1. Weatherman
A clever student studied the weather record in Toronto and compared it with those predicted by a local TV station:

<table>
<thead>
<tr>
<th>Predict</th>
<th>Actual</th>
<th>Rain</th>
<th>No Rain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rain</td>
<td>1/8</td>
<td>3/16</td>
<td></td>
</tr>
<tr>
<td>No Rain</td>
<td>1/16</td>
<td>10/16</td>
<td></td>
</tr>
</tbody>
</table>

He discovered that the weatherman is right only 12/16 of the time, but could be right 13/16 of the time by always predicting no rain. The student explains the situation and applies for the weatherman’s job. But the boss of the TV station, who is an information theorist, turns him down. Why?

2. Channel Capacity
Consider the discrete memoryless channel \( Y = X + Z \) (mod 11), where

\[
Z = \begin{pmatrix}
1, & 2, & 3 \\
1/3, & 1/3, & 1/3
\end{pmatrix}
\]

and \( X \in \{0, 1, \ldots, 10\} \). Assume that \( Z \) is independent of \( X \).

(a) Find the capacity.
(b) What is the maximizing \( p^*(x) \)?

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

\[
Q = \begin{bmatrix}
1 & 0 \\
1/2 & 1/2
\end{bmatrix}
\]

\( x, y \in \{0, 1\} \)

Find the capacity of the Z-channel and the maximizing input probability distribution.
Find the channel capacity of the following discrete memoryless channel:

\[
\begin{array}{c}
Z \\
\downarrow \\
X + \\
\rightarrow \\
Y
\end{array}
\]

where \( \Pr\{Z = 0\} = \Pr\{Z = a\} = \frac{1}{2} \). The alphabet for \( x \) is \( X = \{0, 1\} \). Assume that \( Z \) is independent of \( X \). (Observe that the channel capacity depends on the value of \( a \).)

5. Cascade of Binary Symmetric Channels.
Show that a cascade of \( n \) identical binary symmetric channels,

\[
X_0 \rightarrow \text{BSC} \ #1 \rightarrow X_1 \rightarrow \cdots \rightarrow X_{n-1} \rightarrow \text{BSC} \ #n \rightarrow X_n
\]
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 \( p \neq 0, 1 \). No encoding or decoding takes place at the intermediate terminals \( X_1, \ldots, X_{n-1} \). Thus the capacity of the cascade tends to zero.

6. Using Two Channels At Once.
Consider two discrete memoryless channels \((X_1, p(y_1 | x_1), Y_1)\) and \((X_2, p(y_2 | x_2), Y_2)\) with capacities \( C_1 \) and \( C_2 \) respectively. A new channel \((X_1 \times X_2, p(y_1 | x_1) \times p(y_2 | x_2), Y_1 \times Y_2)\) is formed in which \( x_1 \in X_1 \) and \( x_2 \in X_2 \), are simultaneously sent, resulting in \( y_1, y_2 \). Find the capacity of this channel.

7. Using Two Channels One At A Time
Find the capacity \( C \) of the union of 2 channels \((X_1, p(y_1 | x_1), Y_1)\) and \((X_2, p(y_2 | x_2), Y_2)\) where, at each time, one can send a symbol over channel 1 or over channel 2 but not both. Assume the output alphabets are distinct and do not intersect.

\[
X \begin{cases} 
X_1 \rightarrow p_1(y_1 | x_1) \rightarrow Y_1 \\
X_2 \rightarrow p_2(y_2 | x_2) \rightarrow Y_2 
\end{cases}
\]

(a) Show \( 2^C = 2^{C_1} + 2^{C_2} \).
(b) One can consider \( 2^{C_1} \) as the effective alphabet size of a noiseless channel with capacity \( C_1 \). Interpret (a) in terms of the effective number of noise-free symbols. (Compare this with problem 19, chapter 2, of the textbook where \( 2^H = 2^{H_1} + 2^{H_2} \).
(c) Suppose \( C_1 > C_2 \), why is it not optimal to use channel 1 all the time?
8. Can Signal Alternatives Lower Capacity?
Show that adding a row to a channel transition matrix does not decrease capacity.

9. 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 \( X_i, Y_i \in \{0, 1\} \).
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 \( Z^n \) is independent
of the input \( X^n \). Let \( C = 1 - H(p, 1 - p) \). Show that
\[
\max_{p(x_1, x_2, \ldots, x_n)} \ I(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n) \geq nC.
\]

10. Time-Varying Channels.
Consider a time-varying discrete memoryless channel. Let \( Y_1, Y_2, \ldots, Y_n \) be conditionally
independent given \( X_1, X_2, \ldots, X_n \), with conditional distribution given by \( p(y_i | x_i) = \prod_{i=1}^{n} p_i(y_i | x_i) \).
\begin{center}
\begin{tikzpicture}
  \node (0) at (0,0) {0};
  \node (1) at (0,-1) {1};
  \node (2) at (1,0) {0};
  \node (3) at (1,-1) {1};
  \draw[->] (0) -- node[above]{$1-p_i$} (2);
  \draw[->] (0) -- node[below]{$p_i$} (3);
  \draw[->] (1) -- node[above]{$1-p_i$} (2);
  \draw[->] (1) -- node[below]{$p_i$} (3);
\end{tikzpicture}
\end{center}
Let \( X = (X_1, X_2, \ldots, X_n), Y = (Y_1, Y_2, \ldots, Y_n) \). Find \( \max_{p(x)} I(X; Y) \).