Université Côte d'Azur — L3 Info, L3 MI, L3 ST
Cryptographie et calcul symbolique — 2023-24

TD n° 1

Exercices introductifs sur les codes

Exercice 1) Assurez-vous que vous avez bien compris les notions de *système de communication*, de *source*, de *canal*, de *bruit*, de *codage de source* et de *codage de canal* en citant quelques exemples.

Exercice 2) De l'impossibilité de tout compresser sans perte

1. Combien y a-t-il de fichiers distincts de taille exactement N bits ?
2. Combien y a-t-il de fichiers distincts de taille strictement inférieure à N bits ?
3. Que peut-on en conclure ?

Exercice 3) Algorithme de Sardinas-Paterson Déroulez cet algorithme sur chacun des langages suivants afin de décider si c'est un code (uniquement déchiffrable) ou non :

1. $L = \{01, 11, 100, 101, 110, 0010\}$;
2. $K = \{01, 11, 110\}$.

Exercice 4) Codage du fax Dans une version simplifiée, une page de fax est décomposée en lignes de 1728 pixels. Pour chaque ligne, on transmet la suite des couples contenant le nombre de pixels consécutifs de même couleur et la dite couleur.

1. Comment coderiez-vous une ligne blanche ?
2. Ecrivez un algorithme pour coder une ligne quelconque.
3. Voyez-vous des inconvénients à cette méthode ? Comment l'améliorer ?

Exercice 5) Code de César Scipion reçoit le message suivant, aidez-le à le déchiffrer :

$$IDFLOH\ OD\ FUBSWR\ DXA\ WHPSV\ GHV\ URPDLQV\ !$$

Exercice 6) Chiffre de Vigenère Ce chiffre généralise le codage de César au moyen d'une clef permettant de varier le décalage. On applique ainsi à chaque lettre à chiffrer le décalage correspondant au caractère de la clef.

```
abcdefghijklmnopqrstuvwxyz abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ NOPQRSTUVWXYZABCDEFGHIJKLM
BCDEFGHIJKLMNOPQRSTUVWXYZA OPQRSTUVWXYZABCDEFGHIJKLMN
CDEFGHIJKLMNOPQRSTUVWXYZAB PQRSTUVWXYZABCDEFGHIJKLMNO
DEFGHIJKLMNOPQRSTUVWXYZABC QRSTUVWXYZABCDEFGHIJKLMNOP
EFGHIJKLMNOPQRSTUVWXYZABCD RSTUVWXYZABCDEFGHIJKLMNOPQ
FGHIJKLMNOPQRSTUVWXYZABCDE STUVWXYZABCDEFGHIJKLMNOPQR
GHIJKLMNOPQRSTUVWXYZABCDEF TUVWXYZABCDEFGHIJKLMNOPQRS
HIJKLMNOPQRSTUVWXYZABCDEFG UVWXYZABCDEFGHIJKLMNOPQRST
IJKLMNOPQRSTUVWXYZABCDEFGH VWXYZABCDEFGHIJKLMNOPQRSTU
JKLMNOPQRSTUVWXYZABCDEFGHI WXYZABCDEFGHIJKLMNOPQRSTUV
KLMNOPQRSTUVWXYZABCDEFGHIJ XYZABCDEFGHIJKLMNOPQRSTUVW
LMNOPQRSTUVWXYZABCDEFGHIJK YZABCDEFGHIJKLMNOPQRSTUVWX
MNOPQRSTUVWXYZABCDEFGHIJKL ZABCDEFGHIJKLMNOPQRSTUVWXY
```

1. Codez le message `chiffrement` à l'aide du carré de Vigenère. La clef réside dans le mot `CRYPTO` que l'on répète au besoin.
2. Décodez le message $KSYSSGTUUTZXVKMZ$ chiffré avec la clef $VENUS$.

Exercice 7) Chiffrement affine On définit le chiffrement affine suivant (les lettres `a`, `b`, `c` ... sont préalablement codées $0, 1, 2 \ldots$ et tout autre caractère sera reproduit tel quel) :

$$\begin{array}{ccc} \mathbb{F}_{26} & \to & \mathbb{F}_{26} \\ x & \mapsto & (15\ x + 7) \mod 26 \end{array}$$

1. Chiffrez le message clair : `hello !`
2. Déchiffrez le message reçu : `ehlxqp !`
3. Explicitez la fonction de déchiffrement en fonction des paramètres $a = 15$ et $b = 7$.
4. A votre avis, peut-on choisir librement les paramètres (a, b) ?
5. Sans connaître les paramètres (a, b) du chiffrement affine $ax + b \mod n$, est-il possible de déchiffrer un message ?

Exercice 8) Transformée de Burrows-Wheeler Cette transformation vise à diminuer l'entropie en triant les lettres d'une chaîne avant transmission ... Comment retrouver la chaîne initiale ? Voici la marche à suivre :

- on crée tous les décalages d'une lettre de la chaîne à raison de un par ligne ;
- on trie les lignes par ordre lexicographique en appellant F la 1re colonne et L la dernière ;
- on calcule un *vecteur de transformation* H entre F et L selon :

$$\forall j \quad L[\ H[j]\ ] = F[j]$$

- l'*index primaire* ip est l'indice de la ligne où apparaît la 1re lettre du mot.

1. Calculez F, L et H sur la chaîne `COMPRESSE`.
2. Concevez un algorithme qui, à partir de F, H et ip seulement permet de reconstruire la chaîne initiale.

Exercice 9) Code d'Huffman Soit la source à coder sur l'alphabet binaire :

Symbole	Proba. d'apparition
a	0,30
b	0,10
c	0,28
d	0,20
e	0,12

1. Construisez l'arbre d'Huffman afin de trouver le mot de code de chacun des symboles ;
2. Expérimentez le codage et le décodage d'un message source sur une courte chaîne sur l'alphabet $\{a, b, c, d, e\}$;
3. Quid de l'optimalité du code d'Huffman obtenu ? *(Vous devez calculer son entropie et sa longueur moyenne pondérée.)*
4. Pourquoi le décodage se fait-il à partir de l'arbre de Huffman et non de la table de codage, comme vu au Cours n° 2 ?

Exercice 10) Codage de source Soit une source qui fournit comme information l'une des quatre lettres a_1, a_2, a_3 et a_4. Supposons que le codage transforme cette information en symboles binaires. Dans la table suivante, nous donnons deux codages différents de cette source. Dans la première méthode, deux symboles binaires sont générés pour chaque lettre émise, alors que dans la seconde, le nombre de symboles est variable.

Méthode 1	Méthode 2
$a_1 \rightsquigarrow 00$	$a_1 \rightsquigarrow 0$
$a_2 \rightsquigarrow 01$	$a_2 \rightsquigarrow 10$
$a_3 \rightsquigarrow 10$	$a_3 \rightsquigarrow 110$
$a_4 \rightsquigarrow 11$	$a_4 \rightsquigarrow 111$

1. Calculez l'entropie pour chacune des deux distributions de probabilité.
2. Donnez des jeux de probabilités d'apparition des lettres rendant tour à tour meilleur l'un ou l'autre codage en termes de longueur moyenne.

Exercices complémentaires

Exercice 11) Transformée de Burrows-Wheeler (décodage) L'information est ordonnée avant d'être transmise de sorte à réduire l'entropie. On transmet la colonne de lettres triée **F** et le vecteur-colonne **H** qui code la correspondance entre la première colonne **F** et dernière colonne **L** dans le tableau des permutations circulaires. Une fois **L** reconstituée, l'*index primaire* permet de retrouver le mot selon l'algorithme de l'Exercice n° 7.

F
C
C
E
E
I
N
S

H
4
3
5
6
2
1
0

L
...
...
...
...
...
...
...

1. Commencez par retrouver **L** à partir de la colonne **F** et du vecteur de correspondance **H** ;
2. Sachant que l'index primaire $ip = 0$ ici, appliquez l'algorithme permettant de retrouver le mot de départ à partir de **L** et de **H**.

Exercice 12) Pile ou Face pour jouer au 421 ? On souhaiterait jouer aux dés au seul moyen d'une pièce de monnaie. On va donc chercher à coder un dé à 6 faces en binaire.

1. Retrouvez l'entropie du dé déjà calculée au TP 1.
2. On propose tout d'abord un codage binaire sur 3 bits en suivant la convention d'écriture : 0 pour `Face` et 1 pour `Pile` :

$1 \rightsquigarrow 001$	$4 \rightsquigarrow 100$
$2 \rightsquigarrow 010$	$5 \rightsquigarrow 101$
$3 \rightsquigarrow 011$	$6 \rightsquigarrow 110$

   Pourquoi un tel codage n'est pas optimal ?
3. Proposez à la place un codage *optimal* du dé par une pièce de monnaie.
4. Pourquoi cela ne suffit pas pour simuler le jeu d'un dé avec une pièce de monnaie ?