1

(a)

true:

$$I(X;Y) = H(X) - H(X|Y)$$

If $I(X;Y) = 0$ then $H(X) = H(X|Y)$. We can write:

$$I(X;Y) = D(\ P_{x,y}(x,y)\ ||\ P_x(x)P_y(y)\ ) = 0$$

$D(Q||P) = 0$ iff $P_x(x) = Q_x(x)\ \forall x$, therefore $P_{x,y}(x,y) = P_x(x)P_y(y)$ for every x, y and as result $X \perp Y$.

(b)

true:

$$X - Y - Z \ \Rightarrow\ I(X;Y) \geq I(X;Z)$$

As result:

$$H(X) - H(X|Y) \geq H(X) - H(X|Z) \Rightarrow H(X|Y) \leq H(X|Z)$$

(c)

true:
Using the concave property of the divergence function:

$$D\left(\lambda P + (1-\lambda)Q\ ||\ Q\right) \leq \lambda D(P\ ||\ Q) + (1-\lambda)D(Q\ ||\ Q)$$

Assigning $\lambda = \frac{1}{2}$, and since $D(Q||Q) = 0$:

$$D\left(\frac{1}{2}P + \frac{1}{2}Q\ ||\ Q\right) \leq \frac{1}{2}D(P||Q)$$

(d)

true:

$$H(X+Y) \geq H(X+Y|Y) \overset{(a)}{=} H(X)$$

(a) - since X is independent of Y.

(e)

true:

$$\begin{aligned} I(X;Y) - I(X;Y|Z) &= H(X) - H(X|Y) - [H(X|Z) - H(X|Y,Z)] \\ &= \underbrace{H(X) - H(X|Z)}_{I(X;Z)} - \underbrace{[H(X|Y) - H(X|Y,Z)]}_{\geq 0} \\ &\leq I(X;Z) \\ &= H(Z) - \underbrace{H(Z|X)}_{\geq 0} \\ &\leq H(Z) \end{aligned}$$

(f)

false:
We know that $\frac{1}{n}\log|A_n| \geq H(X) - \varepsilon$ for n sufficiently large (*theorem 3.3.1* in the text book and as proved in class). Since $\lim_{n\to\infty} \Pr(A_n) = 1$ and $\lim_{n\to\infty} \Pr(B_n) = 1$ we can say that also $\lim_{n\to\infty} \Pr(A_n \cap B_n) = 1$ (it was also shown in class) and therefore:

$$\lim_{n\to\infty} \frac{1}{n} \log|A_n \cap B_n| \geq H(X) - \varepsilon$$

But since ε is as small as we like, we cannot say that:

$$\lim_{n\to\infty} \frac{1}{n} \log|A_n \cap B_n| < H(X)$$

(g)

false:
Assuming that the file is already optimally compressed, it cannot be compressed any further. Also, if the entropy rate of the bits in the file is 1 for some reason, it cannot be compressed.

For example, if the bits in the file are *Bernoulli*$(\frac{1}{2})$ distributed, the file cannot be compressed anymore.

(h)

true:
It has shown in class that $R_X \leq H(X) + 1$, $R_{X|Y} \leq H(X|Y) + 1$ and therefore:

$$R_X - R_{X|Y} \leq I(X;Y) + 1$$

(i)

true:
If $X \perp Y$ then $p(x) = p(x|y)$ and $R_X = R_{X|Y}$.

(j)

true:
Increasing the distortion allows rate reduction.

(k)

true:

$$\log|\hat{\mathcal{X}}| \overset{(a)}{\geq} H(\hat{X}) \geq H(\hat{X}) - H(\hat{X}|X) = I(\hat{X};X) \overset{(b)}{\geq} R(D)$$

(a) - equality if $\hat{X}$ is equally distributed.
(b) - equality if $p(\hat{x}|x)$ brings the mutual information into minimum under distortion constraint

2

(a)

Huffman code:

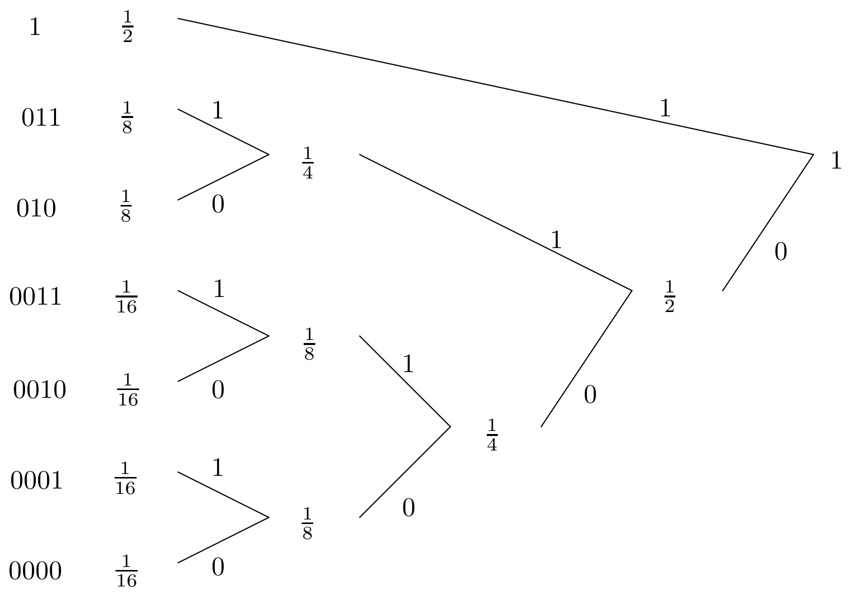

Figure 1: Huffman

(b)

Huffman code is optimal code and achieves the entropy for dyadic distribution. If the distribution of the digits is not $Bernoulli(\frac{1}{2})$ you can compress it further. The binary digits of the data would be equally distributed after applying the Huffman code and therefore $p_0 = p_1 = \frac{1}{2}$.

The expected length would be:

$$E[l] = \frac{1}{2} \cdot 1 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 3 + \frac{1}{16} \cdot 4 + \frac{1}{16} \cdot 4 + \frac{1}{16} \cdot 4 + \frac{1}{16} \cdot 4 = 2.25$$

Therefore, the expected length of 1000 symbols would be 2250 bits.

3

(a)

We can use the solution of the home work:

$$C = \log\left(2^{C_1} + 2^{C_2} + 2^{C_3}\right)$$

Now we need to calculate the capacity of each channel:

$$C_1 = \max_{p(x)} I(X;Y) = H(Y) - H(Y|X) = 0 - 0 = 0$$

$$C_2 = \max_{p(x)} I(X;Y) = H(Y) - H(Y|X) = 1 - 1 = 0$$

$$\begin{aligned} C_3 = \max_{p(x)} I(X;Y) &= \max_{p(x)} \{H(Y) - H(Y|X)\} \\ &= \max_{p(x)} \left[-\frac{1}{2}p_2 \log\left(\frac{1}{2}p_2\right) - \left(\frac{1}{2}p_2 + p_3\right) \log\left(\frac{1}{2}p_2 + p_3\right)\right] - p_2 \end{aligned}$$

Assigning $p_3 = 1 - p_2$ and derive against p_2:

$$\frac{dI(X;Y)}{dp_2} = -\frac{p_2}{2} \cdot \frac{1}{2} \cdot \frac{1}{\frac{p_2}{2}} - \frac{1}{2}\log\left(\frac{p_2}{2}\right) + \frac{2-p_2}{2} \cdot \frac{1}{2} \cdot \frac{1}{\frac{2-p_2}{2}} + \frac{1}{2}\log\left(\frac{2-p_2}{2}\right) - 1 = 0$$

And as result $p_2 = \frac{2}{5}$:

$$C_3 \approx 0.322$$

And, finally:

$$C = \log(2^0 + 2^0 + 2^{0.322}) \approx 1.7$$

(b)

(b)

Encoding: You just use ternary representation of the message and send using 0,1,2 but no 3 (or 0,1,3 but no 2) of the input channel.
Decoding: map the ternary output into the message.

4

(a)+(b)

We can simplify those two schemes to a system in which:

$$Z_i \sim N(0, \sigma^2)$$

Now we can write that:

$$I(X;Y) = h(Y) - h(Y|X)$$

Where:

$$h(Y) \leq \frac{1}{2}\log(2\pi e E[Y^2])$$

And:

$$\begin{aligned} E[Y^2] &= E\left[\left(\sum_{i=1}^{K}(X+Z_i)\right)^2\right] \\ &\overset{(a)}{=} K^2E[X^2] + KE[Z_i^2] \\ &\leq K^2P + K\sigma^2 \end{aligned}$$

(a) - Z_i i.i.d

As a result we have:

$$h(Y) \leq \frac{1}{2}\log[2\pi e(K^2P + K\sigma^2)]$$

And the conditional entropy would be (since Y is sum of K independent Gaussian noises):

$$h(Y|X) = \frac{1}{2}\log(2\pi eK\sigma^2)$$

Therefore:

$$\begin{aligned} I(X;Y) &\leq \frac{1}{2}\log[2\pi e(K^2P + K\sigma^2)] - \frac{1}{2}\log(2\pi eK\sigma^2) \\ &= \frac{1}{2}\log\left(\frac{K^2P + K\sigma^2}{K\sigma^2}\right) \\ &= \frac{1}{2}\log\left(1 + \frac{KP}{\sigma^2}\right) \end{aligned}$$

And the capacity would be:

$$C = \frac{1}{2}\log\left(1 + \frac{KP}{\sigma^2}\right)$$

(c)

This time:

$$\begin{aligned} E[Y^2] &= E\left[(X+Z)^2\right] \\ &= E[X^2] + E[Z_i^2] \\ &\leq KP + \sigma^2 \end{aligned}$$

And:

$$h(Y) \leq \frac{1}{2}\log[2\pi e(KP + \sigma^2)]$$

$$h(Y|X) = \frac{1}{2}\log(2\pi e \sigma^2)$$

As a result:

$$I(X;Y) \leq \frac{1}{2}\log\left(1 + \frac{KP}{\sigma^2}\right)$$

And the capacity would be:

$$C = \frac{1}{2}\log\left(1 + \frac{KP}{\sigma^2}\right)$$

It seems that spatial diversity and time diversity are just like increasing the transmitted signal power.