Références sciences

Évolution de la cryptographie à travers les âges

Cours, exercices corrigés, algorithmes en Scratch et Python

Jean-Michel Masereel

Valérie Nachef

Emmanuel Volte

ellipses

Collection Références sciences

dirigée par Paul de Laboulaye
paul.delaboulaye@editions-ellipses.fr

Retrouvez tous les livres de la collection et des extraits sur www.editions-ellipses.fr

ISBN 9782340-093942
©Ellipses Édition Marketing S.A., 2024
8/10 rue la Quintinie 75015 Paris

DANGER LE PHOTOCOPILLAGE TUE LE LIVRE

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5.2° et 3°a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective », et d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).
Cette représentation ou reproduction, par quelque procédé que ce soit constituerait une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

www.editions-ellipses.fr

Table des matières

Introduction **13**
0.1. La cryptographie 13
0.2. Cryptographie symétrique / asymétrique 15
0.2.1. La cryptographie symétrique 15
0.2.2. La cryptographie asymétrique 15
0.3. Systèmes de chiffrement et clef 15
0.3.1. Coder et chiffrer 15
0.3.2. Quelques propriétés indispensables pour bien chiffrer . 16
0.4. Les apports de la cryptographie 16
0.5. Présentation de l'ouvrage 18
0.6. Exercices 18

I. De l'Antiquité à la fin de la Seconde Guerre mondiale **25**

1. La cryptographie dans l'Antiquité **29**
1.1. Le plus ancien document chiffré de l'Antiquité 29
1.2. Petite histoire de l'alphabet 30
1.2.1. L'évolution de l'alphabet 30
1.2.2. L'alphabet et la mythologie 31
1.3. Le procédé « Atbash », 500 av. J.-C. 31
1.4. La scytale spartiate, 400 av. J.-C. 33
1.5. Le carré de Polybe, 200 à 125 av. J.-C. 34
1.6. Le code de Jules César, 60 av. J.-C. 35
1.7. Exercices 35

2. La cryptographie au Moyen Âge et à la Renaissance **39**
2.1. Le Moyen Âge 39
2.1.1. Les méthodes de chiffrement 39
2.1.2. Les premières cryptanalyses 40

2.1.3. Les premiers nomenclateurs 43
2.2. La Renaissance 44
2.2.1. Le XV^e siècle 44
2.2.2. Le XVI^e siècle 46
2.3. Quelques exemples de nomenclateurs 52
2.3.1. Marie de Guise 52
2.3.2. Marie Stuart 55
2.4. Exercices 58

3. La cryptographie de l'Ancien Régime 63
3.1. Le chiffre de Louis XIV 63
3.2. Le chiffre de Marie-Antoinette 64
3.3. Exercices 68

4. Les chiffrements affine, de Playfair et de Hill 71
4.1. Le chiffrement affine 71
4.2. Le chiffrement de Playfair 72
4.2.1. Les chiffrements polygrammiques 72
4.2.2. Description du chiffre de Playfair 72
4.2.3. Cryptanalyse 74
4.2.4. Utilisation du chiffrement de Playfair 76
4.3. Le chiffrement de Hill 76
4.3.1. Historique 76
4.3.2. Description 76
4.3.3. Cryptanalyse 77
4.3.4. L'influence du chiffrement de Hill sur la cryptographie 78
4.4. Exercices 78

5. La cryptographie des guerres mondiales 81
5.1. La Première Guerre mondiale 81
5.1.1. Le télégramme de Zimmermann 82
5.1.2. Le chiffre ADFGVX 84
5.2. La Seconde Guerre mondiale 86
5.2.1. La machine Enigma et les cryptanalystes de Bletchley Park 86
5.2.2. Le code Navajo 92
5.3. Exercices 93

II. La cryptographie symétrique à clef secrète 95

6. La génération de nombres aléatoires 99

6.1. Définitions . . . 100
6.2. Sécurité des générateurs pseudo-aléatoires . . . 101
6.3. Générateurs de bits purement aléatoires . . . 102
6.3.1. Les générateurs matériels . . . 102
6.3.2. Les générateurs logiciels . . . 103
6.3.3. La correction du désalignement . . . 104
6.4. Générateurs pseudo-aléatoires non cryptographiquement sûrs . 105
6.4.1. Les générateurs congruentiels linéaires - LCG . . . 105
6.4.2. Mersenne Twister . . . 108
6.4.3. Les générateurs basés sur les fonctions à sens unique . 109
6.5. Générateurs cryptographiquement sûrs . . . 109
6.6. Tests statistiques . . . 111
6.6.1. Test du monobit et application . . . 111
6.6.2. Tests du poker . . . 112
6.6.3. Run test . . . 113
6.6.4. Tests graphiques . . . 121
6.7. Exercices . . . 126
6.7.1. Autour des vitraux de Richter . . . 126
6.7.2. Autour de la racine de 2 . . . 131
6.7.3. Autre exercice . . . 133

7. Le chiffrement à flot 135

7.1. Le masque jetable ou One-Time Pad . . . 135
7.1.1. Description . . . 135
7.1.2. Historique . . . 136
7.1.3. Utilisations . . . 136
7.1.4. Le masque jetable et la cryptographie quantique . . . 137
7.2. LFSR . . . 139
7.2.1. Introduction . . . 139
7.2.2. Définition - Exemples - Propriétés . . . 139
7.2.3. Algorithme de Berlekamp-Massey . . . 141
7.2.4. Améliorations des LFSR . . . 142
7.3. Exemples de chiffrements à flot . . . 144
7.3.1. RC4 . . . 144
7.3.2. A5/1 . . . 146

7.3.3. E0 . . . 147
7.3.4. SNOW 2.0 . . . 147
7.4. Exercices . . . 149

8. Chiffrement par blocs 151
8.1. Introduction . . . 151
8.2. Modes opératoires . . . 152
8.2.1. Le mode ECB (Electronic Code Book) . . . 152
8.2.2. Le mode CBC (Cipher Block Chaining) . . . 154
8.2.3. Le mode OFB (Output FeedBack) . . . 155
8.2.4. Le mode CFB (Cipher FeedBack) . . . 157
8.2.5. Le mode CTS (CipherText Stealing) . . . 159
8.3. DES . . . 161
8.3.1. Introduction . . . 161
8.3.2. Fonctionnement . . . 162
8.3.3. Avenir . . . 163
8.4. AES . . . 163
8.4.1. Principe . . . 164
8.4.2. Description . . . 164
8.5. Exercices . . . 169
8.5.1. DES . . . 169
8.5.2. AES . . . 176

9. Les schémas de Feistel 179
9.1. Les différents types de schémas . . . 179
9.1.1. Définitions et notations . . . 179
9.1.2. Les schémas de Feistel classiques . . . 180
9.1.3. Les schémas de Feistel contractants . . . 181
9.1.4. Les schémas de Feistel expansifs . . . 182
9.2. Les attaques . . . 183
9.2.1. Présentation . . . 183
9.2.2. Sur les schémas classiques . . . 185
9.2.3. Sur les schémas de Feistel contractants . . . 188
9.2.4. Sur les schémas de Feistel expansifs . . . 190
9.3. Exercices . . . 195

10. Les fonctions de hachage 199
10.1. Utilisation des fonctions de hachage 200
10.1.1. L'intégrité des données 200
10.1.2. L'engagement 200
10.1.3. Les mots de passe 200
10.1.4. La signature 201
10.1.5. Les structures de données 201
10.2. Définitions 201
10.3. Collisions sur les fonctions de hachage 202
10.3.1. Paradoxe des anniversaires 202
10.3.2. Application aux fonctions de hachage 204
10.4. Construction pratique d'une fonction de hachage en Python . . 204
10.4.1. Quelques fonctions utiles 204
10.4.2. Construction d'une fonction de compression 205
10.5. Construction de fonctions de hachage par le procédé de Merkle-Damgård 205
10.5.1. Complétion du message (padding) 206
10.5.2. Algorithme de Merkle-Damgård 206
10.5.3. Attaque par « length extension » 207
10.6. Conclusion 208
10.7. Exercices 208

11. La stéganographie 211
11.1. Stéganographie dans l'histoire 211
11.1.1. Dans l'Antiquité, 500 av. J.-C. 211
11.1.2. La grille de Cardan, 1550 212
11.1.3. Stéganographie dans la littérature 213
11.1.4. Stéganographie dans les journaux 213
11.1.5. Stéganographie moderne 214
11.2. Cacher un texte dans un arbre de choix 214
11.3. Exercices 218

III. La cryptographie asymétrique à clef publique 223

12. Les origines de la cryptographie à clef publique 227
12.1. La complexité des algorithmes 227

12.2. Le problème du sac à dos (1972) 228
12.2.1. Introduction 228
12.2.2. Formulation mathématique 229
12.2.3. Les méthodes de résolution 229
12.2.4. Le problème du sac à dos en cryptographie : le cryptosystème de Merkle-Hellman (1978) [17] 231
12.3. Le protocole d'échange de clefs de Diffie-Hellman (1976) . . 232
12.3.1. Le protocole 232
12.3.2. L'attaque de l'homme du milieu 233
12.3.3. Le problème calculatoire de Diffie-Hellman 233
12.4. Le chiffrement de Rabin (1979) 234
12.4.1. Présentation du chiffrement 234
12.4.2. Déchiffrement 234
12.4.3. Sécurité du chiffrement 235
12.5. Le chiffrement d'El Gamal (1984) 235
12.5.1. Chiffrement 235
12.5.2. Déchiffrement 236
12.5.3. Sécurité 236
12.6. Exercices 236

13. Le RSA (1977) 243
13.1. Présentation du RSA 243
13.1.1. Le RSA en confidentialité 243
13.1.2. Démonstration de la propriété de base du RSA 245
13.1.3. Le RSA en authentification 246
13.1.4. Le RSA en signature 246
13.1.5. Le RSA en confidentialité/signature 246
13.1.6. Le RSA pour tirer à Pile ou Face sur Internet 247
13.1.7. La factorisation et l'indicatrice d'Euler 248
13.2. Quelques erreurs à ne pas faire avec le RSA 248
13.2.1. Attaques par messages prévisibles 248
13.2.2. Plusieurs envois du même message 248
13.2.3. Deux personnes distinctes doivent avoir deux modules distinctes 248
13.2.4. Les attaques en signature 249
13.3. Le RSA et les cartes bancaires 249
13.3.1. Code confidentiel 249
13.3.2. Authentification hors ligne 250

13.3.3. Authentification en ligne 250
13.4. Exercices 251

14. Les preuves à divulgation nulle de connaissance 255
14.1. Deux exemples simples 255
14.1.1. La caverne d'Ali Baba 255
14.1.2. Les bonbons d'Halloween 257
14.2. Définitions 258
14.2.1. Les propriétés des protocoles 258
14.2.2. Les schémas de mise en gage 258
14.3. Les schémas à divulgation nulle de connaissance avec mise en gage 259
14.3.1. Le problème des 3 couleurs 259
14.3.2. PKP : Permuted Kernel Problem 262
14.4. Exercices 267

15. Les courbes elliptiques 273
15.1. Courbes elliptiques sur $\mathbb{R}$ 273
15.1.1. Définition 273
15.1.2. La loi de groupe 273
15.2. Courbes elliptiques modulo un nombre premier 276
15.2.1. Définition 276
15.2.2. Addition sur les courbes elliptiques modulo un nombre premier 276
15.2.3. Exemple 277
15.3. Application à la cryptographie 278
15.3.1. Fonction exponentielle sur les courbes elliptiques . . . 278
15.3.2. Le système de Menezes-Vanstone 279
15.4. Exercices 279

IV. Notions mathématiques et informatiques 281

16. Pré-requis : rappels 283
16.1. Relations binaires 283
16.2. Propriétés sur les entiers naturels 284
16.3. Raisonnement par récurrence 285
16.4. Division euclidienne dans $\mathbb{Z}$ 285

17. Structures algébriques **287**
17.1. Loi interne sur un ensemble 287
17.2. Groupes 287
17.2.1. Sous-groupes 289
17.3. Anneaux 290
17.4. Corps 291
17.5. Espaces vectoriels 292
17.6. Algèbres 292
17.7. Les groupes finis et le théorème de Lagrange 293

18. Arithmétique **295**
18.1. Algèbre dans $\mathbb{Z}$ 295
18.1.1. Sous-groupes de $\mathbb{Z}$ 295
18.1.2. PGCD 295
18.1.3. Nombres premiers entre eux 298
18.1.4. PPCM 299
18.1.5. Nombres premiers 301
18.2. Algèbre dans $\mathbb{Z}/n\mathbb{Z}$ 302
18.2.1. Définitions - Propriétés 302
18.2.2. Théorème des restes chinois 307
18.2.3. Indicatrice d'Euler 308
18.2.4. Les théorèmes de Fermat et d'Euler-Fermat 309

19. Le logarithme discret **311**
19.1. L'algorithme Baby-step giant-step 311
19.2. Méthode rho de Pollard pour le logarithme discret 312

20. Arithmétique des polynômes **315**
20.1. Définitions - Propriétés 315
20.2. Division euclidienne 316
20.3. Racines d'un polynôme 316
20.4. PGCD, PPCM, polynômes premiers entre eux 317
20.4.1. Plus grand commun diviseur 317
20.4.2. Polynômes premiers entre eux 319
20.4.3. Plus petit commun multiple 320
20.5. Polynômes irréductibles 321
20.5.1. Définition - Propriétés 321
20.5.2. Décomposition en facteurs irréductibles 321

20.6. Algèbre $\mathbb{K}[X]/\langle P\rangle$ 322
20.6.1. Les idéaux de $\mathbb{K}[X]$ 322
20.6.2. Algèbre $\mathbb{K}[X]/\langle P\rangle$ 322
20.6.3. Représentation de l'algèbre $\mathbb{K}[X]/\langle P\rangle$ 323
20.6.4. Calculs dans $\mathbb{K}[X]/\langle P\rangle$ 324

21. Corps finis **325**
21.1. Introduction - Exemple 325
21.2. Construction des corps finis 325
21.3. Caractéristique d'un corps fini 326
21.4. Arithmétique dans $\mathbb{F}_2[X]/\langle P\rangle$ 326
21.5. Application à l'AES 327
21.5.1. L'addition 328
21.5.2. La multiplication 328
21.5.3. L'inverse 328

22. Primalité **331**
22.1. Introduction - Premiers tests 331
22.1.1. Un algorithme élémentaire 331
22.1.2. Le crible d'Eratosthène 331
22.1.3. Test de primalité de Fermat 332
22.2. Le test de Solovay-Strassen 332
22.2.1. Les résidus quadratiques 332
22.2.2. Les symboles de Legendre et Jacobi 333
22.2.3. Le test de primalité d'Euler 335
22.2.4. Le test de Solovay-Strassen 335
22.3. Le test de Miller-Rabin 337
22.4. Deux algorithmes de factorisation 338
22.4.1. La méthode $p-1$ de Pollard 338
22.4.2. La méthode rho de Pollard 339

23. Informatique appliquée **341**
23.1. Introduction rapide à Scratch et Python 341
23.2. Scratch 341
23.2.1. Installation de Scratch, documentation 341
23.2.2. Les types de variables 341
23.2.3. Opérations sur les nombres et les chaînes de caractères 342
23.2.4. Les listes 343

23.2.5. Les blocs 343
23.3. Python 343
23.3.1. Installation de Python, documentation 343
23.3.2. Vue d'ensemble de Python 344
23.3.3. Principaux types d'objets 346
23.3.4. Tranchage 350
23.3.5. La programmation 351
23.3.6. Les fonctions 352
23.3.7. Les dictionnaires 354
23.4. Exemple de problème 356
23.4.1. Énoncé 356
23.4.2. Correction 356

V. Solutions des exercices **359**

Bibliographie **417**

Index **419**

Introduction

0.1. La cryptographie

Étymologiquement, « **cryptographie** » signifie « l'écriture cachée ». On peut d'ailleurs encore de nos jours définir la cryptographie comme la science et l'étude de l'écriture « cachée », c'est-à-dire de la façon de pouvoir écrire ou lire quelque chose grâce à un secret.

Le but de cette « **écriture secrète** » est d'assurer, soit la « confidentialité », soit « l'intégrité » des messages, soit les deux.

Dans le cas de la « **confidentialité** », on souhaite qu'une personne non autorisée ne puisse pas découvrir le contenu réel du message. C'est-à-dire que l'on envoie des messages qui semblent complètement inintelligibles aux personnes qui ne possèdent pas le secret ; et ces messages doivent rester inintelligibles, même si plusieurs messages chiffrés tombent entre les mains d'un ennemi.

Dans le cas de « **l'intégrité** », on souhaite pouvoir être sûr de l'identité d'une personne, ou être sûr qu'un message donné provient bien d'une personne, ou qu'il n'a pas été modifié. On ne cherche donc pas comme ci-dessus à faire passer un message inintelligible, mais à prouver que l'on est bien la personne que l'on prétend être, ou qu'un message donné est bien le message voulu.

Ainsi dans le cas de la « confidentialité », un observateur extérieur ne sait pas de quoi les messages traitent, alors que dans le cas de « l'intégrité », il peut éventuellement le savoir mais il ne peut pas envoyer de message en se faisant passer pour quelqu'un d'autre. Dans les cas de la preuve d'intégrité d'un message, on parle de « **signature** », et dans le cas d'une personne qui souhaite prouver son identité, on parle « **d'authentification** ».

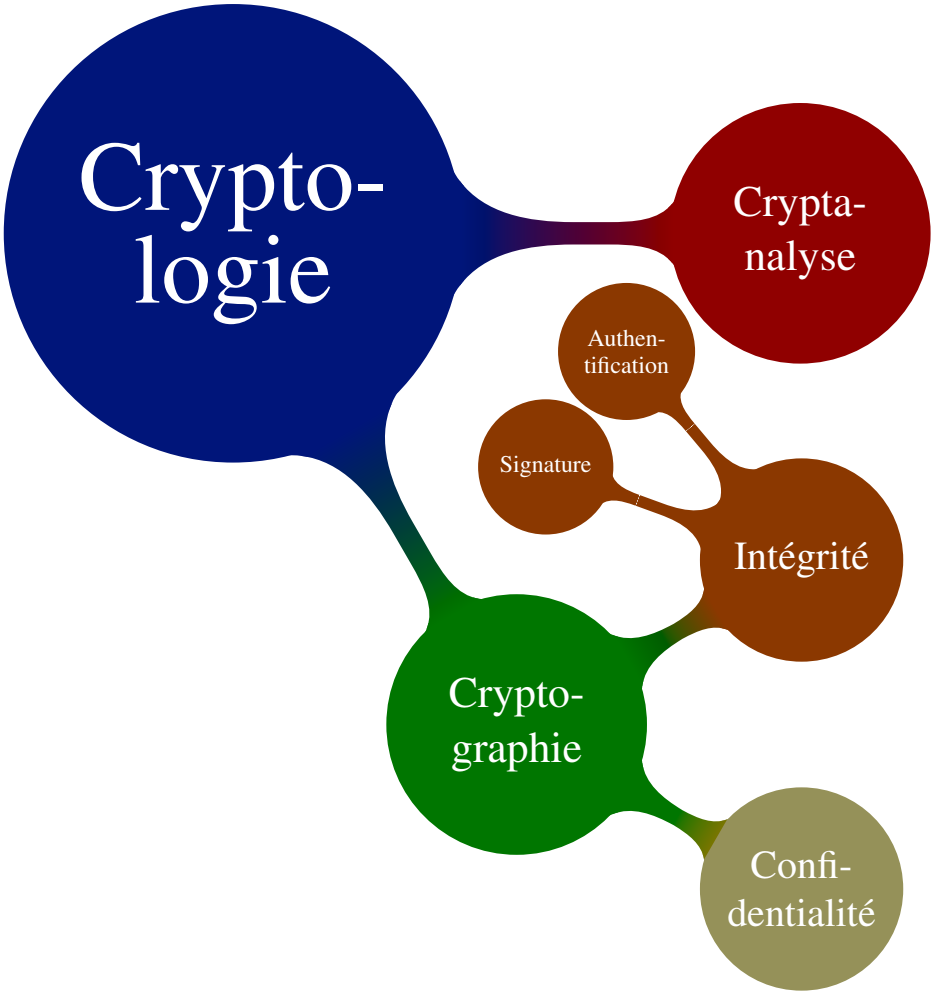
Les signatures peuvent avoir été écrites bien avant leurs vérifications alors que les authentifications ont lieu en général en temps réel, avec éventuellement des échanges de questions/réponses entre la personne qui souhaite s'authentifier et celle qui souhaite vérifier son identité.

Signalons aussi que l'on appelle « **cryptanalyse** » la science et l'étude des méthodes pour « casser » les algorithmes cryptographiques. En cryptanalyse,

on cherche donc à pouvoir faire la même chose que si on avait le secret alors qu'on ne l'a pas. Par exemple, on cherche à s'authentifier sans le secret, ou bien on cherche à découvrir le secret. Une bonne méthode cryptographique est donc une méthode dont la cryptanalyse doit être difficile.

Notons que le verbe « **chiffrer** » est utilisé pour désigner l'opération qui permet de passer du message clair au message inintelligible, et que le verbe « **déchiffrer** » désigne l'opération inverse (effectuée en connaissant le secret). Par contre, le verbe « **décrypter** » est réservé à l'opération de cryptanalyse qui consiste à obtenir le message clair à partir du chiffré sans connaître le secret.

Enfin, on nomme souvent « **cryptologie** » la science qui regroupe la cryptographie, la cryptanalyse, et leurs applications.



0.2. Cryptographie symétrique / asymétrique

La cryptographie est donc la science qui étudie comment écrire, ou lire quelque chose grâce à un secret.

0.2.1. La cryptographie symétrique

Si la connaissance du secret sert à la fois pour la lecture et l'écriture du message, on parle de cryptographie symétrique.

Deux personnes qui souhaitent communiquer avec de la cryptographie symétrique doivent donc toutes les deux connaître un même secret.

0.2.2. La cryptographie asymétrique

Il existe une autre cryptographie : la cryptographie asymétrique. Dans ce cas, la connaissance du secret ne sert que pour la lecture ou que pour l'écriture des messages. D'où le mot « asymétrique » : dans un cas, on a besoin de connaître le secret (pour déchiffrer, par exemple), mais pas dans l'autre (tout le monde peut chiffrer par exemple).

Cette propriété est vraiment étonnante, et la cryptographie asymétrique semble bien plus « magique » que la cryptographie symétrique. Elle nous permettra de faire des choses impossibles en cryptographie symétrique. Cependant, elle est souvent plus difficile à mettre en œuvre, plus délicate, plus lente, et plus chère… Il faudra donc bien savoir quand on souhaite l'utiliser et quand il vaut mieux utiliser la cryptographie symétrique.

0.3. Systèmes de chiffrement et clef

0.3.1. Coder et chiffrer

Au début de l'écriture, le chiffrement n'était pas nécessaire puisque trop peu de personnes savaient lire. Deux personnes éduquées pouvaient s'écrire sans risque d'être lues par une troisième personne. En suivant le même principe, les premiers chiffrements consistaient en un codage différent de l'écriture usuelle. Il fallait alors que les correspondants s'entendent sur ce codage et possèdent chacun un exemplaire du code. On pourrait prendre comme exemple le code du Morse (inventé par Samuel Morse et Alfred Vail en 1830). Lorsqu'on chiffre ainsi, avec un livre de code, on parle plutôt de « coder ».

Coder présente deux inconvénients majeurs : il faut posséder un livre de code assez gros, et en cas de découverte du livre de code, il faut trouver un nouveau code, ce qui peut prendre beaucoup de temps.

Pour les chiffrements, on distinguera le système de chiffrement, de la clef de chiffrement.

Par exemple, remplacer une lettre par une autre est un système de chiffrement qu'on appelle le chiffrement par substitution, et la clef est une permutation des lettres de l'alphabet.

0.3.2. Quelques propriétés indispensables pour bien chiffrer

Il est nécessaire de respecter certaines conditions de base

1. Les fonctions de chiffrement et de déchiffrement doivent être calculables efficacement.
2. Un opposant observant le texte chiffré doit être incapable de déterminer ni la clef utilisée, ni le texte clair.

La seconde propriété définit de manière informelle la sécurité. On note que si un attaquant peut trouver la clef, il peut alors retrouver le texte clair. Donc retrouver la clef est toujours au moins aussi difficile que de retrouver le texte clair.

On suppose toujours qu'un attaquant connaît le système de chiffrement mais pas la clef utilisée. C'est le principe de Kerchoff (1836-1903). Si le système est suffisamment bien conçu, le secret de la seule clef K suffit à assurer la confidentialité du message.

La méthode qui consiste à tester toutes les clefs possibles s'appelle la **recherche exhaustive de clef** ou **l'attaque brutale**. Il est donc nécessaire que l'espace des clefs soit suffisamment grand pour que cette méthode ne puisse pas être utilisée. Nous verrons dans cet ouvrage ce que signifie « suffisamment grand » pour faire face à la puissance de calcul des ordinateurs actuels. Ce n'est pas une condition suffisante pour la sécurité d'un système.

0.4. Les apports de la cryptographie

La cryptographie est intéressante à plusieurs points de vue. Son implication dans l'histoire est primordiale, et en particulier pendant les nombreuses

guerres que l'humanité a traversées. Mais on peut aussi considérer que la cryptographie, en elle même, a donné naissance aux premiers ordinateurs et à un nouveau domaine dans les sciences : l'informatique. Elle a également redonné du sens à cette partie des mathématiques que l'on nomme algèbre, et que Fermat lui même n'imaginait pas qu'on puisse un jour l'appliquer à des problèmes concrets. Force est de reconnaître maintenant que son théorème (le petit) est utilisé des millions de fois chaque jour dans nos cartes bancaires.

Du point de vue de l'enseignant de mathématiques, nous pouvons encore trouver d'autres intérêts à ce domaine des sciences.

D'abord il y a un intérêt interdisciplinaire évident. Avec l'histoire, par exemple, il est très facile d'illustrer à peu près n'importe quelle période historique avec un chiffrement. Avec le français, soit tout simplement à travers l'étude d'un texte faisant mention de procédé cryptographique, soit à travers une étude plus détaillée de l'utilisation des lettres dans la langue française, ce qui peut favoriser et rendre ludique un travail sur l'orthographe des mots et leur prononciations, ou bien tout simplement de travailler sur du vocabulaire. Nous conseillons d'ailleurs à tout le monde de s'essayer à la main de retrouver un texte dont les lettres ont été remplacées par d'autres (substitution mono-alphabétique). La joie que procure la découverte du texte, après de nombreux essais infructueux, mérite d'être vécue ! De la même manière, un travail peut être fait en collaboration avec le professeur de langue étrangère, à travers le chiffrement d'un texte dans une autre langue. Avec la physique, on peut également imaginer un certain nombre d'activités. La cryptographie quantique est peut être la meilleure illustration du lien qui existe entre ces deux domaines, mais on peut également imaginer des procédés mécaniques de chiffrement ou de stéganographie. Pourquoi ne pas construire un modèle simplifié de la machine Enigma, avec l'aide du professeur de technologie et des toutes récentes imprimantes en 3d ? Un tel projet pourrait être mené par toute une classe, tout au long de l'année.

0.5. Présentation de l'ouvrage

Nous avons choisi l'histoire de la cryptographie comme fil conducteur principal de l'ouvrage. Ainsi, dans cet ouvrage, nous traitons la cryptographie de l'antiquité à l'époque actuelle. Nous expliquons les différents schémas de chiffrement à travers les âges : chiffrement de Polybe, de César, de Vigenère, chiffrements à base de nomenclateurs comme ceux de Marie de Guise et de sa fille Marie Stuart, chiffrement affine et chiffrement de Hill ainsi que le chiffrement poly-alphabétique utilisé par Marie-Antoinette dans sa correspondance avec Axel de Fersen. Les codes utilisés pendant les deux guerres mondiales sont présentés : le code ADGFX et la machine ENIGMA. Les ordinateurs et la croissance de leur puissance ont amené de grands changements dans la manière d'utiliser la cryptographie. Nous expliquons ensuite les deux types de chiffrement principaux de la cryptographie moderne : la cryptographie à clef secrète avec l'exemple du DES et des schémas de Feistel et la cryptographie à clef publique avec plusieurs schémas de chiffrement : RSA, El Gamal ou encore courbes elliptiques.

À la fin de chaque chapitre, des exercices sont proposés : des chiffrements « à la main » pour les schémas les plus anciens aux chiffrements en Python ou Scratch, pour les schémas plus récents. Certains de ces exercices ont été calibrés pour être utilisés par les futurs enseignants préparant le CAPES de mathématiques ou de NSI, et par les candidats à l'agrégation interne de mathématiques. La plupart des solutions sont données à la fin de l'ouvrage.

Des rappels de notions mathématiques et de programmation sont présentés à la fin du livre.

0.6. Exercices

Quelques idées de codage.

Exercice 0.1 (Collège)**.** En débranché, vous pouvez coder un message en Morse sur un papier, mais en modifiant les symboles pour les codes longs et courts, et demander de le déchiffrer. Cela peut être par exemple une énigme dans un escape-game. On peut proposer comme indication le déchiffrage d'une partie du message.

Par exemple, quelle patisserie célèbre est cachée derrière ce message ?



Exercice 0.2 (college, lycée)**.** Avec Scratch, on peut proposer de terminer une activité qui va jouer le code Morse avec un instrument de musique, pour un message rentré par l'utilisateur. Le plus simple pour les débutants en informatique est de traiter tous les cas un par un, avec des tests successifs (27 tests en comptant les espaces). Une façon plus conforme au principe de séparation des données et du code serait l'utilisation d'une liste où seraient mis les codes du Morse.

Le programme, que ce soit l'une ou l'autre version, demande une habileté peu commune chez les élèves du secondaire. Ce pourrait donc être un projet pour un ou plusieurs élèves doués dans ce domaine. On peut également proposer un prolongement consistant à écrire un autre programme qui déchiffrerait le code morse entendu.

Pour une activité, certes difficile, mais abordable, on peut demander de compléter un programme.

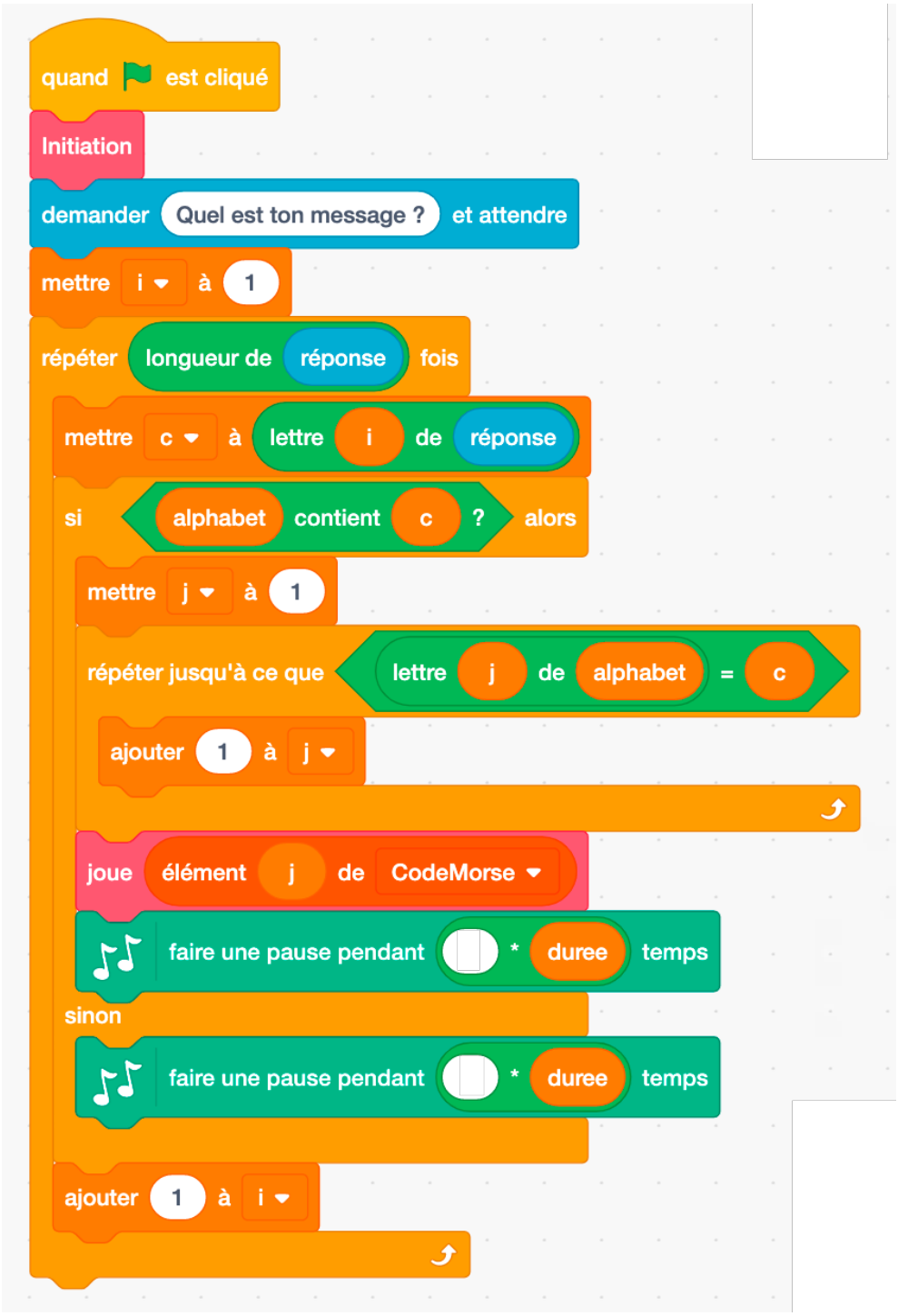
Durée des sons et des pauses en Morse :

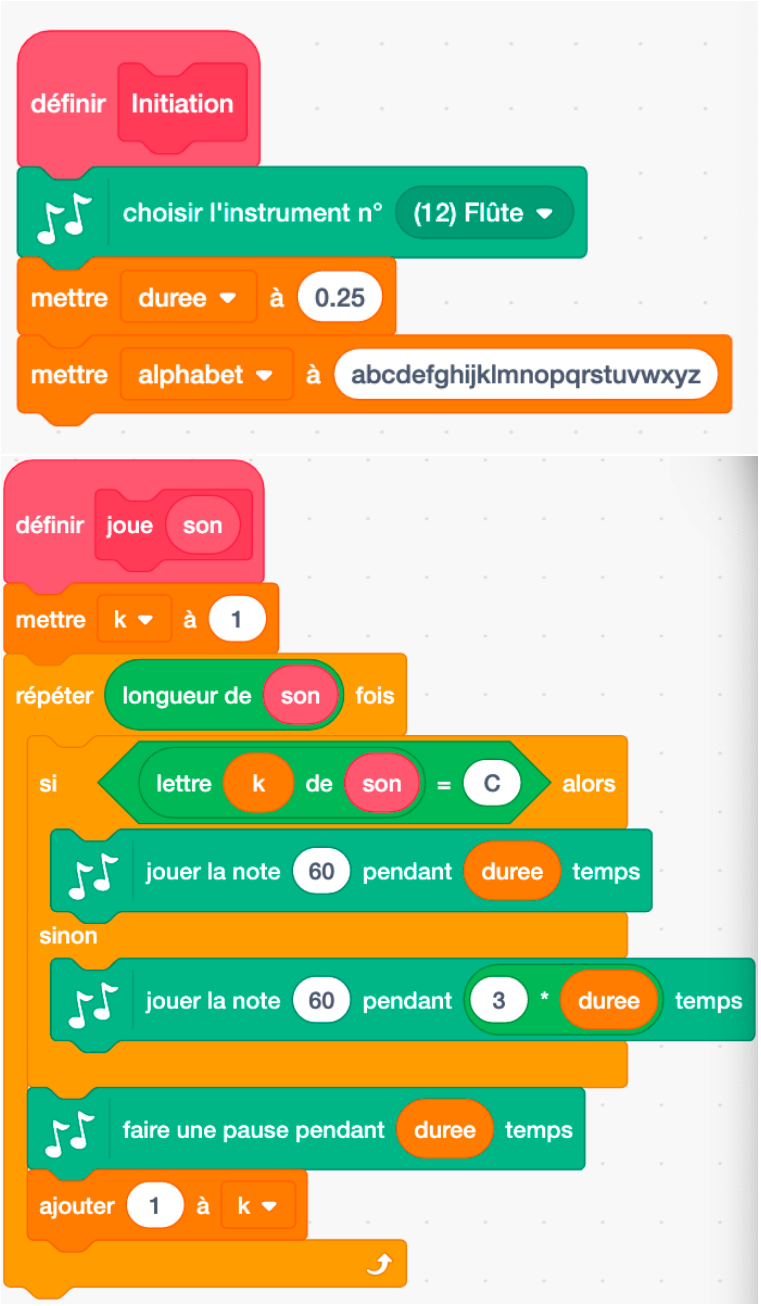
Son court	1 unité
Son long	3 unités
Pause entre chaque son	1 unité
Pause entre deux lettres	3 unités
Pause après chaque mot	7 unités

Comment compléter le code suivant pour que les durées des pauses soient respectées ?

```
A : CL
B : LCCC
C : LCLC
D : LCC
E : C
F : CCLC
G : LLC
H : CCCC
I : CC
J : CLLL
K : LCL
L : CLCC
M : LL
N : LC
O : LLL
P : CLLC
Q : LLCL
R : CLC
S : CCC
T : L
U : CCL
V : CCCL
W : CLL
X : LCCL
Y : LCLL
Z : LLCC
```




quand est cliqué
Initiation
demander Quel est ton message ? et attendre
mettre i à 1
répéter longueur de réponse fois
mettre c à lettre i de réponse
si alphabet contient c ? alors
mettre j à 1
répéter jusqu'à ce que lettre j de alphabet = c
ajouter 1 à j
joue élément j de CodeMorse
faire une pause pendant * duree temps
sinon
faire une pause pendant * duree temps
ajouter 1 à i



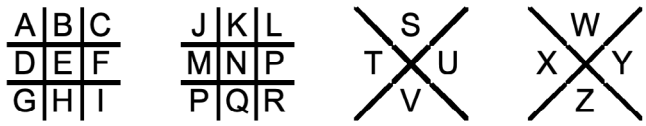
Exercice 0.3 (lycée, NSI)**.** Écrire un programme Python qui prend en entrée la durée d'une unité (correspondant au son court) et un message, et qui renvoie la durée de l'émission du code Morse. Attention, la pause après le dernier mot est nécessaire et doit être comptée dans la durée.

Pour information, voici la durée des sons en Morse :

Son court	1 unité
Son long	3 unités
Pause entre chaque son	1 unité
Pause entre deux lettres	3 unités
Pause après chaque mot	7 unités

Par exemple, 'SOS FANTOMES' correspondrait aux sons suivants :
`*.*.*...***.***.***...*.*.*.....*.*.***.*...*.***...***.*...***...***.***.***...***.***...*.....` où '*' est un son d'une unité, et '.' une pause d'une unité. En tout ce message a une durée de 100 unités.

Activité 0.4 (collège, voire primaire)**.** Codage Scout. On écrit les lettres de l'alphabet en commençant par un damier 3×3, puis un autre damier, puis dans les 4 angles d'une croix $\times$ et enfin les 4 dernières dans une autre croix.



Chaque lettre est codée en dessinant uniquement la case correspondante, avec un point rajouté si il s'agit du deuxième damier ou de la deuxième croix. Par exemple le A est codée par ⌟, et le J par ⋱⌟ S est codée par V.

On peut faire une activité en demandant aux élèves de chiffrer un message (mieux vaut l'imposer : pourquoi pas des citations célèbres ?), puis on échange les messages entre deux élèves.

Exercice 0.5 (lycée, NSI)**.** Codage d'un texte en nombre.

On se donne un alphabet de lettres autorisées. Par exemple `alphabet = " ABCDEFGHIJKLMNOPQRSTUVWXYZ"`. On remarquera que l'alphabet commence par un espace. En effet l'espace joue le même rôle que le 0 dans l'écriture en base 10 des nombres entiers : on supprime les zéros initiaux sauf pour le nombre 0.

On peut alors considérer qu'un message écrit dans cet alphabet est comme un nombre écrit en base 27.

1. Écrire une fonction Python `message2int(m)` qui, à partir du message `m` (chaîne de caractères utilisant les lettres de `alphabet`), renvoie le nombre entier correspondant.

2. Écrire une fonction Python `code(n)` qui, à partir d'un nombre entier, renvoie une chaîne de caractères correspondant au nombre écrit en base 26 à l'aide des lettres majuscules. Par exemple, 0 donnerait « A », 25 donnerait « Z », 26 donnerait « BA », 777 donnerait « BDX ». Ce n'est pas la fonction inverse de la précédente !
3. Quel titre de film célèbre à suspens devient « RVBPNNODCZA », après passage par ces deux fonctions successivement ?

Exercice 0.6 (lycée)**.** Dans le chiffrement par substitution, où la clef est l'alphabet mis dans le désordre, combien y a-t-il de clefs possibles ?

Première partie

De l'Antiquité à la fin de la Seconde Guerre mondiale

Durant l'Antiquité, les premiers schémas de chiffrement utilisés étaient le chiffrement pas transposition (mélange de l'ordre des lettres) ou le chiffrement par substitution (remplacement des lettres par des chiffres ou des symboles). La stéganographie qui consiste à dissimuler le texte par différents moyens au lieu de le transformer était aussi un moyen de transmettre des messages secrets.
Au Moyen Âge, la cryptographie évolue peu. Cependant, la cryptanalyse commence à émerger. Dans les chiffrements par transposition et substitution, une même lettre est toujours chiffrée à l'aide de la même lettre ou du même symbole et l'apparition de l'analyse des fréquences d'apparition des lettres dans le texte chiffré permet de casser ces types de schémas.
Les chiffrements ont alors évolué pour éviter l'analyse des fréquences. À la Renaissance, deux types de chiffrement vont alors faire leur apparition : les nomenclateurs et le chiffrement poly-alphabétique. Avec les nomenclateurs, une même lettre pourra être chiffrée pas plusieurs lettres et symboles, comme indiqué dans le nomenclateur. Parfois des symboles ou lettres peuvent aussi représenter des mots et non des lettres. Cependant, si le nomenclateur est perdu ou volé, il faut le changer. Le chiffrement poly-alphabétique permet également de ne pas toujours chiffrer une même lettre à l'aide de la même lettre ou du même symbole, en utilisant un mot clé qui entraîne un changement dans l'utilisation de l'algorithme. S'il y a suspicion de vol d'un mot clé, il est facile de le changer. Citons par exemple le chiffrement de Vigénère qui a résisté pendant 300 ans.
Un changement s'effectue lors de la Première Guerre mondiale avec les transmissions des messages par radio. C'est la première fois où la cryptographie joue un rôle dans l'issue d'un conflit. L'utilisation de la cryptographie s'accentue pendant la Seconde Guerre mondiale. On assiste à la mécanisation du chiffrement, du déchiffrement et de la cryptanalyse.

1. La cryptographie dans l'Antiquité

Les écritures secrètes semblent être nées spontanément dès que, dans un pays, une partie importante de la population a su lire. En effet, tant que la lecture a été réservée à une petite partie de la population, il n'y avait pas un réel besoin de cryptographie. Ceci a été le cas en Égypte et en Chine. En 1900 av. J.-C., dans une ville nommée Menet Khufu, au bord du Nil, un scribe égyptien a employé des hiéroglyphes non conformes à la langue correcte dans une inscription sur une stèle funéraire relatant la vie de son maître. C'est parfois présenté comme un premier exemple de cryptographie. Cependant le but de ce procédé n'était pas de rendre le texte incompréhensible mais plutôt de le rendre plus solennel. En Chine, c'est la stéganographie qui a été davantage utilisée : le message n'est pas rendu incompréhensible par un procédé de chiffrement mais écrit sur du papier ou de la soie, roulée en boule et recouvert de cire. Le porteur dissimulait la sphère de cire sur lui ou avalait celle-ci. La stéganographie a été également utilisée aux alentours de 600 av. J.-C. par Nabuchodonosor, roi de Babylone : il écrivait le message sur le crâne rasé de ses esclaves, attendait que leurs cheveux aient repoussé, et il les envoyait à ses généraux. Il suffisait ensuite de raser à nouveau le messager pour lire le texte.

1.1. Le plus ancien document chiffré de l'Antiquité

Le premier « document » chiffré connu remonte à l'Antiquité. Il s'agit d'une tablette d'argile, retrouvée en Irak, et datant du XVI^{e} siècle av. J.-C. Un potier y avait gravé sa recette secrète pour appliquer du vernis sur les poteries en supprimant des consonnes et en modifiant l'orthographe des mots.



FIGURE 1.1. – Premier document chiffré

1.2. Petite histoire de l'alphabet

1.2.1. L'évolution de l'alphabet

Les premiers alphabets phonétiques qui ont été utilisés du XX^e^ au VIII^e^ siècle av. J.-C. étaient des systèmes syllabiques. Par exemple, l'écriture des Assyriens utilisait un syllabaire de 500 signes cuneïformes. Puis apparaissent les alphabets consonantiques, avec vers 1100 av. J.-C., un alphabet de 22 consonnes, probablement élaboré à Byblos. Le timbre des voyelles était imposé par le rôle du mot dans la phrase. L'alphabet des Phéniciens se diffuse ensuite dans tout le bassin méditerranéen à partir du VIII^e^ siècle av. J.-C. : d'abord à Chypre et à Carthage, puis en Afrique du Nord et en Espagne. Inspiré de l'alphabet phénicien, l'alphabet grec est le premier à introduire les voyelles. On estime qu'il a dû être conçu vers 900 av. J.-C. Les Grecs ont repris les signes phéniciens pour noter les consonnes et ont ajouté des signes pour noter les voyelles. Cet alphabet a ensuite servi de modèle aux futurs alphabets latin et cyrillique.

Ce sont les Étrusques, en contact avec les Grecs, qui conçoivent un alphabet vers 700 av. J.-C., qui sera utilisé en Italie et répandu par les Romains dans le monde méditerranéen. Dans cet alphabet latin, les débuts de lettres grecques sont conservés : alpha devient a, béta devient b (prononcé bé). Certains signes ont été abandonnés (consonnes aspirées) ou transformés. La lettre G a été créée pour la différencier de la lettre C selon la prononciation. Le Y et le Z ont été repris pour les mots d'origine grecques. On en était alors à 23 lettres.

Au XVI^e^ siècle, on a 23 lettres dans l'alphabet. Les lettres J, V et W n'existent pas ainsi que les accents, le tréma et la cédille. De plus, il y a peu de signes de ponctuation. En 1542, le grammairien Louis Meigret propose d'introduire la lettre J pour faire une distinction entre les sons « i » et « j » qui jusque là étaient représentés tous les deux par la lettre i. Puis, en 1548, Ervé Fayad,

propose de distinguer les sons « u » et « v » qui étaient représentés par la lettre u uniquement. La lettre v apparaît.
En 1762, dans la quatrième édition de son dictionnaire, l'Académie sépare i et j, ainsi que u et v.
L'introduction du W a été plus compliquée. Il a été crée par doublement du V pour représenter le [w] germanique. Dans les premières éditions du dictionnaire de l'Académie, aucun mot orthographié avec la lettre W n'est présent. Les premiers mots ayant un W sont isolés dans le dictionnaire de 1878 car on considère que W est une lettre étrangère. En 1964, il est écrit dans le Robert que W est la 23^{e} lettre de l'alphabet, ce qui donne 26 lettres dans l'alphabet.

1.2.2. L'alphabet et la mythologie

Europe était une princesse phénicienne, fille d'Agénor roi de Tyr. D'après la légende, alors qu'elle se promenait au bord de la mer, elle fut enlevé par Zeus qui s'était transformé en taureau. Ils parvinrent en Crète où Zeus s'unit à Europe. Trois fils naquirent ce cette union : Minos, Sarpedon et Rhadamante. Un des frères d'Europe, Cadmos, partit à sa recherche avec interdiction par son père de revenir tant qu'il n'aurait pas trouvé Europe. Il alla en Grèce interroger l'orale de Delphes. L'oracle lui conseilla de suivre une génisse errante et de fonder une ville à l'endroit où elle se coucherait épuisée. C'est ainsi qu'il se rendit au site de Thèbes où il fonda Cadmée. Selon la légende, pour retrouver sa sœur, il offrit également aux Grecs l'alphabet inventé par les Phéniciens. N'ayant pas retrouvé Europe, il ne rentra jamais à Tyr.

1.3. Le procédé « Atbash », 500 av. J.-C.

Au V^{e} siècle av. J.-C., des scribes hébreux mettant par écrit le livre de Jérémie ont employé un chiffrement de substitution simple connu sous le nom d'Atbash. Ce procédé consiste à faire correspondre l'alphabet classique avec l'alphabet inversé, et d'associer ainsi à chaque lettre la lettre correspondante en position dans l'alphabet inversé. Par exemple, avec l'alphabet latin, le A est chiffré en Z, le B en Y et ainsi de suite. Le nom du système dérive de son fonctionnement, puisqu'il est constitué à partir des lettres aleph, tau, beth et shin, qui sont les deux premières (aleph et beth) et les deux dernières (shin et tau) lettres de l'alphabet hébreux. Le chiffrement Atbash est représenté dans la table 1.1.

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

TABLE 1.1. – Chiffrement « Atbash »

Il existe deux procédés de substitution très similaires qualifiés de chiffre « Albam » et chiffre « Atbah ». Dans le procédé Albam (table 1.2), il y a un décalage des lettres de 13 positions.

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

TABLE 1.2. – Chiffrement « Albam »

Dans le procédé Atbah (table 1.3), on regroupe les lettres par groupes de 4. Les lettres A,B,C, D sont associées aux lettres F, G, H, I et ordonnées à l'envers. Il en est de même pour les lettres J, K, L, M et les lettres O, P, Q, R puis les lettres S, T, U, V et les lettres W,X,Y,T. Les lettres E et N sont échangées.

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

TABLE 1.3. – Chiffrement « Atbah »

1.4. La scytale spartiate, 400 av. J.-C.

Le premier dispositif de cryptographie militaire connu, la scytale spartiate, remonte au Ve siècle av. J.-C.. La scytale consiste en un bâton de bois autour duquel est entourée une bande de cuir ou de parchemin, comme le montre la figure 1.2. L'expéditeur écrit son message sur toute la longueur de la scytale et déroule ensuite la bande qui 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. Ce procédé de chiffrement est un chiffrement par transposition. On bouleverse l'ordre des lettres de façon à rendre le texte incompréhensible mais on ne remplace pas les lettres du message par d'autres lettres ou symboles. Le texte clair et le texte chiffré auront donc la même fréquence de lettres.

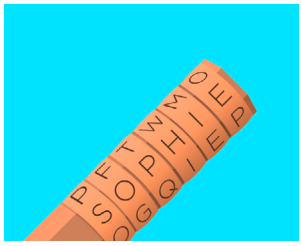


FIGURE 1.2. – Scytale spartiate

Ce dispositif a été utilisé par Lysandre de Sparte [26]. Lors de la bataille d'Aigos Potamos en 405 av. J.-C., la flotte spartiate commandée par Lysandre de Sparte avait complètement détruit la flotte athénienne. Pharnabaze, satrape perse, en charge des côtes d'Asie Mineure avait soutenu Lysandre de Sparte durant cette guerre contre Athènes. Cependant, Lysandre, installé dans le nord de Sparte à Sestos, se demandait s'il devait toujours considérer les Perses comme des alliés. Des troubles avaient éclaté dans des villes et il pensait que la Perse n'y était peut-être pas étrangère. Il devait savoir quel parti prendre et quelle décision le gouvernement de Sparte attendait de lui. Fallait-il déclencher une guerre contre Sparte sans y être préparé ou laisser les villes prendre une expansion qui ne pourrait plus être maîtrisée ? En 404 av. J.-C., Lysandre vit arriver un messager ensanglanté. Il était le seul rescapé d'un groupe de quatre porteurs de messages envoyés de Sparte. Le messager tendit sa ceinture à Lysandre qui l'enroula autour de sa scytale. Il apprit ainsi que Pharnaze de Perse s'apprêtait à l'attaquer. Il put alors se préparer et repousser l'attaque.

C'est le premier emploi de la cryptographie mentionné dans l'histoire et qui a permis de sauver un général et son empire.

1.5. Le carré de Polybe, 200 à 125 av. J.-C.

Polybe est un historien grec qui a vécu aux environs de 200-125 av. J.-C. Il est à l'origine du premier chiffrement par substitution. On utilise un carré de 25 cases (5×5). On peut agrandir ce carré si on veut ajouter des chiffres ou si l'alphabet utilisé comprend davantage de lettres. Le carré de Polybe est représenté dans la table 1.4. En français, on supprime la lettre « W ». En anglais, on regroupe le I et le J. Chaque lettre est représentée par un groupe de deux chiffres : celui de sa ligne et celui de sa colonne. Ainsi E = 15, U = 51, N = 34,... Polybe proposait de transmettre ces nombres au moyen de torches. Une torche à droite et cinq à gauche pour transmettre la lettre « E » par exemple. Ce procédé permettait donc de transmettre des messages sur de longues distances. On peut aussi transmettre les coordonnées des lettres en tapant des coups sur un mur, sur la tuyauterie, etc.

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I	J
3	K	L	M	N	O
4	P	Q	R	S	T
5	U	V	X	Y	Z

TABLE 1.4. – Carré de Polybe

Ce mode de chiffrement peut être amélioré en utilisant un mot-clef partagé par l'expéditeur et le destinataire. On commence à remplir le carré à l'aide des lettres du mot-clef et on continue le remplissage à l'aide des lettres de l'alphabet écrites dans l'ordre alphabétique mais en supprimant les lettres du mot-clef. Ainsi, avec le mot-clef POLYBE, le carré obtenu est donné dans la table 1.5. L'analyse des fréquences permet de casser ce code qui ne fut pas utilisé car il était trop compliqué. Cependant comme nous le verrons plus tard, des améliorations de ce code seront utilisées par l'armée allemande durant la Première Guerre mondiale.

	1	2	3	4	5
1	P	O	L	Y	B
2	E	A	C	D	F
3	G	H	I	J	K
4	M	N	Q	R	S
5	T	U	V	X	Z

TABLE 1.5. – Carré de Polybe avec mot-clef

1.6. Le code de Jules César, 60 av. J.-C.

Durant l'Antiquité, le premier procédé de chiffrement utilisé par les Romains, fut inventé par Jules César pour une utilisation militaire. Pour coder les messages, il suffisait de décaler les lettres de l'alphabet de trois rangs. Ainsi, la lettre A est codée par la lettre D, la lettre B par la lettre E et ainsi de suite. Ce mode de chiffrement fut encore utilisé par les officiers sudistes durant la guerre de Sécession et aussi par l'armée russe en 1915.

On remarque cependant, que si l'on sait que l'on a utilisé un chiffrement par décalage, il suffit d'essayer tous les décalages possibles jusqu'à ce que l'on obtienne un texte qui a du sens. Donc la recherche exhaustive de clefs est, dans ce cas, très facile. Cependant, à cette époque, le nombre de personnes sachant lire était très restreint et donc cela pouvait apporter une certaine sécurité au code.

Les chiffrements Atbash, Albam, Atbah, ainsi que celui de Jules César, font tous partie de la famille des chiffrements par substitution.

1.7. Exercices

Activité 1.1 (College)**.** Pour faciliter le chiffrement par décalage, on peut proposer une activité consistant à écrire toutes les lettres de l'alphabet, de façon régulière autour de deux cercles de rayon 5 cm et 7 cm par exemple. L'exercice n'est pas si facile car 360 n'est pas divisible par 26, et des erreurs qui se cumulent peuvent donner un décalage non négligeable à la fin. Pour simplifier, on peut omettre les lettres J et W (en les remplaçant respectivement par I et V), ce qui donne 24 lettres, et il faut donc placer une lettre tous les 15 degrés.

Une fois les cercles construits, et éventuellement coloriés de couleurs différentes, on les découpe, on met le plus petit sur le plus grand et on les attache au milieu à l'aide d'une attache parisienne.
On peut remarquer que le même outil peut servir pour le chiffrement et le déchiffrement, sans changer la position des cercles, mais en l'utilisant d'une façon différente. C'est déjà une façon de représenter une fonction et de chercher image et antécédent, sans forcément employer le vocabulaire.
Remarque : Une illustration, réalisée avec Scratch est donnée au chapitre suivant (disque d'Alberti)

Activité 1.2 (primaire, collège)**.** On peut se lancer dans la fabrication d'une scytale. Le plus simple est d'utiliser des rouleaux vides de papier absorbant, ou des rouleaux pour du papier cuisson. Il est important que ces rouleaux soit assez longs et suffisamment solides. Un tuyau en PVC pourrait également convenir.

Le plus délicat est de fabriquer une bande. On peut par exemple découper des bandes de 2 cm de largeur et les scotcher les unes après les autres afin d'obtenir une bande assez grande. Une autre idée serait de récupérer une lanière de cuir assez régulière mais il est alors compliqué de l'utiliser plusieurs fois.

Dans tous les cas, on s'aperçoit qu'il n'est pas si facile d'enrouler de manière régulière la bande autour du rouleau. Il faut écrire le message en espaçant bien les lignes.

Voici différentes situations pédagogiques pour utiliser la scytale :

— Demander à un petit groupe de faire un exposé autour de la scytale et de fabriquer une scytale, par exemple pour un projet de fin d'année.
— Dans un escape game. Il faudra que le rouleau à utiliser soit bien visible, ou bien s'arranger avec un jeu de couleur pour que l'association entre la bande et le rouleau soit évidente.

Exercice 1.3. Le texte suivant est un extrait d'un roman qui a été chiffré par un chiffrement mono-alphabétique. Quel est le titre du roman (dernière phrase) ?

XZCLQJLC AVN YTCHZFL BZC BZVYNYT. WT C'TQLNQ EVT UZANQTBBT. YNTC RT QZVQ WL CT O'NOUZYQT, UTCBLNQ-NA, XVBEV'L WT EV'NA BT YTCRT WZOUQT, LHTW VCT TWZTVYLCQT WTYQNQVRT, L EVTA UZNCQ W'TQLNQ GLVK. NA CT HZVALNQ ULB BT YT-

QYZVHTY L BFRCTF BLCB VC BZV, NA CT HZVALNQ ULB LHZNY L WJ-TYWJTY RV QYLHLNA, ULB NOOTRNLQTOTCQ. NA LHLNQ DTBZNC RT QTOUB UZVY YTGATWJNY. NA BZVYNQ RV ONTVK EV'NA UVQ LV OL-QTAZQ QZVXZVYB MVNAATYTQ.

- XT Q'LN BVNHN, RNQ-NA. OZCQYT QZC XTV.

ATCQTOTCQ, LHTW VCT QJTLQYLANQT RTANDTYTT, AT OLYNC OZC-QYL BZC XTV.

WLYYT RT WNCE.

XZCLQJLC BZVUNYL, GNQ RT BZC ONTVK UZVY GZYWTY VC YNYT, RTWZVHYNQ BTB WLYQTB TQ UZVBBL A'TCZYOT UNAT R'LYMTCQ L A'LVQYT DZVQ RT AL QLDAT.

- DZCCT WJLCWT, RNQ-NA TC BT ATHLCQ.

AT OLYNC, YLHN, TOUZWJLNQ BTB MLNCB.

- Q'LDLCRZCCTB ? AVN RTOLCRL-Q-NA.

- ZVN. XT BVNB GLVWJT. X'LNOTYLNB UZVHZNY OLCMTY

RTOLNC.

TKQYLNQ RT : WZZP, PTCCTQJ.

AT QYTBZY RT AL DLNT RTB ZYEVTB.

10 lettres les plus fréquentes : A : 4,9%, B : 5,6%, Z : 6,4%, V : 6,6%, Y : 6,9%, C : 7,1%, Q : 8,2%, N : 8,7%, L : 9,1%, T : 15%.
10 doublons les plus fréquents : TW : 1,6%, TC, TQ, TY, YT : 1,7%, NA : 1,9%, RT : 2%, LN : 2,2%, NQ : 2,3%, ZV : 2,5%

Exercice 1.4 (6-Sup)**.** Le texte suivant a été chiffré à l'aide d'un chiffrement par décalage (du type de celui de César). Retrouvez le texte original. Peut-on trouver le genre du dernier mot ?

HGMJJWMKKAJMFTSFIMWLADXSMLXAUWDWJDWKTSJVWK

Exercice 1.5 (NSI)**.** Dans cet exercice nous allons nous intéresser à la programmation en Python de chiffrements par substitution en utilisant des dictionnaires. On pourra utiliser ces fonctions pour les chiffrements donnés dans le cours, comme Atbash.

1. Écrire une fonction `dico(s1, s2)` qui prend en paramètre deux chaînes de caractères de même longueur et qui renvoie le dictionnaire dont les clefs sont les lettres de `s1` et les valeurs sont les lettres de `s2` correspondantes.
2. Écrire une fonction `substitution(m, alph, clef)` où `m` est une chaîne de caractères n'utilisant que les caractères dans `alph`, et `clef` contient les lettres de `alph`, dans un ordre arbitraire. La première lettre de l'alphabet sera codée à l'aide de la première lettre de la clef, et ainsi de suite. La fonction renvoie le message `m` chiffré de cette manière.

   Par exemple, si
   `alpha1 = "ABCDEFGHIJKLMNOPQRSTUVWXYZ"` et
   `clef1 = "ZYXWVUTSRQPONMLKJIHGFEDCBA"`, alors
   `substitution("LOYAL", alpha1, clef1)`
   renvoie `"OLBZO"`.
3. Comment coder en une ligne la fonction inverse de la précédente :
   `inv_substitution(x, alph, clef)`
   où `x` représente le texte chiffré ?

2. La cryptographie au Moyen Âge et à la Renaissance

Au Moyen Âge, les méthodes de chiffrement sont principalement la substitution et la transposition. Le cryptanalyse par l'analyse des fréquences proposée par Al-kindi au IX^{e} siècle va faire évoluer la cryptographie afin qu'une même lettre ne soit pas toujours chiffrée de la même manière. Deux directions apparaissent :

— Les nomenclateurs qui indiquent comment chiffrer lettres et mots. Nous donnerons les exemples des nomenclateurs utilisés par Marie de Guise et sa fille Marie Stuart.

— Le chiffrement poly-alphabétique qui se présente sous différentes formes : le cadran d'Alberti, les chiffrements de Porta et Vigenère par exemple.

2.1. Le Moyen Âge

2.1.1. Les méthodes de chiffrement

Durant le Moyen Âge la cryptologie évolue faiblement car elle est peu pratiquée. Seuls les moines en Europe utilisent cette science plus par jeu que par nécessité, influencés par le chiffrement « Atbash » décrit au chapitre 1.

En 855, Abu Bakr ben Wahshiyya publie plusieurs alphabets secrets utilisés à des fins de magie, dans son livre *Kitab shauk almustaham fi ma'arifat rumuz al aklam* (Le livre de la connaissance longuement désirée des alphabets occultes enfin dévoilée).

À partir de 1226, on trouve dans les archives de Venise, des textes politiques où des points ou des croix remplacent les voyelles dans quelques mots épars.

Le moine Roger Bacon (1214-1294) philosophe, théologien et savant anglais a non seulement décrit plusieurs chiffres, mais il a aussi écrit dans sa *Lettre sur les prodiges de la nature et sur la nullité de la magie* vers 1260 : « Il est fou celui qui écrit un secret de toute autre manière que celle qui le soustrait à la connaissance du vulgaire ».

En 1392, dans un ouvrage intitulé *L'équatorial des Planètes*, qui décrit le fonctionnement d'un instrument astronomique, Geoffrey Chaucer, a incorporé six courts cryptogrammes écrits de sa propre main.

2.1.2. Les premières cryptanalyses

Jusqu'au Moyen Âge, il n'y a eu aucune recherche suivie en matière de cryptanalyse. Celle-ci naquît chez les Arabes. Ils découvrirent les méthodes de déchiffrement et les consignèrent par écrit.

Nous avons déjà vu que dans les différentes méthodes de chiffrement proposées dans le chapitre 1, il y a deux manières différentes de chiffrer. Dans la première on modifie l'ordre des lettres, par exemple pour la scytale spartiate : c'est la méthode de transposition. Dans la seconde, on remplace une lettre par une autre lettre, mais qui est toujours la même : c'est la méthode de substitution. Certaines méthodes de substitution sont plus efficaces que d'autres. Par exemple, si on sait que c'est le chiffrement par décalage (type chiffrement de César) qui a été utilisé, il suffit d'essayer tous les décalages possibles jusqu'à ce que l'on trouve un texte qui a du sens. C'est ce que l'on appelle la recherche exhaustive de clef. Ici le secret est la valeur du décalage (3 pour le chiffrement de César). On appelle aussi cela la « clef de chiffrement ». Il faut donc que le nombre de clefs d'une méthode de chiffrement soit suffisamment important pour éviter la recherche exhaustive.

On suppose que l'on a un alphabet avec 26 lettres. À chaque lettre, on attribue un chiffre de 0 à 25. Les lettres du message chiffré sont transformées en chiffres en suivant cette règle. L'ensemble des clefs est l'ensemble des permutations des 26 nombres $0, 1, \ldots, 25$. La permutation est appliquée sur chaque chiffre. La chaîne ainsi obtenue est à nouveau transformée en lettres. Pour le déchiffrement, on utilise la permutation inverse.

Dans le cas du chiffrement par substitution, il y a $26! = 26 \times 25 \times \ldots \times 2 \times 1 \simeq 2^{88}$ clefs possibles. C'est un très grand nombre, et ce nombre représente l'effort à faire pour casser le système si on essayait toutes les clefs possibles (recherche exhaustive ou attaque brutale). Cependant, en pratique, il est possible de casser ce système bien plus facilement. C'est au IX^{e} siècle

qu'Abu Yusuf Ya'qub ibn Is-haq ibn as-Sabbah Oòmran ibn Ismaïl al-Kindi rédige le plus ancien texte connu décrivant la technique de déchiffrement appelée « **analyse des fréquences** », qui permet le déchiffrement dans le cas du codage par substitution.



FIGURE 2.1. – *Al-kindi*

Détaillons le principe. Si un texte est suffisamment long, les lettres apparaissent avec une fréquence bien précise qui dépend de la langue d'écriture. Remarquons alors que dans le cas d'un chiffrement par substitution, une même lettre sera toujours chiffrée de la même manière. Ainsi, la lettre la plus fréquente dans le texte chiffré correspond très certainement à la lettre la plus fréquente dans la langue du texte clair. De même pour les autres lettres, on fait correspondre les fréquences dans le texte chiffré aux fréquences de la langue. On peut aussi étudier les bigrammes et les trigrammes.

Les fréquences d'apparition des lettres dans la langue française sont données dans la table 2.1, et pour la langue anglaise dans la table 2.2.

Les 20 bigrammes les plus fréquents en français sont dans l'ordre décroissant :
ES, DE, LE, EN, RE, NT, ON, ER, TE, EL, AN, SE, ET, LA, AI, IT, ME, OU, EM, IE.

Les 20 trigrammes les plus fréquents en français sont dans l'ordre décroissant :
ENT, LES, EDE, DES,QUE, AIT, LLE, SDE, ION, EME, ELA, RES, MEN, ESE, DEL, ANT, TIO, PAR, ESD, TDE.

Lettre	Fréquence	Lettre	Fréquence
A	8,4%	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%

TABLE 2.1. – Fréquence d'apparition des lettres dans la langue française

Lettre	Fréquence	Lettre	Fréquence
A	8,08%	N	7,38%
B	1,67%	O	7,47%
C	3,18%	P	1,91%
D	3,99%	Q	0,09%
E	12,56%	R	6,42%
F	2,17%	S	6,59%
G	1,80%	T	9,15%
H	5,27%	U	2,79%
I	7,24%	V	1,00%
J	0,14%	W	1,89%
K	0,63%	X	0,21%
L	4,04%	Y	1,65%
M	3,60%	Z	0,07%

TABLE 2.2. – Fréquence d'apparition des lettres dans la langue anglaise