

Solutionnaire du travail pratique #1

Réponses

Note: les réponses aux différentes questions ne sont pas toujours uniques.

- (a) À chaque nuit, Anthropy réalise l'une de $15\,000 \cdot 7$ oeuvres artistiques. C'est comme lancer un dé à 105 000 faces équiprobables. L'entropie d'un lancer est de $\log_2 105\,000 \approx 16.68$ bits.

- (b) Il faut d'abord calculer la probabilité de chaque paire de numéros possible. Il y a deux cas pour une paire $i-j$: soit $i = j$, soit $i < j$.

Considérons d'abord le cas $i = j$. Il faut d'abord tirer une carte de numéro i ; la carte peut être de l'une des 4 sortes possibles; ce tirage réussit avec probabilité $4/52$. Ensuite, il faut tirer à nouveau une carte de numéro i ; la carte peut être de l'une des 3 sortes restantes; ce tirage réussit avec probabilité $3/51$. La probabilité d'une paire de numéros $i-j$, $i = j$, est donc de $1/221$.

Considérons maintenant le cas $i < j$. Il y a deux (groupes de) façons de produire cette paire: tirer i avant j et tirer j avant i . Ces deux façons sont disjointes l'une de l'autre. On devra donc additionner les probabilités respectives de ces façons. Considérons d'abord le cas où i est tiré avant j . Il faut tout d'abord tirer une carte de numéro i ; la carte peut être de n'importe quelle sorte; ce tirage réussit avec probabilité $4/52$. Ensuite, il faut tirer une carte de numéro j ; la carte peut à nouveau être de n'importe quelle sorte; ce tirage réussit avec probabilité $4/51$. La probabilité de tirer i puis j est donc de $4/663$. Par un raisonnement similaire, on a une probabilité de $4/663$ de tirer j puis i . En somme, la paire $i-j$, $i < j$, a la probabilité $8/663$ de survenir.

En résumé, la probabilité des paires est définie ainsi:

$$P(i-j) = \begin{cases} 1/221, & \text{si } i = j; \\ 8/663, & \text{si } i < j. \end{cases}$$

En principe, on doit pouvoir vérifier l'identité suivante: $\sum_{\text{paire } i-j} P(i-j) = 1$.

Ce qui nous intéresse ici, c'est l'entropie du processus de l'obtention des paires. Soit X une variable aléatoire qui génère les paires de numéros de Patrick.

$$\begin{aligned} H(X) &= \sum_{\text{paire } i-j} P(i-j) \cdot \log_2 \frac{1}{P(i-j)} \\ &= \sum_{i=1}^{13} \sum_{j=i}^{13} P(i-j) \cdot \log_2 \frac{1}{P(i-j)} \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^{13} \left(P(i-i) \cdot \log_2 \frac{1}{P(i-i)} + \sum_{j=i+1}^{13} P(i-j) \cdot \log_2 \frac{1}{P(i-j)} \right) \\
&= \sum_{i=1}^{13} \left(\frac{1}{221} \cdot \log_2 \frac{1}{1/221} + \sum_{j=i+1}^{13} \frac{8}{663} \cdot \log_2 \frac{1}{8/663} \right) \\
&= \sum_{i=1}^{13} \left(\frac{1}{221} \cdot \log_2 221 + \sum_{j=i+1}^{13} \frac{8}{663} \cdot \log_2 \frac{663}{8} \right) \\
&\approx \sum_{i=1}^{13} \left(0.0352 + \sum_{j=i+1}^{13} 0.0769 \right) \\
&= \sum_{i=1}^{13} (0.0352 + (13-i) \cdot 0.0769) \\
&\approx \sum_{i=1}^{13} (1.0349 - i \cdot 0.0769) \\
&= \left(\sum_{i=1}^{13} 1.0349 \right) - 0.0769 \cdot \sum_{i=1}^{13} i \\
&= 13 \cdot 1.0349 - 0.0769 \cdot \frac{14 \cdot 13}{2} \\
&\approx 6.4558
\end{aligned}$$

- (c) Calculons d'abord les probabilités stationnaires de voir Kao se trouver sur les différents nénuphars. Par abus de langage, je vais parler des probabilités stationnaires *des nénuphars*. Soient p_A la probabilité du nénuphar A, p_B la probabilité du nénuphar B, etc. Par symétrie du problème, on sait qu'on a $p_A = p_C = p_G = p_I$ et $p_B = p_D = p_F = p_H$.

Commençons par poser une équation pour déterminer p_A . Pour pouvoir se retrouver sur le nénuphar A, il fallait que Kao soit sur le nénuphar B ou sur le nénuphar D et saute vers A. On a donc l'équation suivante.

$$p_A = p_B \cdot \frac{1}{3} + p_D \cdot \frac{1}{3}$$

Posons une équation pour déterminer p_B . Pour pouvoir se retrouver sur le nénuphar B, il fallait que Kao soit sur le nénuphar A, sur le nénuphar C ou sur le nénuphar E et saute vers B. On a donc l'équation suivante.

$$p_B = p_A \cdot \frac{1}{2} + p_C \cdot \frac{1}{2} + p_E \cdot \frac{1}{4}$$

Posons une équation pour déterminer p_E . Pour pouvoir se retrouver sur le nénuphar E, il fallait que Kao soit sur le nénuphar B, sur le nénuphar D, sur

le nénuphar F ou sur le nénuphar H et saute vers E. On a donc l'équation suivante.

$$p_E = p_B \cdot \frac{1}{3} + p_D \cdot \frac{1}{3} + p_F \cdot \frac{1}{3} + p_H \cdot \frac{1}{3}$$

Enfin, il faut poser une équation qui force la somme des probabilités à atteindre 1.

$$p_A + p_B + p_C + p_D + p_E + p_F + p_G + p_H + p_I = 1$$

À partir de ce point, nous pouvons profiter des symétries pour simplifier les équations en le système suivant.

$$\begin{aligned} p_A &= 2 \cdot p_B \cdot \frac{1}{3} \\ p_B &= 2 \cdot p_A \cdot \frac{1}{2} + p_E \cdot \frac{1}{4} \\ p_E &= 4 \cdot p_B \cdot \frac{1}{3} \\ 4 \cdot p_A + 4 \cdot p_B + p_E &= 1 \end{aligned}$$

Solutionner ce système nous indique qu'on a $p_A = 1/12$, $p_B = 1/8$ et $p_E = 1/6$. L'entropie du saut de Kao *hors* de A (ou de C, de G ou de I) est de $\log_2 2 = 1$ bit. L'entropie du saut de Kao *hors* de B (ou de D, de F ou de H) est de $\log_2 3 \approx 1.585$ bits. L'entropie du saut de Kao *hors* de E est de $\log_2 4 = 2$ bits. En conséquence, l'entropie d'un saut en général est la suivante.

$$\begin{aligned} &H(\text{saut en général}) \\ &= \sum_{n \in \{A, \dots, I\}} p_n \cdot H(\text{saut hors de } n) \\ &= \left(\sum_{n \in \{A, C, G, I\}} p_n \cdot H(\text{saut hors de } n) \right) \\ &= + \left(\sum_{n \in \{B, D, F, H\}} p_n \cdot H(\text{saut hors de } n) \right) \\ &\quad + \left(\sum_{n \in \{E\}} p_n \cdot H(\text{saut hors de } n) \right) \\ &= \left(\sum_{n \in \{A, C, G, I\}} \frac{1}{12} \cdot \log_2 2 \right) + \left(\sum_{n \in \{B, D, F, H\}} \frac{1}{8} \cdot \log_2 3 \right) + \left(\sum_{n \in \{E\}} \frac{1}{6} \cdot \log_2 4 \right) \\ &= 4 \cdot \frac{1}{12} \cdot \log_2 2 + 4 \cdot \frac{1}{8} \cdot \log_2 3 + \frac{1}{6} \cdot \log_2 4 \\ &\approx 1.459 \text{ bits} \end{aligned}$$

2. **Code 1** : il est uniquement décodable car il est un code préfixe.

Code 2 : il est uniquement décodable. Utilisons l'algorithme de test de décodabilité unique. On commence par initialiser DS. On constate que le mot de code 000 est un préfixe du mot de code 00010, menant au *dangling suffix* 10. On découvre le *dangling suffix* 10 à de nombreuses autres reprises mais ça demeure le seul. Ainsi, DS est initialisé à {10}. Lors de la première itération, on compare les mots de code et les *dangling suffixes* pour trouver de nouveaux *dangling suffixes*. En particulier, on voit que le *dangling suffix* 10 est un préfixe du mot de code 1010. Toutefois, ceci ne mène pas à la découverte de nouveaux *dangling suffixes*. À l'inverse, aucun mot de code n'est préfixe d'un *dangling suffix*. Donc, DS est stabilisé. Aucun *dangling suffix* n'est un mot de code. On conclut que le code est uniquement décodable.

Code 3 : il n'est pas uniquement décodable car:

$$C(\mathbf{k}) \cdot C(\mathbf{e}) = 10101 \cdot 0101 = 1010 \cdot 10101 = C(\mathbf{j}) \cdot C(\mathbf{k}) .$$

Code 4 : il est uniquement décodable. Malheureusement, la trace de l'algorithme de test est longue. À la fin, on obtient une longue liste de *dangling suffixes* mais aucun n'est un mot de code. Voici la liste des *dangling suffixes* et la façon dont chacun peut être découvert.

- 1 est un d. s. entre les m. de c. 00 et 001 .
- 10 est un d. s. entre les m. de c. 0011 et 001110 .
- 101 est un d. s. entre les m. de c. 0011 et 0011101 .
- 1011 est un d. s. entre les m. de c. 0011 et 00111011.
- 11 est un d. s. entre les m. de c. 00 et 0011 .
- 110 est un d. s. entre les m. de c. 001 et 001110 .
- 1101 est un d. s. entre les m. de c. 001 et 0011101 .
- 11011 est un d. s. entre les m. de c. 001 et 00111011.
- 111 est un d. s. entre les m. de c. 001 et 001111 .
- 1110 est un d. s. entre les m. de c. 00 et 001110 .
- 11101 est un d. s. entre les m. de c. 00 et 0011101 .
- 111011 est un d. s. entre les m. de c. 00 et 00111011.
- 1111 est un d. s. entre les m. de c. 00 et 001111 .
- 11111 est un d. s. entre les m. de c. 00 et 0011111 .
- 111111 est un d. s. entre les m. de c. 00 et 00111111.

Code 5 : il est uniquement décodable. Utilisons l'algorithme de test de décodabilité unique. On commence par initialiser DS. On constate que le mot de code 00 est un préfixe des mots de code 0010 et 001010, menant aux *dangling suffixes* 10 et 1010. On peut redécouvrir ces *dangling suffixes* à de nombreuses autres reprises. Ainsi, DS est initialisé à {10, 1010}. Lors de la première itération, on peut redécouvrir les mêmes deux *dangling suffixes* en comparant les mots de code et les *dangling suffixes* ensemble. Toutefois, ceci ne mène pas à la découverte de nouveaux *dangling suffixes*. Donc, DS est stabilisé. Aucun *dangling suffix* n'est un mot de code. On conclut que le code est uniquement décodable.

3. (a) Oui, il y a le code préfixe suivant.

a	100		b	00
b	00		j	01
c	1100		a	100
d	1101		g	101
e	11110		c	1100
f	111110	ou, présenté autrement,	d	1101
g	101		h	1110
h	1110		e	11110
i	111111		f	111110
j	01		i	111111

(b) Il n'y a pas de tel code préfixe car les longueurs des mots de code ne respectent pas l'inégalité de Kraft-McMillan.

$$\begin{aligned}
 & \frac{1}{2^5} + \frac{1}{2^6} + \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^6} + \frac{1}{2^4} + \frac{1}{2^6} + \frac{1}{2^4} + \frac{1}{2^5} + \frac{1}{2^3} \\
 = & \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \frac{1}{2^4} + \frac{1}{2^5} + \frac{1}{2^5} + \frac{1}{2^6} + \frac{1}{2^6} + \frac{1}{2^6} \\
 > & \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \frac{1}{2^4} \\
 = & 1
 \end{aligned}$$

4. (a) On trie les symboles par fréquences d'apparition. Le code que vous avez construit peut être passablement différent; certains groupes de symboles peuvent être divisés de deux façons différentes à cause d'égalités dans les sommes de fréquences.

```

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

```

Donc, le code obtenu est le suivant.

l : 00000		a : 110
b : 00001		b : 00001
i : 0001		c : 011
h : 0010		d : 101
f : 0011		e : 0101
g : 0100	ou, présenté autrement:	f : 0011
e : 0101		g : 0100
c : 011		h : 0010
j : 100		i : 0001
d : 101		j : 100
a : 110		k : 111
k : 111		l : 00000

$$\begin{aligned}
 L_{\text{moy}} &= \frac{11 * 3 + 2 * 5 + 8 * 3 + 10 * 3 + 7 * 4 + 5 * 4 + 6 * 4 + 4 * 4 + 3 * 4 + 9 * 3 + 12 * 3 + 1 * 5}{78} \\
 &= \frac{(11 + 8 + 10 + 9 + 12) * 3 + (7 + 5 + 6 + 4 + 3) * 4 + (2 + 1) * 5}{78} \\
 &= \frac{50 * 3 + 25 * 4 + 3 * 5}{78} \\
 &= \frac{150 + 100 + 15}{78} \\
 &= \frac{265}{78}
 \end{aligned}$$

(b) L'algorithme de Huffman doit joindre des symboles pour en faire des frères. Par exemple, à la première étape ci-bas, les symboles **l** et **b** sont joints. Nous dénotons par $A \equiv l \swarrow \searrow b$ le fait de joindre ces deux symboles et de les remplacer par un nouveau symbole artificiel nommé **A**. Cette notation signifie que le mot de code de **l** sera formé en ajoutant 0 au mot de code de **A** et que le mot de code de **b** sera formé en ajoutant 1 au mot de code de **A**. De plus, la fréquence de **A** est la somme des fréquences de **l** et de **b**.

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

$A \equiv l \swarrow \searrow b$

$i : 3, A : 3, h : 4, f : 5, g : 6, e : 7, c : 8, j : 9, d : 10, a : 11, k : 12$

$B \equiv i \swarrow \searrow A$

$h : 4, f : 5, g : 6, B : 6, e : 7, c : 8, j : 9, d : 10, a : 11, k : 12$

$C \equiv h \swarrow \searrow f$

$g : 6, B : 6, e : 7, c : 8, j : 9, C : 9, d : 10, a : 11, k : 12$

$D \equiv g \swarrow \searrow B$

$e : 7, c : 8, j : 9, C : 9, d : 10, a : 11, k : 12, D : 12$

$E \equiv e \swarrow \searrow c$

$j : 9, C : 9, d : 10, a : 11, k : 12, D : 12, E : 15$

$F \equiv j \swarrow \searrow C$

Donc, le code obtenu est le suivant.

$e : 000$	$a : 101$	$d : 10, a : 11, k : 12, D : 12, E : 15, F : 18$
$c : 001$	$b : 111111$	$G \equiv d \swarrow \searrow a$
$j : 010$	$c : 001$	$k : 12, D : 12, E : 15, F : 18, G : 21$
$h : 0110$	$d : 100$	$H \equiv k \swarrow \searrow D$
$f : 0111$	$e : 000$	$E : 15, F : 18, G : 21, H : 24$
$d : 100$	ou,	$I \equiv E \swarrow \searrow F$
$a : 101$	présenté	$G : 21, H : 24, I : 33$
$k : 110$	autrement:	$J \equiv G \swarrow \searrow H$
$g : 1110$	$i : 11110$	$I : 33, J : 45$
$i : 11110$	$j : 010$	$K \equiv I \swarrow \searrow J$
$l : 111110$	$k : 110$	$K : 78$
$b : 111111$	$l : 111110$	

$$\begin{aligned}
 L_{\text{moy}} &= \frac{11 * 3 + 2 * 6 + 8 * 3 + 10 * 3 + 7 * 3 + 5 * 4 + 6 * 4 + 4 * 4 + 3 * 5 + 9 * 3 + 12 * 3 + 1 * 6}{78} \\
 &= \frac{(11 + 8 + 10 + 7 + 9 + 12) * 3 + (5 + 6 + 4) * 4 + 3 * 5 + (2 + 1) * 6}{78} \\
 &= \frac{57 * 3 + 15 * 4 + 3 * 5 + 3 * 6}{78} \\
 &= \frac{171 + 60 + 15 + 18}{78} \\
 &= \frac{264}{78}
 \end{aligned}$$

5. Pour construire le mot de code pour le nombre i , on doit déterminer le numéro du groupe de naturels, $q = \lfloor i/5 \rfloor$, et identifier le membre du groupe, $r = i - 5q \in \{0, 1, 2, 3, 4\}$. Le numéro de groupe est encodé à l'aide du code unaire vu en classe. Le numéro q se traduit en le mot de code 1^q0 , en $q + 1$ bits. Le numéro de membre est encodé à l'aide d'un code préfixe. Les numéros $r \in \{0, 1, 2, 3, 4\}$ se traduisent en les mots de code 00, 01, 10, 110 et 111, respectivement.

0 : 000	3 : 0110	6 : 1001	9 : 10111
1 : 001	4 : 0111	7 : 1010	10 : 11000
2 : 010	5 : 1000	8 : 10110	11 : 11001

Soit $l(i)$ la longueur du mot de code pour le naturel i . La longueur $l(i)$ vient de la somme de la longueur $l_1(q)$ du mot de code pour le quotient q et de la longueur $l_2(r)$ du mot de code pour le reste r . Nous avons donc que $l_1(q) = q + 1$ et $l_2(r) = \lfloor \frac{r+12}{5} \rfloor$. Maintenant, soit l_{moy} la longueur moyenne du mot de code pour un naturel tiré au hasard.

$$\begin{aligned}
l_{moy} &= \sum_{i=0}^{\infty} p_i \cdot l(i) \\
&= \sum_{q=0}^{\infty} \sum_{r=0}^4 p_{5q+r} \cdot l(5q+r) \\
&= \sum_{q=0}^{\infty} \sum_{r=0}^4 p_{5q+r} \cdot (l_1(q) + l_2(r)) \\
&= \sum_{q=0}^{\infty} \sum_{r=0}^4 p_{5q+r} \cdot \left(q + 1 + \left\lfloor \frac{r+12}{5} \right\rfloor \right) \\
&= \sum_{q=0}^{\infty} \left(p_{5q} \cdot \left(q + 1 + \left\lfloor \frac{12}{5} \right\rfloor \right) + p_{5q+1} \cdot \left(q + 1 + \left\lfloor \frac{13}{5} \right\rfloor \right) + \right. \\
&\quad \left. p_{5q+2} \cdot \left(q + 1 + \left\lfloor \frac{14}{5} \right\rfloor \right) + p_{5q+3} \cdot \left(q + 1 + \left\lfloor \frac{15}{5} \right\rfloor \right) + \right. \\
&\quad \left. p_{5q+4} \cdot \left(q + 1 + \left\lfloor \frac{16}{5} \right\rfloor \right) \right) \\
&= \sum_{q=0}^{\infty} \left(p_{5q} \cdot (q+3) + p_{5q+1} \cdot (q+3) + \right. \\
&\quad \left. p_{5q+2} \cdot (q+3) + p_{5q+3} \cdot (q+4) + \right. \\
&\quad \left. p_{5q+4} \cdot (q+4) \right) \\
&= \sum_{q=0}^{\infty} \left(\frac{1}{8} \cdot \left(\frac{7}{8} \right)^{5q} \cdot (q+3) + \frac{1}{8} \cdot \left(\frac{7}{8} \right)^{5q+1} \cdot (q+3) + \right. \\
&\quad \left. \frac{1}{8} \cdot \left(\frac{7}{8} \right)^{5q+2} \cdot (q+3) + \frac{1}{8} \cdot \left(\frac{7}{8} \right)^{5q+3} \cdot (q+4) + \right. \\
&\quad \left. \frac{1}{8} \cdot \left(\frac{7}{8} \right)^{5q+4} \cdot (q+4) \right) \\
&= \frac{1}{8} \cdot \sum_{q=0}^{\infty} \left(\left(\frac{7}{8} \right)^{5q} \cdot (q+3) + \left(\frac{7}{8} \right)^{5q+1} \cdot (q+3) + \right. \\
&\quad \left. \left(\frac{7}{8} \right)^{5q+2} \cdot (q+3) + \left(\frac{7}{8} \right)^{5q+3} \cdot (q+4) + \right. \\
&\quad \left. \left(\frac{7}{8} \right)^{5q+4} \cdot (q+4) \right) \\
&= \frac{1}{8} \cdot \sum_{q=0}^{\infty} \left(\left(\frac{7}{8} \right)^{5q} \cdot q + \left(\frac{7}{8} \right)^{5q+1} \cdot q + \left(\frac{7}{8} \right)^{5q+2} \cdot q + \right. \\
&\quad \left. \left(\frac{7}{8} \right)^{5q+3} \cdot q + \left(\frac{7}{8} \right)^{5q+4} \cdot q \right) + \\
&= \frac{1}{8} \cdot \sum_{q=0}^{\infty} \left(\left(\frac{7}{8} \right)^{5q} \cdot 3 + \left(\frac{7}{8} \right)^{5q+1} \cdot 3 + \left(\frac{7}{8} \right)^{5q+2} \cdot 3 + \right. \\
&\quad \left. \left(\frac{7}{8} \right)^{5q+3} \cdot 4 + \left(\frac{7}{8} \right)^{5q+4} \cdot 4 \right)
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{8} \cdot \sum_{q=0}^{\infty} \left(\frac{7}{8}\right)^{5q} \cdot q \cdot \left(\left(\frac{7}{8}\right)^0 + \left(\frac{7}{8}\right)^1 + \left(\frac{7}{8}\right)^2 + \left(\frac{7}{8}\right)^3 + \left(\frac{7}{8}\right)^4\right) + \\
&= \frac{1}{8} \cdot \sum_{q=0}^{\infty} \left(\frac{7}{8}\right)^{5q} \cdot \left(\left(\frac{7}{8}\right)^0 \cdot 3 + \left(\frac{7}{8}\right)^1 \cdot 3 + \left(\frac{7}{8}\right)^2 \cdot 3 + \left(\frac{7}{8}\right)^3 \cdot 4 + \left(\frac{7}{8}\right)^4 \cdot 4\right) \\
&= \frac{1}{8} \cdot \left(\left(\frac{7}{8}\right)^0 + \left(\frac{7}{8}\right)^1 + \left(\frac{7}{8}\right)^2 + \left(\frac{7}{8}\right)^3 + \left(\frac{7}{8}\right)^4\right) \cdot \sum_{q=0}^{\infty} \left(\frac{7}{8}\right)^{5q} \cdot q + \\
&= \frac{1}{8} \cdot \left(\left(\frac{7}{8}\right)^0 \cdot 3 + \left(\frac{7}{8}\right)^1 \cdot 3 + \left(\frac{7}{8}\right)^2 \cdot 3 + \left(\frac{7}{8}\right)^3 \cdot 4 + \left(\frac{7}{8}\right)^4 \cdot 4\right) \cdot \sum_{q=0}^{\infty} \left(\frac{7}{8}\right)^{5q} \\
&= \frac{1}{8} \cdot \left(\left(\frac{7}{8}\right)^0 + \left(\frac{7}{8}\right)^1 + \left(\frac{7}{8}\right)^2 + \left(\frac{7}{8}\right)^3 + \left(\frac{7}{8}\right)^4\right) \cdot \sum_{q=0}^{\infty} \left(\frac{7^5}{8^5}\right)^q \cdot q + \\
&= \frac{1}{8} \cdot \left(\left(\frac{7}{8}\right)^0 \cdot 3 + \left(\frac{7}{8}\right)^1 \cdot 3 + \left(\frac{7}{8}\right)^2 \cdot 3 + \left(\frac{7}{8}\right)^3 \cdot 4 + \left(\frac{7}{8}\right)^4 \cdot 4\right) \cdot \sum_{q=0}^{\infty} \left(\frac{7^5}{8^5}\right)^q \\
&= \frac{1}{8} \cdot \left(\left(\frac{7}{8}\right)^0 + \left(\frac{7}{8}\right)^1 + \left(\frac{7}{8}\right)^2 + \left(\frac{7}{8}\right)^3 + \left(\frac{7}{8}\right)^4\right) \cdot \frac{\frac{7^5}{8^5}}{\left(1 - \frac{7^5}{8^5}\right)^2} + \\
&= \frac{1}{8} \cdot \left(\left(\frac{7}{8}\right)^0 \cdot 3 + \left(\frac{7}{8}\right)^1 \cdot 3 + \left(\frac{7}{8}\right)^2 \cdot 3 + \left(\frac{7}{8}\right)^3 \cdot 4 + \left(\frac{7}{8}\right)^4 \cdot 4\right) \cdot \frac{1}{1 - \frac{7^5}{8^5}} \\
&= \frac{1}{8} \cdot \left(\left(\frac{7}{8}\right)^0 + \left(\frac{7}{8}\right)^1 + \left(\frac{7}{8}\right)^2 + \left(\frac{7}{8}\right)^3 + \left(\frac{7}{8}\right)^4\right) \cdot \frac{7^5 \cdot 8^5}{(8^5 - 7^5)^2} + \\
&= \frac{1}{8} \cdot \left(\left(\frac{7}{8}\right)^0 \cdot 3 + \left(\frac{7}{8}\right)^1 \cdot 3 + \left(\frac{7}{8}\right)^2 \cdot 3 + \left(\frac{7}{8}\right)^3 \cdot 4 + \left(\frac{7}{8}\right)^4 \cdot 4\right) \cdot \frac{8^5}{8^5 - 7^5} \\
&\approx 4.375
\end{aligned}$$

6. On fait croître le dictionnaire de mots sources tant qu'on reste à l'intérieur de la limite de 25 entrées.

a	3/10
b	2/10
c	1/10
d	4/10

a	30/100
b	20/100
c	10/100
da	12/100
db	8/100
dc	4/100
dd	16/100

aa	9/100
ab	6/100
ac	3/100
ad	12/100
b	20/100
c	10/100
da	12/100
db	8/100
dc	4/100
dd	16/100

aa	9/100
ab	6/100
ac	3/100
ad	12/100
ba	6/100
bb	4/100
bc	2/100
bd	8/100
c	10/100
da	12/100
db	8/100
dc	4/100
dd	16/100

aa	90/1000
ab	60/1000
ac	30/1000
ad	120/1000
ba	60/1000
bb	40/1000
bc	20/1000
bd	80/1000
c	100/1000
da	120/1000
db	80/1000
dc	40/1000
dda	48/1000
ddb	32/1000
ddc	16/1000
ddd	64/1000

aa	90/1000
ab	60/1000
ac	30/1000
ada	36/1000
adb	24/1000
adc	12/1000
add	48/1000
ba	60/1000
bb	40/1000
bc	20/1000
bd	80/1000
c	100/1000
da	120/1000
db	80/1000
dc	40/1000
dda	48/1000
ddb	32/1000
ddc	16/1000
ddd	64/1000

aa	90/1000
ab	60/1000
ac	30/1000
ada	36/1000
adb	24/1000
adc	12/1000
add	48/1000
ba	60/1000
bb	40/1000
bc	20/1000
bd	80/1000
c	100/1000
daa	36/1000
dab	24/1000
dac	12/1000
dad	48/1000
db	80/1000
dc	40/1000
dda	48/1000
ddb	32/1000
ddc	16/1000
ddd	64/1000

aa	90/1000
ab	60/1000
ac	30/1000
ada	36/1000
adb	24/1000
adc	12/1000
add	48/1000
ba	60/1000
bb	40/1000
bc	20/1000
bd	80/1000
ca	30/1000
cb	20/1000
cc	10/1000
cd	40/1000
daa	36/1000
dab	24/1000
dac	12/1000
dad	48/1000
db	80/1000
dc	40/1000
dda	48/1000
ddb	32/1000
ddc	16/1000
ddd	64/1000

7. Le tableau suivant indique l'état de l'intervalle avant et après le traitement de chaque caractère de l'entrée à coder. L'état du compteur $E_3.count$ est aussi indiqué. Enfin, l'intervalle résultant du codage du symbole est redimensionné si nécessaire. Les éventuels bits émis par le redimensionnement sont indiqués.

# SYMBOLE	I	II	III	IV
INT./ $E_3.count$ DÉBUT	$[\frac{0}{20}, \frac{20}{20})/0$	$[\frac{0}{100}, \frac{80}{100})/0$	$[\frac{110}{500}, \frac{470}{500})/1$	$[\frac{1180}{2500}, \frac{1900}{2500})/0$
INT. POUR a	$[\frac{0}{20}, \frac{4}{20})$	$[\frac{0}{100}, \frac{16}{100})$	$[\frac{110}{500}, \frac{182}{500})$	$[\frac{1180}{2500}, \frac{1324}{2500})$
INT. POUR b	$[\frac{4}{20}, \frac{9}{20})$	$[\frac{16}{100}, \frac{36}{100})$	$[\frac{182}{500}, \frac{272}{500})$	$[\frac{1324}{2500}, \frac{1504}{2500})$
INT. POUR c	$[\frac{9}{20}, \frac{18}{20})$	$[\frac{36}{100}, \frac{72}{100})$	$[\frac{272}{500}, \frac{434}{500})$	$[\frac{1504}{2500}, \frac{1828}{2500})$
INT. POUR d	$[\frac{18}{20}, \frac{20}{20})$	$[\frac{72}{100}, \frac{80}{100})$	$[\frac{434}{500}, \frac{470}{500})$	$[\frac{1828}{2500}, \frac{1900}{2500})$
INT./ $E_3.count$ MILIEU	$[\frac{0}{20}, \frac{4}{20})/0$	$[\frac{36}{100}, \frac{72}{100})/0$	$[\frac{434}{500}, \frac{470}{500})/1$	$[\frac{1324}{2500}, \frac{1504}{2500})/0$
REDIM. & BITS	$E_1 : 0$	$E_3 : \epsilon$	$E_2 : 10$	$E_2 : 1$
INT./ $E_3.count$	$[\frac{0}{20}, \frac{8}{20})/0$	$[\frac{22}{100}, \frac{94}{100})/1$	$[\frac{368}{500}, \frac{440}{500})/0$	$[\frac{148}{2500}, \frac{508}{2500})/0$
REDIM. & BITS	$E_1 : 0$		$E_2 : 1$	$E_1 : 0$
INT./ $E_3.count$	$[\frac{0}{20}, \frac{16}{20})/0$		$[\frac{236}{500}, \frac{380}{500})/0$	$[\frac{296}{2500}, \frac{1016}{2500})/0$
REDIM. & BITS				$E_1 : 0$
INT./ $E_3.count$				$[\frac{592}{2500}, \frac{2032}{2500})/0$
INT./ $E_3.count$ FIN	$[\frac{0}{20}, \frac{16}{20})/0$	$[\frac{22}{100}, \frac{94}{100})/1$	$[\frac{236}{500}, \frac{380}{500})/0$	$[\frac{592}{2500}, \frac{2032}{2500})/0$

Cette fois-ci, l'intervalle est géré en nombres entiers. L'usage d'aussi peu que 16 tuiles mène à de sévères distorsions des probabilités. Il faut faire attention à ne pas assigner un intervalle vide à l'un des symboles sources.

# SYMBOLE	I	II	III	IV
INT./ $E_3.count$ DÉBUT	$[0, 15]/0$	$[0, 11]/0$	$[2, 13]/1$	$[0, 15]/0$
INT. POUR a	$[0, 2]$	$[0, 1]$	$[2, 3]$	$[0, 2]$
INT. POUR b	$[3, 6]$	$[2, 4]$	$[4, 6]$	$[3, 6]$
INT. POUR c	$[7, 13]$	$[5, 10]$	$[7, 12]$	$[7, 13]$
INT. POUR d	$[14, 15]$	$[11, 11]$	$[13, 13]$	$[14, 15]$
INT./ $E_3.count$ MILIEU	$[0, 2]/0$	$[5, 10]/0$	$[13, 13]/1$	$[3, 6]/0$
REDIM. & BITS	$E_1 : 0$	$E_3 : \epsilon$	$E_2 : 10$	$E_1 : 0$
INT./ $E_3.count$	$[0, 5]/0$	$[2, 13]/1$	$[10, 11]/0$	$[6, 13]/0$
REDIM. & BITS	$E_1 : 0$		$E_2 : 1$	
INT./ $E_3.count$	$[0, 11]/0$		$[4, 7]/0$	
REDIM. & BITS			$E_1 : 0$	
INT./ $E_3.count$			$[8, 15]/0$	
REDIM. & BITS			$E_2 : 1$	
INT./ $E_3.count$			$[0, 15]/0$	
INT./ $E_3.count$ FIN	$[0, 11]/0$	$[2, 13]/1$	$[0, 15]/0$	$[6, 13]/0$