PROBLEM 1. Suppose $L : \mathbb{R}^K \to \mathbb{R}^N$ is a linear function and $g : \mathbb{R}^N \to \mathbb{R}$ is a concave function. Show that $f : \mathbb{R}^K \to \mathbb{R}$ defined as $f(x) = g(L(x))$ is concave.

PROBLEM 2. From the notes on Lempel-Ziv algorithm, we know that the total length n of c distinct binary strings satisfies

$$n > c \log_2(c/8)$$

The same technique, when applied to the non-binary strings yields

$$n > c \log_K(c/K^3)$$

where K is the size of the alphabet the letters of the string belong to. This inequality lower bounds n in terms of c. We will now show that n can also be upper bounded in terms of c.

(a) Show that, if $n \geq \frac{1}{2}m(m-1)$, then $c \geq m$.

(b) Find a sequence for which the bound in (a) is met with equality.

(c) Show now that $n < \frac{1}{2}c(c+1)$.

PROBLEM 3. Let X be the channel input. Assume that the channel output Y is passed through a data processor in such a way that no information is lost. That is,

$$I(X;Y) = I(X;Z)$$

where Z is the processor output. Find an example where $H(Y) > H(Z)$ and find an example where $H(Y) < H(Z)$.
Hint: The data processor does not have to be deterministic

PROBLEM 4. A "K-ary erasure channel with erasure probability p" is described as follows: the input U belongs to the alphabet $\{1, \dots, K\}$, the output V belongs to the alphabet $\{1, \dots, K\} \cup \{?\}$, and if u is the input, the output V equals u with probability $1-p$, and equals ? with probability p. Note that $\Pr(V = ?) = p$ regardless of the input distribution.

(a) Show that $\Pr(U = u | V = ?) = p_U(u)$.

(b) Show that $I(U;V) = (1-p)H(U)$.

(c) Find the capacity of this channel and the input distribution that maximizes the mutual information.

PROBLEM 5. We are given a memoryless stationary binary symmetric channel BSC(ϵ). Namely, if $X_1, \dots, X_n \in \{0,1\}$ are the input of this channel and $Y_1, \dots, Y_n \in \{0,1\}$ are the output, we have:

$$P(Y_i|X_i, X^{i-1}, Y^{i-1}) = P(Y_i|X_i) = \begin{cases} 1-\epsilon & \text{if } Y_i = X_i, \\ \epsilon & \text{otherwise.} \end{cases}$$

Let W be a random variable that is uniform in $\{0,1\}$ and consider a communication system with feedback which transmits the value of W to the receiver as follows:

- At time $t = 1$, the transmitter sends $X_1 = W$ through the channel.
- At time $t = i + 1 \leq n$, the transmitter gets the value of Y_i from the feedback and sends $X_{i+1} = Y_i$ through the channel.

(a) Give the capacity C of the channel in terms of ϵ, and show that $C = 0$ when $\epsilon = \frac{1}{2}$.

(b) Show that if $\epsilon = \frac{1}{2}$, $I(X^n; Y^n) = n - 1$. This means that $I(X^n; Y^n) \leq nC$ does not hold for this system.

(c) Show that although $I(X^n; Y^n) > nC$ when $\epsilon = \frac{1}{2}$, we still have $I(W; Y^n) \leq nC$.

Note that since W is the useful information that is being transmitted, it is the value of $I(W; Y^n)$ that we are interested in when we want to compute the amount of information that is shared with the receiver.