**CODES AND CRYPTOGRAPHY – Example Sheet 2**

TKC Michaelmas 2007

1. In a binary symmetric channel we usually take the probability $p$ of error to be less than $1/2$. Why do we not consider $1 \geqslant p \geqslant 1/2$? What happens if $p = \frac{1}{2}$?
2. Books are identified by the ISBN-10 code. For example: ISBN 0-521-404568 This consists of 9 decimal digits $x_1$-$x_2x_3x_4$-$x_5x_6x_7x_8x_9$ which identify the publisher, author and title. The final digit $x_{10} \in \{0, 1, 2, \ldots, 9, X\}$ is a check digit which satisfies

$$10x_1 + 9x_2 + 8x_3 + \ldots + 2x_9 + x_{10} \equiv 0 \pmod{11} .$$

   Check the ISBN number above is valid. Show that altering any single digit or transposing any two adjacent digits in an ISBN-10 code will always result in an invalid code. Hence such errors can be detected.

   How do check digits work in ISBN-13? Do they detect all single digit errors?
3. Use Stirling's formula $\left[n! \sim \sqrt{2\pi n}\frac{n^n}{e^n}\right]$ to describe the behaviour of $\binom{N}{r}$ in terms of the entropy $h(r/N) = H(1 - (r/N), r/N)$.
4. Show that, if we connect two time-independent, memoryless channels in parallel or series, then the result is another time-independent, memoryless channel. Compute the transition matrices for the resulting channels when the original channels are binary symmetric channels with error probabilities $p$ and $q$. What are the capacities in this case?
5. A binary symmetric channel with error probability $p = \frac{1}{3}$ is used to send codewords 1100, 0110, 0001, 1111 with probabilities $\frac{1}{4}, \frac{1}{2}, \frac{1}{12}, \frac{1}{6}$. How would you decode 1001 using the (a) ideal observer rule, or (b) maximum likelihood rule, or (c) minimum distance rule?
6. Suppose we use eight hole tape with the standard paper tape code (*ie.* 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 $1-(21\times10^{-8})$ and the probability that the entire program will be correctly decoded is better than 99.6%.
7. Determine the set of integers $n$ for which the repetition code of length $n$ is perfect for the binary alphabet $\{0, 1\}$
8. We can model the situation where bits may be lost in transmission by the *binary erasure channel*. This has input $0, 1$ and output $0, *, 1$. A bit is left unchanged with probability $1 - \alpha$ and changed to $*$ with probability $\alpha$. Find the information capacity of this channel. .
9. Define $A(N, \delta)$ as the maximum size of a binary code of length $N$ with minimum distance $\delta$. Prove that

$$\frac{2^N}{V(N, \delta)} \leqslant A(N, \delta) \leqslant \frac{2^N}{V(N, \frac{1}{2}\delta)} .$$

   Find the (trivial) values of $A(N, 1), A(N, 2), A(N, N - 1)$ and $A(N, N)$.
10. Can equality occur in Fano's inequality?
11. Consider a ternary alphabet and a channel that has transition matrix

$$\begin{pmatrix} 1-2\alpha & \alpha & \alpha \\ \alpha & 1-2\alpha & \alpha \\ \alpha & \alpha & 1-2\alpha \end{pmatrix} .$$

    Calculate the information capacity of the channel.
12. Two code words are chosen independently at random from $\mathbb{F}_2^N$ with each string word equally likely. What is the expected Hamming distance between them?

13. Consider the code $c : \{0,1\}^4 \to \{0,1\}^8$ obtained by using the Hamming code for the first 7 bits and then a check digit

$$x_8 \equiv x_1 + x_2 + \ldots + x_6 + x_7 \pmod 2$$

for the last bit. Find the minimum distance for this code. How many errors can it detect or correct?

14. If there is a perfect, $e$-error correcting binary code of length $N$, show that $V(N, e + \frac{1}{2})$ divides $2^N$. This condition is not sufficient for such a code to exist. Prove this by establishing the following results.

(a) $2^{90}/V(90, \frac{5}{2}) = 2^{78}$.

Suppose now that there is a perfect, 2-error correcting binary code of length 90 and let $\mathcal{C}$ be the set of code words.

(b) We may assume that the zero word $\mathbf{0}$ is in $\mathcal{C}$. For any $\boldsymbol{x} \in \mathbb{F}_2^{90}$ with $d(\mathbf{0}, \boldsymbol{x}) = 3$, there is an unique code word $c(\boldsymbol{x})$ with $d(\boldsymbol{x}, c(\boldsymbol{x})) = 2$. Furthermore, $d(\mathbf{0}, c(\boldsymbol{x})) = 5$.

(c) Write $\boldsymbol{e}_j$ for the vector in $\mathbb{F}_2^{90}$ with 1 in the $j$th place and 0 everywhere else. The set

$$X = \{\boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_j : 2 < j \leqslant 90\}$$

has 88 elements. Exactly 3 of the vectors $\boldsymbol{x} \in X$ have the same value for $c(\boldsymbol{x})$. So $3|88$, which is untrue.

15. Codewords 00 and 11 are sent with equal probability through a binary symmetric channel 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.

16. For random variables $X, Y, Z$ show that:

(a) $H(Y,Z) \leqslant H(X,Y,Z) = H(X|(Y,Z)) + H(Y,Z)$.

(b) $H(X|(Y,Z)) \leqslant H(X|Y)$.

Deduce that $H(X|Z) \leqslant H(X|Y) + H(Y|Z)$.

Prove that $\Delta(X,Y) = H(X|Y) + H(Y|X)$ satisfies the triangle inequality. Is it a metric?

17. **Data Processing Inequality**

Consider two independent channels in series. A random variable $X$ is sent through channel 1 and received as $Y$. This is then sent through channel 2 and received as $Z$. Our aim is to prove that $I(X,Z) \leqslant I(X,Y)$, so the further processing of the second channel can only reduce the mutual information.

The independence of the channels means that, if we condition on the value of $Y$, then $(X|Y=y)$ and $(Z|Y=y)$ are independent. Deduce that

$$H(X,Z|Y) = H(X|Y) + H(Z|Y) \ .$$

By writing the conditional entropies as $H(A|B) = H(A,B) - H(B)$, show that

$$H(X,Y,Z) + H(Y) = H(X,Y) + H(Y,Z) \ .$$

Define $I(X,Y|Z)$ as $H(X|Z) + H(Y|Z) - H(X,Y|Z)$ and show that

$$I(X,Y|Z) = I(X,Y) - I(X,Z) \ .$$

Deduce from this the *data processing inequality*:

$$I(X,Z) \leqslant I(X,Y) \ .$$

When is there equality?

---

*Please send any comments or corrections to me at: t.k.carne@dpmms.cam.ac.uk .*