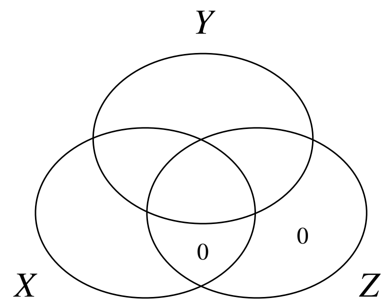
1. (3 pts) Find a necessary and sufficient condition for each of the following, where $g(\cdot)$ denotes a function.

   (a) (3 pts) $H(X|Y) = H(X|g(Y))$. Hint: Let $Z = g(Y)$.

   (b) (3 pts) $H(X|Y) = H(g(X)|Y)$.

Solution:

(a)

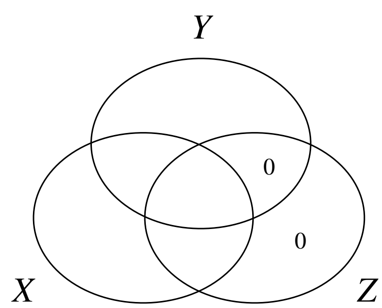


Let $Z = g(Y)$. Then $H(Z|Y) = 0$, which implies $H(Z|X, Y) = 0$ and $I(X; Z|Y) = 0$. From the information diagram, we have

$$\begin{aligned} H(X|Y) &= H(X|Y, Z) \\ H(X|Z) &= H(X|Y, Z) + I(X; Y|Z). \end{aligned}$$

Thus $H(X|Y) = H(X|Z)$ if and only if $I(X; Y|Z) = 0$, or $X \to Z \to Y$.

(b)



Let $Z = g(X)$. Then $H(Z|X) = 0$, which implies $H(Z|X, Y) = 0$ and $I(Y; Z|X) = 0$. From the information diagram, we have

$$\begin{aligned} H(X|Y) &= H(X|Y, Z) + I(X; Z|Y) \\ H(Z|Y) &= I(X; Z|Y). \end{aligned}$$

Thus $H(X|Y) = H(Z|Y)$ if and only if $H(X|Y,Z) = 0$.

2. Let X, Y, Z, and T be discrete random variables. Consider the conditions

   C1: $X \to Y \to Z \to T$.

   C2: $X \to Y \to Z$ and $Y \to Z \to T$.

   Answer each of the following questions and explain.

   (a) (1 pts) Does C1 imply C2?

   (b) (3 pts) Does C2 imply C1? Hint: C1 implies $I(X;T|Y,Z) = 0$.

   Solution:

   (a) C1 implies C2 because the two Markov chains in C2 are subchains of the Markov chain in C1.

   (b) Based on the hint, let $X = T = U$ and $Y = Z =$ constant, where U is some discrete random variable such that $H(U) > 0$. Obviously, C2 but not C1 is satisfied. Hence C2 does not imply C1.

3. Let n be the length of a sequence $\mathbf{x}$. The empirical distribution $q_{\mathbf{x}}$ of the sequence $\mathbf{x}$ is also called the *type* of $\mathbf{x}$.

   (a) (3 pts) Assuming that $\mathcal{X}$ is finite, the total number of distinct types $q_{\mathbf{x}}$ is given by

   $$T(n) = \binom{n + |\mathcal{X}| - 1}{n}.$$

   Show that $T(n)$ is upper bounded by $(n+1)^{|\mathcal{X}|-1}$.

   (b) (2 pts) Show that

   $$\lim_{n\to\infty} \frac{1}{n} \log T(n) = 0.$$

Solution:

(a) Consider

$$\begin{aligned}
T(n) &= \binom{n+|\mathcal{X}|-1}{n} \\
&= \frac{(n+|\mathcal{X}|-1)!}{n!\,(|\mathcal{X}|-1)!} \\
&= \frac{(n+1)(n+2)\cdots(n+|\mathcal{X}|-1)}{(1)(2)\cdots(|\mathcal{X}|-1)} \\
&= (n+1)\left(\frac{n}{2}+1\right)\cdots\left(\frac{n}{|\mathcal{X}|-1}+1\right) \\
&\le (n+1)^{|\mathcal{X}|-1}.
\end{aligned}$$

(b) Consider

$$\begin{aligned}
0 &\le \lim_{n\to\infty}\frac{1}{n}\log T(n) \\
&\le \lim_{n\to\infty}\frac{1}{n}\log(n+1)^{|\mathcal{X}|-1} \\
&= (|\mathcal{X}|-1)\lim_{n\to\infty}\frac{1}{n}\log(n+1) \\
&= 0,
\end{aligned}$$

where the limit can be evaluated by L' Hôpital's rule. Thus

$$\lim_{n\to\infty}\frac{1}{n}\log T(n) = 0.$$

4. (a) (3 pts) For any integer $n \ge 2$, let $n = a_n + b_n$, where

$$a_n = 2^{\lceil \log_2 \frac{n}{2} \rceil} \quad \text{and} \quad b_n = n - a_n.$$

Determine a_n and b_n for $n = 2, 3, 4, 5$.

(b) (**bonus** 3 pts) Show that $a_n \le n-1$ and $b_n \le n-1$.

(c) For every integer $n \ge 2$, let

$$f(n) = l_1(n), l_2(n), \ldots, l_n(n)$$

be a list of n positive integers defined recursively as follows:

$$f(2) = 1, 1$$

and for $n \geq 3$,

$$f(n) = \begin{cases} f(a_n)+1, 1 & \text{if } b_n = 1 \\ f(a_n)+1, f(b_n)+1 & \text{if } b_n > 1 \end{cases}$$

where for a list s, $s+1$ denotes the list obtained from s by adding 1 to each element in s. For example, if $s = 1, 3, 2$, then $s+1 = 2, 4, 3$.

i. (0 pt) Familiarize yourself with the notation by verifying that

$$f(n) = \begin{cases} l_1(a_n)+1, \cdots, l_{a_n}(a_n)+1, 1 & \text{if } b_n = 1 \\ l_1(a_n)+1, \cdots, l_{a_n}(a_n)+1, l_1(b_n)+1, \cdots, l_{b_n}(b_n)+1 & \text{if } b_n > 1 \end{cases}$$

You do not have to show your work.

ii. (3 pts) Determine $f(n)$ for $n = 3, 4, 5$.

(d) (4 pts) Show by induction that for all $n \geq 2$, $f(n)$ satisfies the Kraft inequality with equality, i.e.,

$$\sum_{i=1}^{n} 2^{-l_i(n)} = 1.$$

Hint: Use Part (b).

(e) (2 pts) Let an information source with uniform distribution on an alphabet of n symbols be encoded by a binary prefix code with codeword lengths specified by $f(n)$, where $n \geq 2$. Let $L(n)$ be the expected length of this prefix code. Show that

$$\sum_{i=1}^{n} l_i(n) = nL(n).$$

(f) (3 pts) Prove the following recursive formula for $L(n)$ for $n \geq 3$:

$$L(n) = \begin{cases} \frac{1}{n}[a_n L(a_n) + a_n + 1] & \text{if } b_n = 1 \\ \\ \frac{1}{n}[a_n L(a_n) + b_n L(b_n) + a_n + b_n] & \text{if } b_n > 1. \end{cases}$$

Solution:

(a)

$$\begin{aligned} a_2 &= 1; \quad b_2 = 1 \\ a_3 &= 2; \quad b_3 = 1 \\ a_4 &= 2; \quad b_4 = 2 \\ a_5 &= 4; \quad b_5 = 1 \end{aligned}$$

(b) To show that $a_n \le n-1$, consider

$$a_n = 2^{\lceil \log_2 \frac{n}{2} \rceil} < 2^{\log_2 \frac{n}{2}+1} = 2^{\log_2 n} = n,$$

where we have used $\lceil x \rceil < x+1$ for any real number x. Then $a_n < n$ which implies $a_n \le 1$ because a_n is an integer.
To show that $b_n \le n-1$, consider $a_2 = 1$ and a_n increasing in n, so that $a_n \ge 1$ for all $n \ge 2$. Therefore,

$$b_n = n - a_n \le n - 1.$$

(c) ii.

$$\begin{aligned} f(3) &= 1+f(2), 1 \\ &= 2, 2, 1 \\ f(4) &= 1+f(2), 1+f(2) \\ &= 2, 2, 2, 2 \\ f(5) &= 1+f(4), 1 \\ &= 3, 3, 3, 3, 1 \end{aligned}$$

(d) For $n = 2$, $f(2) = 1, 1$, and we check that

$$2^{-1} + 2^{-1} = 1.$$

Thus the Kraft inequality is satisfied with equality. We claim that for all $n \ge 2$, $f(n)$ satisfies the Kraft inequality with equality. Assume that the claim is true for all $2 \le n \le m-1$ for some $m \ge 3$. We now show that the claim is true for m. If $b_m = 1$, consider

$$\sum_{i=1}^{m} 2^{-l_i(m)} = \sum_{i=1}^{a_m} 2^{-(l_i(a_m)+1)} + 2^{-1}.$$

By Part (b), $a_m \le m-1$, so by the induction hypothesis, we have

$$\sum_{i=1}^{a_m} 2^{-(l_i(a_m)+1)} = 2^{-1} \sum_{i=1}^{a_m} 2^{-l_i(a_m)} = 2^{-1} \cdot 1 = 2^{-1}.$$

Therefore,

$$\sum_{i=1}^{m} 2^{-l_i(m)} = 2^{-1} + 2^{-1} = 1.$$

If $b_m > 1$, consider

$$\begin{aligned}
\sum_{i=1}^{m} 2^{-l_i(m)} &= \sum_{i=1}^{a_m} 2^{-(l_i(a_m)+1)} + \sum_{i=1}^{b_m} 2^{-(l_i(b_m)+1)} \\
&= 2^{-1} \sum_{i=1}^{a_m} 2^{-l_i(a_m)} + 2^{-1} \sum_{i=1}^{b_m} 2^{-l_i(b_m)} \\
&= 2^{-1} \cdot 1 + 2^{-1} \cdot 1 \\
&= 1,
\end{aligned}$$

where in the 3rd equality above, we have invoked the induction hypothesis in light of $b_m \leq m-1$ from Part (b).

(e) Assume the uniform distribution over the alphabet of n symbols. For every $n \geq 2$, we have

$$L(n) = \sum_{i=1}^{n} n^{-1} \cdot l_i(n) = n^{-1} \sum_{i=1}^{n} l_i(n),$$

or

$$\sum_{i=1}^{n} l_i(n) = nL(n). \tag{1}$$

(f) We now prove the recursive formula for $L(n)$ for $n \geq 3$. If $b_n = 1$, we have

$$\begin{aligned}
L(n) &= n^{-1} \sum_{i=1}^{n} l_i(n) \\
&= \frac{1}{n} \left[\sum_{i=1}^{a_n} (l_i(a_n) + 1) + 1 \right] \\
&= \frac{1}{n} \left[\sum_{i=1}^{a_n} l_i(a_n) + a_n + 1 \right] \\
&= \frac{1}{n} \left[a_n L(a_n) + a_n + 1 \right],
\end{aligned}$$

where we have used (1). If $b_n > 1$, we have

$$\begin{aligned}
L(n) &= n^{-1} \sum_{i=1}^{n} l_i(n) \\
&= \frac{1}{n} \left[\sum_{i=1}^{a_n} (l_i(a_n) + 1) + \sum_{i=1}^{b_n} (l_i(b_n) + 1) \right] \\
&= \frac{1}{n} \left[a_n L(a_n) + b_n L(b_n) + a_n + b_n \right].
\end{aligned}$$