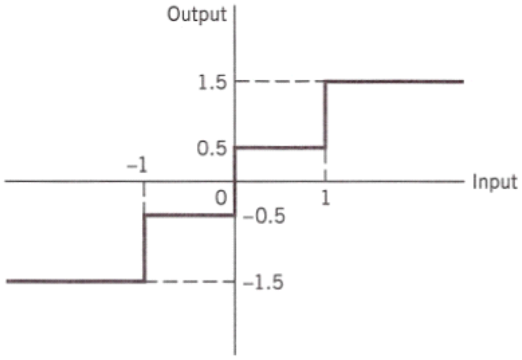
Information Theory and Coding Course

Problem Set #1

1. A source emits one of four symbols s_0, s_1, s_2, and s_3 with probabilities 1/3, 1/6, 1/4, and 1/4, respectively. The successive symbols emitted by the source are statistically independent. Calculate the entropy of the source.

2. Let X represents the outcome of a single roll of a fair die. What is the entropy of X?

3. The sample function of a Gaussian process of zero mean and unit variance is uniformly sampled and then applied to a uniform quantizer having the input-output amplitude characteristic shown in the Figure below. Calculate the entropy of the quantizer output.



4. Consider a discrete memoryless source with source alphabet : $S=\{s_0, s_1, s_2\}$ and source statistics $\{0.7, 0.15, 0.15\}$.

(a) Calculate the entropy of the source.

(b) Calculate the entropy of the second-order extension of the source. (two symbols)

-

Problem Set # 2

1. Consider a sequence of letters of the English alphabet with their probabilities of occurrence as given here:

Letter	*a*	*i*	*l*	*m*	*n*	*o*	*p*	*y*
Probability	0.1	0.1	0.2	0.1	0.1	0.2	0.1	0.1

Compute two different Huffman codes for this alphabet. In one case, move a combined symbol in the coding procedure as high as possible, and in the second case, move it as low as possible. Hence, for each of the two codes, find the average code-word length and the variance of the average code-word length over the ensemble of letters.

2. A discrete memoryless source has an alphabet of seven symbols whose probabilities of occurrence are as described here:

Symbol	s_0	s_1	s_2	s_3	s_4	s_5	s_6
Probability	0.25	0.25	0.125	0.125	0.125	0.0625	0.0625

Compute the Huffman code for this source, moving a "combined" symbol as high as possible. Explain why the computed source code has an efficiency of 100 percent.

3. Consider a discrete memoryless source with alphabet $\{s_0, s_1, s_2\}$ and statistics {0.7, 0.15, 0.15} for its output.

(a) Apply the Huffman algorithm to this source. Hence, show that the average code-word length of the Huffman code equals 1.3 bits/symbol.

(b) Let the source be extended to order two. Apply the Huffman algorithm to the resulting extended source, and show that the average code-word length of the new code equals 1.1975 bits/symbol.

(c) Compare the average code-word length calculated in part (b) with the entropy of the original source.

4. Consider the following binary sequence

11101001100010110100...

Use the Lempel-Ziv algorithm to encode this sequence. Assume that the binary symbols 0 and 1 are already in the codebook.

5. A discrete memoryless source with $A = \{ a, b, c \}$ emits the following string.

bccacbcccccccccccaccca

Using the Lempel-Ziv algorithm, encode this sequence and find the code dictionary and the transmitted sequence.

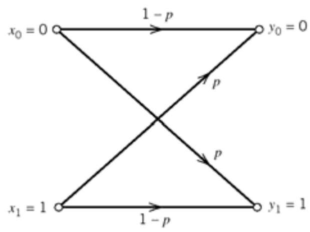
6. A source with $A = \{a, b, c\}$ is encoded using the Lempel-Ziv algorithm. The transmitted code word sequence is 2,3,3,1,3,4,5,10,11,6,10.

Construct the dictionary and decode this sequence.

Problems Set #3

Binary Symmetric Channel:

FIGURE 1



1. Consider the transition probability diagram of a binary symmetric channel shown in the Figure. The input binary symbols 0 and 1 occur with equal probability. Find the probabilities of the binary symbols 0 and 1 appearing at the channel output.

2. Repeat the calculation in Problem 1, assuming that the input binary symbols 0 and 1 occur with probabilities 1/4 and 3/4, respectively.

Mutual Information and Channel Capacity:

3. Consider a binary symmetric channel characterized by the transition probability *p*. Plot the mutual information of the channel as a function of P1, the *a priori* probability of symbol 1 at the channel input; do your calculations for the transition probability *p*=0, 0.1, 0.2, 0.3, 0.5.

4. Two binary symmetric channels are connected in cascade, as shown in the Figure below. Find the overall channel capacity of the cascaded connection, assuming that both channels have the same transition probability diagram shown in the Figure 1.

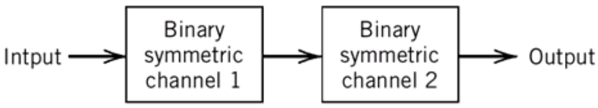


FIGURE 2

5. The *binary erasure channel* has two inputs and three outputs as described in Figure 3. The inputs are labeled 0 and 1, and the outputs are labeled 0, 1, and *e*. A fraction α of the incoming bits are erased by the channel. Find the capacity of the channel.

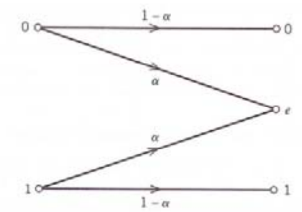


FIGURE 3

Problem Set # 4

1. Golay code (23,12) is a perfect code. Determine the error correction capability of the code.

2. Confirm the possibility of an (18,7) code that can correct up to three errors.

3. Consider a (7,4) Hamming code with the parity check matrix H given by:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

   a) Construct the **G** matrix.
   b) Find the codeword for the information sequence [1 1 0 0].
   c) If the word [0101100] is received, what is the decoded codeword?
   d) What action will the decoder take for the following scenarios of error patterns:
      i) two errors in the first and second positions.
      ii) three errors in the first, fourth and seventh positions.
      iii) four errors in the first, fifth, sixth and seventh positions.

4. For a (6,3) systematic linear block code, the three parity-check bits b_0, b_1, and b_2 are given by:

$$b_0 = m_0 \oplus m_1 \oplus m_2$$
$$b_1 = m_0 \oplus m_1$$
$$b_2 = m_0 \oplus m_2$$

   a) Construct the appropriate generator matrix **G**.
   b) Construct the code generated by this matrix.
   c) Determine the error correcting capabilities of this code.
   d) Prepare a suitable decoding table.
   e) Decode the following received codewords: 101100, 000110, 101010.

5. Consider a generator matrix for a nonsystematic (6,3) code:

$$\mathbf{G} = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

   a) Construct the code for this **G**.
   b) Find the minimum distance and therefore the error correcting capability of the code.
   c) Prepare a code table for this code.
   d) Prepare a suitable decoding table.

Problem Set # 5

1. Perform the following calculation in GF(2)[x]

a) $(1+x)(1+x^2)+x^3$

b) $x+x^4 \bmod x^2+1$

c) $1+x+x^2 \bmod 1+x$

2. For polynomials in GF(2)[x], show that

$$\left(1+x^n\right)^2 = 1+x^{2n}$$

3. Given that $X^9+1=(X+1)(X^2+X+1)(X^6+X^3+1)$ determine the cyclic codes with block length 9.

4. Determine the parity-check polynomial of the (15,5) cyclic code with generator polynomial given by:

$$g(X)=1+X+X^2+X^4+X^5+X^8+X^{10}$$

5. Show that the following linear code with generator matrix **G** is not cyclic:

$$\mathbf{G} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

6. The generator polynomial of a (15,11) Hamming code is defined by:

$$g(X)=1+X+X^4$$

Develop the encoder and syndrome calculator for this code, using a systematic form for the code.

7. Consider the (7,4) Hamming code defined by the generator polynomial:

$$g(X)=1+X+X^3$$

The code word 0111001 is sent over a noisy channel, producing the received word 0101001 that has a single error. Determine the syndrome polynomial s(X) for this received word, and show that it is identical to the error polynomial e(X).

8. Construct a systematic (7,3) cyclic code.

Problem Set # 6

Convolutional Codes

10.15 A convolutional encoder has a single-shift register with two stages, (i.e., constraint length $K = 3$), three modulo-2 adders, and an output multiplexer. The generator sequences of the encoder are as follows:

$$g^{(1)} = (1, 0, 1)$$
$$g^{(2)} = (1, 1, 0)$$
$$g^{(3)} = (1, 1, 1)$$

Draw the block diagram of the encoder.

Note: For Problems 10.16–10.23, *the same message sequence* 10111 . . . *is used so that we may compare the outputs of different encoders for the same input.*

10.16 Consider the rate $r = 1/2$, constraint length $K = 2$ convolutional encoder of Fig. P10.16. The code is systematic. Find the encoder output produced by the message sequence 10111. . . .

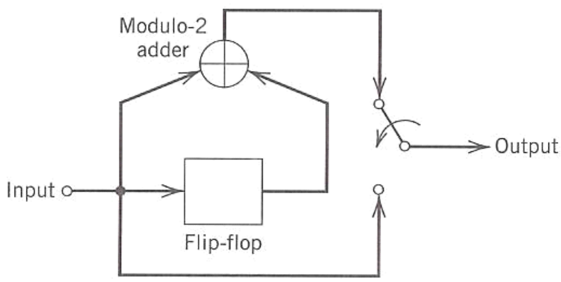


FIGURE P10.16

10.17 Figure P10.17 shows the encoder for a rate $r = 1/2$, constraint length $K = 4$ convolutional code. Determine the encoder output produced by the message sequence 10111. . . .

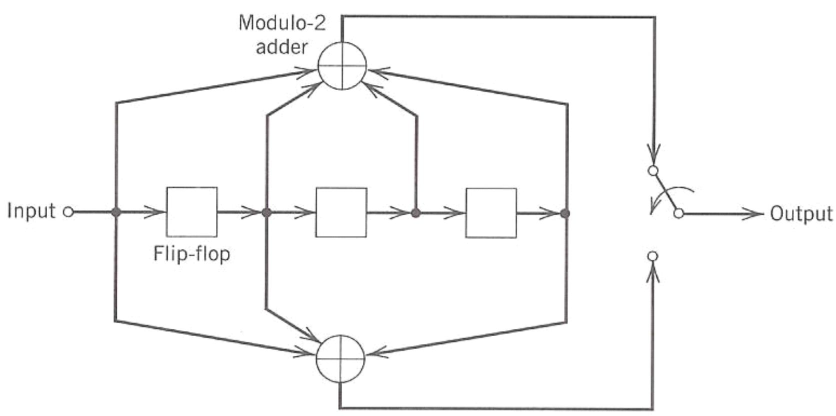


FIGURE P10.17

10.19 Construct the code tree for the convolutional encoder of Fig. P10.16. Trace the path through the tree that corresponds to the message sequence 10111 . . . , and compare the encoder output with that determined in Problem 10.16.

10.20 Construct the code tree for the encoder of Fig. P10.17. Trace the path through the tree that corresponds to the message sequence 10111. . . . Compare the resulting encoder output with that found in Problem 10.17.

10.21 Construct the trellis diagram for the encoder of Fig. P10.17, assuming a message sequence of length 5. Trace the path through the trellis corresponding to the message sequence 10111. . . . Compare the resulting encoder output with that found in Problem 10.17.

10.25 The trellis diagram of a rate-1/2, constraint length-3 convolutional code is shown in Figure P10.25. The all-zero sequence is transmitted, and the received sequence is 100010000. . . . Using the Viterbi algorithm, compute the decoded sequence.

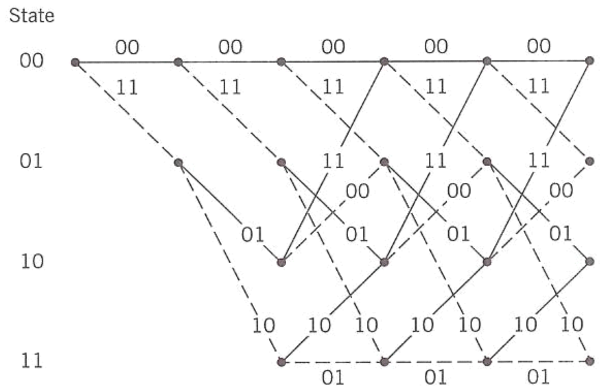


FIGURE P10.25