Information Theory I

Fall Semester 2017

Prof. Dr. A. Lapidoth

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

Exercise 4 of October 11, 2017

http://www.isi.ee.ethz.ch/teaching/courses/it1.html

Problem 1 — *Slackness in Kraft's Inequality*

Let $l_1, l_2, \ldots, l_m$ denote the codeword lengths of a prefix-free code over the D-ary alphabet $\mathcal{D} = \{0, 1, \ldots, \mathsf{D}-1\}$.

a) For $\mathsf{D} = 2$, i.e. binary codes, show that if

$$\sum_{i=1}^{m} 2^{-l_i} < 1,$$

then there exists another prefix-free binary code that is deterministically better than the original code. Deterministically better means that for each source symbol, the corresponding codeword is no longer than the codeword from the original code, and that for at least one source symbol, the corresponding codeword is shorter.

b) For $\mathsf{D} > 2$, (e.g. ternary codes, quaternary codes, ...) with

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

does the same conclusion hold?

Problem 2 — *Shannon Code*

Consider the following method for generating a code for a random variable X which takes on m values $\{1, 2, \ldots, m\}$ with positive probabilities $p_1, p_2, \ldots, p_m$. Assume that the probabilities are ordered so that $p_1 \geq p_2 \geq \cdots \geq p_m$. Define

$$\mathsf{F}_i = \sum_{k=1}^{i-1} p_k,$$

the sum of the probabilities of all symbols less than i. Then the codeword for i is the binary representation[1] of the number $\mathsf{F}_i \in [0, 1)$ truncated to l_i bits, where $l_i = \left\lceil \log \frac{1}{p_i} \right\rceil$.

a) Construct the code for the probability distribution $(0.5, 0.25, 0.125, 0.125)$.

b) Show that the code constructed by this process is prefix-free and the average length satisfies

$$H(X) \leq \mathsf{L} < H(X) + 1.$$

[1] In cases where the binary representation is not unique we always take the terminating expansion. For example, we represent $1/2$ by 0.1 instead of $0.0111\ldots$

Problem 3 ***Optimal Code Lengths that Require One Bit above Entropy***

The source coding theorem asserts that an optimal code for a chance variable X has expected length less than $H(X) + 1$. Show that this upper bound cannot be improved by constructing, for every $\epsilon > 0$, a chance variable for which the least expected length is larger than $H(X) + 1 - \epsilon$.