Information Theory – SS 2024

Exercise sheet 08 – 06.06.2024

Exercise 33
Let $(X_n)_{n\geq 0}$ be a time-homogeneous Markov chain with state space $\mathcal{X} = \{0, 1\}$ and transition probabilities: $p(0|0) = 0.3$, $p(1|0) = 0.7$, $p(0|1) = 0.2$, $p(1|1) = 0.8$ and suppose that $X_1 = 0$.

(a) Find an integer n and a word $w \in \mathcal{X}^n$ such that w lies in the typical set $A^{(n)}_{0.005}$.

(b) Given $\epsilon > 0$, provide an idea on how to construct a word w lying in the typical set $A^{(n)}_{\epsilon}$ for some integer n.

(Recall that the stationary distribution $\nu = \left(\frac{2}{9}, \frac{7}{9}\right)$ and the entropy rate h is $\approx 0.757\ldots$)

Exercise 34
Let $(X_n)_{n\geq 1}$ be a stochastic process which satisfies the AEP with entropy rate h and suppose $r > h$. Show that there is a sequence of block codes $C^{(n)}: \mathcal{X}^n \to \{0, 1\}^m$ with rate $r = \frac{m}{n}$ such that $p^{(n)}_{\text{err}} \to 0$.
(Hint: Define an injective encoding of the elements of the typical set $A^{(n)}_{\epsilon}$ as in the proof of Theorem 4.2)

Exercise 35
Let $(X_n)_{n\geq 1}$ be a stochastic process which satisfies the AEP with entropy rate h and suppose $r < h$. Show that there is some $\epsilon > 0$ which depends only on r and h such that for sufficiently large n and any block code $C^{(n)}: \mathcal{X}^n \to \{0, 1\}^m$ with rate $r = \frac{m}{n}$, $p^{(n)}_{\text{err}} > \epsilon$.
(Hint: Consider the Markovian triple $(X_1, \ldots, X_n) \to C^{(n)}(X_1, \ldots, X_n) \to (\hat{X}_1, \ldots, \hat{X}_n)$ given by the source, code and estimate and apply Fano's inequality)

Exercise 36
Let $\mathcal{X} = \{a, b, c, d\}$ and

$$C(a) = 01, \qquad C(b) = 11, \qquad C(c) = 00, \qquad C(d) = 001.$$

(a) Show that this code is not prefix free but uniquely decodable.

(b) Show that there are arbitrarily long words $x_1 \cdots x_n$ such that $C(x_1 \cdots x_n)$ can only be decoded at the very end, that is, that the initial element x_1 cannot be recovered from any proper prefix of $C(x_1 \ldots x_n)$.

Exercise 37
Let $X : \Omega \to \mathcal{X} = \{a, b, c, d\}$ be a random variable with

$$\mathbb{P}[X = a] = \frac{5}{8}, \qquad \mathbb{P}[X = b] = \frac{1}{8}, \qquad \mathbb{P}[X = c] = \frac{1}{8}, \qquad \mathbb{P}[X = d] = \frac{1}{8}.$$

The elements of $\mathcal{X}$ are encoded as follows:

$$C(a) = 00, \qquad C(b) = 10, \qquad C(c) = 11, \qquad C(d) = 110.$$

(a) Is the code C (i) non-singular, (ii) prefix-free, (iii) uniquely decodable?

(b) Calculate the entropy $H(X)$ and the expected code length $\mathbb{E}(\ell(C))$.

(c) Give a better code for this random variable (prefix-free, shorter expected length).

Exercise 38
Let $C : \mathcal{X} \to \Sigma = \{0, 1, \ldots, D - 1\}$ be a D-ary prefix code on a set $\mathcal{X}$ of cardinality m and suppose that the code word lengths $l_1, \ldots, l_m$ satisfy the strict inequality

$$\sum_{i=1}^{m} D^{-l_i} < 1.$$

Show that there are arbitrary long strings $w \in \Sigma^*$ which are not code words, that is, w is not the code word of any string in $\mathcal{X}^*$.