

MATHEMATICAL TRIPOS PART II (2006–07)

Coding and Cryptography - Example Sheet 1 of 4

T.A. Fisher

- 1) In a Binary Symmetric Channel (BSC) we usually take the probability p of error to be less than $1/2$. Why do we not consider $1 \geq p \geq 1/2$? What if $p = 1/2$?
- 2) Show that if we connect two Discrete Memoryless Channels (DMC's) in series or in parallel then the result is again a DMC. How are the channel matrices related? Illustrate in the case of two BSC's with error probabilities p and q .
- 3) (i) Give an example of a decipherable code which is not prefix-free. (Hint: What happens if you reverse all the codewords in a prefix-free code?)
(ii) Give an example of a non-decipherable code which satisfies the Kraft inequality.
(iii) Check directly that comma codes satisfy the Kraft inequality.
- 4) For a code $f : \Sigma_1 \rightarrow \Sigma_2^*$ and a code $f' : \Sigma'_1 \rightarrow \Sigma'_2^*$ the product code is $g : \Sigma_1 \times \Sigma'_1 \rightarrow (\Sigma_2 \cup \Sigma'_2)^*$ given by $g(x, y) = f(x)f'(y)$. Show that the product of two prefix-free codes is prefix-free, but that the product of a decipherable code and a prefix-free code need not even be decipherable.
- 5) Jensen's inequality states that if $f : \mathbb{R} \rightarrow \mathbb{R}$ is a convex function and p_1, \dots, p_n is a probability distribution (*i.e.* $0 \leq p_i \leq 1$ and $\sum p_i = 1$) then $f(\sum p_i x_i) \leq \sum p_i f(x_i)$ for any $x_1, \dots, x_n \in \mathbb{R}$. Deduce Gibbs' inequality from Jensen's inequality applied to the convex function $f(x) = -\log x$.
- 6) Show that $H(p_1, p_2, p_3) \leq H(p_1) + (1 - p_1)$ and determine when equality occurs.
- 7) Use the methods of Shannon-Fano and Huffman to construct prefix-free binary codes for messages μ_1, \dots, μ_5 emitted (i) with equal probabilities, or (ii) with probabilities 0.3, 0.3, 0.2, 0.15, 0.05. Compare the expected word lengths in each case.
- 8) Messages μ_1, \dots, μ_5 are emitted with probabilities 0.4, 0.2, 0.2, 0.1, 0.1. Determine whether there are optimal binary codings with (i) all but one codeword of the same length, or (ii) each codeword a different length.
- 9) Show that if an optimal binary code has word lengths s_1, \dots, s_m then

$$m \log m \leq s_1 + \dots + s_m \leq (m^2 + m - 2)/2.$$

- 10) Suppose that a gastric infection is known to originate in exactly one of m restaurants, the probability it originates in the j^{th} being p_j . A health inspector has samples from all of the m restaurants and by testing the pooled samples from a set A of them can determine with certainty whether the infection originates in A or its complement. Let $N(p_1, \dots, p_m)$ denote the minimum expected number of such tests needed to locate the infection. Show that $H(p_1, \dots, p_m) \leq N(p_1, \dots, p_m) \leq H(p_1, \dots, p_m) + 1$, and determine when the lower bound is attained.

Further Problems

Note: the examples above are minimal to cover the course; you are encouraged to do those below also.

- 11) (For those who did IB Optimisation.) Use a Lagrange multiplier to solve the following constrained optimisation problem: Given $p_i > 0$ with $\sum_{i=1}^m p_i = 1$ find real numbers s_1, \dots, s_m to minimise $\sum_{i=1}^m p_i s_i$ subject to $\sum_{i=1}^m a^{-s_i} \leq 1$.
- 12) Extend the definition of entropy to a random variable taking values in the non-negative integers. Compute the expected value $E(X)$ and entropy $H(X)$ of a random variable X with $P(X = k) = p(1-p)^k$. Show that among non-negative integer valued random variables with the same expected value, X achieves the maximum possible entropy.
- 13) In a horse race with m horses the probability that the i^{th} horse wins is p_i . The odds offered on each horse are a_i -for-1, *i.e.* a bet of x pounds on the i^{th} horse will yield $a_i x$ pounds if the horse wins, and nothing otherwise. A gambler bets a proportion b_i of his wealth on horse i , with $\sum_{i=1}^m b_i = 1$. He seeks to maximise $W = \sum_{i=1}^m p_i \log(a_i b_i)$. Suggest a motivation for this choice. Solve to find the b_i that maximise W . Show that in the case of even odds this maximum and the entropy $H(p_1, \dots, p_m)$ sum to a constant.
- 14) A source emits messages μ_1, \dots, μ_m with non-zero probabilities p_1, \dots, p_m . Let S be the codeword length random variable for a decipherable code $f : \Sigma_1 \rightarrow \Sigma_2^*$ where $\Sigma_1 = \{\mu_1, \dots, \mu_m\}$ and $|\Sigma_2| = a$. Show that the minimum possible value of $E(a^S)$ satisfies

$$\left(\sum_{i=1}^m \sqrt{p_i} \right)^2 \leq E(a^S) < a \left(\sum_{i=1}^m \sqrt{p_i} \right)^2.$$

(Hint: The Cauchy-Schwarz inequality.)

- 15) (i) In lectures we only described Huffman coding in the binary case, *i.e.* $a = 2$. In general we add extra messages of probability zero so that the number of messages m satisfies $m \equiv 1 \pmod{a-1}$. Then at each stage we group together the a smallest probabilities. Carry this out for a ternary coding of a source with probabilities 0.2, 0.2, 0.15, 0.15, 0.1, 0.1, 0.05, 0.05.
(ii) Show that if a ternary decipherable code of size m meets the lower bound in the noiseless coding theorem then m is odd.