**PROBLEM 1.** Channels with memory have higher capacity. Consider a binary symmetric channel with $Y_i = X_i \oplus Z_i$, where $\oplus$ is mod 2 addition, and $X_i, Y_i \in \{0, 1\}$.

(a) Suppose that $\{Z_i\}$ has constant marginal probabilities $\Pr\{Z_i = 1\} = p = 1 - \Pr\{Z_i = 0\}$, but that $Z_1, Z_2, \ldots, Z_n$ are not necessarily independent. Assume that $(Z_1, \ldots, Z_n)$ is independent of the input $(X_1, \ldots, X_n)$. Let $C = \log 2 - H(p, 1 - p)$. Show that
\[
\max_{p_{X_1, X_2, \ldots, X_n}} I(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n) \geq nC.
\]

(b) Suppose that the $\{Z_i\}$ are generated as follows – $\Pr(Z_1 = 0) = \Pr(Z_1 = 1) = \frac{1}{2}$ and for $i \geq 1, \Pr(Z_{i+1} \neq Z_i) = q$.

(i) What is the marginal probability – $\Pr(Z_i = 1)$?

(ii) Justify the following sequence of steps:
\[
I(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n) = H(Y_1, \ldots, Y_n) - H(Y_1, \ldots, Y_n|X_1, \ldots, X_n)
\]
\[
\overset{a}{=} H(Y_1, \ldots, Y_n) - H(Z_1, \ldots, Z_n)
\]
\[
\overset{b}{=} H(Y_1, \ldots, Y_n) - (H(Z_1) + \sum_{i=1}^{n} H(Z_{i+1}|Z_i))
\]
\[
\overset{c}{\leq} (n - 1)(1 - h(q))
\]

Find the distribution on $X^n$, $p_{X_1, X_2, \ldots, X_n}$ that achieves the upper bound. This shows that the capacity of the channel with such a noise sequence is $1 - h(q)$.

**PROBLEM 2.** Consider two discrete memoryless channels. The input alphabet, output alphabet, transition probabilities and capacity of the $k$’th channel is given by $X_k$, $Y_k$, $p_k$ and $C_k$ respectively ($k = 1, 2$). The channels operate independently. A communication system has access to both channels, that is, the effective channel between the transmitter and receiver has input alphabet $X_1 \times X_2$, output alphabet $Y_1 \times Y_2$ and transition probabilities $p_1(y_1|x_1)p_2(y_2|x_2)$. Find the capacity of this channel.

**PROBLEM 3.** Show that a cascade of $n$ identical binary symmetric channels,

\[
X_0 \to \text{BSC } \#1 \to X_1 \to \cdots \to X_{n-1} \to \text{BSC } \#n \to X_n
\]
each with raw error probability $p$, is equivalent to a single BSC with error probability $\frac{1}{2}(1 - (1 - 2p)^n)$ and hence that $\lim_{n \to \infty} I(X_0; X_n) = 0$ if $p \neq 0, 1$. Thus, if no processing is allowed at the intermediate terminals, the capacity of the cascade tends to zero.

**PROBLEM 4.** Consider a memoryless channel with transition probability matrix $P_{Y|X}(y|x)$, with $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. For a distribution $Q$ over $\mathcal{X}$, let $I(Q)$ denote the mutual information between the input and the output of the channel when the input distribution is $Q$. Show that for any two distributions $Q$ and $Q'$ over $\mathcal{X}$,
(a) 
\[ I(Q') \leq \sum_{x \in \mathcal{X}} Q'(x) \sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) \log \left( \frac{P_{Y|X}(y|x)}{\sum_{x' \in \mathcal{X}} P_{Y|X}(y|x') Q(x')} \right) \]

(b) 
\[ C \leq \max_x \sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) \log \left( \frac{P_{Y|X}(y|x)}{\sum_{x' \in \mathcal{X}} P_{Y|X}(y|x') Q(x')} \right) \]

where \( C \) is the capacity of the channel. Notice that this upper bound to the capacity is independent of the maximizing distribution.