

CSE 490 GZ
Assignment 1
Due Friday, January 18, 2002

1. Consider the alphabet $\{a, b, c, d, e, f\}$ with fixed prefix code

a	00
b	01
c	100
d	101
e	110
f	111

Use the adaptive Huffman code to decode the following binary string.

10110000101100111

Show the decoder's tree after each symbol is decoded.

2. The Kraft-McMillan inequality

$$\sum_{i=1}^n 2^{-l_i} \leq 1.$$

provides a condition for the existence of a prefix code for n symbols where symbol a_i has length l_i . Design a *recursive* algorithm that constructs for an input sequence of positive integers l_1, l_2, \dots, l_n an output binary tree with leaves labeled a_1, a_2, \dots, a_n and the depth of the leaf labeled a_i is l_i . (Hint: there is a construction described on pages 33-34 of Sayood which may be helpful. However, it may be more useful to look at slides 11 and 21 of Lecture 2.)

Your algorithm should start with n nodes thought of as a forest of n trees each with one node. In each recursive call, the quantity $s = \sum_{i=1}^n l_i$ is reduced by one. Generally, there are three cases to consider if there is a recursive call. If there are at least two symbols then either $l_1 = l_2$ and $l_1 > l_2$. The third case is when there is just one symbol left and $s > 0$.

Demonstrate your algorithm for the sequence 4, 3, 3, 2, 2.

3. In this problem we consider how to build Huffman codes for one symbol contexts. Consider the four symbols $\{a, b, c, d\}$ with conditional probabilities $P(x, y)$ defined by the table.

P	a	b	c	d
a	.1	.5	.1	.3
b	.4	.4	.1	.1
c	.1	0	.2	.7
d	.2	.3	.2	.3

$P(x, y)$ is the conditional probability that the symbol immediately following symbol x is symbol y . The x is indicated by a row and the y by a column. Thus, each row must sum to 1. Define $P^*(x)$ as the probability of x in the long run. It can be seen, for example, that the probability of the symbol a is

$$P^*(a) = .1P^*(a) + .4P^*(b) + .1P^*(c) + .2P^*(d)$$

which is the sum for each symbol of the probability of the symbol times the conditional probability that the next symbol is a .

- (a) Calculate the probability of each symbol. That is, calculate $P^*(x)$ for each x . To do this you need to solve a system of linear equations.
- (b) Construct the Huffman trees for these contexts. (See slides 26-27 of Lecture 2)
- (c) Calculate the bit rate of these trees as a code. This can be done using $P^*(x)$ for $x \in \{a, b, c, d\}$ and the bit rate for each Huffman tree.