Information Theory SS 2025

Institut für Diskrete Mathematik (5050), TU Graz

Exercise sheet 6 (05.06.2025)

Exercise 6.1. Consider again the Markov Chain $(X_n)_{n\geq 1}$ from Exercise 5.6 (last session).

(a) Calculate the entropy rate h.

(b) Given that the Markov chain starts at $X_1 = 0$, find whether the sequences $w_1, w_2 \in \mathcal{X}^{20}$ are in the typical set $A_\varepsilon^{(20)}$, for $\varepsilon = 0.05$:

$$w_1 = 01001010101101100100,$$
$$w_2 = 01111001110111110111.$$

Exercise 6.2. Let $X : \Omega \to \mathcal{X} = \{a, b, c, d\}$ be a random variable with

$$\mathbb{P}[X = a] = \frac{3}{8}, \qquad \mathbb{P}[X = b] = \frac{2}{8}, \qquad \mathbb{P}[X = c] = \frac{2}{8}, \qquad \mathbb{P}[X = d] = \frac{1}{8}.$$

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

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

(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 6.3. Let $C : \mathcal{X} \to \Sigma = \{0, 1, \dots, D-1\}$ be a D-ary prefix code on a set $\mathcal{X}$ of cardinality m and suppose that the code word lengths $l_1, \dots, l_m$ 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 $\mathcal{X}^*$.

Exercise 6.4. Let $\mathcal{X} = \{a, b, c, d, e, f, g, h\}$.

(a) Give an example of a prefix code $C : \mathcal{X} \to \{0, 1\}^*$ with code word lengths

$$\ell(C(a)) = \ell(C(b)) = \ell(C(c)) = 2, \ \ell(C(d)) = 3,$$
$$\ell(C(e)) = \ell(C(f)) = \ell(C(g)) = 5, \ \ell(C(h)) = 6.$$

(b) Show that the Kraft inequality is a strict inequality for this code, that is,

$$\sum_{x \in \mathcal{X}} 2^{-\ell(C(x))} < 1.$$

(c) Find a binary sequence of length 6 which cannot be decoded.

(d) Find a prefix code $C : \mathcal{X} \to \{0,1\}^*$ such that we have equality in the Kraft inequality.

(e) Find a probability distribution for a random variable $X : \Omega \to \mathcal{X}$ such that your code is optimal and compare the expected code length with $H(X)$.

Exercise 6.5. Consider the random variable $X : \Omega \to \{x_1, \ldots, x_{10}\}$ with probability distribution

$$p = (\, 0.25,\ 0.12,\ 0.13,\ 0.15,\ 0.03,\ 0.05,\ 0.04,\ 0.13,\ 0.08,\ 0.02\,).$$

In a D-ary Huffman code at each step we choose D nodes with the lowest probability and construct their parent with probability the sum of the probabilities of the children. Only at the first step we may merge less than D children: given that $|\mathcal{X}| = N$, at the first step we build the parent of $k = 2 + (N-2) \bmod (D-1)$ children. For example, for $N = 19$ and $D = 5$ we get $k = 2 + 17 \bmod 4 = 3$.

(a) Build a binary Huffman code.

(b) Compute the expected code length of your code and compare it to $H(X)$.

(c) Build a ternary Huffman code over the alphabet $\Sigma = \{0, 1, 2\}$.

(d) Compute the expected code length of your code and compare it to

$$H_3(X) = -\sum_{x \in \mathcal{X}} p(x) \log_3(p(x)).$$

Exercise 6.6. Let X, Y be random variables with values in $\mathcal{X} = \mathcal{Y} = \{0, 1\}$ and let $p_X(0) = q$, $p_X(1) = 1 - q$. Let $\mathcal{C} = (\mathcal{X}, P, \mathcal{Y})$ be the channel with input X, output Y and transition matrix

$$P = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}.$$

(a) Compute $I(X, Y)$ in terms of p and q.

(b) For which values of p is $I(X; Y)$ maximal?

(c) For which values of p is $I(X; Y)$ minimal?