11/03/21

1 Exercises

Exercise 1. Consider a source $\mathcal{S}$ that can emit five symbols s_i, $i \in \{1, \ldots, 5\}$ and two possible encodings (*bin1* and *oct1*) for these symbols (described below). $\mathcal{S}$ emits a message of two symbols $s_1 x$ (x is unknown) with a compression rate (from *oct1* to *bin1*) equals to 1. What is the symbol x?

Name	code words associated to					alphabet
	s_1	s_2	s_3	s_4	s_5	
bin1	0	01	001	0001	00001	$\{0,1\}$
oct1	0	1	5	3	7	$\{0,1,\ldots,7\}$

compression rate $= \dfrac{\bar{n}_1}{\bar{n}_2} \cdot \dfrac{\log q_1}{\log q_2}$ (from: $\bar{n}_1$; To: $\bar{n}_2$)

$q_{oct1} = 8$ $\quad S_1 x$

$q_{bin1} = 2$

$n_{oct1} = 1 + 1$

$n_{bin1} = 1 + n^x_{bin1}$

Thus $1 = \dfrac{2}{1 + n^x_{bin1}} \cdot \dfrac{\log_2 8}{\log_2 2}$

$= \dfrac{2}{1 + n^x_{bin1}} \cdot \dfrac{3}{1}$

$\Leftrightarrow n^x_{bin1} = 5 \Rightarrow \boxed{x = s_5}$

Exercise 2. Are the following source stationary and/or memoryless? Justify.

(a) The source $\mathcal{S}_1$ can emit $Q = 26$ different symbols $s_k (k \in \{1, \dots, 26\})$. We denote an emitted message by the source by $s^1 s^2 s^3 \dots$. Each new symbol s^i is sampled in the following way : the probability that the symbol s_k is emitted corresponds to the occurence frequency of the k^{th} letter of the alphabet on the page i of a given book.

(b) Each symbol emitted by the source $\mathcal{S}_2$ corresponds to the sum of results of three throws of a dice (perfectly balanced).

(c) A binary source $\mathcal{S}_3$ which emits, each minute, a "1" if a car has entered the *tunnel de Cointe* this minute, and emits a "0" otherwise.

(d) A source $\mathcal{S}_4$ emits each day the difference between the temperature of the day and the seasonal average.

(a) not stationary : letter frequencies vary at each page
memoryless : next emitted symbol prob. does not depend on previously emitted symbols.

(b) stationary :
memoryless :

(d) not stationary : the seasonal average changes over time
not memoryless : depends on your justificat°.

Exercise 3. [5.8] Is the code $\{1, 101\}$ uniquely decodable?

not prefix free

Any received sequence can be decoded without ambiguity

1 101 → m_1 m_2

1 101 1 → m_1 m_2 m_1

Yes

Exercise 4. [5.19] Is the code $\{00, 11, 0101, 111, 1010, 100100, 0110\}$ uniquely decodable?

No

11111

m_2 m_4

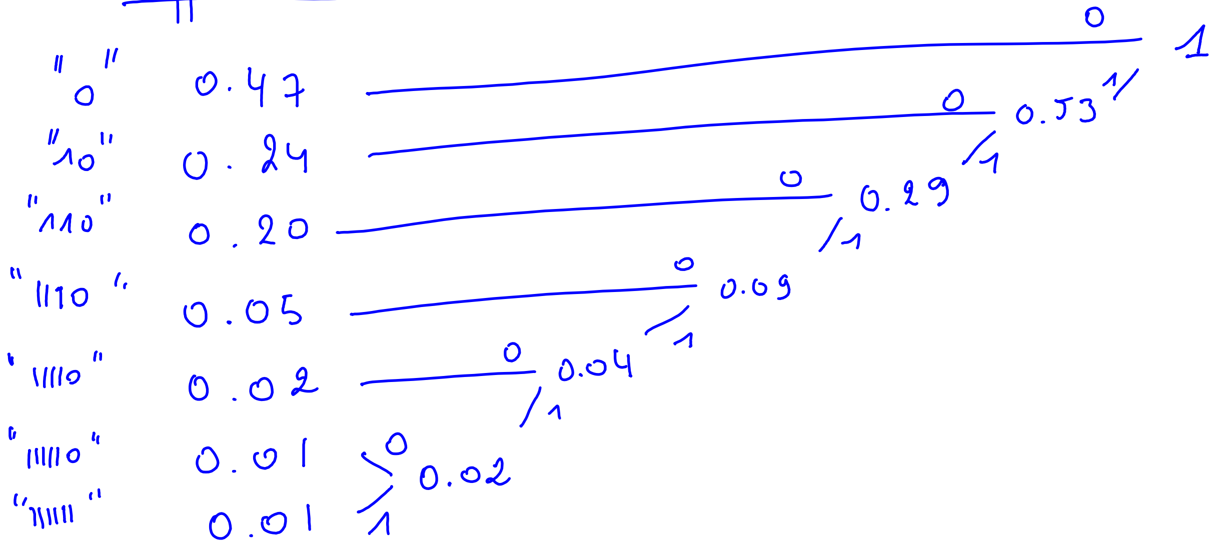
m_4 m_2

Exercise 5. [5.20] Is the code $\{00, 012, 0110, 0112, 100, 201, 212, 22\}$ uniquely decodable?

prefix - free => uniquely decodable

Exercise 6. [5.18] Find the optimal binary symbol code for a source $\mathcal{S} = \{a, b, c, d, e, f, g\}$ and $P(\mathcal{S}) = [0.01, 0.24, 0.05, 0.20, 0.47, 0.01, 0.02]$.

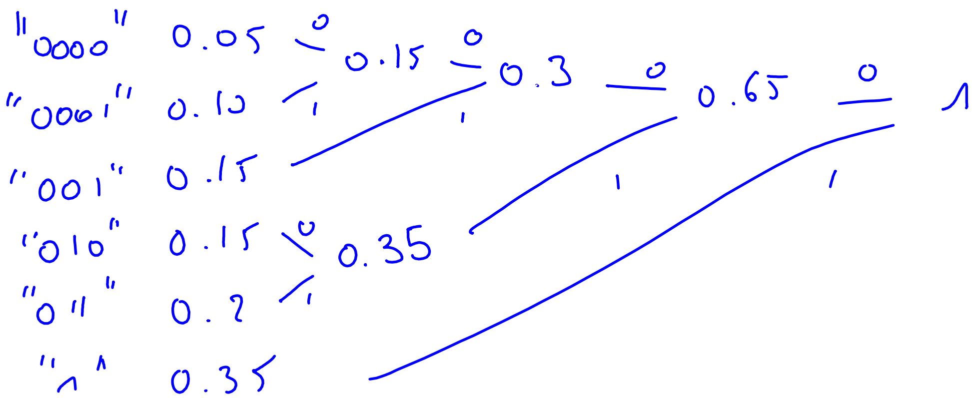
Huffman code



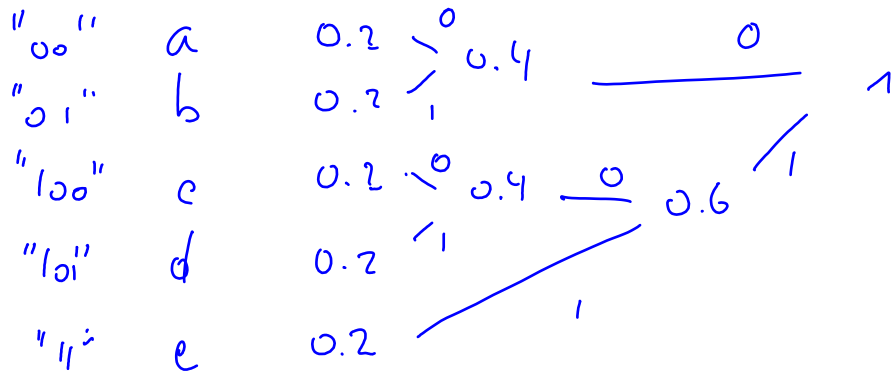
Exercise 7. Let a source be stationary and memoryless emitting six different symbols. Their probabilities are given by the following vector :

$$P(\mathcal{S}) = [0.05\ 0.10\ 0.15\ 0.15\ 0.2\ 0.35]$$

Use the Huffman algorithm to find a binary code for that source.



Exercise 8. Using the Huffman algorithm, find a symbol code for a source $\mathcal{S} = \{a, b, c, d, e\}$ and $P(\mathcal{S}) = [0.2, 0.2, 0.2, 0.2, 0.2]$. Is the Kraft inequality verified by this symbol code? What are the properties of this code?



Kraft $\sum_{i=1}^{Q} q^{-n_i} = 2^{-2} \times 3 + 2^{-3} \times 2 = 1$

- complete code
- prefix-free
- uniquely decodable
- instantaneous
- regular.

$\bar{n} = 0.2 \times 2 \times 3 + 0.2 \times 3 \times 2 = 2,4$

$H(S) = -5 \, \frac{1}{5} \log_2 (1/5) = 2,32$

$H(S) \leq \bar{n} < H(S) + 1$

- optimal
- not absolutely optimal $H(S) \neq \bar{n}$

Exercise 9. [5.21] Make Huffman codes for X^2, X^3 and X^4 where the alphabet is $\{0, 1\}$ and $P(X) = [0.9, 0.1]$. Compute their expected lengths and compare them with entropies $H(X^2)$, $H(X^3)$ and $H(X^4)$.
Repeat this exercise for X^2 and X^4 where $P(X) = [0.6, 0.4]$.

Exercise 10. [5.22] Find a probability distribution $\{p_1, p_2, p_3, p_4\}$ such that there are *two* optimal codes that assign different lengths $\{l_i\}$ to the four symbols.

Exercise 11. Let a random variable $\mathcal{X}$ with the following values and probability distribution:

x_1	x_2	x_3	x_4	x_5	x_6	x_7
0.49	0.26	0.12	0.04	0.04	0.03	0.02

(a) Find a binary Huffman code and the corresponding average length.

(b) Find a ternary Huffman code and the corresponding average length.

(c) Compare both average lengths with the theoretical bounds predicted by the first Shannon theorem.

Exercise 12. Among the following codes, which ones can not be binary Huffman codes? Justify.

(a) $\{0, 10, 11\}$

(b) $\{00, 01, 10, 110\}$

(c) $\{01, 10\}$

Exercise 13. Let a Markov source (with three states) be characterized by the following transition matrix:

$$\begin{bmatrix} S_{n-1} \backslash S_n & | & s_1 & s_2 & s_3 \\ \hline s_1 & | & 1/2 & 1/4 & 1/4 \\ s_2 & | & 1/4 & 1/2 & 1/4 \\ s_3 & | & 0 & 1/2 & 1/2 \end{bmatrix}$$

(Thus the probability that s_1 follows s_3 is zero)

(a) Build a code made of three instantaneous binary codes C_1,C_2,C_3 to optimally encode the Markov chain?

(b) What is the average length of the following symbol, conditionally to the previous state?

(c) What is the unconditional average length ?

(d) Find the link with the entropy per symbol of the Markov chain.

Exercise 14. One is given six bottles of wine. It is known that precisely one bottle has gone bad (tastes terrible). From inspection of the bottles it is determined that the probability p_i that the i^{th} bottle is bad is given by $(p_1, p_2, \ldots, p_6) = (\frac{8}{23}, \frac{6}{23}, \frac{4}{23}, \frac{2}{23}, \frac{2}{23}, \frac{1}{23})$. Tasting will determine the bad wine. Suppose that you taste the wines one at a time. Choose the order of tasting to minimize the expected number of tastings required to determine the bad bottle. Remember, if the first five wines pass the test, you don't have to taste the last.

(a) What is the expected number of tastings required?

(b) Which bottle should be tasted first?

Now you get smart. For the first sample, you mix some of the wines in a fresh glass and sample the mixture. You proceed, mixing and tasting, stopping when the bad bottle has been determined.

(a) What is the minimum expected number of tastings required to determine the bad wine?

(b) What mixture should be tasted first?

Exercise 15. How many different Huffman codes can you build to code the messages emitted by a stationary and memoryless source (with a source alphabet size of 4) with the following probability distribution

s_i	10	01	23	03	30
p_i	0.4	0.2	0.2	0.19	0.01

Exercise 16. If $X_1 \to X_2 \to X_3 \to \cdots \to X_n$ forms a Markov chain, what is $I(X_1; X_2, \ldots, X_n)$? Simplify as much as possible.

Exercise 17. State and prove the chain rule for entropy (without using the definition of mutual information).

Exercise 18. Prove the following inequality and find conditions for equality:

$$I(\mathcal{X};\mathcal{Z}|\mathcal{Y}) \geq I(\mathcal{Z};\mathcal{Y}|\mathcal{X}) - I(\mathcal{Z};\mathcal{Y}) + I(\mathcal{X};\mathcal{Z})$$