

Travail pratique #1 Introduction – codage arithmétique

Questions

Donnez le développement pour vos réponses, en particulier quand le calcul est long.

1. (10 points.) **Entropie.**

(a) L'artiste punk Anthropy fait un graffiti sur un mur de sa ville chaque nuit. À chaque fois, il choisit un mot au hasard dans le dictionnaire français. Disons que le dictionnaire français contient exactement 15 000 mots. Aussi, Anthropy écrit le mot choisi à l'aide de l'une de ses 7 cannettes de peinture de couleurs variées. Quelle est, en bits, l'entropie associée à chacun de ses graffitis?

(b) Patrick souffre d'insomnie. Tant qu'à passer de longues heures à ne pas dormir, il traverse l'épreuve le plus agréablement possible en jouant aux cartes. Ce qu'il fait est très simple: à chaque tour, il mélange tout le paquet, il tire deux cartes, il remarque les numéros de ces deux cartes (mais pas leurs couleurs) et il inscrit la paire de numéros en ordre croissant. Quelle est l'entropie de la paire de numéros qu'il inscrit à chaque tour?

Supposons que son paquet ne contient pas les jokers; i.e. il contient 52 cartes. De plus, faites attention au fait que le tirage de la première carte a une influence sur le tirage de la deuxième carte: il ne reste alors que 51 cartes dans le paquet et il est un peu moins probable de tirer une carte ayant le même numéro que celle du premier tirage. Voici trois exemples de tours effectués par Patrick.

- Il tire la dame de coeur, puis le 4 de pique, donc il inscrit la paire 4–12.
- Il tire le 6 de trèfle, puis le 9 de pique, donc il inscrit la paire 6–9.
- Il tire l'as de carreau, puis l'as de trèfle, donc il inscrit la paire 1–1.

- (c) La grenouille Kao adore sauter d'un nénuphar à l'autre dans son étang. Il y a 9 nénuphars disposés en une grille 3×3 comme suit.

A	B	C
D	E	F
G	H	I

Malgré ce que suggère son nom, Kao agit d'une manière particulièrement ordonnée. Elle saute exactement une fois à chaque seconde. Elle ne saute qu'à un nénuphar immédiatement voisin, dans l'une des quatre directions nord, sud, est, ouest. Elle ne saute pas ailleurs que sur l'un de ses 9 nénuphars. Avant chacun de ses sauts, elle dresse la liste des nénuphars vers lesquels elle pourrait sauter. Ça signifie qu'elle obtient une liste de 2, 3 ou 4 destinations, tout dépendant d'où elle se trouve. Elle choisit dans la liste le nénuphar vers lequel elle saute avec probabilité uniforme entre les destinations possibles. Voyons trois exemples.

- Si Kao se trouve sur le nénuphar A, alors elle pourrait sauter vers B ou D avec probabilité $1/2$ pour chaque destination.
- Si Kao se trouve sur le nénuphar B, alors elle pourrait sauter vers A, C ou E avec probabilité $1/3$ pour chaque destination.
- Si Kao se trouve sur le nénuphar E, alors elle pourrait sauter vers B, D, F ou H avec probabilité $1/4$.

La façon de sauter de Kao forme un modèle de Markov d'ordre 1. Calculez quelle est l'entropie d'un saut en général.

2. (20 points.) **Codes uniquement décodables.** Pour chacun des codes suivants, indiquez s'il est uniquement décodable. Si oui, donnez une justification (par exemple, une trace de l'algorithme vu en classe). Si non, donnez aussi une justification (par exemple, une chaîne de bits qui est décodable de deux façons différentes).

Code 1	Code 2	Code 3	Code 4	Code 5
a: 00	a: 000	a: 000	a: 00	a: 00
b: 010	b: 00010	b: 00011	b: 001	b: 0010
c: 0110	c: 001	c: 001	c: 0011	c: 001010
d: 0111	d: 00110	d: 0100	d: 001110	d: 010
e: 100000	e: 010	e: 0101	e: 0011101	e: 01010
f: 100001	f: 01010	f: 011	f: 00111011	f: 011
g: 10001	g: 011	g: 01111	g: 001111	g: 01110
h: 1001	h: 01110	h: 100	h: 0011111	h: 0111010
i: 1010	i: 1010	i: 10011	i: 00111111	i: 101010
j: 10110	j: 11	j: 1010	j: 01	j: 110
k: 10111	k: 1110	k: 10101	k: 011	k: 11010
l: 11		l: 1011	l: 0111	l: 1101010
				m: 111
				n: 11110
				o: 1111010

3. (10 points.) **Codes préfixes.** Soit l'alphabet $\{a, b, c, d, e, f, g, h, i, j\}$. Dans chacun des cas si-dessous, indiquez s'il existe un code préfixe dont la longueur du mot de code de chaque symbole est celle indiquée dans le tableau. Si oui, donnez un tel code préfixe. Si non, expliquez pourquoi.

(a)

a	b	c	d	e	f	g	h	i	j
3	2	4	4	5	6	3	4	6	2

(b)

a	b	c	d	e	f	g	h	i	j
5	6	1	2	6	4	6	4	5	3

4. (10 points.) **Codes préfixes.** Considérez l'alphabet Σ suivant ainsi que les fréquences d'apparition associées.

a	b	c	d	e	f	g	h	i	j	k	l
11	2	8	10	7	5	6	4	3	9	12	1

- (a) Construisez un code préfixe pour Σ en utilisant l'algorithme de Shannon-Fano. Calculez la longueur moyenne des mots de code.
- (b) Construisez un code préfixe pour Σ en utilisant l'algorithme de Huffman. Calculez la longueur moyenne des mots de code.
5. (10 points.) **Codes de Golomb.** Considérons l'alphabet des nombres naturels. La probabilité du nombre i , $i \geq 0$, est:

$$p_i = \frac{1}{8} \cdot \left(\frac{7}{8}\right)^i$$

Considérez le code de Golomb de paramètre $m = 5$. Montrez la forme des mots de code pour les divers naturels. Aussi, donnez explicitement les mots de code pour les naturels 0 à 11. Enfin, calculez la longueur moyenne du mot de code d'un naturel tiré aléatoirement en fonction des probabilités.

Voici deux identités qui peuvent vous être utiles.

$$\sum_{i=0}^{\infty} r^i = \frac{1}{1-r} \qquad \sum_{i=0}^{\infty} i \cdot r^i = \frac{r}{(1-r)^2}$$

6. (10 points.) **Codes de Tunstall.** Construisez le code de Tunstall dont le dictionnaire a la capacité de contenir 25 mots. L'alphabet source et les probabilités des symboles sont les suivants.

a	b	c	d
$\frac{3}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{4}{10}$

7. (15 points.) **Codage arithmétique.** Soient l'alphabet et les probabilités des symboles suivants.

a	b	c	d
$\frac{4}{20}$	$\frac{5}{20}$	$\frac{9}{20}$	$\frac{2}{20}$

Vous devez décrire la gestion de l'intervalle de travail au cours de l'encodage de la chaîne **acdb**, en commençant par l'intervalle complet. Il y a cinq moments distincts: le moment précédant l'encodage du premier symbole et chacun des moments qui suivent immédiatement l'encodage d'un symbole. Vous devez faire les redimensionnements de l'intervalle de travail (à l'aide de E_1 , E_2 et E_3) au fur et à mesure. Faites deux fois l'exercice: une fois en manipulant des fractions exactes et une autre fois en nombres entiers, où l'intervalle complet comporte 16 tuiles.

Remise des travaux

Vous devez remettre le travail dans la boîte de dépôt de **monPortail** au plus tard le **25 février**. Veuillez ignorer la date de remise inscrite dans le plan de cours. Les autres modalités de remise sont inscrites dans le plan de cours.