The following problems are from the NIT book (referenced by chapter and problem number, or by page number) and Cover and Thomas book (CT).

1. Problem 2.4 parts (a), (d), and (e).
3. Problem 2.14 parts (a) and (d).
5. Problem 3.5.
6. Problem 3.7.

7. Binary multiplier channel (CT).
   
   (a) Consider the channel $Y = XZ$, where $X$ and $Z$ are independent binary random variables that take on values 0 and 1, with $P\{Z = 1\} = \alpha$. Find the capacity of this channel and the maximizing distribution on $X$.
   
   (b) Now suppose that the receiver can observe $Z$ as well as $Y$. What is the capacity?

8. Cascade of binary symmetric channels (CT). Show that a cascade of $n$ identical independent binary symmetric channels, each with crossover probability $p$ with no encoding or decoding at the intermediate terminals $X_1, \ldots, X_{n-1}$, is equivalent to a single BSC with error probability $(1/2)(1 - (1 - 2p)^n)$.

$$X_0 \rightarrow \text{BSC} \rightarrow X_1 \rightarrow \cdots \rightarrow X_{n-1} \rightarrow \text{BSC} \rightarrow X_n$$

Hence $\lim_{n \to \infty} I(X_0; X_n) = 0$ as $p \neq 0, 1$, and the capacity of the cascade tends to zero as the number of stages approaches infinity.

Extra problems. Do not submit solutions.

1. Entropy of a function of a random variable. Let $X \sim p(x)$ and $g(x)$ be a function of $x \in \mathcal{X}$. Show that $H(g(X)) \leq H(X)$. Under what condition does equality hold? Interpret the result. (Hint: Use the data processing inequality.)

2. Inequalities (CT). Which of the following inequalities are generally $\geq, =, \leq$? Label each with $\geq, =, \text{or} \leq$.

   (a) $H(5X)$ vs. $H(X)$. 


(b) $I(g(X); Y)$ vs. $I(X; Y)$. Under what conditions does equality hold?
(c) $I(X; Z|Y)$ vs. $I(Z; Y|X) - I(Z; Y) + I(X; Z)$.
(d) $I(X; Y, Z)$ vs. $I(X; Z) + I(Y; Z)$ if $X, Y$ are independent and $Z$ is an arbitrary random variable.

3. **Conditional mutual information vs. unconditional mutual information (CT).** Give examples of joint random variables $X$, $Y$, and $Z$ such that
   (a) $I(X; Y|Z) < I(X; Y)$
   (b) $I(X; Y|Z) > I(X; Y)$

4. **Entropy of a disjoint mixture (CT).** Let $X_0$ and $X_1$ be discrete random variables drawn over the respective alphabets $X_0 = \{1, \ldots, m\}$ and $X_1 = \{m+1, \ldots, n\}$. Let $S$ be a binary random variable with $P\{S = 1\} = p_1$, $0 \leq p_1 \leq 1$. Let
   $$X = \begin{cases} 
   X_0 & \text{if } S = 0, \\
   X_1 & \text{otherwise}.
   \end{cases}$$
   (a) Find $H(X)$ in terms of $H(X_0)$, $H(X_1)$, and $p_1$.
   (b) Maximize over $p_1$ to show that
   $$2^{H(X)} \leq 2^{H(X_0)} + 2^{H(X_1)}.$$
   (c) Find $H(X|S)$ in terms of $H(X_0)$, $H(X_1)$, and $p_1$.

5. **Data processing (CT).** Let $X_1 \to X_2 \to X_3 \to \cdots \to X_n$ form a Markov chain in this order. Reduce $I(X_1; X_2, \ldots, X_n)$ to its simplest form.

6. **Typical average lemma and typicality.** Prove the typical average lemma and the four properties following it on p.26 of the NIT book.
7. **Joint typicality.** Prove properties 1(c), 3, and 4 on p.27 of the NIT book.
8. **Problem 3.2.**
9. **Erasures and errors in a binary channel (CT).** Consider a channel with binary inputs that has both erasures and errors (see figure below). Let the probability of error be $\epsilon$ and the probability of erasure be $\alpha$.
   (a) Find the capacity of this channel.
   (b) Specialize to the case of the binary symmetric channel ($\alpha = 0$).
   (c) Specialize to the case of the binary erasure channel ($\epsilon = 0$).