TCOM 370 - 1999

Home Work 6 and Reading Assignment/Problems

Problems 1, 2, 3, and 10 to be turned in for grading on March 30, 1999

Problem 1 (from 1997 Exam 2)

A source produces independent symbols from an alphabet of three letters. Each source symbol can be A, B, or C with respective probability 0.5, 0.25 and 0.25.

(a) What is the source entropy H ?

(b) Find a Huffman code for the individual letters of the source alphabet. What is the average number of bits per symbol for this code?

(c) You are asked to design a code for *blocks of two symbols* from the source at a time. The alphabet for this extended source is of size 9.

- (i) Can such a code provide better performance than the one in part (a)?
- (ii) Find a best uniquely-decodable code for encoding blocks of two symbols at a time.
- (iii) What is the average number of bits per source symbol for your code?

Problem 2: A source produces independent symbols from a five-letter alphabet $\{a_0, a_1, a_2, a_3, a_4\}$, the letters having respective probabilities {0.4, 0.2, 0.2, 0.1, 0.1}. You will design binary Huffman codes for this source, to work on single symbols.

(a) In combining lowest-probability letters at each stage in the code tree and re-sorting, whenever equal-probability groups of symbols (combined or individual) are created follow the natural procedure of retaining the previous stage ordering as far as possible. What is the average length of the Huffman codeword?

(b) This time place combined symbols as high up as possible in the re-sorted list at any stage. What is the average length of the Huffman codeword?

(c) Which code would you prefer in transmission over a fixed rate link?

Problem 3 (Exam 3, 1998 #2)

Source I produces independent symbols from an alphabet of two letters, A and B. For each symbol the probabilities of the letters A and B are 0.6 and 0.4 respectively.

(a) Design a Huffman code operating on *blocks of two output symbols* from Source I.

(b) What is the average number of bits per symbol for your code? What is the minimum number of bits per symbol that the best code should be able to give you for Source I?

(c) Source II is also a source with a two-letter alphabet {C, D} with P{C}=0.8 and P{D}=0.2. A *composite source* is formed by choosing each time at random with equal probability between Source I and Source II, for its output letters.
Find the Huffman code for single letters for the *composite source,* and the average bits/letter.

Problem 4 (Exam 3, 1998 #1)

(a) Define the term *"prefix-condition"* in describing a code assigning binary codewords to the letters of a finite-alphabet source.

(b) Consider the following binary code for source alphabet {X, Y, Z}:

X --> 11
Y --> 10
Z --> 100

(i) Is this a prefix-condition code?
(ii) For any sequence of source letters encoded using this code, would you be able to decide from the binary coded stream the original sequence of source letters (i.e., is the code uniquely decodable)? Justify your answer briefly.

Problem 5 A source is producing a sequence of independent symbols from a binary alphabet {A, B}. The source produces letter A with probability 0.9 and letter B with probability 0.1. You are to design Huffman codes and find the average number of bits required per source symbol.

(a) Find the source entropy H.

(b) What is the Huffman code designed for the individual letters of the alphabet? What is the average number of bits per symbol in this case? (The answers are trivial).

(c) Consider encoding blocks of **three** symbols from the source at a time. The alphabet for this extended source is now of size 8 (all length-3 combinations of the 2 letters). Obtain a Huffman code for this case. Compute the average number of bits per source symbol, compare your answer with the results of part (a) and (b) above.

Problem 6 (from Final Exam, 1997)

Consider a standard telephone-quality PCM scheme for voice digitization at 64 Kbps. An engineer determines that for a certain type of speech the entropy of each quantized sample is 6 bits/sample, and that adjacent samples are statistically *dependent.*

(a) Is it possible to do a lossless compression after the PCM encoding that results in a bit rate of 50 Kbps for the speech? Explain your answer clearly, and if it is possible, explain what you would need to do to achieve this.

(b) Explain briefly one practical technique to decrease the bit rate to significantly below 64 Kbps without suffering major quality degradation compared to standard PCM speech.

Problem 7

Exer. 3.16, Halsall

Problem 8

Exer. 3.17, Halsall

Reading Assignment and Problem 9

Read Section 3.5.6 on Facsimile; focus on Group 3 (T4) fax standard.

A fax transmission produces the following bit sequence (this is the first part of some page). What is the corresponding decoded pattern of pixels?

00000000000100110101000110001101100100110101001100000000000001

How much compression has been obtained in this example?

(Given: Transmission always starts with an EOL code; the first run is always a run of whites; each line ends with an EOL code).

Reading Assignment and Problem 10

Read Section 3.7 first, focusing on the ideas in 3.7.1 and 3.7.2. Do not worry about details of implementations.

In a statistical multiplexing situation, suppose the following parameters are defined:

N = number of terminals
R = data rate of each terminal, bps, when it is transmitting
α = average fraction of time each terminal is transmitting
M = effective capacity of multiplexed transmission line, bps, taking into account overhead bits introduced by multiplexer

$K = \frac{M}{NR}$ = ratio of multiplexed line capacity to total maximum input rate

(a) Explain why we have $\alpha<K<1$ for a useful system. What does the value K=1 correspond to?

(b) The inputs to the multiplexer can be modeled as coming in to a "server" or buffer which transmits the data if the line is free or, if the line is busy, buffers it for transmission when the line becomes free. Under a simple model of arriving inputs (random arrivals described by the Poisson distribution), a formula for the utilization factor ρ, or fraction of time the server is busy (line is transmitting) when an input arrives, is obtained as

$$\rho = \frac{\alpha}{K}$$

Also, the average number of inputs that are in the queue (buffer) is given by

$$q = \frac{\rho^2}{2(1-\rho)} + \rho$$

Plot q as a function of K for fixed $\alpha=0.2$. What does this suggest about the trade-off between buffer size/delay on the one hand, and the speed of the line (M) on the other hand?