S-72.340 INFORMATION THEORY

Exercise 5, October 18, 2000 Solutions

1. At first, we notice that the joint distribution can be factored as

$$P(x_1, \dots, x_n) = P(x_1)P(x_2|x_1)P(x_3|x_2, x_1) \cdots P(x_n|x_{n-1}, \dots, x_1) \,.$$

By definition, the joint entropy of these n random variables is given by

$$\begin{aligned}
H(X_1, \dots, X_n) &= -\sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \cdots \sum_{j_n=1}^{N_n} P(x_{j_1}, x_{j_2}, \dots, x_{j_n}) \log P(x_{j_1}, x_{j_2}, \dots, x_{j_n}) \\
&= -\sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \cdots \sum_{j_n=1}^{N_n} P(x_{j_1}, x_{j_2}, \dots, x_{j_n}) \log P(x_{j_1}) \\
&\quad -\sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \cdots \sum_{j_n=1}^{N_n} P(x_{j_1}, x_{j_2}, \dots, x_{j_n}) \log P(x_{j_2}|x_{j_1}) - \cdots \\
&\quad -\sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \cdots \sum_{j_n=1}^{N_n} P(x_{j_1}, x_{j_2}, \dots, x_{j_n}) \log P(x_{j_n}|x_{j_{n-1}}, \dots, x_{j_1}) \\
&= -\sum_{j_1=1}^{N_1} P(x_{j_1}) \log P(x_{j_1}) \\
&\quad -\sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} P(x_{j_1}, x_{j_2}) \log P(x_{j_2}|x_{j_1}) \\
&\quad - \cdots - \sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \cdots \sum_{j_n=1}^{N_n} P(x_{j_1}, x_{j_2}, \dots, x_{j_n}) \log P(x_{j_n}|x_{j_{n-1}}, \dots, x_{j_1}) \\
&= H(X_1) + H(X_2|X_1) + \cdots + H(X_n|X_{n-1}, \dots, X_1) \\
&= \sum_{i=1}^{n} H(X_i|X_{i-1}, \dots, X_1) \,.
\end{aligned}$$

2. The conditional average mutual information $I(X;Y|Z)$ is defined as

$$I(X;Y|Z) = H(X|Z) - H(X|Y,Z) = \sum_{x,y,z} P(x,y,z) \log \frac{P(x,y|z)}{P(x|z)P(y|z)} \,.$$

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

So, the inequality appearing in the problem seems to be an equality after all.

3. Only sequences with an even number of 1's are possible. Thus, there are 2^{n-1} 'admissible' sequences. By the definition,

$$
\begin{aligned}
I(X_1;X_2) &= \sum_{x_1,x_2} P(x_1,x_2)\log\frac{P(x_1|x_2)}{P(x_1)} \\
&= P(0,0)\log\frac{P(0|0)}{P(0)} + P(0,1)\log\frac{P(0|1)}{P(0)} \\
&+ P(1,0)\log\frac{P(1|0)}{P(1)} + P(1,1)\log\frac{P(1|1)}{P(1)}\,.
\end{aligned}
$$

If $n > 2$, we see that the knowledge of X_2 doesn't give any new information about X_1, and we get

$$
I(X_1;X_2) = 4\cdot\frac{1}{4}\log\frac{1/2}{1/2} = 0\,.
$$

But, if $n = 2$,

$$
I(X_1;X_2) = \frac{1}{2}\log\frac{1}{1/2} + 0 + 0 + \frac{1}{2}\log\frac{1}{1/2} = \log 2 = 1\,\text{bit}.
$$

In the same way, for $1 \le k < n-2$:

$$
\begin{aligned}
I(X_{k+1};X_{k+2}|X_1,\ldots,X_k) &= \sum_{x_1,\ldots,x_{k+2}} P(x_1,\ldots,x_{k+2})\log\frac{P(x_{k+1},x_{k+2}|x_1,\ldots,x_k)}{P(x_{k+1}|x_1,\ldots,x_k)P(x_{k+2}|x_1,\ldots,x_k)} \\
&= \sum_{x_1,\ldots,x_{k+2}} P(x_1,\ldots,x_{k+2})\cdot\log\frac{1/4}{1/4} = 0\,.
\end{aligned}
$$

But,

$$
\begin{aligned}
I(X_{n-1};X_n|X_1,\ldots,X_{n-2}) &= \sum_{\text{admissible}} P(x_1,\ldots,x_n)\log\frac{1/2}{1/4} \\
&= 2^{n-1}\cdot\frac{1}{2^{n-1}}\cdot\log 2 = 1\,\text{bit}.
\end{aligned}
$$

4. It is easy to see that

$$P(x|y) = \begin{cases} 1-p & \text{when } x = y \\ \frac{p}{M-1} & \text{when } x \neq y\,. \end{cases}$$

(You can check this by using the Bayes theorem.) The average mutual information is given by

$$\begin{aligned}
I(X;Y) &= \sum_{x,y} P(x,y) \log \frac{P(x|y)}{P(x)} \\
&= P(0,0) \log \frac{P(0|0)}{P(0)} + P(0,1) \log \frac{P(0|1)}{P(0)} \\
&+ P(0,2) \log \frac{P(0|2)}{P(0)} + \cdots \\
&+ P(M-1, M-1) \log \frac{P(M-1|M-1)}{P(M-1)} \\
&= \frac{1}{M} \cdot \left[M(1-p) \log \frac{1-p}{1/M} + (M^2 - M) \frac{p}{M-1} \log \frac{p/(M-1)}{1/M} \right] \\
&= (1-p) \log[M(1-p)] + p \log \left[\frac{M}{M-1} p \right] \\
&= p \log p + (1-p) \log(1-p) + \log M - p \log(M-1)\,.
\end{aligned}$$

5. The non-zero conditional probabilities $P(x|y)$ are given by

$$\begin{aligned}
&P(0|0) = P(0|1) = P(1|1) = P(1|2) = P(2|0) = P(2|2) = \frac{1}{2} \\
&P(0|3) = P(1|3) = P(2|3) = \frac{1}{3}\,.
\end{aligned}$$

$$\begin{aligned}
I(X;Y) &= \sum_{x,y} P(x,y) \log \frac{P(x|y)}{P(x)} \\
&= \frac{1}{9} \left(6 \log \frac{1/2}{1/3} + 3 \log \frac{1/3}{1/3} \right) = \frac{2}{3} \log \frac{3}{2}\,.
\end{aligned}$$