TD5: CDC

Exercice 2.1 : ***Codage de Huffman***

On considère la source d'information définie par l'alphabet $\mathcal{S} = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8\}$ avec la distribution de probabilités suivante f :

s	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
$p(s)$	0,4	0,18	0,1	0,1	0,07	0,06	0,05	0,04

Q 1.1 Calculez la quantité d'information contenue dans chacun des huit symboles, puis l'entropie de la source d'information $(\mathcal{S}, f)$.

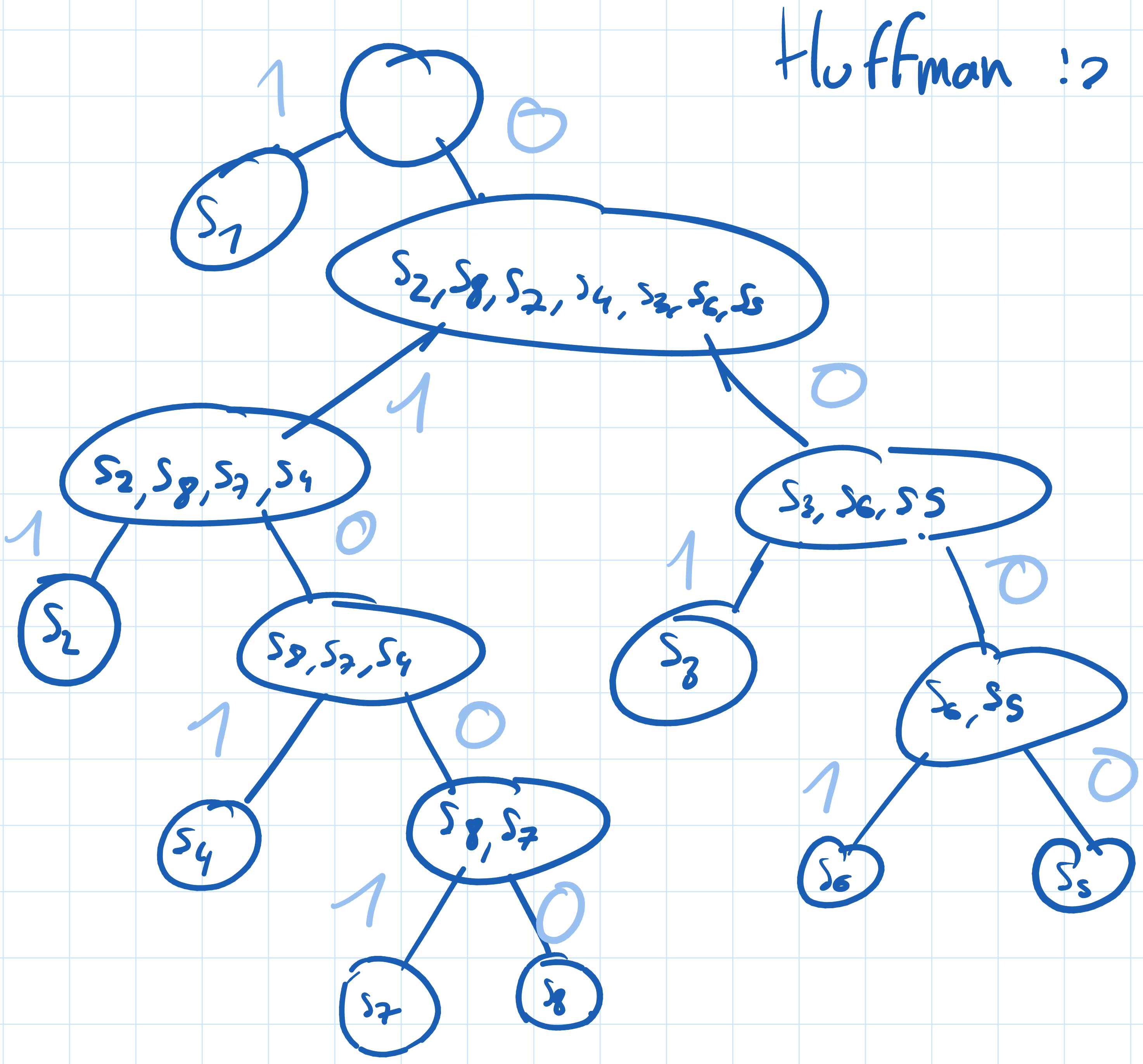
Q 1.2 Calculez un codage de Huffman de cette source, puis la longueur moyenne de ce codage. Comparez cette longueur moyenne à l'entropie.

$I(s) = -\log_2(P(s))$, donc ez.

Entropie de S: $H(S) = \sum_{i=1}^{8} P(s_i) * I(s_i) = 2,5524$ bit.

$\underbrace{,(s_2, s_8, s_7, s_4)}_{0,37}, \underbrace{(s_3\ s_6, s_5)}_{0,23}, s_1$

Huffman :



$\bar{n}_c = 1 \times 0,4 + 3 \times (0,1 + 0,18) + 4(0,1 + 0,07, 0,06) + 5 \times (0,05 + 0,04)$

$= 2,61$ bits/symbole.

qui vérifie le théorème de Shannon sans bruit.

$$\underset{2,53}{H(S)} \leq \underset{2,61}{\bar{n}_c} < \underset{3,53}{H(S)+1}$$

Exercice 2.3 :

Q 3.1 Pour coder sept symboles, existe-t-il un codage binaire optimal comprenant

1. trois mots de longueur 2, deux mots de longueur 3 et deux mots de longueur 4 ?
2. un mot de longueur 2, cinq mots de longueur 3 et un mot de longueur 4 ?

Codage optimal => Somme de Kraft = 1.

$K(L) \neq 1$ => codage non optimal ou pas un code.

« code » C1 ?

$$K(C_1) = \frac{3}{2^2} + \frac{2}{2^3} + \frac{2}{2^4} = \frac{9}{8} > 1$$ => pas un code :c

« code » C_2 ?

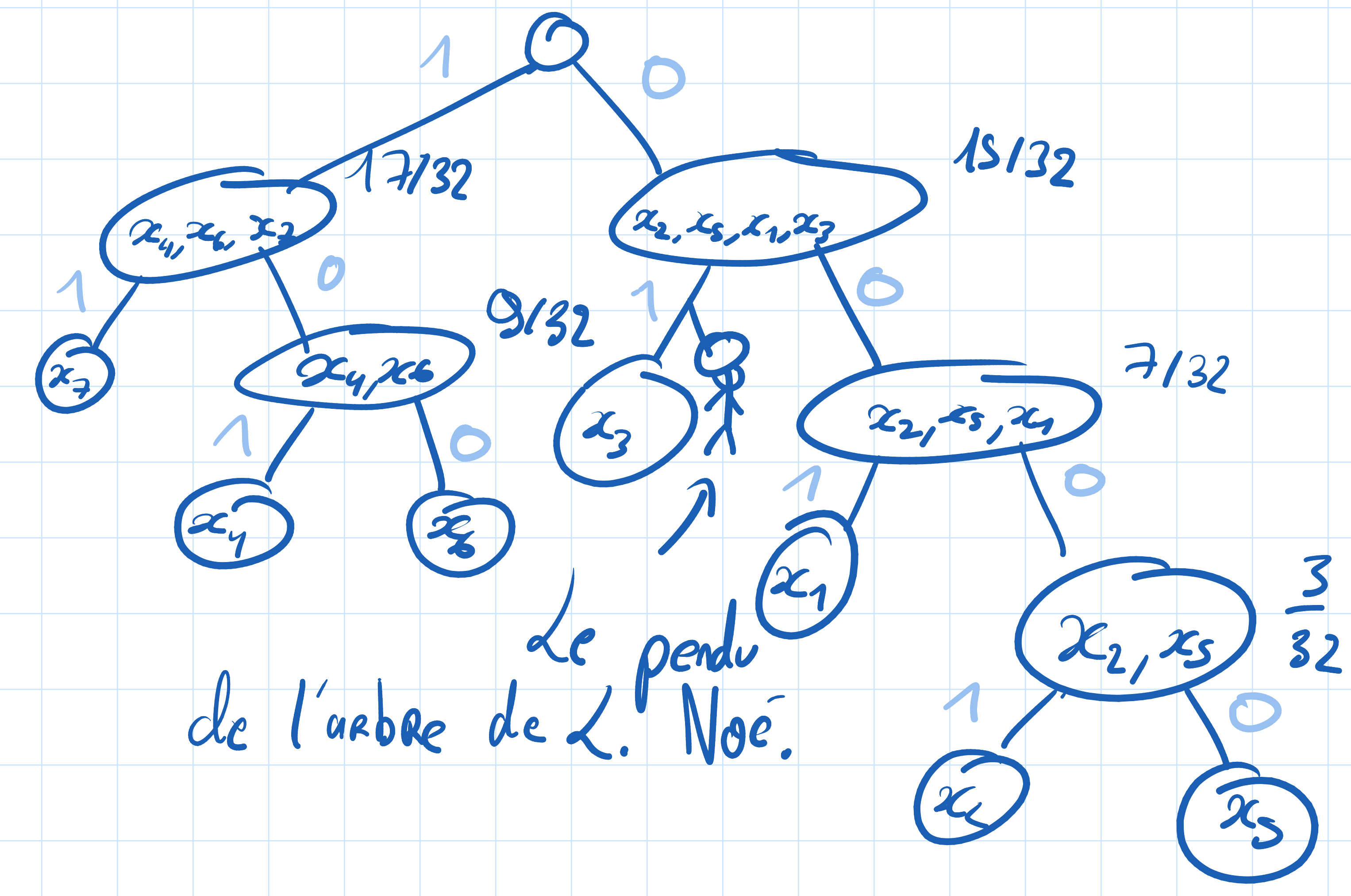
$$K(C_2) = \frac{1}{2^2} + \frac{5}{2^3} + \frac{1}{2^4} = \frac{4}{16} + \frac{10}{16} + \frac{1}{16} = \frac{15}{16}$$ -> pas égal à 1 :c

Soit $\mathcal{S}$ la source munie de la distribution de fréquences f suivante :

x_1	x_2	x_3	x_4	x_5	x_6	x_7
1/8	1/32	1/4	1/8	1/16	5/32	1/4

Q 3.2 Construisez un codage binaire optimal pour $\mathcal{S}$.

$\cancel{x_1}$	$\cancel{x_2}$	$\cancel{x_3}$	$\cancel{x_4}$	$\cancel{x_5}$	$\cancel{x_6}$	$\cancel{x_7}$
$\cancel{\frac{4}{32}}$	$\cancel{\frac{1}{32}}$	$\cancel{\frac{8}{32}}$	$\cancel{\frac{4}{32}}$	$\cancel{\frac{2}{32}}$	$\cancel{\frac{5}{32}}$	$\cancel{\frac{8}{32}}$



Exercice 2.5 : ***Fichier aléatoire***

Supposons qu'un fichier comportant n symboles au total, composé de 256 symboles différents, chacun codé sur 8 bits, ait été généré avec un générateur pseudo-aléatoire. Dans ce cas, la fréquence de chaque symbole est comparable. On pourra d'ailleurs considérer que la fréquence minimale d'un symbole dans ce fichier est supérieure à la moitié de la fréquence maximale dans ce même fichier.

Q 5.1 Quelle sera la taille du fichier après utilisation de l'algorithme de Huffman ?

Exercice 2.6 : ***Stratégie***

On a lancé deux fois une pièce de monnaie.

Q 6.1 Quelle est la quantité d'information contenue dans les affirmations suivantes :

1. PILE n'est pas apparu ?
2. PILE est apparu une seule fois ?
3. PILE est apparu deux fois ?

Q 6.2 Quelle est l'entropie de la source d'information indiquant le nombre de PILE ?

Q 6.3 Vous désirez connaître le nombre de PILE sortis lors de ces deux tirages. Pour cela vous pouvez poser autant de questions à réponses oui ou non que vous voulez. Quelle est la meilleure stratégie vous permettant de connaître ce nombre ? Par meilleure stratégie, on entend celle qui permet d'obtenir le nombre en un minimum de questions en moyenne.

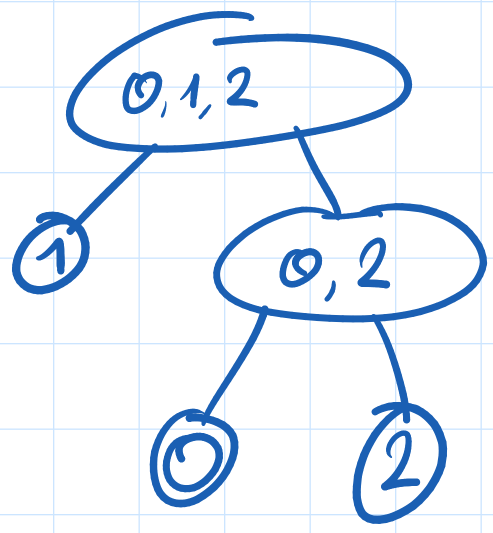
Q 6.4 Reprendre le même exercice avec 3 lancers.

$k \to$ nombre de pile

K	0	1	2
$P(K)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
$I(k)$ bit	2	1	2

$H(s) \longrightarrow = \frac{3}{2}$ bits/symbole.

(0,2) 1



$\frac{1}{8}$ $\frac{1}{8}$ $\frac{3}{8}$ $\frac{3}{8}$

0, 3, 1, 2