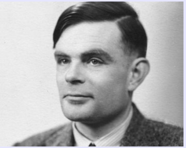
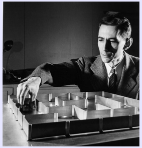
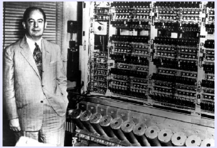




# Information, Calcul et Communication

## Compléments de cours

J.-C. Chappelier

EPFL

# Leçon II.3 – Examen final 2018 Q1.1

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 $a + b\log_2(3)$.

$$\frac{1}{2} + \frac{1}{3}\log_2(3) + \frac{1}{6}\log_2(12) = \mathbf{\frac{5}{6}} + \mathbf{\frac{1}{2}}\log_2(3)$$

# Leçon II.3 – Examen 2 2016 Q5

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]** $N-1$

# Leçon II.3 – Examen final 2018 Q3.1

Quelle est l'entropie (telle que définie en cours) d'un mot $X$ constitué de $n$ fois la lettre 'F' et $m$ fois la lettre 'G' : « $\underbrace{\texttt{F......F}}_{n}\underbrace{\texttt{G......G}}_{m}$ » ?

Soit $p = \dfrac{n}{n+m}$.

Utilisez $p$ pour exprimer cette entropie comme une fonction de $p$ uniquement :

$$H(X) = -\frac{n}{n+m}\,\log_2\!\left(\frac{n}{n+m}\right) - \frac{m}{n+m}\,\log_2\!\left(\frac{m}{n+m}\right) = p\,\log_2\!\left(\frac{1}{p}\right) + (1-p)\,\log_2\!\left(\frac{1}{1-p}\right)$$

# Leçon II.3 – Examen 2 2016 Q6

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

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

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

$$h(\frac{1}{8}) \leq H(X) \leq 2 \text{ bit}$$

avec :

$$\begin{aligned} h(\frac{1}{8}) &= \frac{1}{8}\log_2 8 + \frac{7}{8}\log_2\frac{8}{7} \\ &= \frac{3}{8} + \frac{21}{8} - \frac{7}{8}\log_2 7 \\ &\simeq 0.54 \text{ bit} \end{aligned}$$

# Leçon II.3 – Examen 2 2016 Q4

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

# Leçon II.4 (compression) – Étude de cas

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.

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) ?

entre $10 \times \frac{5.51}{8}$ et $10 \times \frac{6.51}{8}$

(soit entre environ 7 et 8 Ko)

# Leçon II.4 (compression) – Étude de cas

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

**A]** entropie ?

**B]** code de Shannon-Fano ?

**C]** code de Hufmann ?

**D]** conclure

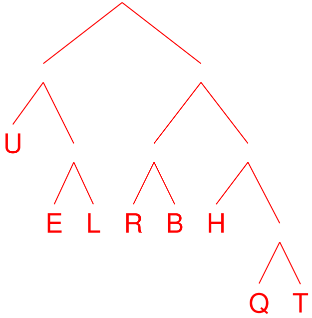
# Leçon II.4 (compression) – Étude de cas

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

$$\begin{aligned}
H(X) &= \frac{1}{4}\log 4 + 3\times\frac{3}{20}\log\frac{20}{3} + 2\times\frac{1}{10}\log 10 + 2\times\frac{1}{20}\log 20 \\
&= \frac{1}{2} + \frac{9}{20}(2+\log 5 - \log 3) + \frac{3}{10}(1+\log 5) + \frac{1}{10} \\
&= \frac{9}{5} + \frac{3}{4}\log 5 - \frac{9}{20}\log 3 \\
&\simeq 2.83 \text{ bit}
\end{aligned}$$

# Leçon II.4 (compression) – Étude de cas

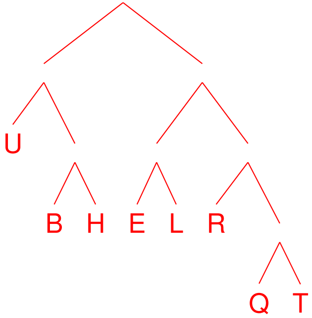
| U | E | L | R | B | H | Q | T | |
|---|---|---|---|---|---|---|---|---|
| 5 | 3 | 3 | 3 | 2 | 2 | 1 | 1 | 20 |




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

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

# Leçon II.4 (compression) – Étude de cas




| 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 |
|---|---|---|---|---|---|---|---|
| 0.25 | 0.15 | 0.15 | 0.15 | 0.1 | 0.1 | 0.05 | 0.05 |
| 2 | 3 | 3 | 3 | 3 | 3 | 4 | 4 |
| 0.5 | 0.45 | 0.45 | 0.45 | 0.3 | 0.3 | 0.2 | 0.2 |

$$L(C_{\mathrm{H}}) = 2.85$$

On a donc ici (mais ce n'est pas forcément toujours le cas) :

$$H(X) < L(C_{\mathrm{H}}) = L(C_{\text{S-F}})$$