MATHEMATICAL TRIPOS PART II (2004–05)

Coding and Cryptography - Example Sheet 2 of 4 **T.A. Fisher**

21) Consider a Binary Symmetric Channel with error probability $p = 1/3$. Suppose we send codewords 1100, 0110, 0001, 1111 with probabilities $1/4, 1/2, 1/6, 1/12$. How would you decode 1001 using (i) ideal observer decoding, or (ii) maximum likelihood decoding.

22) (i) Determine the number of codes with parameters $[n, 2, n]$ for $n \geq 2$.
(ii) Find the largest possible information rate of a 1-error detecting code of size 2^n.

23) Suppose we use eight hole tape with the standard paper tape code (i.e. the simple parity check code of length 8) and the probability that an error occurs at a particular place on the tape (i.e. a hole occurs where it should not or fails to occur where it should) is 10^{-4}. A program requires about 10 000 lines of tape (each line containing eight places) using the paper tape code. Using the Poisson approximation, direct calculation (possible with a hand calculator but really no advance on the Poisson method) or otherwise show that the probability that the tape will be accepted as error free by the decoder is less than .04%.

Suppose now that we use the Hamming scheme (making no use of the last place in each line). Explain why the program requires about 17 500 lines of tape but that any particular line will be correctly decoded with probability about $1 - (21 \times 10^{-8})$ and the probability that the entire program will be correctly decoded is better than 99.6%.

24) Show that the repetition code of length n is perfect if and only if n is odd.

25) Let C be the $[11, 12]$-code consisting of the word 10111000100 and its cyclic shifts (that is 01011100010, 00101110001 and so on) together with the zero code word. Is C linear? Show that C has minimum distance 5.

26) We show that, even if $2^n/V(n, e)$ is an integer, no perfect code may exist.
(i) Verify that $2^{90}/V(90, 2) = 2^{78}$.
(ii) Suppose that C is a perfect 2-error correcting code of length 90 and size 2^{78}. Explain why we may suppose without loss of generality that $0 \in C$.
(iii) Let C be as in (ii) with $0 \in C$. Consider the set

$$X = \{x \in \mathbb{F}_2^{90} : x_1 = 1,\ x_2 = 1,\ d(0, x) = 3\}.$$

Show that corresponding to each $x \in X$ we can find a unique $c(x) \in C$ such that $d(c(x), x) = 2$.
(iv) Continuing with the argument of (iii) show that $d(c(x), 0) = 5$ and that $c(x)_i = 1$ whenever $x_i = 1$. If $y \in X$ find the number of solutions to the equation $c(x) = c(y)$ with $x \in X$ and, by considering the number of elements of X, obtain a contradiction.
(v) Conclude that there is no perfect $[90, 2^{78}]$-code.

27) Prove that $A(n, d) \leq 2A(n - 1, d)$. Construct a $[7, 8, 4]$-code from Hamming's code.

28) Show that if d is even then $A(n - 1, d - 1) = A(n, d)$. Hence compute $A(5, 4)$.

29) (i) Show that $H(X|Y) \geq 0$ with equality if and only if X is a function of Y.
(ii) Give an example where $H(X|Y=y) > H(X)$, even though $H(X|Y) \leq H(X)$.

30) Players A and B play a (best of) 5 set tennis match. Let X be the number of sets won by A, and let Y be the total number of sets played. Assuming that the players are equally matched and the outcome of each set is independent, compute the conditional entropies $H(X|Y)$, $H(Y|X)$ and the mutual information $I(X;Y)$.

31) Show that $H(X,Z|Y) \leq H(X|Y) + H(Z|Y)$ and deduce that $H(X|Y,Z) \leq H(X|Y)$. Can you find the condition for equality?

32) Two BSC's with error probability p are connected in series. Compute the capacity of the new channel.

33) Consider two DMC's of capacity C_1 and C_2 with each having input alphabet Σ_1 and output alphabet Σ_2. Connecting in parallel gives the product channel with input alphabet $\Sigma_1 \times \Sigma_1$, output alphabet $\Sigma_2 \times \Sigma_2$, and channel probabilities given by

$$P(y_1 y_2 \text{ received} \,|\, x_1 x_2 \text{ sent}) = P(y_1 \text{ received} \,|\, x_1 \text{ sent}) P(y_2 \text{ received} \,|\, x_2 \text{ sent}).$$

Show that the product channel has capacity $C = C_1 + C_2$.

34) Show that the capacity of a DMC with channel matrix

$$\begin{pmatrix} 1-\alpha-\beta & \alpha & \beta \\ \alpha & 1-\alpha-\beta & \beta \end{pmatrix}$$

is given by $C = (1-\beta)(1-\log(1-\beta)) + (1-\alpha-\beta)\log(1-\alpha-\beta) + \alpha \log \alpha$.

Further Problems

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

35) Let C be an $[n,m,d]$-code. Show that $m(m-1)d \leq \sum\sum d(c_i,c_j) \leq \frac{1}{2}nm^2$ where the sum is over all codewords c_i and c_j of C. Use this to give an upper bound on $A(n,d)$ in the case $n < 2d$.

36) Let C be a perfect code with minimum distance 7. Show that C either has length 7 or has length 23. (You are *not* required to construct a code of length 23.)

37) Codewords 00 and 11 are sent with equal probability through a BSC with error probability p. Compute the mutual information between the codeword sent and the first digit received as output. Show that the extra mutual information to accrue on receipt of the second digit is $H(2p(1-p)) - H(p)$ bits.

38) If a channel matrix, with output alphabet of size n, is such that the set of entries in any row is the set $\{p_1, \ldots, p_n\}$, and the set of entries in each column is the same, show that its information capacity C is given by

$$C = \log n + \sum_{i=1}^{n} p_i \log p_i.$$

39) For random variables X and Y we define $\Delta(X,Y) = H(X|Y) + H(Y|X)$. Show that Δ satisfies the triangle inequality. Is it a metric?

40) Rephrase your solution to Question 20 (on coin weighing) in terms of conditional entropy.