2003 Paper 7 Question 8

Information Theory and Coding

(a) Describe the types of data which are amenable to lossy compression with an explanation of why they are amenable to lossy compression. Describe the mechanisms that are used to perform lossy compression with an explanation of why they improve the compression rate. [8 marks]

(b) Two coding schemes are proposed for the binary coding of a four-symbol alphabet:

Symbol	Code 1	Code 2
a	00	0
b	01	10
c	10	110
d	11	111

Under what probability distributions would Code 2 be a more efficient code than Code 1? You may assume that $p(a) \geq p(b) \geq p(c) \geq p(d)$. [6 marks]

(c) For an alphabet consisting of m equiprobable symbols encoded using a binary prefix code, prove that the average length per symbol of the binary code is greater than or equal to $\log_2(m)$ bits, for any possible prefix code. [6 marks]