1. (数据处理不等式)$H(Y \mid X) \leqslant H(Y \mid f(X))$

   解析：

   存在马尔可夫链

   $$Y \to X \to f(X)$$

   由数据处理不等式知

   $$I(X;Y) \geqslant I(f(X);Y)$$

   即

   $$H(Y) - H(Y \mid X) \geqslant H(Y) - H(Y \mid f(X))$$

   于是有

   $$H(Y \mid X) \leqslant H(Y \mid f(X))$$

2. (平稳分布)$\{X_i\}_{i=-\infty}^{+\infty}$ 为平稳过程，证明

   $$H\left(X_0 \mid X_{-1}, X_{-2}, \cdots, X_{-n}\right) = H\left(X_0 \mid X_1, X_2, \cdots, X_n\right)$$

   解析：

   $$\begin{aligned} H\left(X_0 \mid X_{-1}, \ldots, X_{-n}\right) &= H\left(X_0, X_{-1}, \ldots, X_{-n}\right) - H\left(X_{-1}, \ldots, X_{-n}\right) \\ &= H\left(X_0, X_1, X_2, \ldots, X_n\right) - H\left(X_1, X_2, \ldots, X_n\right) \\ &= H\left(X_0 \mid X_1, X_2, \ldots, X_n\right) \end{aligned}$$

3. (平稳的马尔可夫链) 设 $\{X_n\}$ 为平稳的马尔可夫链。则存在 $k = (\quad)$ 满足

   $$H\left(X_{-n} \mid X_0, X_1\right) = H\left(X_k \mid X_0, X_1\right) \quad (n > 1)$$

   解析：

   由马尔可夫性可知

   $$H\left(X_{-n} \mid X_0, X_1\right) = H\left(X_{-n} \mid X_1\right)$$

   由平稳知

   $$H\left(X_{-n} \mid X_1\right) = H\left(X_n \mid X_0\right)$$

   由马尔可夫性可知

   $$H\left(X_n \mid X_0\right) = H\left(X_n \mid X_0, X_{-1}\right)$$

   由平稳知

   $$H\left(X_n \mid X_0, X_{-1}\right) = H\left(X_{n+1} \mid X_0, X_1\right)$$

所以 $k=1$。

4. (哈夫曼编码) 下列哪一组码字长度可以成为三元哈夫曼码的码字长度?

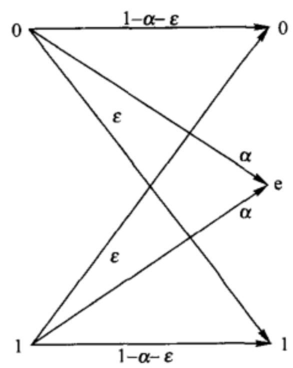
   (a) $(1,2,2,2,2)$

   (b) $(2,2,2,2,2,2,2,2,2,3,3,3)$

   解析：

   (b) 可以。判断方法：使得 Kraft 不等式取等号。另外，如果给的是一组码让判断是否是哈夫曼编码，则需要判断两点。一个是码长是否恰满足 Kraft 不等式，一个是是否是前缀码。

5. (擦除 + 出错信道) 考虑一个既有擦除又有出错的二元输入信道。设出错的概率为 ε，擦除概率为 α。信道模型如图所示：



   求解信道容量。

   解析：

$$
\begin{aligned}
C &= \max_{p(x)} I(X;Y) \\
&= \max_{p(x)} (H(Y) - H(Y \mid X)) \\
&= \max_{p(x)} H(Y) - H(1-\varepsilon-\alpha, \alpha, \varepsilon)
\end{aligned}
$$

   设输入概率分布为 $(\pi, 1-\pi)$。于是有

$$
\begin{aligned}
H(Y) &= H(\pi(1-\varepsilon-\alpha)+(1-\pi)\varepsilon, \alpha, (1-\pi)(1-\varepsilon-\alpha)+\pi\varepsilon) \\
&= H(\alpha) + (1-\alpha) H\left(\frac{\pi+\varepsilon-\pi\alpha-2\pi\varepsilon}{1-\alpha}, \frac{1-\pi-\varepsilon+2\varepsilon\pi-\alpha+\alpha\pi}{1-\alpha}\right) \\
&\leqslant H(\alpha) + (1-\alpha)
\end{aligned}
$$

   等号成立当且仅当

$$\frac{\pi+\varepsilon-\pi\alpha-2\pi\varepsilon}{1-\alpha}=\frac{1}{2}$$

即 $\pi=\dfrac{1}{2}$。于是有

$$\begin{aligned}
C &= H(\alpha)+1-\alpha-H(1-\alpha-\varepsilon,\alpha,\varepsilon)\\
&= H(\alpha)+1-\alpha-H(\alpha)-(1-\alpha)H\left(\frac{1-\alpha-\varepsilon}{1-\alpha},\frac{\varepsilon}{1-\alpha}\right)\\
&= (1-\alpha)\left(1-H\left(\frac{1-\alpha-\varepsilon}{1-\alpha},\frac{\varepsilon}{1-\alpha}\right)\right)
\end{aligned}$$

6. 添加一行到信道转移矩阵不会降低容量。

$$\begin{aligned}
C_{m+1} &= \max_{p(x_1,\ldots,x_{m+1})} I(X;Y)\\
&\geqslant \max_{p(x_1,\ldots,x_m,0)} I(X;Y)\\
&= C_m.
\end{aligned}$$

7. 假定信道 $\mathcal{P}$ 的容量为 C, 其中 $\mathcal{P}$ 表示一个 $m\times n$ 的信道矩阵。计算信道

$$\tilde{\mathcal{P}}=\begin{bmatrix}\mathcal{P} & 0\\ 0 & 1\end{bmatrix}$$

的容量。

解析：实际上这是一个并联信道模型，利用结论即可得到

$$\tilde{C}=\log\left(2^0+2^C\right)=\log\left(1+2^C\right)$$

8. (信道容量) 胶片是由碘酸银晶体按照泊松分布组成，每平方英寸的粒子密度 λ 已知。在不知道碘化银粒子位置的情况下对该胶片随机地进行光照。当碘酸银粒子感光 (接受光照) 后，其所在的区域被点亮，没有碘酸银粒子的区域和有碘酸银粒子但没有感光的区域未被点亮。我们做如下的假设，将胶片的区域划分为若干小格子，大小为 dA。假设每个格子中至多一个碘酸银粒子并且不在格子的边界上。于是，胶片可以看作是一系列具有交叉概率 $1-\lambda dA$ 的并联二元非对称信道。通过必要的近似，计算该胶片的信道容量 (量纲为比特/平方英寸)。
解析：

本题的关键在于理解题意并构建信道模型。对于每一个小块，其输入 X 为是否接受光照，记 $X=0$ 表示未接受光照，$X=1$ 表示接受光照；其输出 Y 为

是否被点亮，记 $Y=0$ 表示未被点亮，$Y=1$ 表示被点亮。设 $P(X=0)=\alpha$，$\beta=\lambda dA$。由题意可得信道转移概率矩阵

$$p(Y \mid X)=\begin{bmatrix} 1 & 0 \\ 1-\beta & \beta \end{bmatrix}$$

则有

$$H(Y \mid X)=P(X=0)H(0)+P(X=1)H(\beta)=\alpha H(\beta)$$

和

$$H(Y)=H(P(Y=1))=H(\alpha\beta)$$

于是可得

$$I(X;Y)=H(Y)-H(Y \mid X)=H(\alpha\beta)-\alpha H(\beta)$$

将互信息作为 α 的函数并对 α 求导来求解最大值。为了求导方便，熵使用奈特的量纲。求导可得

$$\frac{d}{d\alpha}I(X;Y)=\beta\ln\frac{1-\alpha\beta}{\alpha\beta}-H_e(\beta)$$

令导数为 0 可得

$$\alpha=\frac{1}{\beta}\frac{1}{1+e^{\frac{H_e(\beta)}{\beta}}}$$

记

$$\gamma=\alpha\beta=\frac{1}{1+e^{\frac{H_e(\beta)}{\beta}}}$$

可得

$$\bar{\gamma}=1-\gamma=\frac{e^{\frac{H_e(\beta)}{\beta}}}{1+e^{\frac{H_e(\beta)}{\beta}}}=\gamma e^{\frac{H_e(\beta)}{\beta}}$$

和

$$\ln\bar{\gamma}-\ln\gamma=\frac{H_e(\beta)}{\beta}$$

所以有

$$
\begin{aligned}
I(X;Y) &= H_e(\alpha\beta) - \alpha H_e(\beta) \\
&= H_e(\gamma) - \frac{1}{1+e^{\frac{H_e(\beta)}{\beta}}}\frac{H_e(\beta)}{\beta} \\
&= -\gamma\ln\gamma - \bar{\gamma}\ln\bar{\gamma} - \gamma(\ln\bar{\gamma} - \ln\gamma) \\
&= -\ln\bar{\gamma} \\
&= \ln\left(1+e^{-\frac{H_e(\beta)}{\beta}}\right)
\end{aligned}
$$

进而可得

$$
C = \lim_{dA\to 0}\frac{I(X;Y)}{dA} = \lim_{dA\to 0}\frac{e^{-\frac{H(\beta)}{\beta}}}{dA} = \lim_{dA\to 0}\frac{\beta}{e\cdot dA} = \frac{\lambda}{e}\text{ 奈特/平方英寸} = \log_2 e\cdot\frac{\lambda}{e}\text{ 比特/平方英寸}
$$