Fundamentals of Information Science: Homework 5

March 19, 2025

Problem 1.
Consider the LZW compression and decompression algorithms. Assume that the scheme has an initial table with codewords 0 through 255 corresponding to the 8-bit ASCII characters; character "a" is 97 and "b"is 98. The receiver gets the following sequence of codewords, each of which is 10 bits long:

97 97 98 98 257 256

(a)What was the original message sent by the sender?

(b)By how many bits is the compressed message shorter than the original message(each character in the original message is 8 bits long)?

(c)What is the first string of length 3 added to the compression table?(If there's no such string,your answer should be "None".)

Problem 2.
Z-channel. The Z-channel has binary input and output alphabets and transition probabilities $p(y|x)$ given by the following matrix:

$$Q = \begin{pmatrix} 1 & 0 \\ 3/4 & 1/4 \end{pmatrix} \quad x, y \in \{0, 1\}$$

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

Problem 4.
The weight of a codeword in a linear block code over $\mathcal{F}_2$ is the number of 1's in the word. Show that any linear block code must either: (1) have only even weight codewords, or (2) have an equal number of even and odd weight codewords. Hint: Proof by contradiction.

Problem 2. The Matrix Reloaded.

Neo receives a 7-bit string, $D_1D_2D_3D_4P_1P_2P_3$ from Morpheus, sent using a code, $\mathcal{C}$, with parity equations

$$P_1 = D_1 + D_2 + D_3$$

$$P_2 = D_1 + D_2 + D_4$$

$$P_3 = D_1 + D_3 + D_4$$

(a)Write down the generator matrix, G,for $\mathcal{C}$.

(b)Write down the parity check matrix, H ,for $\mathcal{C}$.

(c)If Neo receives 1000010 and does maximum-likelihood decoding on it,what would his estimate of the data transmission $D_1D_2D_3D_4$ from Morpheus be? For your convenience, the syndrome s_i corresponding to data bit D_i being wrong are given below,for $i = 1, 2, 3, 4$:

$$s_1 = (111)^T, s_2 = (110)^T, s_3 = (101)^T, s_4 = (011)^T.$$

(d)If Neo uses syndrome decoding for error correction, how many syndromes does he need to compute and store for this code, including the syndrome with no errors?