Information Theory
Exercises

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

$$abcabcbabcabc$$

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 for c.)

Sketch of Solution:

Using the algorithm we learnt one obtains the code 1,2,3,4,3,2,7,7. The dictionary becomes: 1: a; 2: b; 3: c; 4: ab; 5: bc; 6: ca; 7: abc; 8: cb; 9: ba; 10: abca.

2. Let X be a source whose outcome is given by the state of a binary stationary Markov chain. The transition probabilities of the Markov chain are given by

$$P(X_2 = 0|X_1 = 0) = p,\ P(X_2 = 1|X_1 = 0) = 1 - p,$$

$$P(X_2 = 0|X_1 = 1) = \frac{1-p}{2},\ P(X_2 = 1|X_1 = 1) = \frac{1+p}{2}.$$

What is $Prob(X_i = 0)$? Give the entropy of this source.

(Note that since the Markov chain is stationary, it is also homogenous and we have that $Prob(X_i = 0)$ does not depend on i.)

Sketch of Solution:

Let $Prob(X_i = 0) = q, Prob(X_i = 1) = 1 - q$. Then we have

$$q = qp + (1 - q)\frac{1-p}{2}.$$

From here we get

$$(1 - p)q = \frac{1}{2}(1 - p)(1 - q).$$

Here $(1 - p)$ cancels out and we obtain $q = \frac{1}{3}$. Thus

$$Prob(X_i = 0) = \frac{1}{3}.$$

Now we can express the source entropy:

$$H(X) = H(X_{i+1}|X_i) = \frac{1}{3}h(p) + \frac{2}{3}h\left(\frac{1-p}{2}\right).$$

3. 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.

Sketch of Solution:

We know by a theorem that

$$H(X) = \lim_{n\to\infty} H(X_n|X_1, X_2, \ldots, X_{n-1}).$$

In this case

$$\lim_{n\to\infty} H(X_n|X_1, X_2, \ldots, X_{n-1}) = \lim_{n\to\infty} H(X_n|X_{n-1}, X_{n-2}),$$

since the equality or non-equality of X_{n-1} and X_{n-2} determines the distribution for X_n. We can think about the system as a Markov chain with two states A and B, where A means that the last two outputs were equal, B means that they were not. Then the transition probabilities are: from A the system goes to B with probability 1, from B it goes to both A and B with probability 1/2. Calculating the stationary distribution the usual way we get that $Prob(A) = 1/3, Prob(B) = 2/3$. Thus

$$\lim_{n\to\infty} H(X_n|X_{n-1}, X_{n-2}) =$$

$$\lim_{n\to\infty} (Prob(X_{n-1} = X_{n-2})H(X_n|X_{n-1} = X_{n-2})+$$
$$+Prob(X_{n-1} \neq X_{n-2})H(X_n|X_{n-1} \neq X_{n-2})) =$$
$$1/3 \cdot 0 + 2/3 \cdot h(1/2) = 2/3.$$

4. Find the capacity of the discrete memoryless channel with input alphabet $\{0, 1\}$, output alphabet $\{a, b, c\}$ and the following conditional probabilities characterizing the channel: $W(a|0) = W(a|1) = 1/3, W(b|0) = W(c|1) = 1/2, W(b|1) = W(c|0) = 1/6$.

Sketch of Solution:

We need to find $C = \max I(X,Y) = \max\{H(Y) - H(Y|X)\}$. Observe that $H(Y|X = 0) = H(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}) = H(Y|X = 1)$ and thus $H(Y|X)$ is also equal to this common value $H(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}) = \frac{1}{2}\log 2 + \frac{1}{3}\log 3 + \frac{1}{6}\log 6 = \frac{1}{2} + \frac{1}{6} + \log 3(\frac{1}{3} + \frac{1}{6}) = \frac{2}{3} + \frac{1}{2}\log 3$.

Thus this last term is a constant, so the maximum of $I(X,Y) = H(Y) - H(Y|X)$ is attained when $H(Y)$ is maximal. Clearly, $H(Y) \leq \log 3$ (as Y has 3 possible values), and this can be attained if and only if Y can have a uniform distribution. Now observe, that if we set $Prob(X = 0) = Prob(X = 1) = \frac{1}{2}$, then all the three outputs have probability $\frac{1}{3}$, so the above maximum is indeed attained. Thus

$$C = \log 3 - \left(\frac{2}{3} + \frac{1}{2}\log 3\right) = \frac{1}{2}\log 3 - \frac{2}{3}.$$