Chapter 4 : Solutions to the exercises

Exercise 4.1 Solution

$$H(X) = 2.4087 \text{ bits/symbol.}$$

In the case where the symbols are equally probable

$$H(X) = \log_2(6) = 2.585 \text{ bits/symbol.}$$

Exercise 4.2 Solution

a) Defining $\alpha = p(1-\varepsilon) + (1-p)\varepsilon$

$$H(X) = H_b(p) = -p\log_2(p) - (1-p)\log_2(1-p).$$

$$H(Y) = H_b(\alpha) = -\alpha\log_2(\alpha) - (1-\alpha)\log_2(1-\alpha).$$

$$H(Y|X) = H_b(\varepsilon).$$

$$H(X,Y) = H_b(p) + H_b(\varepsilon).$$

$$H(X|Y) = H_b(p) + H_b(\varepsilon) - H_b(\alpha)$$

$$I(X,Y) = H_b(\alpha) - H_b(\varepsilon).$$

b) $p = \frac{1}{2}$

c) $\varepsilon = \frac{1}{2}$

Exercise 4.3 Solution

a) $H(X) = 2.5282$ bits/symbol. For equiprobable symbols

$$H(X) = \log_2(7) = 2.8074 \text{ bits/symbol.}$$

b) $H(Y) = 1.4060$ bits/symbol. For equiprobable symbols

$$H(Y) = \log_2(3) = 1.585 \text{ bits/symbol}$$

c)

$$H(X,Y) = 2.5282 \text{ bits/symbol.}$$

$$H(X|Y) = 1.1222 \text{ bits/symbol.}$$

$$H(Y|X) = 0 \text{ bits/symbol.}$$

$$I(X,Y) = 1.4060 \text{ bits/symbol.}$$

Exercise 4.4 Solution

The channel capacity is

$$C = (1-\varepsilon) \text{ bits/usage},$$

and that this capacity is achieved for some a priori probabilities

$$p_X(0) = p_X(1) = \frac{1}{2}.$$

The capacity is represented as a function of ε in Figure 4.1.

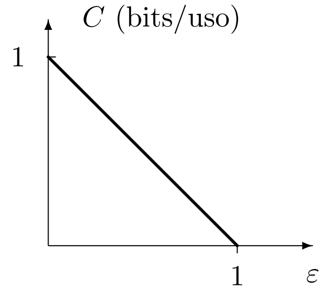


Figure 4.1: Channel capacity as a function of ε, Exercise 4.4.

Exercise 4.5 Solution

a) $C = \varepsilon$, for $p_X(x_0) = p_X(x_1) = \frac{1}{2}$.

b)

$$H(X,Y) = H_b(p) + H_b(\varepsilon)$$

$$H(X|Y) = (1-\varepsilon)H_b(p)$$

Exercise 4.6 Solution

a) For $\varepsilon = 0$

$$C = \log_2(3) = 1.585 \text{ bits/use}.$$

The input probabilities for which the channel capacity is reached are those that satisfy

$$p_X(x_1) = p_X(x_3) = \frac{1}{3}, \quad [p_X(x_o) + p_X(x_2)] = \frac{1}{3}.$$

For $\varepsilon = 1$

$$C = \log_2(2) = 1 \text{ bit/use}.$$

The input probabilities for which the channel capacity is reached are those that satisfy

$$[p_X(x_0) + p_X(x_2)] = \frac{1}{2}, \quad [p_X(x_1) + p_X(x_3)] = \frac{1}{2}.$$

b) $\varepsilon = 0.75$

$$H(X|Y) = 0.931 \text{ bits/symbol}.$$

c) The probabilities for which the channel capacity is obtained are

$$p_X(x_0) = 0.403, \quad p_X(x_1) = 0.339, \quad p_X(x_2) = 0.258.$$

The channel capacity is the maximum of the mutual information

$$C = H(Y) - H(Y|X) = 1.308 \text{ bits/use}$$

Exercise 4.7 Solution

a) The channel capacity is

$$C = \log_2(3) = 1.585 \text{ bits/use}$$

for

$$p_X(x_0) = p_X(x_1) = p_X(x_2) = \frac{1}{3}.$$

b) Any of these two values is possible

$$\varepsilon = \begin{cases} 0.25 \\ 0.75 \end{cases}.$$

$$H(Y) = \begin{cases} 1.5546, & \varepsilon = 0.25 \\ 1.2807, & \varepsilon = 0.75 \end{cases}.$$

$$I(X,Y) = \begin{cases} 1.0137, & \varepsilon = 0.25 \\ 0.7398, & \varepsilon = 0.75 \end{cases}.$$

$$H(X|Y) = H(X,Y) - H(Y) = \begin{cases} 0.5712, & \varepsilon = 0.25 \\ 0.8451, & \varepsilon = 0.75 \end{cases}.$$

Exercise 4.8 Solution

a) The channel capacity is $C = 1$ bit/usage and is reached for

$$p_X(x_1) = 1/2, \quad p_X(x_0) + p_X(x_2) = 1/2.$$

b) $H(X|Y) = 0$ and $H(X,Y) = H_b(p) + p\, H_b(\varepsilon)$. For $\varepsilon = \frac{1}{2}$, $H(X,Y) = H_b(p) + p$.

Exercise 4.9 Solution

a) $C = 1$ bit/use, for $p_X(x_0) = \frac{1}{2}$ and $p_X(x_1) + p_X(x_2) = \frac{1}{2}$.

b)

$$H(X,Y) = 2p + H_b(2p) + (1-2p)H_b(\varepsilon)$$

$$H(X|Y) = 2p$$

Exercise 4.10 Solution

a)

$$\mathbf{P}^{DMC} = \begin{bmatrix} 1-\varepsilon & \varepsilon & 0 & 0 \\ \varepsilon & 1-2\varepsilon & \varepsilon & 0 \\ 0 & \varepsilon & 1-2\varepsilon & \varepsilon \\ 0 & 0 & \varepsilon & 1-\varepsilon \end{bmatrix}$$

$$\mathbf{P}^{Const} = \begin{bmatrix} 1-Q\left(\frac{1}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right)-Q\left(\frac{3}{\sqrt{N_0/2}}\right) & Q\left(\frac{3}{\sqrt{N_0/2}}\right)-Q\left(\frac{5}{\sqrt{N_0/2}}\right) & Q\left(\frac{5}{\sqrt{N_0/2}}\right) \\ Q\left(\frac{1}{\sqrt{N_0/2}}\right) & 1-2Q\left(\frac{1}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right)-Q\left(\frac{3}{\sqrt{N_0/2}}\right) & Q\left(\frac{3}{\sqrt{N_0/2}}\right) \\ Q\left(\frac{3}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right)-Q\left(\frac{3}{\sqrt{N_0/2}}\right) & 1-2Q\left(\frac{1}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right) \\ Q\left(\frac{5}{\sqrt{N_0/2}}\right) & Q\left(\frac{3}{\sqrt{N_0/2}}\right)-Q\left(\frac{5}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right)-Q\left(\frac{3}{\sqrt{N_0/2}}\right) & 1-Q\left(\frac{1}{\sqrt{N_0/2}}\right) \end{bmatrix}$$

The approximation that has been made is that only the probability of error between symbols that are at a minimum distance is considered, neglecting the probability of making errors with symbols at $2d_{min}$ or $3d_{min}$. That means

$$Q\left(\frac{k}{\sqrt{N_0/2}}\right) \approx 0, \text{ for } k > 1,$$

and

$$\varepsilon = Q\left(\frac{1}{\sqrt{N_0/2}}\right).$$

b)

$$H(Y|X) = \frac{1}{2}H_b(\varepsilon) + \frac{1}{2}H_b(2\varepsilon) + \varepsilon$$

$$H(X|Y) = \frac{1}{2}H_b(\varepsilon) + \frac{1}{2}H_b(2\varepsilon) + \varepsilon$$

$$H(X,Y) = 2 + \frac{1}{2}H_b(\varepsilon) + \frac{1}{2}H_b(2\varepsilon) + \varepsilon$$

$$I(X,Y) = 2 - \frac{1}{2}H_b(\varepsilon) - \frac{1}{2}H_b(2\varepsilon) - \varepsilon$$

The value of ε that minimizes $H(Y|X)$ and $H(X|Y)$ and maximizes $I(X,Y)$ is $\varepsilon = 0$ (in which case $H(X|Y) = H(Y|X) = 0$, and $I(X,Y) = 2$).

c) The minimum and maximum values that ε can take are $\varepsilon = 0$ and $\varepsilon = \frac{1}{2}$, respectively.

i) For $\varepsilon = 0$, $C = 2$ bits/use for $p_X(x_i) = \frac{1}{4}$, $i = 0, 1, 2, 3$.

ii) For $\varepsilon = \frac{1}{2}$, $C = 1$ bits/use for $p_X(x_0) = p_X(x_3)$ and $p_X(x_1) = p_X(x_2)$.

Exercise 4.11 Solution

a) The channel matrix for the DMC is

$$\mathbf{P}^{DMC} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-\varepsilon & \varepsilon & 0 \\ 0 & \varepsilon & 1-\varepsilon & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

And for the constellation

$$\mathbf{P}^{Const.} = \begin{bmatrix} 1-Q\left(\frac{2}{\sqrt{N_0/2}}\right) & Q\left(\frac{2}{\sqrt{N_0/2}}\right)-Q\left(\frac{5}{\sqrt{N_0/2}}\right) & Q\left(\frac{5}{\sqrt{N_0/2}}\right)-Q\left(\frac{8}{\sqrt{N_0/2}}\right) & Q\left(\frac{8}{\sqrt{N_0/2}}\right) \\ Q\left(\frac{2}{\sqrt{N_0/2}}\right) & 1-Q\left(\frac{1}{\sqrt{N_0/2}}\right)-Q\left(\frac{2}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right)-Q\left(\frac{4}{\sqrt{N_0/2}}\right) & Q\left(\frac{4}{\sqrt{N_0/2}}\right) \\ Q\left(\frac{4}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right)-Q\left(\frac{4}{\sqrt{N_0/2}}\right) & 1-Q\left(\frac{1}{\sqrt{N_0/2}}\right)-Q\left(\frac{2}{\sqrt{N_0/2}}\right) & Q\left(\frac{2}{\sqrt{N_0/2}}\right) \\ Q\left(\frac{8}{\sqrt{N_0/2}}\right) & Q\left(\frac{5}{\sqrt{N_0/2}}\right)-Q\left(\frac{8}{\sqrt{N_0/2}}\right) & Q\left(\frac{2}{\sqrt{N_0/2}}\right)-Q\left(\frac{5}{\sqrt{N_0/2}}\right) & 1-Q\left(\frac{2}{\sqrt{N_0/2}}\right) \end{bmatrix}.$$

The approximation that has been made is that the probability of transition with symbols further away from distance 2 is zero, or what is the same, that errors only occur between the symbols $\mathbf{a}_1$ and $\mathbf{a}_2$, and with this assumption

$$Q\left(\frac{k}{\sqrt{N_0/2}}\right) \approx 0, \text{ for } k > 1.$$

Therefore

$$\varepsilon = Q\left(\frac{1}{\sqrt{N_0/2}}\right).$$

b)

$$H(X,Y) = 2 + \frac{1}{2}\, H_b(\varepsilon)$$

$$H(Y|X) = \frac{1}{2}\, H_b(\varepsilon).$$

$$H(X|Y) = \frac{1}{2}\, H_b(\varepsilon).$$

$$I(X,Y) = 2 - \frac{1}{2}\, H_b(\varepsilon).$$

The representation of these functions is shown in Figure 4.2.

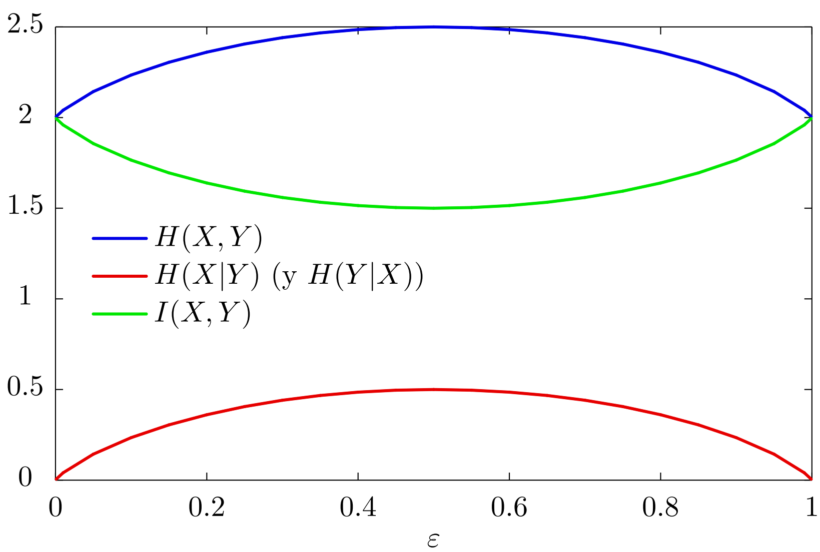


Figure 4.2: Plot of $H(X,Y)$, $H(X|Y)$ (equal to $H(Y|X)$) and $I(X,Y)$ as a function of ε.

The maximum values of $H(X,Y)$, $H(X|Y)$, and $H(Y|X)$ are obtained for $\varepsilon = 1/2$, the same as the minimum value of $I(X,Y)$.

c) The channel capacity is

$$C = 1 + H_b\left(\frac{2^{H_b(\varepsilon)}}{1+2^{H_b(\varepsilon)}}\right) - \left(\frac{1}{1+2^{H_b(\varepsilon)}}\right)\ H_b(\varepsilon),$$

value obtained for

$$p_X(x_0) = p_X(x_3) = \frac{2^{H_b(\varepsilon)-1}}{1+2^{H_b(\varepsilon)}},\ p_X(x_1) = p_X(x_2) = \frac{1}{2\ (1+2^{H_b(\varepsilon)})}.$$

Exercise 4.12 Solution

1. $a = d = 0$, $b = c = \varepsilon_1$.

$$H(Y) = 2 + \frac{1}{2}\left[(1-\varepsilon_0)\log_2\frac{1}{1-\varepsilon_0} + (1+\varepsilon_0)\log_2\frac{1}{1+\varepsilon_0}\right].$$

If we take into account that $\frac{1}{2}(1+\varepsilon_0) = 1 - \frac{1}{2}(1-\varepsilon_1)$, we could regroup terms and write

$$H(Y) = 1 + H_b\left(\frac{1}{2}(1-\varepsilon_0)\right).$$

$$H(X,Y) = 2 + \frac{1}{2}H_b(\varepsilon_0) + \frac{1}{2}H_b(\varepsilon_1)$$

$$H(Y|X) = \frac{1}{2}H_b(\varepsilon_0) + \frac{1}{2}H_b(\varepsilon_1)$$

$$H(X|Y) = 1 + \frac{1}{2}H_b(\varepsilon_0) + \frac{1}{2}H_b(\varepsilon_1) - H_b\left(\frac{1}{2}(1-\varepsilon_0)\right)$$

$$I(X,Y) = 1 + H_b\left(\frac{1}{2}(1-\varepsilon_0)\right) - \frac{1}{2}H_b(\varepsilon_0) - \frac{1}{2}H_b(\varepsilon_1)$$

2. The channel capacity is

$$C = 1 + H_b\left(p(1-\varepsilon_0)\right) - pH_b(\varepsilon_0) - (1-p)H_b(\varepsilon_1) \text{ bits/usage}$$

for

$$p = \frac{1}{(1-\varepsilon_0)\ \left[1 + 2^{\frac{H_b(\varepsilon_0)-H_b(\varepsilon_1)}{1-\varepsilon_0}}\right]}.$$

The distribution that allow to achieve the capacity is

$$p_X(x_0) = p_X(x_3) = \frac{p}{2},\ p_X(x_1) = p_X(x_2) = \frac{1-p}{2}$$

Exercise 4.13 Solution

a)

$$H(Y) = \frac{2}{3}(1-\varepsilon_0+\varepsilon_1) + H_b\left(\frac{2}{3}(1-\varepsilon_0+\varepsilon_1)\right)$$

$$H(Y|X) = \frac{2}{3}H_b(\varepsilon_0) + \frac{1}{3}H_b(2\varepsilon_1) + \frac{2}{3}\varepsilon_1.$$

$$H(X,Y) = \log_2(3) + \frac{2}{3}H_b(\varepsilon_0) + \frac{1}{3}H_b(2\varepsilon_1) + \frac{2}{3}\varepsilon_1.$$

$$H(X|Y) = H(X,Y) - H(Y)$$

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

b) In this case the channel is the *binary erasure channel* (BEC)

$$C = (1-\varepsilon_0) \text{ bits/symbol, for } p_X(x_0) = p_X(x_2) = \frac{1}{2}.$$

c)

$$C = p(1-\varepsilon_0) + (1-p)2\varepsilon_1 + H_b\left(p(1-\varepsilon_0) + (1-p)2\varepsilon_1\right) - pH_b(\varepsilon_0) - (1-p)\left(H_b(2\varepsilon_1) + 2\varepsilon_1\right)$$

for the value

$$p = \frac{(1-2\varepsilon_1)\, 2^{\frac{H_b(\varepsilon_0)-H_b(2\varepsilon_1)-2\varepsilon_1}{2\varepsilon_1+\varepsilon_0-1}} - \varepsilon_1}{(1-2\varepsilon_1-\varepsilon_0)\, 2^{\frac{H_b(\varepsilon_0)-H_b(2\varepsilon_1)-2\varepsilon_1}{2\varepsilon_1+\varepsilon_0-1}} + \frac{1}{2}\,(1-\varepsilon_0-2\varepsilon_1)}$$

being the input distribution

$$p_X(x_0) = p_X(x_2) = \frac{p}{2},\ p_X(x_1) = 1-p$$

d) If we have that $\varepsilon_0 = \varepsilon_1$, it is very easy to see that for $\varepsilon_0 = \varepsilon_1 = 0$ we have an ideal channel, in which $H(Y|X) = 0$, and capacity is achieved with equiprobable input symbols, which generate equiprobable output symbols and therefore the channel capacity is $\log_2(3)$, which is the maximum possible value for a system with three inputs and three outputs.

Exercise 4.14 Solution

a) $a = b = c = d = 0$. They are two BSC channels in parallel with different error probabilities. Therefore it would be a good model for a system that is transmitting bits through two independent subsystems.

b)

$$H(X) = 1 + H_b(p).$$
$$H(Y) = 1 + H_b(p)$$
$$H(X,Y) = 1 + H_b(p) + pH_b(\varepsilon_1) + (1-p)H_b(\varepsilon_0)$$
$$H(Y|X) = pH_b(\varepsilon_1) + (1-p)H_b(\varepsilon_0)$$
$$H(X|Y) = pH_b(\varepsilon_1) + (1-p)H_b(\varepsilon_0)$$

c)

$$C = 1 + H_b(p) - pH_b(\varepsilon_1) - (1-p)H_b(\varepsilon_0)$$

with

$$p = \frac{1}{1 + 2^{H_b(\varepsilon_1)-H_b(\varepsilon_0)}}$$

being the input distribution

$$p_X(x_0) = p_X(x_1) = \frac{p}{2},\ p_X(x_0) = p_X(x_1) = \frac{1-p}{2}$$

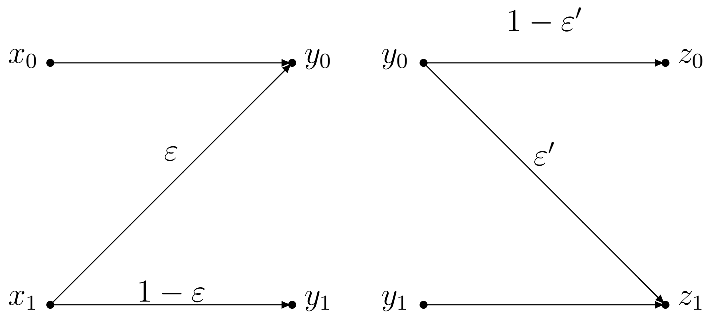
Exercise 4.15 Solution

a)

$$H(X) = H_b(\alpha),\ H(Y) = H_b((1-\alpha)(1-\varepsilon)),\ H(Z) = H_b\left(\beta(1-\varepsilon')\right)$$
$$H(X,Y) = H_B(\alpha) + (1-\alpha)H_b(\varepsilon)$$
$$H(Y|Z) = H_b(\beta) + \beta H_b(\varepsilon') - H_b(\beta(1-\varepsilon'))$$

b) The transition probabilities, given in the channel matrix, are

$$\mathbf{P}_{Y|X} = \begin{bmatrix} 1 & 0 \\ \varepsilon & 1-\varepsilon \end{bmatrix}, \ \mathbf{P}_{Z|Y} = \begin{bmatrix} 1-\varepsilon' & \varepsilon' \\ 0 & 1 \end{bmatrix}$$



c) $I(X,Y) = H_b((1-\alpha)(1-\varepsilon)) - (1-\alpha)H_b(\varepsilon)$

d)

$$\mathbf{P}_{Z|X} = \mathbf{P}_{Y|X} \times \mathbf{P}_{Z|Y} = \begin{bmatrix} 1-\varepsilon' & \varepsilon' \\ \varepsilon(1-\varepsilon') & \varepsilon\varepsilon' + 1 - \varepsilon \end{bmatrix}$$

e) The concatenation will be a BSC if $\varepsilon = \frac{\varepsilon'}{1-\varepsilon'}$

Exercise 4.16 Solution

a) The channel matrix is

$$\mathbf{P} = \begin{bmatrix} 1 & 0 \\ \varepsilon & 1-\varepsilon \\ 0 & 1 \end{bmatrix}.$$

b) Entropy values in the two scenarios

i)

$$H(X) = \log_2 3 = 1.585 \text{ bits/symbol.}$$
$$H(Y) = H_b\left(\frac{1}{3}(1+\varepsilon)\right) = H_b\left(\frac{1}{3}(2-\varepsilon)\right) \text{ bits/symbol.}$$
$$H(Y|X) = p_1\ H_b(\varepsilon) = \frac{1}{3}\ H_b(\varepsilon) \text{ bits/symbol.}$$
$$H(X,Y) = \log_2 3 + \frac{1}{3}\ H_b(\varepsilon) \text{ bits/symbol.}$$
$$H(X|Y) = \log_2 3 + \frac{1}{3}\ H_b(\varepsilon) - H_b\left(\frac{1}{3}(1+\varepsilon)\right) \text{ bits/symbol.}$$
$$I(X,Y) = H_b\left(\frac{1}{3}(1+\varepsilon)\right) - \frac{1}{3}\ H_b(\varepsilon) \text{ bits/symbol.}$$

ii)

$$H(X) = 0 \text{ bits/symbol,}$$
$$H(Y) = H_b(\varepsilon) \text{ bits/symbol.}$$
$$H(Y|X) = H_b(\varepsilon) \text{ bits/symbol.}$$
$$H(X,Y) = H_b(\varepsilon) \text{ bits/symbol.}$$
$$H(X|Y) = 0 \text{ bits/symbol,}$$
$$I(X,Y) = 0 \text{ bits/symbol.}$$

c) The channel capacity, for each case, is

i) $C = \max_{p_X(x_i)} I(X,Y) = 1$ bit/use, value that is reached for $p = 0$, which means that

$$p_0 = \frac{1}{2},\ p_1 = 0,\ p_2 = \frac{1}{2}.$$

ii)

$$C = H_b\left(\frac{1}{1+2^{\frac{H_b(\varepsilon)}{1-\varepsilon}}}\right) - \frac{1}{(1-\varepsilon)\ \left(1+2^{\frac{H_b(\varepsilon)}{1-\varepsilon}}\right)}\ H_b(\varepsilon),$$

what is obtained for

$$p_X(x_0) = 1-p = 1 - \frac{1}{(1-\varepsilon)\ \left(1+2^{\frac{H_b(\varepsilon)}{1-\varepsilon}}\right)},\ p_X(x_1) = p = \frac{1}{(1-\varepsilon)\ \left(1+2^{\frac{H_b(\varepsilon)}{1-\varepsilon}}\right)}.$$

Exercise 4.17 Solution

a) The joint entropy is

$$\begin{aligned} H(X,Y) &= -\sum_{i=0}^{1}\sum_{j=0}^{1} p_{X,Y}(x_i,y_j)\log_2 p_{X,Y}(x_i,y_j) \\ &= -0.1\log_2 0.1 - 0.2\log_2 0.2 - 0.3\log_2 0.3 - 0.4\log_2 0.4 = 1.8464 \text{ bits/symbol} \end{aligned}$$

b) The probabilities of the random variable X are

$$\begin{cases} p_X(x_0) = p_{X,Y}(x_0,y_0) + p_{X,Y}(x_0,y_1) = 0.1 + 0.2 = 0.3 \\ p_X(x_1) = p_{X,Y}(x_1,y_0) + p_{X,Y}(x_1,y_1) = 0.3 + 0.4 = 0.7 \end{cases}$$

so the entropy is

$$H(X) = -0.3\log_2 0.3 - 0.7\log_2 0.7 = H_b(0.3) = H_b(0.7) = 0.8813 \text{ bits/symbol}$$

The probabilities of the random variable Y are

$$\begin{cases} p_Y(y_0) = p_{X,Y}(x_0,y_0) + p_{X,Y}(x_1,y_0) = 0.1 + 0.3 = 0.4 \\ p_Y(y_1) = p_{X,Y}(x_0,y_1) + p_{X,Y}(x_1,y_1) = 0.2 + 0.4 = 0.6 \end{cases}$$

so the entropy is

$$H(Y) = -0.4\log_2 0.4 - 0.6\log_2 0.6 = H_b(0.4) = H_b(0.6) = 0.971 \text{ bits/symbol}$$

The conditional entropy $H(X|Y)$ can be calculated in several ways. One is by taking advantage of the relationship

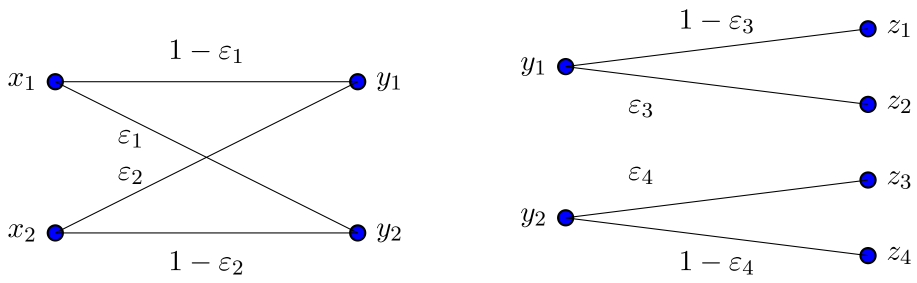
$$H(X|Y) = H(X,Y) - H(Y) = 1.8464 - 0.971 = 0.8754 \text{ bits/symbol}$$

c) Mutual information can also be calculated in several ways. One of them is

$$I(X,Y) = H(X) - H(X|Y) = 0.8813 - 0.8754 = 0.0059 \text{ bits/symbol}$$

Exercise 4.18 Solution

a) The schematic representation of the two DMCs is shown in the figure



b) The entropy of this variable, being binary, is

$$H(X) = \Omega(\alpha) \text{ bits/symbol.}$$

The probabilities of the exit symbols are

$$\begin{aligned} P_Y(y_1) =& \alpha(1-\varepsilon_1) + (1-\alpha)\varepsilon_2 \equiv \gamma \\ P_Y(y_2) =& \alpha\varepsilon_1 + (1-\alpha)(1-\varepsilon_2) \equiv 1-\gamma \end{aligned}$$

so since it is also a binary variable it has entropy

$$H(Y) = \Omega(\gamma) \text{ bits/symbol}$$

For the joint entropy it is convenient to calculate the conditional

$$H(Y|X) = \alpha\Omega(\varepsilon_1) + (1-\alpha)\Omega(\varepsilon_2) \text{ bits/symbol, since } \begin{cases} H(Y|X=x_1) = \Omega(\varepsilon_1) \\ H(Y|X=x_2) = \Omega(\varepsilon_2) \end{cases}$$

Now the joint entropy is

$$H(X,Y) = H(X) + H(Y|X) = \Omega(\gamma) + \alpha\Omega(\varepsilon_1) + (1-\alpha)\Omega(\varepsilon_2) \text{ bits/symbol}$$

c) The probabilities of Z are now

$$\begin{aligned} P_Y(z_1) =& \beta(1-\varepsilon_3) \\ P_Y(z_2) =& \beta\varepsilon_3 \\ P_Y(z_2) =& (1-\beta)\varepsilon_4 \\ P_Y(z_1) =& (1-\beta)(1-\varepsilon_4) \end{aligned}$$

so the entropy of Z is

$$\begin{aligned} H(Z) =& -\beta(1-\varepsilon_3)\log_2 \beta(1-\varepsilon_3) - \beta\varepsilon_3 \log_2 \beta\varepsilon_3 \\ & - (1-\beta)\varepsilon_4 \log_2 (1-\beta)\varepsilon_4 - (1-\beta)(1-\varepsilon_4)\log_2(1-\beta)(1-\varepsilon_4) \\ =& -\beta\left[\beta\log_2\beta + (1-\beta)\log_2(1-\beta)\right] - \beta\left[\varepsilon_3\log_2\varepsilon_3 + (1-\varepsilon_3)\log_2(1-\varepsilon_3)\right] \\ & - (1-\beta)\left[\varepsilon_4\log_2\varepsilon_4 + (1-\varepsilon_4)\log_2(1-\varepsilon_4)\right] \\ =& \Omega(\beta) + \beta\Omega(\varepsilon_3) + (1-\beta)\Omega(\varepsilon_4) \text{ bits/symbol} \end{aligned}$$

As for $H(Y|Z)$, knowing the value of Z there is no uncertainty about the value of Y, so

$$H(Y|Z) = 0 \text{ bits/symbol}$$

d) For DMC A

$$I(X,Y) = H(Y) - H(Y|X) = \Omega(\gamma) - \alpha\Omega(\varepsilon_1) - (1-\alpha)\Omega(\varepsilon_2)$$

For DMC B

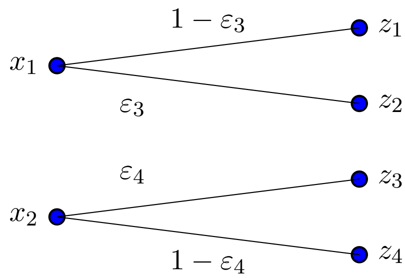
$$I(Y,Z) = H(Y) - H(Y|Z) = \Omega(\beta).$$

It could also be calculated as

$$I(Y,Z) = H(Z) - H(Z|Y),$$

taking into account that for this channel it is very easy to obtain $H(Z|Y) = \beta\Omega(\varepsilon_3) + (1-\beta)\Omega(\varepsilon_4)$.

e) The concatenated channel, for $\varepsilon_1 = \varepsilon_2 = 0$, with its transition probabilities, is shown in the figure



f) The mutual information between input and output of this concatenated channel is

$$I(X,Z) = H(X) - H(X|Z) = \Omega(\alpha),$$

because for this channel $H(X|Z) = 0$. Since the channel capacity is the maximum value that this mutual information can take depending on the input probabilities, in this case

$$C = 1 \text{ bit/use, for } p_X(x_1) = p_X(x_2) = \frac{1}{2}$$

since the maximum value of $\Omega(\alpha)$ is one and is obtained for $\alpha = \frac{1}{2}$.