20. **Run length coding.** Let $X_1, X_2, \ldots, X_n$ be (possibly dependent) binary random variables. Suppose one calculates the run lengths $\mathbf{R} = (R_1, R_2, \ldots)$ of this sequence (in order as they occur). For example, the sequence $\mathbf{X} = 0001100100$ yields run lengths $\mathbf{R} = (3, 2, 2, 1, 2)$. Compare $H(X_1, X_2, \ldots, X_n)$, $H(\mathbf{R})$ and $H(X_n, \mathbf{R})$. Show all equalities and inequalities, and bound all the differences.

Solution: *Run length coding.* Since the run lengths are a function of $X_1, X_2, \ldots, X_n$, $H(\mathbf{R}) \leq H(\mathbf{X})$. Any X_i together with the run lengths determine the entire sequence

$X_1, X_2, \ldots, X_n$. Hence

$$\begin{aligned} H(X_1, X_2, \ldots, X_n) &= H(X_i, \mathbf{R}) && (2.36) \\ &= H(\mathbf{R}) + H(X_i|\mathbf{R}) && (2.37) \\ &\leq H(\mathbf{R}) + H(X_i) && (2.38) \\ &\leq H(\mathbf{R}) + 1. && (2.39) \end{aligned}$$

34. **Entropy of initial conditions.** Prove that $H(X_0|X_n)$ is non-decreasing with n for any Markov chain.

Solution: *Entropy of initial conditions.* For a Markov chain, by the data processing theorem, we have

$$I(X_0; X_{n-1}) \geq I(X_0; X_n). \tag{2.68}$$

Therefore

$$H(X_0) - H(X_0|X_{n-1}) \geq H(X_0) - H(X_0|X_n) \tag{2.69}$$

or $H(X_0|X_n)$ increases with n.

11. **Stationary processes.** Let $\ldots, X_{-1}, X_0, X_1, \ldots$ be a stationary (not necessarily Markov) stochastic process. Which of the following statements are true? Prove or provide a counterexample.

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

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

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

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

Solution: *Stationary processes.*

(a) $H(X_n|X_0) = H(X_{-n}|X_0)$.
This statement is true, since

$$H(X_n|X_0) = H(X_n, X_0) - H(X_0) \tag{4.76}$$

$$H(X_{-n}|X_0) = H(X_{-n}, X_0) - H(X_0) \tag{4.77}$$

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

(b) $H(X_n|X_0) \geq H(X_{n-1}|X_0)$.
This statement is not true in general, though it is true for first order Markov chains. A simple counterexample is a periodic process with period n . Let $X_0, X_1, X_2, \ldots, X_{n-1}$ be i.i.d. uniformly distributed binary random variables and let $X_k = X_{k-n}$ for $k \geq n$. In this case, $H(X_n|X_0) = 0$ and $H(X_{n-1}|X_0) = 1$, contradicting the statement $H(X_n|X_0) \geq H(X_{n-1}|X_0)$.

(c) $H(X_n|X_1^{n-1}, X_{n+1})$ is non-increasing in n .
This statement is true, since by stationarity $H(X_n|X_1^{n-1}, X_{n+1}) = H(X_{n+1}|X_2^n, X_{n+2}) \geq H(X_{n+1}|X_1^n, X_{n+2})$ where the inequality follows from the fact that conditioning reduces entropy.