CIE 6020 Assignment 4

WU, Chenhao 117010285

April 14, 2019

2. Assume that $p_X(1) = \lambda$, and then we have

$$p_Y(0) = -\lambda + \lambda p + 1$$
$$p_Y(1) = \lambda - \lambda p$$

$$H(Y|X) = -p\lambda \log p - (1-p)\lambda \log(1-p)$$
$$= p\lambda \log \frac{1-p}{p} - \lambda \log(1-p)$$
$$H(Y) = -(1-p)\lambda \log \lambda(1-p) - (1-(1-p)\lambda)\log(1-(1-p)\lambda)$$
$$= (1-p)\lambda \log \frac{1-(1-p)\lambda}{(1-p)\lambda} - \log(1-(1-p)\lambda)$$
$$I(X;Y) = H(Y) - H(Y|X)$$
$$= \lambda \log(1-p)\frac{1-(1-p)\lambda}{(1-p)\lambda} - p\lambda \frac{1-p}{p}\frac{1-(1-p)\lambda}{(1-p)\lambda} - \log(1-(1-p)\lambda)$$
$$= \lambda \log \frac{1-(1-p)\lambda}{\lambda} - p\lambda \log \frac{1-(1-p)\lambda}{p\lambda} - \log(1-(1-p)\lambda)$$

To find the λ that gives the maximum $I(X;Y)$, we obtain the derivative of mutual information as

$$\frac{d}{d\lambda}I(X;Y) = -\frac{((p-1)\lambda + 1)\left(p\log\left(\frac{(p-1)\lambda+1}{p\lambda}\right) - \log\left(\frac{(p-1)\lambda+1}{\lambda}\right)\right) - 2p + 2}{(p-1)\lambda + 1}$$
$$= (p-1)\log \frac{(p-1)\lambda + 1}{p\lambda} + \frac{2(p-1)}{(p-1)\lambda + 1}$$

Let the derivative equal to zero, we have

$$\log \frac{p\lambda}{(p-1)\lambda+1} = \frac{2}{(p-1)\lambda+1}$$

to maximize the mutual information of input and output.

When input probability distribution satisfies the equation above, the channel capacity can be achieved as

3.

$$\begin{aligned}
I(X;Y) &= H(Y) - H(Y|X) \\
&= H(Y) - H(Z) \\
&= H(Y) - \log 3 \\
&\leq \log 11 - \log 3
\end{aligned}$$

where the equality holds when Y is an uniform distribution, which implies the input probability distribution X is also an uniform distribution.

4. (a) *Proof:* Assume that $p_X(0) = \lambda$. For the random distribution Y we have

$$\begin{aligned}
p_Y(0) &= p_U(1)p_X(1) + p_U(0)p_X(0) = q(1-\lambda) + (1-q)\lambda \\
p_Y(1) &= p_U(0)p_X(1) + p_U(1)p_X(0) = (1-q)(1-\lambda) + q\lambda
\end{aligned}$$

Since U_i and X_i are independent for any integer $n > 0$, $Y_i(U_i, X_i)$ is also independent, and therefore, the channel Y is a memoryless binary symmetric channel with $\epsilon = q$. And the channel capacity of this symmetric channel is $C = 1 - H(q)$.

(b) *Proof:* The joint distribution of Y_1 and Y_2 is given by

$$\begin{aligned}
p(y_1, y_2|x_1, x_2) &= p(y_1|x_1, x_2)p(y_2|x_1, x_2, y_1) && \text{conditional probability} \\
&= p(u_1|x_1)p(u_2|x_1, x_2, u_1) && \text{definition of noise channel} \\
&= p(u_1|x_1) && \\
&= p(u_1) && \text{independency of U and X}
\end{aligned}$$

2

and at the same time,

$$p(y_1|x_1)p(y_2|x_2) = p(u_1|x_1)p(u_2|x_2) \qquad \text{definition of noise channel}$$
$$= p(u_1)p(u_1)$$
$$= p(u_1)^2$$
$$\neq p(y_1, y_2|x_1, x_2)$$

The output of the channel is not independent from previous outputs, thus it is not a memoryless channel by definition.

(c) Under the condition of (b), we can substitute the alphabet of channel input and channel output by $\mathcal{X}' = \{00, 01, 10, 11\}$ and $\mathcal{Y}' = \{00, 01, 10, 11\}$. By this way, the channel can be equivalent to a DMC by combining two consecutive uses of the channel.

The transition matrix $W(y'|x')$ of the channel is

$$\begin{bmatrix} 1-q & 0 & 0 & q \\ 0 & 1-q & q & 0 \\ 0 & q & 1-q & 0 \\ q & 0 & 0 & 1-q \end{bmatrix}$$

where the capacity of the channel is given by

$$I(X'; Y') = H(Y') - H(Y'|X')$$
$$= H(Y') - \sum_{x \in \mathcal{X}'} p(x)H(Y|x)$$
$$= H(Y') - H(q)$$
$$\leq 1 - H(q)$$

(d) Assume that we have a set of capacity achieving codes for the memoryless BSC under the condition of (a), without losing the generality, we can expand it to the channel under the condition of (b). Notice that, $\mathbf{X}_1 = \{00, 11\}$ can only output $\mathbf{Y}_1 = \{00, 11\}$, and $\mathbf{X}_2 = \{01, 10\}$ can only output $\mathbf{Y}_2 = \{01, 10\}$. With this understanding, we can apply the capacity achieving codes onto the input sets $\mathbf{X}_1$ and $\mathbf{X}_2$ respectively, and the combination of the codes can also achieve the channel capacity under the condition of (b).

5.