Leçon II.3 – Examen final 2018 Q1.1

$$H(X) = \sum_{i=1}^{N} p_i \log_2 \frac{1}{p_i} = - \sum_{i=1}^{N} p_i \log_2 p_i$$

Un sac contient 24 billes de 4 couleurs différentes :
12 billes rouges, 8 billes bleues, 2 billes vertes, et 2 billes oranges.

Quelle est, en bit, l'entropie du jeu consistant à deviner la couleur d'une bille tirée au hasard ?

Donnez votre réponse sous la forme $a + b\log_2(3)$.

p_i $\quad \frac{1}{2} \quad \frac{1}{3} \quad \underbrace{\frac{1}{12} \quad \frac{1}{12}}$

$\frac{1}{2} \times 1 \quad \frac{1}{3}\log 3 \quad \frac{2}{12}\boxed{\log 12} = \log 3 + \underbrace{\log 4}_{2}$

N bits

$\boxed{0}$ 0 1 0 1 ? ? ? N-1

$\underbrace{1}{\log_2 2}$

On utilise une représentation des nombres entiers *signés* sur N bits et on s'en sert pour représenter uniquement les nombres entiers positifs (zéro y compris). Sachant que N est pair, quelle est (en bit) l'entropie maximale (telle que définie en cours) pour une telle séquence (de symboles 0/1)?

A] $\log_2(N-1)$ **B]** $\log_2(N)-1$ **C]** 1 **D]** $N-1$

1 10 - 20 pas assez $\simeq 10$

N − 1 pair

$\boxed{0}$ 0 1 ? 0 1 ? ? ?

N impairs

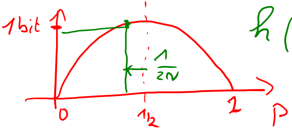
$\longrightarrow$ p les 0

1−p les 1

$\underbrace{(1-p)\,\dfrac{N-1}{2}}_{\text{les } 1}$ $\quad$ $\underbrace{p\,\dfrac{N-1}{2} + 1}_{\text{les } 0}$

$$h(p) = p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p}$$

entropie d'un choix binaire



$$h\left(\frac{N-1}{2N}\right) = h\left(\frac{1}{2} - \frac{1}{2N}\right)$$

$$0 \leq H(X) \leq \log n$$

taille alphabet
nb de lettres

$=$
déterministe
(pas choix)

$=$
équirépartition

$Y = A??????$

Leçon II.3 – Examen 2 2016 Q6

Considérons une séquence de caractères de la forme « $AB\underbrace{\cancel{******}}_{\leftarrow 6 \rightarrow}$ »,

où le symbole '$*$' peut être remplacé par n'importe quelle lettre de l'alphabet $\overbrace{\{A, B, C, D\}}^{n=4}$.

Quelle bornes (en bit) *les plus strictes* pouvez-vous donner pour l'entropie H d'une telle séquence ?

$$0 \leq H(X) \leq \underbrace{\log_2 n}_{2\text{ bit}} \quad n \to 4$$

D

ABCDABCD

ABAAAAAA

$$P_B = \frac{1}{8} \quad P_A = \frac{7}{8} \quad \to h\left(\frac{1}{8}\right)$$

Leçon II.3 – Examen 2 2016 Q4

$S_1 = AB \quad H(S_1) = 1$ bit

Soit H_1 l'entropie d'une séquence (non vide) de lettres S_1 et H_2 l'entropie d'une séquence de lettres S_2 contenant strictement S_1
(c.-à-d. que la séquence S_1 en tant que telle est une sous-séquence de S_2).

A] On a forcément $H_2 > H_1$.

B] On a forcément $H_2 < H_1$.

C] On a forcément $H_2 = H_1$.

D] Aucune des trois autres propositions.

Leçon II.4 (compression) – Points clés

- algorithme (de compression) de Shannon-Fano
- théorème de Shannon (de compression sans perte)
- lien avec la définition intuitive de l'entropie (nombre moyen de questions)
- algorithme (de compression) de Huffman
- optimalité (compression sans perte) des codes de Huffman

$= L(\text{Shannon-Fano})$

$H(X) \leq L < H(X) + 1$

$L - 1 < H(X) \leq L$

$\forall X \quad \forall C$ code de X

↳ sans perte sans prefixe

$$H(X) \leq L(C)$$

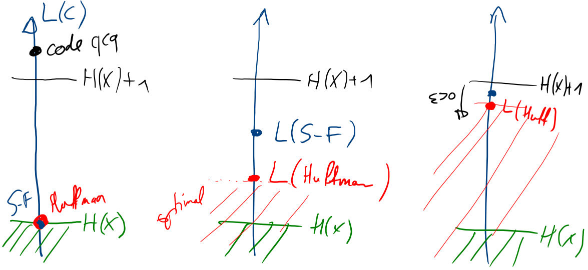
universel

$\forall X \quad \exists$ code ....

Shannon-Fano

Huffman.

$$L(C) < H(X) + 1$$


L(C)
code qcq
H(X)+1
S-F
Huffman
H(X)
H(X)+1
L(S-F)
L(Huffman)
optimal
H(x)
H(X)+1
ε>0
L(Huff)
H(x)

général

$X = \boxed{0\,1\,0\,1\,1\,1\,0\,1\,1}\ \boxed{0\,1\,1\,0\,1\,1\,0}$

Vous souhaitez envoyer un message, qui,

- ~~considéré bit à bit, a une entropie de 0.9 bit,~~
- et considéré octet par octet, une entropie de 5.51 bit.

$$\frac{5.51}{8}\cdot 10 \le T < \frac{6.51}{8}\cdot 10$$

Sachant que ce message (d'origine, complet) a une taille de 10 Ko, quelle taille pouvez-vous espérer après compression (sans autre information que celles fournies ici) ?

8

0	1	2	255

$\simeq 8$ Ko

$\simeq 7$ Ko

$5.51 \le L(C) < 6.51$

$H(X) \le L(C) < H(X)+1$

8 bits

Leçon II.4 (compression) – Étude de cas

Considérons la séquence $X =$« HUBERT QUEL HURLUBERLU » (sans les espaces).

1. entropie ?
2. code de Shannon-Fano ?
3. code de Hufmann ?
4. conclure

H	U	B	

Leçon II.4 (compression) – Étude de cas

Considérons la séquence $X =$« HUBERT QUEL HURLUBERLU » (sans les espaces).

1. entropie ?
2. code de Shannon-Fano ?
3. code de Hufmann ?
4. conclure $\rightarrow L(C)$

U	E	L	R	B	H	Q	T	
5	3	3	3	2	2	1	1	20

$\boxed{nb}$

U	E
$\frac{5}{20}$	$\frac{3}{20}$

$$\log 20 - \frac{1}{20}\left(9 \log 3 + 4 \log 2\right)$$

Leçon II.4 (compression) – Étude de cas

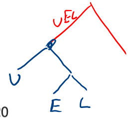
Considérons la séquence $X =$« HUBERT QUEL HURLUBERLU » (sans les espaces).

1. entropie ?
2. code de Shannon-Fano ?
3. code de Hufmann ?
4. conclure

11 | 9

U	E	L	R	B	H	Q	T
5	3	3	3	2	2	1	1

20

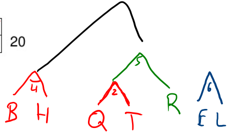


Leçon II.4 (compression) – Étude de cas

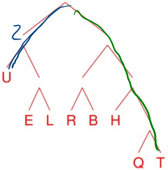
Considérons la séquence $X =$« HUBERT QUEL HURLUBERLU » (sans les espaces).

1. entropie ?
2. code de Shannon-Fano ?
3. code de Hufmann ?   moins probable en 1e
4. conclure

U	E	L	R	B	H	Q	T	
5	3	3	3	2	2	1	1	20



U	E	L	R	B	H	Q	T	
5	3	3	3	2	2	1	1	20



	U	E	L	R	B	H	Q	T
P	0.25	0.15	0.15	0.15	0.1	0.1	0.05	0.05
L	2	3	3	3	3	3	4	4
	0.5	0.45	0.45	0.45	0.3	0.3	0.2	0.2

Somme

$L(C_{\text{S-F}}) = 2.85$

$= \sum_i p_i l_i$