EE/Ma 126a Information Theory - Midterm Exam

Ling Li, ling@cs.caltech.edu

October 25, 2000

M.1 Unique Decodability

(a) Since the binary description of symbol b is 000 and it is a prefix to 0001, the binary description of symbol c, so $\boxed{\text{this code is not instantaneous}}$.

(b) $\boxed{\text{This code is non-singular}}$ since no two codewords are the same.

(c) Kraft inequality for uniquely decodable codes (McMillan) says the codeword lengths of any uniquely decodable code must satisfy

$$\sum_i D^{-l_i} \le 1.$$

In this problem, $D = 2$, $\{l_i\} = \{2, 3, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5\}$. So,

$$\sum_i D^{-l_i} = 2^{-2} + 2^{-3} + 7 \times 2^{-4} + 8 \times 2^{-5} = \frac{17}{16} > 1.$$

Thus $\boxed{\text{this code is not uniquely decodable}}$.

M.2 Best Worst-Case Performance

$$D(p||q) = p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}.$$

(a) $D(p||q) \ge 0$ with equality iff $p = q$. So

$$\boxed{\min_p D(p||q) = 0}$$

and the value of p that minimizes $D(p||q)$ is $\boxed{p = q}$.

(b) While $0 < q < 1$ and q is fixed, $D(p||q)$ is a function of p. For $0 < p < 1$,

$$\begin{aligned} \frac{dD(p||q)}{dp} &= \log \frac{p}{1-p} - \log \frac{q}{1-q}, \\ \frac{d^2 D(p||q)}{dp^2} &= \frac{1}{p(1-p)} > 0. \end{aligned}$$

So $D(p||q)$ is convex in p. Thus

$$\max_p D(p||q)$$

can only be obtained at $p = 0$ or $p = 1$, i.e.,

$$\begin{aligned}\max_p D(p||q) &= \max\{D(p=0||q), D(p=1||q)\} \\ &= \max\{-\log(1-q), -\log q\}.\end{aligned}$$

Therefore

$$\boxed{\max_p D(p||q) = \begin{cases} -\log q, & \text{when } 0 < q \le \frac{1}{2}; \\ -\log(1-q), & \text{when } \frac{1}{2} \le q < 1. \end{cases}}$$

Note that when $q = \frac{1}{2}$, $-\log q = -\log(1-q)$, $\max_p D(p||q)$ can be obtained both at $p = 0$ and $p = 1$. So the value of p that maximizes $D(p||q)$ is

$$\boxed{p = \begin{cases} 1, & \text{when } 0 < q < \frac{1}{2}; \\ 0, 1, & \text{when } q = \frac{1}{2}; \\ 0, & \text{when } \frac{1}{2} < q < 1. \end{cases}}$$

(c) When $0 < q \le \frac{1}{2}$, $\max_p D(p||q) = -\log q$, which decreases when q increases ($\frac{d(-\log q)}{dq} = -\frac{1}{q} < 0$); When $\frac{1}{2} \le q < 1$, $\max_p D(p||q) = -\log(1-q)$, which increases when q increases ($\frac{d(-\log(1-q))}{dq} = \frac{1}{1-q} > 0$). So when $q = \frac{1}{2}$, $\max_p D(p||q)$ gets its minimum, which is

$$\boxed{\min_q \max_p D(p||q) = 1}.$$

M.3 Drunken Walk

Let X_i $(i = 1..n)$ denote the displacement of ith step that the sailor makes. $X_i \in \mathcal{X} = \{2, 1, -4\}$. In this problem, $X_1, X_2, \ldots, X_n$ are i.i.d.$\sim p(x)$, where $p(2) = \frac{1}{4}$, $p(1) = \frac{1}{2}$, and $p(-4) = \frac{1}{4}$.

After n steps, the sailor is $|\sum_{i=1}^n X_i|$ feet away from the center of town. The probability wanted is

$$\Pr\left(\left|\sum_{i=1}^n X_i\right| > n\right).$$

The Chebyshev's inequality states that

$$\Pr\left(\left|\sum_{i=1}^n X_i - n\mu\right| \ge \varepsilon\right) \le \frac{n\sigma^2}{\varepsilon^2},$$

where μ and σ^2 are the mean and variance of X_i. In this problem,

$$\begin{aligned}\mu &= \frac{1}{4} \times 2 + \frac{1}{2} \times 1 + \frac{1}{4} \times (-4) = 0, \\ \sigma^2 &= \frac{1}{4} \times 2^2 + \frac{1}{2} \times 1^2 + \frac{1}{4} \times 4^2 = 5.5.\end{aligned}$$

Let $\varepsilon = n+1$. Since X_i are integers, we have

$$\Pr\left(\left|\sum_{i=1}^{n} X_i\right| > n\right) = \Pr\left(\left|\sum_{i=1}^{n} X_i - n\mu\right| \ge n+1\right) \le \frac{5.5n}{(n+1)^2}.$$

So the upper bound of the probability that after n steps the sailor is more than n feet away from the center of town is $\boxed{\dfrac{5.5n}{(n+1)^2}}$.

M.4 Mixture Distributions

$X_1, X_2, \ldots, X_n$ are independent, so

$$\begin{aligned} H(X_1, X_2, \ldots, X_n) &= H(X_1) + H(X_2) + \cdots + H(X_n) \\ &= -\sum_{i=1}^{n} \sum_{x \in \mathcal{X}} p_i(x) \log p_i(x) \\ &= -\sum_{x \in \mathcal{X}} \left(\sum_{i=1}^{n} p_i(x) \log p_i(x)\right). \end{aligned}$$

$Y_1, Y_2, \ldots, Y_n$ are i.i.d., so

$$\begin{aligned} H(Y_1, Y_2, \ldots, Y_n) &= nH(Y_1) \\ &= -n \sum_{x \in \mathcal{X}} q(x) \log q(x) \\ &= -\sum_{x \in \mathcal{X}} \left(\left(\sum_{i=1}^{n} p_i(x)\right) \log \frac{\sum_{i=1}^{n} p_i(x)}{n}\right). \end{aligned}$$

Using $a_i = p_i(x)$ and $b_i = 1$ in the log sum inequality, we have

$$\sum_{i=1}^{n} p_i(x) \log p_i(x) \ge \left(\sum_{i=1}^{n} p_i(x)\right) \log \frac{\sum_{i=1}^{n} p_i(x)}{n}$$

with equality iff $p_1(x) = p_2(x) = \cdots = p_n(x)$. Thus,

$$\boxed{H(X_1, X_2, \ldots, X_n) \le H(Y_1, Y_2, \ldots, Y_n)}$$

with equality iff for all $x \in \mathcal{X}$, $p_1(x) = p_2(x) = \cdots = p_n(x)$, i.e., $\boxed{X_1, X_2, \ldots, X_n \text{ are i.i.d.}}$.

M.5 Jointly Distributed Random Variables

(a) Since relative mutual information is non-negative,

$$I(X, X+Z; Y) = I(X+Z; Y) + I(X; Y|X+Z) \ge I(X+Z; Y),$$

namely,

$$H(X, X+Z) - H(X, X+Z|Y) \ge H(X+Z) - H(X+Z|Y).$$

Thus

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

(b) Since $(A+B)$ is a function of A and B, and A is a function of $(A+B)$ and B, we have

$$H(A+B|A,B)=H(A|A+B,B)=0.$$

From

$$\begin{aligned}
H(A,A+B|B) &= H(A|B)+H(A+B|A,B)\\
&= H(A+B|B)+H(A|A+B,B)
\end{aligned}$$

we get

$$H(A+B|B)=H(A|B).$$

(c) Similarly with (b),[1] we have (C is any random variable)

$$H(A+B,C|B)=H(A,C|B).$$

Using this equation twice,

$$H(X+Y,X+Y+Z|Y)=H(X,X+Y+Z|Y)=H(X,X+Z|Y).$$

So

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

(d) Z is independent of both X and Y, so when $X+Y$ is given, $X+Y+Z$ is independent of Y. That is $Y\to X+Y\to X+Y+Z$ is a Markov chain, or

$$I(Y;X+Y+Z|X+Y)=0.$$

Thus

$$\begin{aligned}
I(X+Y;X+Y+Z)-I(X+Y;X+Y+Z|Y) &= I(X+Y;X+Y+Z;Y)\\
&= I(Y;X+Y+Z)-I(Y;X+Y+Z|X+Y)\\
&= I(Y;X+Y+Z)\geq 0.
\end{aligned}$$

That is

$$I(X+Y;X+Y+Z|Y)\leq I(X+Y;X+Y+Z).$$

[1] $H(A,A+B,C|B)=H(A,C|B)+H(A+B|A,B,C)=H(A+B,C|B)+H(A|A+B,B,C)$.

(e) From (a), (c) and (d),

$$I(X;X+Z) - I(X;Y) \le I(X+Y;X+Y+Z).$$

With the joint distribution $p(X+Y, X+Y+Z)$,

$X+Y+Z$ \ $X+Y$	0	1	$p(X+Y+Z)$
0	1/3	0	1/3
1	1/6	1/3	1/2
2	0	1/6	1/6
$p(X+Y)$	1/2	1/2	

$$\begin{aligned} I(X+Y;X+Y+Z) &= \frac{1}{3}\log\frac{\frac{1}{3}}{\frac{1}{3}\times\frac{1}{2}} + \frac{1}{6}\log\frac{\frac{1}{6}}{\frac{1}{2}\times\frac{1}{2}} + \frac{1}{3}\log\frac{\frac{1}{3}}{\frac{1}{2}\times\frac{1}{2}} + \frac{1}{6}\log\frac{\frac{1}{6}}{\frac{1}{6}\times\frac{1}{2}} \\ &= \frac{4}{3} - \frac{1}{2}\log 3. \end{aligned}$$

So $I(X;X+Z) - I(X;Y) \le \frac{4}{3} - \frac{\log 3}{2}$. □