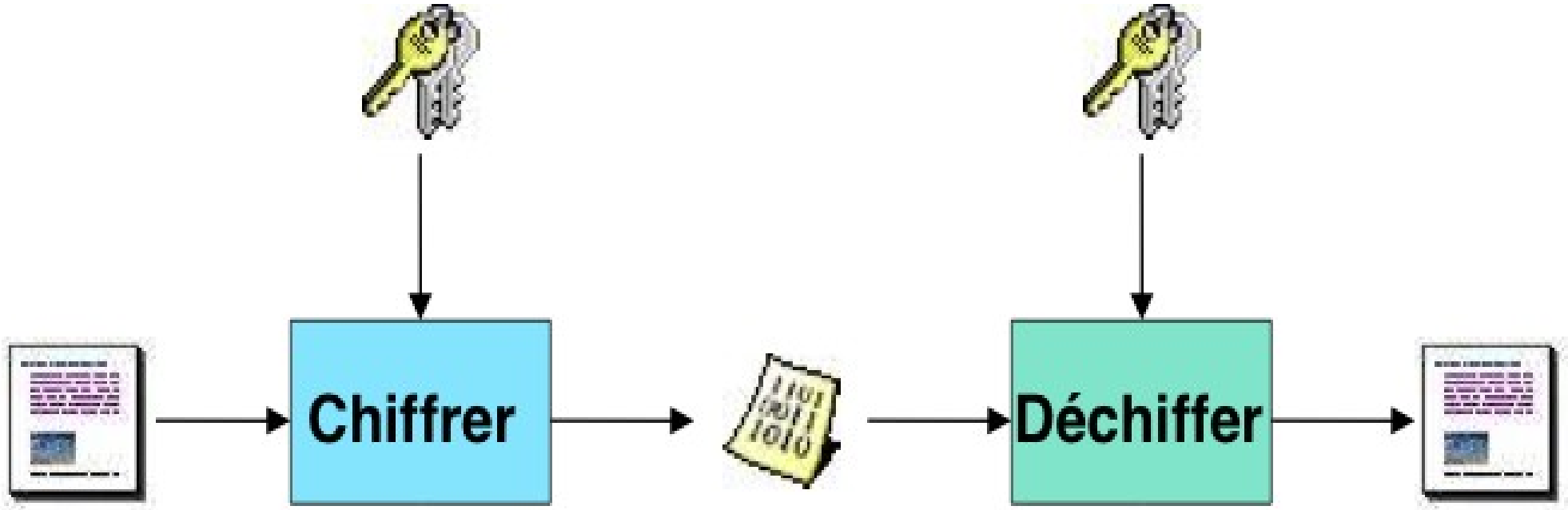
One-time-pad

- Confusion et Diffusion ?
  - Confusion *totale* :
    - Chiffrement de « aaaa … aaa » complètement aléatoire.
  - Diffusion *totale* :
    - Assurée car mot-clé ***jamais*** réutilisé :
    - résultat différent lorsque on *rechiffre* « aaaa … aaa ».

I. Histoire, définition et objectifs de la cryptographie
   I. Définition
   II. Transposition, Substitution
   III. Cryptographie moderne



Cryptographie moderne



Principes de la cryptographie

- Principe de Kerckhoffs : la sécurité repose sur le secret de la clé, et non sur le secret de l'algorithme (19ème siècle).
- Le déchiffrement sans la clé est impossible (*à l'échelle humaine*).
- Trouver la clé à partir du clair et du chiffré est impossible (*à l'échelle humaine*).

Théorie de l'information

- Claude Shannon - 1948

Problématique : A envoie un message M à B au travers un canal C

- Théorème 1 : codage de la source.
- Théorème 2 : code correcteur d'erreur.
- Théorème 3 : chiffrement parfait.

Entropie et incertitude

- **Quantité d'information :** nombre minimal de bits nécessaires pour coder (les significations de) l'information contenue dans un message.

- **Entropie :** permet de mesurer la quantité d'information dans un message *M*, noté *H(M)*.
  - En général, $H(M) = log_2(n)$ si *n* est le nombre de significations possible de *M*.

- **Incertitude :** nombre de bits qui permet de retrouver l'ensemble du message en clair.
  - L'entropie d'un message donne également son incertitude.

Codage de la source

But : trouver le codage le plus économique.

Théorème 1 :

Pour toute source X d'entropie H(X), on peut trouver un code dont la longueur moyenne s'approche de H(X) d'aussi près que l'on veut.

Code correcteur d'erreur

But : caractériser le canal de communication.

Théorème 2 :

Pour toute canal, on peut toujours trouver une famille de codes dont la probabilité d'erreur après décodage tend vers 0.

Chiffrement parfait

Soit *M* un message, *K* une clé et *C* le chiffré.

Définition :

On a un chiffrement parfait lorsque le chiffré *C* ne fournit aucune information sur *M* ou *K*.

$$H(M|C) = H(M) \text{ et } H(K|C) = H(K).$$

Théorème 3 :

Si un chiffrement est parfait, alors il y a au moins autant de clés que de messages :

$$|\mathcal{K}| \geq |\mathcal{M}|$$

$\mathcal{K}$ étant l'ensemble des clés, et $\mathcal{M}$ l'ensemble des messages.

Conséquence :

- Si $|\mathcal{K}| < |\mathcal{M}|$, le chiffrement n'est pas parfait.
- L'entropie d'un chiffrement est fonction de la taille des clés utilisées.

Chiffrement parfait

En pratique :

- Il existe un unique chiffrement parfait (*Vernam – 1917*) :
  - Soit *M*, on choisit *K* aléatoire tel que |*K*| = |*M*| et *K* jamais utilisé : *C* = *M xor K*.

- Plus l'entropie d'un chiffrement est grande, plus l'attaque par recherche exhaustive des clés est difficile.

Principe

- Méthodologie pour analyser la complexité de calcul des algorithmes :
  - Complexité en temps de calcul.
  - Complexité en espace de stockage.

- Complexité exprimée comme fonction de n, la taille du paramètre d'entrée :
  - Algorithme constant, linéaire, polynomial, exponentiel.

Complexité / Temps de calcul

Classe	*Complexité*	*Nombre d'opérations pour n = 10^6*	*Temps pour 10^6 opérations par seconde*
Constant	$O(1)$	1	1µs
Linéaire	$O(n)$	10^6	1s
Quadratique	$O(n^2)$	10^{12}	11,6 jours
Cubique	$O(n^3)$	10^{18}	32000 années
Exponentiel	$O(2^n)$	10^{301030}	10^{301006} fois l'âge de l'univers

Temps de calcul en fonction de la complexité de l'algorithme.

Complexité des problèmes

- Définition :
  - Complexité de l'algorithme permettant de résoudre l'instance la plus difficile du problème.

- Classification des problèmes :
  - Problèmes solubles (polynomial).
  - Problèmes non solubles ou difficiles.
  - Problèmes indécidables.

Application à la cryptographie

- Détermine le niveau de complexité d'une attaque.
  - A comparer avec la recherche exhaustive.

- Idéalement :
  - Chiffrement sûr : toutes les attaques sont de complexité exponentielle.

- En pratique :
  - Chiffrement sûr : toutes les attaques **connues** sont de complexité exponentielle.

Bibliographie

- Schneier Bruce, Cryptographie appliquée, International Thomson Publishing France, Paris, 1997
- http://www.apprendre-en-ligne.net/crypto/
- Dubertret Gilles, Initiation à la cryptographie, Vuibert Informatique, 2000
- Stinson Douglas, Cryptographie - Théorie et pratique, Vuibert Informatique, Paris, 2001
- http://www.supelec-rennes.fr/ren/perso/cbidan/co