Problem 1

The two random variables X and Y take values in $x \in \{0, 1\}$ and $y \in \{0, 1, 2\}$, respectively. Their joint distribution function can be written as

$$P(x, y) = K \cdot (x + y)$$

(a) Set up a table for $K \cdot (x + y)$,

$K(x+y)$		$Y=0$	$Y=1$	$Y=2$
X	0	0	K	$2K$
	1	K	$2K$	$3K$

For this to be $p(x, y)$ we need the requirement that $\sum_{x,y} p(x, y) = 1$, and we se that $K = 1/9$.

(b) We can now derive the probabilities for X and Y as

$K(x+y)$		$Y=0$	$Y=1$	$Y=2$	$p(x)$
X	0	0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{3}$
	1	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{2}{3}$
$P(y)$		$\frac{1}{9}$	$\frac{3}{9}$	$\frac{5}{9}$	

and we get

$$H(X) = h(\tfrac{1}{3}) = \cdots = \log 3 - \frac{2}{3} \approx 0.9183$$

$$H(Y) = H(\tfrac{1}{9}, \tfrac{3}{9}, \tfrac{5}{9}) = \cdots = \frac{15}{9}\log 3 - \frac{5}{9}\log 5 \approx 1.3516$$

(c) The joint entropy is

$$H(X, Y) = H(\tfrac{1}{9}, \tfrac{2}{9}, \tfrac{1}{9}, \tfrac{2}{9}, \tfrac{3}{9}) = \cdots = \frac{15}{9}\log 3 - \frac{4}{9}$$

Hence the mutual information is

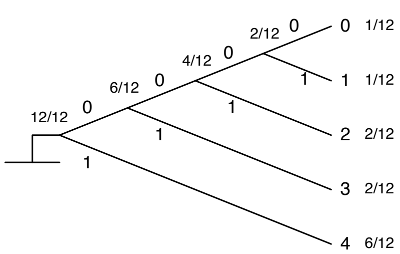
$$I(X; Y) = H(X) + H(Y) - H(X, Y) = \log 3 - \frac{5}{9}\log 5 - \frac{2}{9} \approx 0.0728$$

Problem 2

(a) To start with, the entropy is

$$H(X) = H(\tfrac{1}{12}, \tfrac{1}{12}, \tfrac{2}{12}, \tfrac{2}{12}, \tfrac{6}{12}) = \cdots = \frac{1}{2}\log 3 + \frac{7}{6} \approx 1.9591$$

We can create a Huffman code from the tree below.



X	Y
0	0000
1	0001
2	001
3	01
4	1

$$E[L] = \frac{12+6+4+2}{12} = 2$$

(b) In the Huffman tree above the mapping can be chosen in two ways for each branching in the tree. Since there are four branchings, we get 2^4 different codes. Furthermore, the Huffman algorithm can end up with two other trees as well, both with four branchings. Thus, in total we can get $3 \cdot 2^4 = 48$ different Huffman codes.

Problem 3

(a) Viewing the rooms as states in a Markov chain, we get the state transition matrix

$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

Then the equation system

$$\begin{aligned} \boldsymbol{w}P &= \boldsymbol{w} \\ \sum_w w &= 1 \end{aligned}$$

has a solution in

$$\begin{aligned} \boldsymbol{w} &= \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix} \tilde{Q}^{-1} = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\frac{2}{3} & \frac{1}{3} & 0 & 1 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} & -1 & 1 \\ \frac{1}{3} & 0 & \frac{1}{3} & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\frac{21}{25} & -\frac{2}{5} & \frac{4}{25} & \frac{27}{25} \\ \frac{3}{5} & -2 & -\frac{2}{5} & \frac{9}{5} \\ \frac{3}{25} & -\frac{4}{5} & -\frac{22}{25} & \frac{39}{25} \\ \frac{6}{25} & \frac{2}{5} & \frac{6}{25} & \frac{3}{25} \end{pmatrix} = \begin{pmatrix} \frac{6}{25} & \frac{2}{5} & \frac{6}{25} & \frac{3}{25} \end{pmatrix} \end{aligned}$$

where $\tilde{Q}$ is the matrix $P-I$ where the last column is replaced with all-ones, to represent $\sum w = 1$.

(b) The entropy rate is

$$\begin{aligned} H_\infty(W) &= \frac{6}{25} H(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}) + \frac{2}{5} h(\tfrac{1}{2}) + \frac{6}{25} h(\tfrac{1}{2}) + \frac{3}{25} H(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}) \\ &= \frac{6}{25} \log 3 + \frac{2}{5} + \frac{6}{25} + \frac{3}{25} \log 3 = \frac{9}{25} \log 3 + \frac{16}{25} \approx 1.2106 \end{aligned}$$

Problem 4

(a) The differential entropy can be derived as

$$H(X) = -\int_{\mathbb{R}} f(x) \log f(x) dx = -2\frac{1}{6} \log \frac{1}{6} - 2\frac{1}{3} \log \frac{1}{3} = \cdots = \log 3 + \frac{1}{3} \approx 1.9183$$

(b) For the all-zero vector to among the typical sequences we need that

$$2^{-n(H(X)+\epsilon)} \leqslant f(0 \dots 0) \leqslant 2^{-n(H(X)-\epsilon)}$$

Since the variables $x_1, \dots, x_n$ are i.i.d.

$$2^{-(H(X)+\epsilon)} \leqslant f(0) \leqslant 2^{-(H(X)-\epsilon)}$$

or, derived as numbers,

$$0.2468 \leqslant f(0) \leqslant 0.2836$$

Since $f(0) = 1/3$ this is not fulfilled and the all-zero vector is not a typical vector. To find a typical vector we use that there are only two non-zero levels in $f(x)$, i.e. $1/3$ and $1/6$. If a vector consists of k elements with $f(x) = 1/3$ and $10 - k$ elements with $f(x) = 1/6$, we get the probability

$$f(x_1 \dots x_{10}) = \left(\frac{1}{3}\right)^k \left(\frac{1}{6}\right)^{10-k} = \frac{1}{3^{10}2^{10-k}}$$

For a typical vector we require

$$2^{-n(H(X)+\epsilon)} \leqslant \frac{1}{3^{10}2^{10-k}} \leqslant 2^{-n(H(X)-\epsilon)}$$

or, equivalently,

$$0.2468 = 2^{-H(X)-\epsilon} \leqslant \frac{1}{3\sqrt[10]{2^{10-k}}} \leqslant 2^{-H(X)+\epsilon} = 0.2836$$

Setting up a table we get

k	$\frac{1}{3\sqrt[10]{2^{10-k}}}$
0	0.1667
1	0.1786
2	0.1914
3	0.2052
4	0.2199
5	0.2357
6	0.2526
7	0.2708
8	0.2902
9	0.3110
10	0.3333

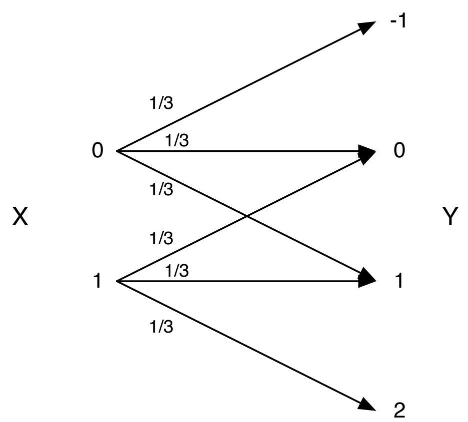
We conclude that it is required six or seven elements with probability $1/3$ and the rest with probability $1/6$. For example the vector

$$\boldsymbol{x} = (1.5, 1.5, 1.5, 1.5, 0, 0, 0, 0, 0, 0)$$

is a typical vector.

Problem 5

(a) We start by drawing the discrete memoryless channel as a picture:



To derive the capacity we start with

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

where we get

$$H(Y|X) = \sum_x H(Y|X=x)P(X=x) = H(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}) = \log 3$$

What we need next is $H(Y)$. For that we let $P(X=0) = p$ and set up the table for the joint distribution:

		Y			
$p(x,y)$		-1	0	1	2
X	0	$p\frac{1}{3}$	$p\frac{1}{3}$	$p\frac{1}{3}$	0
	1	0	$(1-p)\frac{1}{3}$	$(1-p)\frac{1}{3}$	$(1-p)\frac{1}{3}$
$p(y)$		$p\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$(1-p)\frac{1}{3}$

The entropy of Y is

$$\max_p H(Y) = \max_p H(p\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}, (1-p)\tfrac{1}{3}) = \cdots = \max_p \Big\{\frac{1}{3}h(p) + \log 3\Big\} = \frac{1}{3} + \log 3$$

where $p = 1/2$ gives the maximizing distribution of X. Hence,

$$C = \max_{p(x)} I(X;Y) = \frac{1}{3}$$

(b) Following the hint we start with

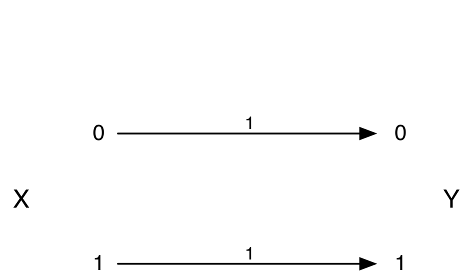
$$\begin{aligned} I(X;Y|W) &= H(Y|W) - H(Y|X,W) \\ &= \sum_w H(Y|W=w)P(W=w) - \sum_w H(Y|X,W=w)P(W=w) \\ &= \sum_w \Big(H(Y|W=w) - H(Y|X,W=w)\Big)P(W=w) \\ &= \sum_w I(X;Y|W=w)P(W=w) \end{aligned}$$

Looking closer into the probabilities we can see that

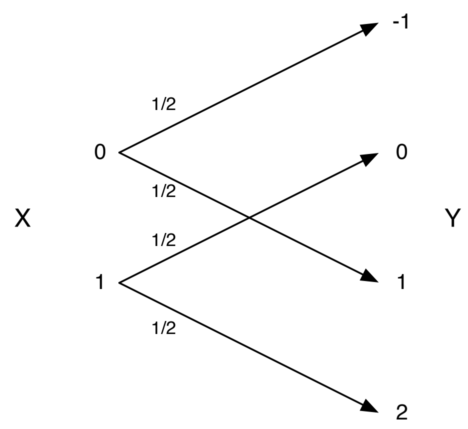
$$I(X;Y|W=w) = \sum_{x,y} p(x,y|w) \log \frac{p(x,y|w)}{p(x|w)p(y|w)}$$

Conditioning on the two alternatives for W we get

$w = 0$, where $P(W = 0) = 1/3$, gives the channel



$w = 1$, where $P(W = 0) = 2/3$, gives the channel



For both of these cases it is clear from the received symbol what was transmitted, i.e. $I(X;Y) = 1$. Hence,

$$I(X;Y|W) = I(X;Y|W=0)P(W=0) + I(X;Y|W=1)P(W=1) = 1\frac{1}{3} + 1\frac{2}{3} = 1$$

and we see that the capacity is 1 bit per channel use.