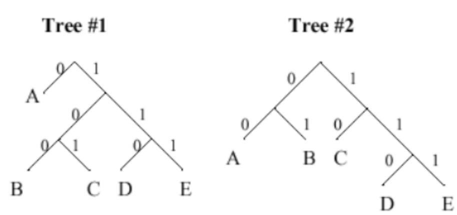
Assignment 2

1 Huffman Coding

1. Consider the following two Huffman decoding trees for a variable-length code involving 5 symbols: A, B, C, D, and E.



   (a) Using Tree #1, decode the following encoded message: "01000111101".

   (b) Suppose we were encoding messages with the following probabilities for each of the 5 symbols: $p(A) = 0.5, p(B) = p(C) = p(D) = p(E) = 0.125$. Which of the two encodings above (Tree #1 or Tree #2) would yield the shortest encoded messages averaged over many messages?

   (c) Using the probabilities from part (b), if Tree #2 is used to encode messages, what is the average length of 100-symbol messages, averaged over many messages?

2. Several people at a party are trying to guess a 3-bit binary number.

   (a) Alice is told that the number is odd.

   (b) Bob is told that it is not a multiple of 3 (i.e., not 0, 3, or 6).

   (c) Charlie is told that the number contains exactly two 1's.

   (d) Deb is given all three of these clues. How much information (in bits) did each player get about the number?

3. Consider messages made up entirely of vowels (A, E, I, O, U) with given probabilities.

I	$p(I)$	$\log_2(1/p(I))$	$p(I) \cdot \log_2(1/p(I))$
A	0.22	2.18	0.48
E	0.34	1.55	0.53
I	0.17	2.57	0.43
O	0.19	2.40	0.46
U	0.08	3.64	0.29
Totals	1.00	12.34	2.19

Table 1: Probability distribution and entropy calculation

(a) Find the number of bits of information received when learning that a particular vowel is either I or U.

(b) Using Huffman's algorithm, construct a variable-length code assuming that each vowel is encoded individually. Draw a Huffman tree and provide encodings.

(c) Using the obtained code, express the expected length in bits of an encoded message transmitting 100 vowels.

(d) Ben Bitdiddle proposes an alternative encoding with a message length of 197 bits per 100 vowels. Evaluate if this is efficient.

4. The following table shows current undergraduate and MEng enrollments for the School of Engineering.

Course (Department)	**# of students**	**Prob.**
I (Civil & Env.)	121	0.07
II (Mech. Eng.)	389	0.23
III (Mat. Sci.)	127	0.07
VI (EECS)	645	0.38
X (Chem. Eng.)	237	0.13
XVI (Aero & Astro)	198	0.12
Total	1717	1.0

Table 2: Student Distribution by Course

(a) Determine which student department provides the least amount of information.

(b) Design a variable-length Huffman code to minimize the average number of bits used to encode department data. Provide the Huffman tree and encodings.

(c) Compute the average message length when encoding departments for groups of 100 randomly chosen students.

2 Channel Coding

1. (a) How does channel coding differ from source coding?
   (b) Define the generator matrix and the parity-check matrix in a linear block code and describe their roles in encoding and error detection.

2. For binary repetition code with block length $n = 5$ (5,1) code:
   (a) Determine the systematic generator matrix G and the corresponding parity-check matrix H.

3. Consider the single parity-check code with block length $n = 6$ (6,5) code.
   (a) Find the systematic generator matrix G and derive the parity check matrix H.

4. For a Hamming code with $m = 4$ and $k = 11$, find the systematic generator matrix G and the parity check matrix H.

5. Given a generator matrix $G = [I_5|\mathbf{1}]$ where $\mathbf{1}$ is an all-one column vector of length 5:
   (a) Determine the rate of this code.
   (b) Describe the set of all codewords and determine the codeword corresponding to the message vector $(1, 0, 0, 1, 0)$.
   (c) Write down a set of parity-check equations and obtain a parity-check matrix for the code.

6. Given a generator matrix $G = [I_3|P]$ for a binary linear block code where

$$P = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

   (a) Find the corresponding parity-check matrix H.