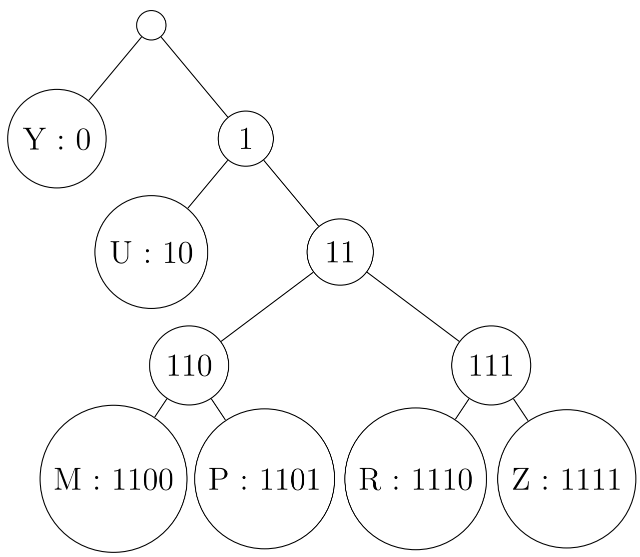
Feuille d'exercices n^{o}7
Corrigé

Exercice 1

1. $H(X) = -\left(\frac{1}{16}\log_2\left(\frac{1}{16}\right) + \frac{1}{16}\log_2\left(\frac{1}{16}\right) + \frac{1}{16}\log_2\left(\frac{1}{16}\right) + \frac{1}{4}\log_2\left(\frac{1}{4}\right) + \frac{1}{2}\log_2\left(\frac{1}{2}\right) + \frac{1}{16}\log_2\left(\frac{1}{16}\right)\right)$ donc $H(X) = 2$.

2. a)



b) Toutes les lettres de probabilité 1/16 sont codées par 4 bits, celle de probabilité 1/4 par 2 bits et celle de probabilité 1/2 par 1 bits. La longueur moyenne est donc :

$$\frac{1}{16}.4.4 + \frac{1}{4}.2 + \frac{1}{2}.1 = 2 = H(X)$$

3. On procède par récurrence sur n. Pour $n = 0$, c'est vrai. Supposons qu'on l'a montré jusqu'à n et montrons-le pour $n + 1$.

- Si $\epsilon_1 = 0$: par l'hypothèse de récurrence, il existe un mot m dont le code commence par $\epsilon_2 ... \epsilon_n$. Alors le code de Ym commence par $\epsilon_1 ... \epsilon_n$.
- Si $\epsilon_1 = 1, \epsilon_2 = 0$: soit m un mot dont le code commence par $\epsilon_3 ... \epsilon_n$. Alors le code de Um commence par $\epsilon_1 ... \epsilon_n$.
- Si $\epsilon_1 = \epsilon_2 = 1$: il existe une lettre l dont le code est $\epsilon_1\epsilon_2\epsilon_3\epsilon_4$. Soit m un mot dont le code commence par $\epsilon_5 ... \epsilon_n$. Alors le code de lm commence par $\epsilon_1 ... \epsilon_n$.

4. a) 0101100

b) Si on change le premier bit, on obtient 1101100. Ceci est le code du mot PUY.

Exercice 2

1. Soit G telle que G est affine sur $[y_{k-1}; y_k]$, pour tout k, et envoie y_{k-1} sur $\tilde{y}_{k-1}$ et y_k sur $\tilde{y}_k$. Sur $]-\infty; y_1]$, on choisit G affine de sorte que $G(y_1) = \tilde{y}_1$ et $G(a_1) = \tilde{a}_1$. De même sur $[y_K; +\infty[$.
Pour tout $k = 2, ..., K-1$, on a $G(a_k) = G\left(\frac{y_{k-1}+y_k}{2}\right) = \frac{G(y_{k-1})+G(y_k)}{2} = \frac{\tilde{y}_{k-1}+\tilde{y}_k}{2} = \tilde{a}_k$.
Pour tout $t \in \mathbb{R}$, si on note k l'entier tel que $t \in [y_{k-1}; y_k]$, on a $G^{-1} \circ \tilde{Q} \circ G(t) = G^{-1}(\tilde{a}_k) = a_k = Q(t)$.
Sur l'intervalle $[y_{k-1}; y_k]$, on a $G'(t) = \frac{\tilde{y}_k - \tilde{y}_{k-1}}{y_k - y_{k-1}}$. C'est en particulier vrai pour $t = a_k$.

2. L'hypothèse que Q et $\tilde{Q}$ sont des quantificateurs haute résolution fait qu'on peut supposer que p est nulle en-dehors de $[y_1; y_{K-1}]$ (et constante sur $[y_{k-1}; y_k]$ pour tout k).

$$\begin{aligned}
D &= E(|X - Q(X)|^2) \\
&= \sum_{k=2}^{K-1} \int_{y_{k-1}}^{y_k} (x - a_k)^2 p(a_k) dx \\
&= \sum_{k=2}^{K-1} \frac{(y_k - y_{k-1})^3}{12} p(a_k) \\
&= \frac{1}{12} \sum_{k=2}^{K-1} \frac{1}{G'(a_k)^3} (\tilde{y}_k - \tilde{y}_{k-1})^3 p(a_k) \\
&= \frac{(\tilde{y}_{K-1} - \tilde{y}_1)^3}{12(K-2)^3} \sum_{k=2}^{K-1} \frac{1}{G'(a_k)^3} p(a_k)
\end{aligned}$$

On a utilisé le fait que $\tilde{y}_k - \tilde{y}_{k-1} = \frac{\tilde{y}_{K-1} - \tilde{y}_1}{K-2}$ car la subdivision $(\tilde{y}_k)_k$ est uniforme.

3.

$$\begin{aligned}
\sum_{k=2}^{K-1} \frac{1}{G'(a_k)^3} p(a_k) &= \sum_{k=2}^{K-1} \frac{1}{\tilde{y}_k - \tilde{y}_{k-1}} \frac{1}{G'(a_k)^2} p(a_k)(y_k - y_{k-1}) \\
&\approx \frac{K-2}{\tilde{y}_{K-1} - \tilde{y}_1} \int_{y_1}^{y_{K-1}} \frac{p(x)}{G'(x)^2} dx
\end{aligned}$$

donc :

$$D \approx \frac{(\tilde{y}_{K-1} - \tilde{y}_1)^2}{12(K-2)^2} \int_{y_1}^{y_{K-1}} \frac{p(x)}{G'(x)^2} dx$$

4. a) Par Hölder, $\int_{y_1}^{y_{K-1}} p(x)^{1/3} dx = \int_{y_1}^{y_{K-1}} \left(\frac{p(x)}{G'(x)^2}\right)^{1/3} (G'(x))^{2/3} dx \leq \left(\int \frac{p(x)}{G'(x)^2} dx\right)^{1/3} \left(\int G'(x) dx\right)^{2/3}$.
Donc $\int_{y_1}^{y_{K-1}} \frac{p(x)}{G'(x)^2} dx \geq \left(\int_{y_1}^{y_{K-1}} p(x)^{1/3} dx\right)^3 \left(\int_{y_1}^{y_{K-1}} G'(x) dx\right)^{-2} = \frac{\left(\int_{y_1}^{y_{K-1}} p(x)^{1/3} dx\right)^3}{(\tilde{y}_{K-1} - \tilde{y}_1)^2}$.
Ainsi :

$$D \geq \frac{1}{12(K-2)^2} \left(\int_{y_1}^{y_{K-1}} p(x)^{1/3} dx\right)^3$$

b) Pour $\tilde{Q}$, on a $D \approx \frac{(\tilde{y}_{K-1} - \tilde{y}_1)^2}{12(K-2)^2}$, ce qui peut être nettement plus grand si p n'est pas à peu près constante sur $[\tilde{y}_1; \tilde{y}_{K-1}]$.

Exercice 3

1. Soit r le réel codant le mot de n lettres.

On décode par récurrence : si $n = 0$, on renvoie le mot vide. Sinon, la première lettre x est celle telle que $r \in [a_{x-1}; a_x[$. La suite du mot est le mot de $n-1$ lettres qui est codé par $\frac{r-a_{x-1}}{a_x-a_{x-1}}$.

2. Pour toute suite de lettres $x_1...x_n$, on note $[y_n^{\inf}(x_1...x_n), y_n^{\sup}(x_1...x_n)[$ l'intervalle correspondant.

On vérifie par récurrence que sa largeur est exactement $p_{x_1}...p_{x_n}$. Cet intervalle contient donc un réel de la forme $M2^{-m}$ où M est un entier et $m = \lceil -\log_2(p_{x_1}...p_{x_n}) \rceil$.

Puisqu'un réel compris entre 0 et 1 de la forme $M2^{-m}$ se code en m bits, la longueur du code de $x_1...x_n$ est au plus :

$$\lceil -\log_2(p_{x_1}...p_{x_n}) \rceil \leq 1 - \sum_{s \leq n} \log_2(p_{x_s})$$

On obtient :

$$\begin{aligned} l_n &\leq \sum_{x_1,\ldots,x_n} p(x_1)...p(x_n) \left(1 - \sum_{s \leq n} \log_2(p_{x_s}) \right) \\ &= 1 - \sum_{s \leq n} \sum_{x_s} p_{x_s} \log_2(p_{x_s}) \\ &= 1 + nH(X) \end{aligned}$$

donc $l_n/n \leq H(X) + 1/n$.

3. a) $H(X) = -(1-\epsilon)\log_2(1-\epsilon) - \epsilon \log_2 \epsilon$

b) 0 est codé par 0 et 1 est codé par 1.

c) Un mot de n lettres est toujours codé par un code de longueur n. On a donc $l_n/n = 1$. Pourtant, si ϵ est proche de 0 ou de 1, $H(X) \ll 1$ (car la fonction $x \to x\log_2(x)$ tend vers 0 en $x = 0$ et en $x = 1$).

Exercice 4

Soit $\eta_1, ..., \eta_n$ une suite de zéros et de uns. Notons k le nombre de uns qu'elle contient. On a :

$$P(X_1 = \eta_1, ..., X_n = \eta_n) = p^{n-k}(1-p)^k = p^n \left(\frac{1-p}{p} \right)^k$$

La suite $(\eta_1, ..., \eta_n)$ est ϵ-typique si :

$$H(X) - \epsilon \leq -\frac{\log_2(P(X_1 = \eta_1, ..., X_n = \eta_n))}{n} \leq H(X) + \epsilon$$

L'entropie de X vaut $H(X) = -p\log_2(p) - (1-p)\log_2(1-p)$. Puisque $\frac{\log_2(P(X_1=\eta_1,\ldots,X_n=\eta_n))}{n} = \log_2(p) + \frac{k}{n}\log_2\left(\frac{1-p}{p}\right)$, la suite $(\eta_1, ..., \eta_n)$ est ϵ-typique si et seulement si :

$$\begin{aligned} & (1-p)\log_2\left(\frac{1-p}{p}\right) - \epsilon \leq \frac{k}{n}\log_2\left(\frac{1-p}{p}\right) \leq (1-p)\log_2\left(\frac{1-p}{p}\right) + \epsilon \\ \Leftrightarrow \quad & n\left((1-p) - \epsilon\left|\log_2\left(\frac{1-p}{p}\right)\right|\right) \leq k \leq n\left((1-p) + \epsilon\left|\log_2\left(\frac{1-p}{p}\right)\right|\right) \end{aligned}$$

En effet, on a pu diviser par $\log_2\left(\frac{1-p}{p}\right)$ car, comme $p \neq 1/2$, ce nombre est non-nul.
Le nombre de uns dans une suite ϵ-typique est donc de l'ordre de $n(1-p)$.

Exercice 5

1.

$$\begin{aligned}
H(X) &= -\sum_{k\geq 1} 2^{-k}\log_2\left(2^{-k}\right) \\
&= \sum_{k\geq 1} k2^{-k} \\
&= \sum_{k\geq 1}\sum_{s\geq k} 2^{-s} \\
&= \sum_{k\geq 1} 2^{-k+1} \\
&= \sum_{k\geq 0} 2^{-k} = 2
\end{aligned}$$

2. On code le nombre k par le mot 1...10, qui contient $(k-1)$ zéros (c'est-à-dire que 1 est codé par 0, 2 par 10, 3 par 110 etc.).
Le nombre moyen de bits pour un symbole est (puisque le symbole k est codé par k bits) :

$$\sum_k P(X=k).k = \sum_k k2^{-k} = H(X)$$

Donc le code est optimal.

3. Dans ce cas, les symboles $1, ..., N$ voient leur code entier envoyé. En revanche, les codes des symboles $N+1, N+2, ...$ sont tronqués : pour un tel symbole, on n'envoie que $x = 1...1$ (N fois le bit 1). On suppose que lorsque le N-uplet reçu est x, on le décode par $N+1$.
Dans ce cas, le nombre de bits moyen est :

$$\begin{aligned}
\sum_{k\geq 1} P(X=k)\min(k,N) &= \sum_{k\geq 1} \min(k,N)2^{-k} \\
&= \sum_{k=1}^{N}\sum_{s\geq k} 2^{-s} \\
&= \sum_{k=1}^{N} 2^{-(k-1)} \\
&= \sum_{k=0}^{N-1} 2^{-k} \\
&= 2.(1-2^{-N})
\end{aligned}$$

La distance entre le symbole k à transmettre et le symbole décodé est 0 si $k \leq N+1$ et

$k-(N+1)$ si $k>N+1$. L'erreur quadratique moyenne est donc :

$$\begin{aligned}
\sum_{k\geq N+1} P(X=k)(k-(N+1))^2 &= \sum_{k\geq N+1} 2^{-k}(k-(N+1))^2 \\
&= 2^{-(N+1)}\sum_{k\geq 0} k^2 2^{-k} \\
&= 2^{-(N+1)}\sum_{k\geq 0}(1+3+5+...+(2k-1))2^{-k} \\
&= 2^{-(N+1)}\sum_{k\geq 1}(2k-1)\left(\sum_{s\geq k} 2^{-s}\right) \\
&= 2^{-(N+1)}\sum_{k\geq 1}(2k-1)2^{-(k-1)} \\
&= 2^{-(N+1)}\left(2.\sum_{k\geq 0} k2^{-k} + \sum_{k\geq 0} 2^{-k}\right) \\
&= 2^{-(N+1)}(2.2+2) = 3.2^{-N}
\end{aligned}$$