Solution of Problem 1

(Chain Rule)

Let $X_1, \ldots, X_n$ and $Y$ be discrete random variables. Proof that

a) $H(X_1, \ldots, X_n) = \sum_{i=1}^{n} H(X_i|X_{i-1}, \ldots, X_1)$.

By the two-variable expansion rule of entropies we have

\[
H(X_1, X_2) = H(X_1) + H(X_2|X_1)
\]
\[
H(X_1, X_2, X_3) = H(X_1) + H(X_2, X_3|X_1) = H(X_1) + H(X_2|X_1) + H(X_3|X_2, X_1)
\]
\[
\vdots
\]
\[
H(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} H(X_i|X_{i-1}, \ldots, X_1).
\]

□

b) $I(X_1, \ldots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y|X_{i-1}, \ldots, X_1)$.

\[
I(X_1, \ldots, X_n; Y) = H(X_1, \ldots, X_n) - H(X_1, \ldots, X_n|Y)
\]
\[
= \sum_{i=1}^{n} H(X_i|X_{i-1}, \ldots, X_1) - \sum_{i=1}^{n} H(X_i|X_{i-1}, \ldots, X_1, Y)
\]
\[
= \sum_{i=1}^{n} [H(X_i|X_{i-1}, \ldots, X_1) - H(X_i|X_{i-1}, \ldots, X_1, Y)]
\]
\[
= \sum_{i=1}^{n} I(X_i; Y|X_{i-1}, \ldots, X_1).
\]

□

Let $Z = \{(X_n, Y_n)\}_{n \in \mathbb{N}}$ be an i.i.d. sequence of pairs of discrete random variables.

c) Show that $I(X_1, \ldots, X_n; Y_1, \ldots, Y_n) = \sum_{i=1}^{n} I(X_i; Y_i)$.

\[
I(X_1, \ldots, X_n; Y_1, \ldots, Y_n) = \sum_{i=1}^{n} I(X_i; Y_1, \ldots, Y_n|X_{i-1}, \ldots, X_1)
\]
\[
= \sum_{i=1}^{n} I(X_i; Y_i|X_{i-1}, \ldots, X_1)
\]
\[
= \sum_{i=1}^{n} I(X_i; Y_i).
\]

□
Solution of Problem 2

*(Stationary Processes)*

Let \(\ldots, X_{-1}, X_0, X_1, \ldots\) be a stationary stochastic process. Which of the following are true? Prove or provide a counterexample.

a) \(H(X_n|X_0) = H(X_{-n}|X_0)\).

This is true since

\[
H(X_n|X_0) = H(X_n, X_0) - H(X_0)
\]

and \(H(X_n, X_0) = H(X_{-n}, X_0) - H(X_0)\) by stationarity.

b) \(H(X_n|X_0) \geq H(X_{n-1}|X_0)\).

This is not true in general. A simple counterexample would be a periodic process with period \(n\). Let \(X_0, \ldots, X_{n-1}\) be i.i.d. uniform random variables and let \(X_k = X_{k-n}\) for \(k \geq n\). In this case \(H(X_n|X_0) = 0\) and \(H(X_{n-1}|X_0) = \log|\mathcal{X}|\), contradicting the statement.

c) \(H(X_n|X_1, X_2, \ldots, X_{n-1}, X_{n+1})\) is non-increasing in \(n\).

This statement is true, since by stationarity

\[
H(X_n|X_1, X_2, \ldots, X_{n-1}, X_{n+1}) = H(X_{n+1}|X_2, X_3, \ldots, X_n, X_{n+2}) \\
\geq H(X_{n+1}|X_1, X_2, X_3, \ldots, X_n, X_{n+2}).
\]

d) \(H(X_n|X_1, \ldots, X_{n-1}, X_{n+1}, \ldots, X_{2n})\) is non-increasing in \(n\).

In the same manner, this statement is true since

\[
H(X_n|X_1, X_2, \ldots, X_{n-1}, X_{n+1}, \ldots, X_{2n}) = H(X_{n+1}|X_2, X_3, \ldots, X_n, X_{n+2}, \ldots, X_{2n+1}) \\
\geq H(X_{n+1}|X_1, X_2, X_3, \ldots, X_n, X_{n+2}, \ldots, X_{2n+1}).
\]

Solution of Problem 3

*(The past has little to say about the future)*

For a stationary stochastic process \(X_1, X_2, \ldots\), show that

\[
\lim_{n \to \infty} \frac{1}{2n} I(X_1, \ldots, X_n; X_{n+1}, \ldots, X_{2n}) = 0.
\]

Since \(X_1, X_2, \ldots\) is stationary we know that \(H(X_{n+1}, \ldots, X_{2n}) = H(X_1, \ldots, X_n)\), then

\[
I(X_1, \ldots, X_n; X_{n+1}, \ldots, X_{2n}) \\
= H(X_1, \ldots, X_n) + H(X_{n+1}, \ldots, X_{2n}) - H(X_1, \ldots, X_n, X_{n+1}, \ldots, X_{2n}) \\
= H(X_1, \ldots, X_n) + H(X_{n+1}, \ldots, X_{2n}) - H(X_1, \ldots, X_n, X_{n+1}, \ldots, X_{2n}) \\
= 2H(X_1, \ldots, X_n) - H(X_1, \ldots, X_n, X_{n+1}, \ldots, X_{2n}).
\]

Thus
\[
\lim_{n \to \infty} \frac{1}{2n} I(X_1, \ldots, X_n; X_{n+1}, \ldots, X_{2n}) \\
= \lim_{n \to \infty} \frac{1}{2n} 2H(X_1, \ldots, X_n) - \lim_{n \to \infty} \frac{1}{2n} H(X_1, \ldots, X_n, X_{n+1}, \ldots, X_{2n}) \\
= \lim_{n \to \infty} \frac{1}{n} H(X_1, \ldots, X_n) - \lim_{n \to \infty} \frac{1}{2n} H(X_1, \ldots, X_n, X_{n+1}, \ldots, X_{2n}) \\
= \lim_{n \to \infty} \frac{1}{n} H(X) - \lim_{n \to \infty} \frac{1}{2n} H(X) \\
= H_\infty(X) = 0. \quad \square
\]

**Solution of Problem 4**

*(Entropy Rate)*

Let \( X = \{X_n\}_{n \in \mathbb{Z}} \) be a stationary sequence of discrete random variables with entropy rate \( H_\infty(X) \).

a) Show that \( H_\infty(X) \leq H(X_1) \).

From Theorem 2.3.4 (d) we get

\[
H_\infty(X) = \lim_{n \to \infty} \frac{1}{n} H(X_n|X_0, \ldots, X_{n-1}) \\
= \lim_{n \to \infty} \frac{1}{n} H(X_1|X_{-(n-1)}, \ldots, X_{-1}, X_0) \\
\leq \lim_{n \to \infty} H(X_1) = H(X_1). \quad \square
\]

b) What are the conditions for equality?

We have equality only if \( X_1 \) is independent of the past \( X_0, X_{-1}, \ldots \), i.e., if and only if \( X_i \) is an i.i.d. process.