GLASGOW COLLEGE UESTC

Exam Paper

Elements of Information Theory (UESTC) (UESTC3021)

DATE： 13th Jan. 2021
Time: 09:30am - 11:30am

Attempt all PARTS. Total 100 marks

Use one answer sheet for each of the questions in this exam.
Show all work on the answer sheet.

Make sure that your University of Glasgow and UESTC Student Identification Numbers are on all answer sheets.

An electronic calculator may be used provided that it does not allow text storage or display, or graphical display.

All graphs should be clearly labelled and sufficiently large so that all elements are easy to read.

The numbers in square brackets in the right-hand margin indicate the marks allotted to the part of the question against which the mark is shown. These marks are for guidance only.

Continued overleaf

Q1 Markov chains have been widely used for approximating a practical information source, such as English text. When zero-order approximation is invoked, the approximated information source randomly outputs English letters with non-sense. However, as the order of the approximated Markov information source increases, meaningful English words can be output. This problem aims for investigating the entropy rate of one-order Markov information source:

(a) Find the entropy rate of the two-state Markov chain with the following transition matrix:

$$P = \begin{bmatrix} 1-p_{01} & p_{01} \\ p_{10} & 1-p_{10} \end{bmatrix}$$ [5]

(b) What values of p_{01}, p_{10} maximize the rate of part (a)? [5]

(c) Find the entropy rate of the two-state Markov chain with the following transition matrix

$$P = \begin{bmatrix} 1-p & p \\ 1 & 0 \end{bmatrix}$$ [5]

(d) Find the maximum value of the entropy rate of the Markov chain of part (c). We expect that the maximizing value of p should be less than 1/2, since the 0 state permits more information to be generated than the 1 state. [5]

Continued overleaf

Q2 Huffman encoding is one of the optimal lossless source coding technique. ZIP is the most widely used data compression toll that uses Huffman encoding as its basis. The latest of the most efficient lossless compression algorithms Brotli Compression, released b Google last month also uses Huffman encoding. Consider the following discrete information source:

$$X = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 \\ 0.49 & 0.26 & 0.12 & 0.04 & 0.04 & 0.03 & 0.02 \end{pmatrix}$$

(a) Find a binary Huffman code for X, its expected code length and its coding efficiency. [8]

(b) Find a ternary Huffman code for X, its expected code length and its coding efficiency. [8]

(c) Find the binary Huffman code for the source Y with probabilities (1/3, 1/5, 1/5, 2/15, 2/15), its expected code length and its coding efficiency. [8]

(d) Argue that the Huffman code obtained in (c) is also optimal for the source with probabilities (1/5, 1/5, 1/5, 1/5, 1/5). [6]

Continued overleaf

Q3 Consider a commander of an army besieged a fort for whom the only means of communication to his allies is a set of carrier pigeons. Assume that each carrier pigeon can carry one letter (8 bits), and assume that pigeons are released once every 5 minutes, and that each pigeon takes exactly 3 minutes to reach its destination.

(a) Assuming all the pigeons reach safely, what is the capacity of this link in bits/hour? [5]

(b) Now assume that the enemies try to shoot down the pigeons, and that they manage to hit a fraction α of them. Since the pigeons are sent at a constant rate, the receiver knows when the pigeons are missing. What is the capacity of this link? [10]

(c) Now assume that the enemy is more cunning, and every time they shoot down a pigeon, they send out a dummy pigeon carrying a random letter (chosen uniformly from all 8-bit letters). What is the capacity of this link in bits/hour?

Set up an appropriate model for the channel in each of the above cases, and indicate how to go about finding the capacity. [10]

Continued overleaf

Q4 (a) The Z-channel has binary input and output alphabets and transition probabilities p(y|x) given by the following matrix:

$$Q = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \end{bmatrix} \qquad x, y \in \{0, 1\}$$

Find the capacity of the Z-channel and the maximizing input probability distribution.[10]

(b) Consider a 5-repetition codes, i.e., "1" is encoded into "11111" and "0" is encoded into "00000", which will pass through a binary symmetric channel (BSC). Assume the BSC has correct transition probability p and error transition probability $\bar{p} = 1 - p$. Please answer following 3 questions.

(i) Firstly, how many types of error transmission can be detected? [5]

(ii) Secondly, how many errors can be decoded? [5]

(iii) If $p = 0.01$, please compute the decoding error probability. [5]

End of question paper