Université de Bourgogne
Théorie de l'information

M1 Informatique
Janvier 2013

Examen *théorie de l'information*
Tous documents autorisés
Le barème est donné à titre indicatif

1(3 pt)

1. Quel est le résultat de l'algorithme *Move to Front* appliqué au message : `aaaccccbbc` ?
2. Calculer l'entropie du système obtenu au point 1. et donner le code de Huffman de celui-ci.
3. Donner le résultat de l'algorithme de compression LZ78 appliqué au message `aaaccccbbc`.
4. Donner le résultat de l'algorithme de décompression LZ78 appliqué au message :
   (0,a) (0,b) (1,a) (3,b) (3,a) (2,a) (5,b)

Correction

1. L'évolution du tableau correspondant au message : `aaaccccbbc` est :

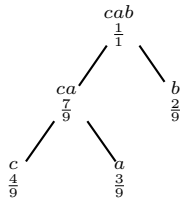
Indice	0	1	2
Etat initial	a	b	c
Tableau conservé par le 1*er a*	a	b	c
Tableau conservé par le 2*èmea*	a	b	c
Tableau conservé par le 3*ème a*	a	b	c
Tableau modifié par le 1*er c*	c	a	b
Tableau conservé par le 2*ème c*	c	a	b
Tableau conservé par le 3*ème c*	c	a	b
Tableau modifié par le 1*ème b*	b	c	a
Tableau conservé par le 2*ème b*	b	c	a
Tableau modifié par le 4*ème c*	c	b	a

Et donc `aaaccccbbc` est code par `000200201` ou, revenant à l'alphabet $\{a, b, c\}$, par `aaacaacab`.

2. L'alphabet est $\{a, b, c\}$ avec les probabilités :

$p(a)$	$p(b)$	$p(c)$
$\frac{3}{9}$	$\frac{2}{9}$	$\frac{4}{9}$,

et l'entropie est : $H = \frac{3}{9} \cdot \log_2 \frac{9}{3} + \frac{2}{9} \cdot \log_2 \frac{9}{2} + \frac{4}{9} \cdot \log_2 \frac{9}{4} = 1,5$.
L'arbre de Huffman est :



et on étiquetant les branches gauches par 0 et les branches droites par 1 on obtient le codage

c	00
a	01
b	1

N.B. L'arbre de Huffman (et donc le codage) n'est pas unique, on aurait pu avoir, par exemple, le codage

c	11
a	10
b	0

3. Le dictionnaire correspondent est :

ditionaire		lexème
0	null	
1	a	(0,a)
2	aa	(1,a)
3	c	(0,c)
4	cc	(3,c)
5	b	(0,b)
6	bc	(5,c)

et la version compressée est : (0,a) (1,a) (0,c) (3,c) (0,b) (5,c).

2 (2 pt) Donner le codage δ de Elias pour le mot hexadécimal $x = 1\underbrace{00\ldots0}_{15 \text{ fois}}A2\underbrace{00\ldots0}_{63 \text{ fois}}$.

Correction

La version compressée de x par le codage δ de Elias est $\mathbf{101F}A2\mathbf{02FF}$

3(3 pt) Soit la matrice génératrice $G = \begin{pmatrix} 100101 \\ 010110 \\ 001011 \end{pmatrix}$ d'un code linéaire.

1. Quels sont les mots du code ?
2. Quelle est la taille et la dimension du code ?
3. Quelle est sa distance minimale?
4. Quel est son taux d'information?
5. Combien détecte-t-il et corrige-t-il d'erreurs ?

Correction

1. Les mots du code sont $x \times G$, avec x des suites binaires de longueur 3.

x	G
000	000 000
001	001 011
010	010 110
011	011 101
100	100 101
101	101 110
110	110 011
111	111 000

2. Taille=6, dimension=3.
3. Distance mini=poids mini=3
4. Taux d'info=50%
5. Le code détecte 2 erreurs et corrige une erreur.

4 (2 pt) Soit le code de longueur 7 obtenu par 3 répétitions doublées par un bit de parité. Exemple : le mot 01 sera codé par 01 01 01 1. Le code ainsi obtenu est un code linaire de dimension 2 et taille 7.

1. Donner la matrice génératrice de ce code;
2. Quels sont les mots du code?
3. Quelle est sa distance minimale?
4. Quel est son taux d'information?
5. Combien corrige-t-il d'erreurs?
6. Vérifier la proposition de Hamming pour ce code.

Correction

1. La matrice génératricematrice génératrice est $G = \begin{pmatrix} 10\,10\,10\,1 \\ 01\,01\,01\,1 \end{pmatrix}$.
2. Les mots du code sont $x \times G$, avec x des suites binaires de longueur 2.

x	G
00	00 00 00 0
01	01 01 01 1
10	10 10 10 1
11	11 11 11 0

3. La distance mini est 4.
4. Le taux d'info est $\frac{2}{7}$.
5. Il detecte 3 erreurs est corrige 1 erreurs.
6. La proposition de Hamming donne une borne supérieur du nombre de mots du code :

$$card(C) \leq \frac{2^n}{\sum_{i=0}^{e} C_n^i}$$

où

- $C : \{0000000, 0101011, 1010101, 1111110\}$,
- $n = 7$ est la taille du code,
- $e = 1$ est le nombre d'erreurs corrigées
- $C_n^i = \frac{n!}{i!(n-i)!}$.

On vérifie facilement que

$$4 \leq \frac{2^7}{C_7^0 + C_7^1}.$$

5 (1 pt) Résoudre les équations suivantes en $\mathbb{F}_7$: $4x + 3 = 1$, $3x + 6 = 0$.

Correction

$4x + 3 = 1 \Rightarrow 4x = 5 \Rightarrow x = 3$ (les calculs sont effectués modulo 7).
$3x + 6 = 0 \Rightarrow 3x = 1 \Rightarrow x = 5$.

6 (4 pt)

1. 2 est un générateur de $\mathbb{F}_7^*$?
2. 3 est un générateur de $\mathbb{F}_7^*$?
3. 5 est un générateur de $\mathbb{F}_7^*$?
4. combien de générateurs possède $\mathbb{F}_7^*$?
5. en $\mathbb{F}_7^*$ calculer $\log_3 4$ et $\log_5 4$
6. Alice et Bob veulent appliquer le protocole de Diffie-Hellman. Ils choisissent en commun $g = 5$ et $p = 7$. Choisir des clés secrètes a (pour Alice) et b (pour Bob) et appliquer le protocole avec ces deux valeurs.

Correction

On construit les fonctions $x \mapsto g^x$ en $\mathbb{F}_7^*$ avec $g \in \{2, 3, 4, 5, 6\}$:

x	1	2	3	4	5	6
2^x	2	4	1	...		
3^x	3	2	6	4	5	1
4^x	4	2	1	...		
5^x	5	4	6	2	3	1
6^x	6	1	...			

1.–4. $\mathbb{F}_7^*$ a deux générateurs : 3 et 5.

5. $\log_3 4 = z \Leftrightarrow 3^z = 4 \Leftrightarrow z = 4$; $\log_5 4 = z \Leftrightarrow 5^z = 4 \Leftrightarrow z = 2$.

6.

• Alice et Bob se mettent d'accord sur un nombre premier $p = 7$ et un générateur de $\mathbb{F}_7^*$, $g = 5$.
• Alice choisit $a \in \mathbb{F}_7^*$, disons $a = 3$ (son secret); calcule $A = g^a = 5^3 = 6$; et envoie $A = 6$ à Bob (les calculs sont effectués modulo $p = 7$)
• Bob choisit $b \in \mathbb{F}_7^*$, disons $b = 5$ (son secret); calcule $B = g^b = 5^5 = 3$; et envoie $B = 3$ à Alice
• Alice calcule $K = B^a = 3^3 = 6$
• Bob calcule $K = A^b = 6^5 = 6$

Ainsi, Alice et Bob partagent un secret, $K = 6$, et quiconque a espionné la ligne ne peut retrouver K.

7 (3 pt) Décrire le protocole RSA avec

1. les données privées d'Alice : $p = 3$, $q = 5$,
2. la clé publique : $(n, e) = (15, 7)$

Correction

• Alice fabrique les clefs :

- choisit deux nombres premiers de 'grande' taille $p = 3$ et $q = 5$;
- choisit $e = 7$ premier avec $(p - 1)(q - 1) = 8$;
- calcule d (son secret) tel que $e \cdot d = 1 \ mod\ 8$, donc $d = 7$ (dans ce cas $e = d$ est une coïncidence);
- publie $(p \cdot q, e) = (15, 7)$;

• Bob veut envoyer $x = 2$ à Alice. Il envoie $y = x^e = 2^7 = 8$, la version chiffrée de x (les calculs sont effectués modulo $n = 15$).
• Alice calcule $y^d = 8^7 = 2 = x$ et retrouve x.

8 (3 pt) Le coffre-fort d'une banque s'ouvre si trois de ses quatre directeurs saisissent leur code (part du secret). Les codes sont :

$$p_1 = 3;\ p_2 = 0;\ p_3 = 6;\ p_4 = 0.$$

Soit $P : \mathbb{F}_7 \to \mathbb{F}_7$ le polynôme correspondent à ce protocole.

1. Donner $P(0)$ (le secret),
2. Donner P.

Correction

2. P est une fonction polynomiale qui passe par les points $(1,3)$, $(2,0)$, $(3,6)$, $(4,0)$. Le polynôme P est déterminé par (n'importe quels) trois points parmi $(1,3)$, $(2,0)$, $(3,6)$, $(4,0)$, donc aura la forme $P(x) = ax^2 + bx + c$. Si on choisit les points $(1,3)$, $(2,0)$, $(3,6)$, on a le système d'équations

$$\begin{cases} P(1) = 3 \\ P(2) = 0 \\ P(3) = 6 \end{cases} \text{ et donc } \begin{cases} a \cdot 1^2 + b \cdot 1 + c = 3 \\ a \cdot 2^2 + b \cdot 2 + c = 0 \\ a \cdot 3^2 + b \cdot 3 + c = 6 \end{cases}$$

avec les solutions $a = b = c = 1$ (les calculs sont effectués modulo 7). Finalement $P(x) = x^2 + x + 1$.

1. Le secret est $P(0) = 1$.