Information Theory
Exercises

1. Let us have a source with alphabet $\mathcal{X} = \{a, b, c\}$. Encode the source sequence

$$cabcbcbcb$$

with the Lempel-Ziv-Welch algorithm. (The dictionary originally contains the codewords (i.e., the indices) for the one character sequences: 1 for $a$, 2 for $b$, and 3 for $c$.)

2. Let us have a source with alphabet $\mathcal{X} = \{a, b\}$. Encode the source sequence

$$abbbbabbbba$$

with the Lempel-Ziv-Welch algorithm. (The dictionary originally contains the codewords (i.e., the indices) for the one character sequences: 1 for $a$, 2 for $b$.)

3. Let us have a source with alphabet $\mathcal{X} = \{a, b, c\}$ (the dictionary originally contains the codewords (i.e., the indices) for the one character sequences: 1 for $a$, 2 for $b$, and 3 for $c$). Decode the sequence

$$3, 4, 5, 6, 7, 1.$$

4. Let $\mathbf{X} = (X_1, X_2, \ldots)$ be a stationary source with entropy $H(\mathbf{X})$. Decide whether the entropy of the following sources exists and determine it if it does.

a) $\mathbf{X}_a = (X_1, X_1, X_2, X_2, X_3, X_3, \ldots)$ (all random variables are repeated once)

b) $\mathbf{X}_b = (X_1, X_1, X_2, X_3, X_3, X_4, X_5, X_5, X_6, \ldots)$ (only the odd indexed random variables are repeated)

c) $\mathbf{X}_c = (X_1, X_2, X_2, X_3, X_3, X_3, X_4, X_4, X_4, X_4, \ldots)$ (the random variable with index $i$ is repeated $i$ times)

5. A homogenous Markov chain has three states: $A$, $B$, and $C$. From state $A$ it goes to state $B$ with probability 1. From state $B$ it goes to state $C$ with probability $1/3$ and stays in state $B$ with probability $2/3$. From state $C$ it goes to state $A$ with probability $2/3$ and stays in state $C$ with probability $1/3$. Determine the entropy of the source formed by this Markov chain. (That is, the source emits a symbol after each state transition of the given Markov chain and the output is simply the new state.)

6. Let $\mathbf{Y} = Y_1, Y_2, \ldots$ be a source, where the $Y_i$'s are independent, identically distributed random variables, each taking the value 0 or 1 with probability $1/2 - 1/2$. (So $Prob(Y_i = 0) = Prob(Y_i = 1) = 1/2$ for all $i$.)

a) The source $\mathbf{X_a}$ is defined as follows. $X_1 = Y_1, X_2 = Y_2$ and for $i \geq 3$ we have $X_i = X_{i-1} + Y_i \pmod 2$. (That is, $X_i = 0$ if the sum $(X_{i-1} + Y_i)$ is even and $X_i = 1$ if $(X_{i-1} + Y_i)$ is odd.)

b) The source $\mathbf{X_b}$ is defined as follows. $X_1 = Y_1, X_2 = Y_2$ and for $i \geq 3$ we have $X_i = X_{i-1}$ if $Y_i = 0$ and $X_i = X_{i-2}$ if $Y_i = 1$.

Give the entropy of the sources $X_a$ and $X_b$ (in case they exist).

7. A source $\mathbf{X} = X_1, X_2, \ldots$ works as follows. Each $X_i$ is equal to either 0 or 1, $Prob(X_1 = 0) = Prob(X_1 = 1) = 1/2$ and similarly $Prob(X_2 = 0) = Prob(X_2 = 1) = 1/2$. For $i \geq 3$ the rule is the following. If $X_{i-1} = X_{i-2}$ then $X_i = 1 - X_{i-1}$ for sure (that is, with probability 1). If $X_{i-1} \neq X_{i-2}$, then $X_i = 0$ and $X_i = 1$ has equal probability, that is $Prob(X_i = 0|X_{i-1} \neq X_{i-2}) = Prob(X_i = 1|X_{i-1} \neq X_{i-2}) = \frac{1}{2}$. Give the entropy of the source $\mathbf{X}$ if it exists.

8. The result of an experiment is a random variable taking one of 7 possible values according to the probability distribution

$$(1/3, 1/3, 1/9, 1/9, 1/27, 1/27, 1/27)$$

. We want to transfer the result via phone and can choose one of two services the telephone company offers for this. In the first type of service information is transferred in binary form and the transfer of each bit costs 200 forints. In the second type of service the information is transferred in ternary form and the transfer of each ternary character costs 325 forints. (The corresponding binary or ternary code can be chosen by us.) Which service will be cheaper for us if the experiment is repeated very many times independently but we have to encode the outcomes one-by-one? What is the answer if we are allowed to encode the outcomes together? ($\log_2 3 \approx 1.585$)