

3. Consider the discrete channel in Figure 3.1.

- Derive an expression for the channel capacity as a function of ϵ .
- Find the maximum and minimum values for the capacity, and sketch a plot of the capacity as ϵ varies, where you show minimum, maximum and end points.

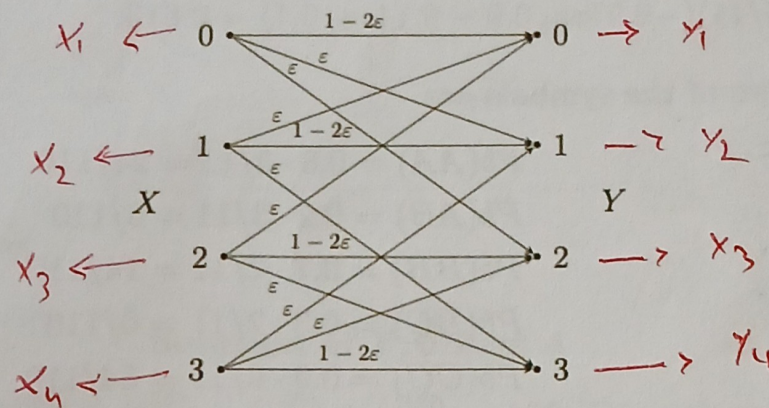


Figure 3.1: A discrete channel

4. Consider a discrete memoryless source generating a sequence of independent equally distributed binary digits 0 and 1 with probability $p(1) = 0.005$ and $p(0) = 0.995$. The digits are taken 100 at a time and a unique binary sequence of the (same) length l is assigned to each such sequence.

- What is the length (in bits) of encoded sequences in case only sequences with at most three 1s are encoded?
- Is this a good compression in Shannons sense or not? Motivate your answer by explaining why it is good or not good. How would you improve encoding of such a source if possible at all.

Corrected Solution to Problem 3

Problem 3

From the figure, the channel transitions are:

$$0 \rightarrow \{0, 1, 3\},$$

$$1 \rightarrow \{0, 1, 2\},$$

$$2 \rightarrow \{1, 2, 3\},$$

$$3 \rightarrow \{0, 2, 3\}.$$

Thus:

- correct transmission probability is $1 - 2\varepsilon$,
- each connected neighboring output has probability ε ,
- all other transitions have probability 0.

Hence the transition probability matrix is

$$P(Y|X) = \begin{bmatrix} 1 - 2\varepsilon & \varepsilon & \varepsilon & 0 \\ \varepsilon & 1 - 2\varepsilon & 0 & \varepsilon \\ \varepsilon & 0 & 1 - 2\varepsilon & \varepsilon \\ 0 & \varepsilon & \varepsilon & 1 - 2\varepsilon \end{bmatrix}.$$

Notice specifically that

$$P(1|3) = 0,$$

so indeed there is no connection between 3 and 1.

(a) Channel Capacity

The channel is weakly symmetric because:

- every row is a permutation of every other row,
- every column sum is identical.

Therefore the capacity-achieving input distribution is uniform:

$$P_X(x) = \frac{1}{4}.$$

The capacity of a weakly symmetric channel is

$$C = \log_2 |\mathcal{Y}| - H(\text{row}).$$

Since there are four outputs,

$$\log_2 4 = 2.$$

Each row contains probabilities

$$(1 - 2\varepsilon, \varepsilon, \varepsilon, 0).$$

Hence the row entropy is

$$H(\text{row}) = -(1 - 2\varepsilon) \log_2(1 - 2\varepsilon) - 2\varepsilon \log_2 \varepsilon.$$

Therefore,

$$C(\varepsilon) = 2 + (1 - 2\varepsilon) \log_2(1 - 2\varepsilon) + 2\varepsilon \log_2 \varepsilon$$

bits/channel use.

Valid range:

$$0 \leq \varepsilon \leq \frac{1}{2}.$$

(b) Maximum and Minimum Values

At $\varepsilon = 0$

The channel is noiseless:

$$C(0) = 2 \text{ bits.}$$

Thus,

$$C_{\max} = 2 \text{ bits}$$

at

$$\varepsilon = 0.$$

Critical Point

Differentiate:

$$\frac{dC}{d\varepsilon} = 2 \log_2 \left(\frac{\varepsilon}{1 - 2\varepsilon} \right).$$

Set derivative to zero:

$$\frac{\varepsilon}{1 - 2\varepsilon} = 1.$$

Thus,

$$\varepsilon = 1 - 2\varepsilon,$$

giving

$$3\varepsilon = 1,$$

so

$$\boxed{\varepsilon = \frac{1}{3}}.$$

At this value,

$$1 - 2\varepsilon = \frac{1}{3},$$

so every nonzero transition probability equals $1/3$.

Then

$$H(\text{row}) = \log_2 3.$$

Hence

$$\boxed{C_{\min} = 2 - \log_2 3 \approx 0.415 \text{ bits.}}$$

At $\varepsilon = \frac{1}{2}$

Now

$$1 - 2\varepsilon = 0.$$

Each row becomes

$$(0, \frac{1}{2}, \frac{1}{2}, 0).$$

Thus

$$H(\text{row}) = 1,$$

and therefore

$$C\left(\frac{1}{2}\right) = 1 \text{ bit.}$$

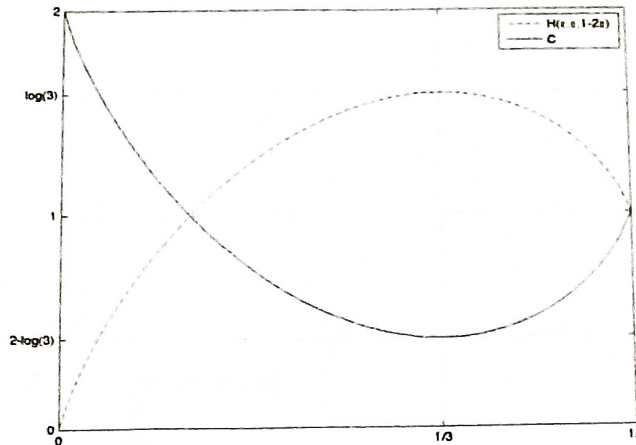
Final Answers

$$C(\varepsilon) = 2 + (1 - 2\varepsilon) \log_2(1 - 2\varepsilon) + 2\varepsilon \log_2 \varepsilon$$

$$C_{\max} = 2 \text{ bits at } \varepsilon = 0$$

$$C_{\min} = 2 - \log_2 3 \text{ at } \varepsilon = \frac{1}{3}$$

$$C\left(\frac{1}{2}\right) = 1 \text{ bit}$$



1.png

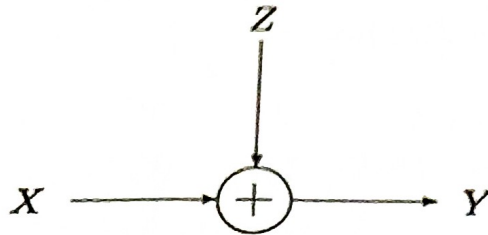
Solution of Problem 4

- (a) There are exactly $\binom{100}{0} + \binom{100}{1} + \binom{100}{2} + \binom{100}{3} = 166,751$ such sequences. These can be encoded using $\lceil \log_2(166751) \rceil = 18$ bits.
- (b) This is **not good** compression since:
- The entropy $H(X) = H(0.005) \approx 0.0454$ bits/symbol
 - The Shannon limit is $nH(X) = 100 \times 0.0454 = 4.54$ bits
 - We're using 18 bits compared to the theoretical limit of 4.54 bits

This part was elementary in my opinion, the explanation why is a bit harder !!

The problem is that $n = 100$ is quite small for this source, and we actually deal with the sequences which are not typical at all. According to the probability distribution we expect a single 1 to appear in a sequence of length 200, whereas encoding nontypical sequences such as those having 3 ones in a sequence of length 100 we spend much bits in encoding these (these are in majority) non typical sequences. One should increase n in order to deal better with this source.

1. An additive noise channel: Find the channel capacity of the following discrete memoryless channel:



where $\Pr\{Z = 0\} = \Pr\{Z = a\} = \frac{1}{2}$. The alphabet for x is $\mathbf{X} = \{0, 1\}$. Assume that Z is independent of X .

Observe that the channel capacity depends on the value of a .

Solution: *A sum channel.*

$$Y = X + Z \quad X \in \{0, 1\}, \quad Z \in \{0, a\}$$

We have to distinguish various cases depending on the values of a .

$a = 0$ In this case, $Y = X$, and $\max I(X; Y) = \max H(X) = 1$. Hence the capacity is 1 bit per transmission.

$a \neq 0, \pm 1$ In this case, Y has four possible values $0, 1, a$ and $1 + a$. Knowing Y , we know the X which was sent, and hence $H(X|Y) = 0$. Hence $\max I(X; Y) = \max H(X) = 1$, achieved for a uniform distribution on the input X .

$a = 1$ In this case Y has three possible output values, $0, 1$ and 2 , and the channel is identical to the binary erasure channel discussed in class, with $a = 1/2$. As derived in class, the capacity of this channel is $1 - a = 1/2$ bit per transmission.

$a = -1$ This is similar to the case when $a = 1$ and the capacity here is also $1/2$ bit per transmission.

Channel Capacity of an Additive Noise Channel

Problem

Consider the discrete memoryless additive noise channel

$$Y = X + Z,$$

where:

$$P(Z = 0) = P(Z = a) = \frac{1}{2},$$

and

$$X \in \{0, 1\}.$$

Assume Z is independent of X .

We determine the channel capacity and show how it depends on a .

Step 1: Possible Outputs

Since

$$Y = X + Z,$$

the possible outputs are:

$$X = 0: Y \in \{0, a\},$$

$$X = 1: Y \in \{1, 1 + a\}.$$

Thus the output alphabet depends on the value of a .

Step 2: Analyze Different Cases

The channel capacity changes depending on whether output values overlap.

Case 1: $a \neq 0, \pm 1$

The four outputs

$$0, a, 1, 1 + a$$

are all distinct.

Hence:

$$X = 0 \Rightarrow Y \in \{0, a\},$$

$$X = 1 \Rightarrow Y \in \{1, 1 + a\},$$

and the receiver can determine X exactly from Y .

Therefore the channel is noiseless.

Hence

$$H(X|Y) = 0.$$

The mutual information is

$$I(X; Y) = H(X).$$

This is maximized by the uniform input distribution:

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

Thus

$$H(X) = 1 \text{ bit.}$$

Therefore the capacity is

$$\boxed{C = 1 \text{ bit}}$$

for

$$\boxed{a \neq 0, \pm 1.}$$

Case 2: $a = 0$

Now

$$Z = 0$$

with probability 1, because both possible noise values are identical.

Hence

$$Y = X.$$

Again the channel is noiseless.

Therefore

$$\boxed{C = 1 \text{ bit.}}$$

Now

$$Y = X + Z, \quad Z \in \{0, 1\}.$$

Possible outputs:

$$X = 0 : Y \in \{0, 1\},$$

$$X = 1 : Y \in \{1, 2\}.$$

The output $Y = 1$ can result from either input, so ambiguity exists.

Assume

$$P(X = 0) = p, \quad P(X = 1) = 1 - p.$$

Output Distribution

We compute:

For $Y=0$ only possibility is $X=Z=0$
 $P(Y=0) = P(X=0) \times P(Z=0)$
 Z and X are ind.

$$P(Y = 0) = \frac{p}{2},$$

$$P(Y = 1) = \frac{1}{2},$$

$$P(Y = 2) = \frac{1-p}{2}.$$

How to get this

When $a=1$, $z \in \{0,1\}$
 $P(z=0) = P(z=1) = 1/2$
and $Y = X + Z$

Hence

$$H(Y) = -\frac{p}{2} \log_2 \frac{p}{2} - \frac{1}{2} \log_2 \frac{1}{2} - \frac{1-p}{2} \log_2 \frac{1-p}{2}.$$

Since the channel noise has entropy

$$H(Z) = 1 \text{ bit},$$

we have

$$H(Y|X) = H(Z) = 1.$$

Therefore

$$I(X; Y) = H(Y) - 1.$$

	X	Z
Y	p	$\frac{1}{2} = p \cdot \frac{1}{2}$
	$p-1$	$\frac{1}{2}$

This is maximized at

$$p = \frac{1}{2}.$$

Then

$$P(Y = 0) = \frac{1}{4}, \quad P(Y = 1) = \frac{1}{2}, \quad P(Y = 2) = \frac{1}{4}.$$

Thus

$$H(Y) = -\frac{1}{4} \log_2 \frac{1}{4} - \frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{4} \log_2 \frac{1}{4}.$$

Compute:

$$H(Y) = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} = \frac{3}{2}.$$

$P(Y=0)$

$$X + Z = 0$$

This is possible only if $X=Z=0$

$$\text{Since } X \text{ and } Z \text{ ind. } P(Y=0) = P(X=0) \times P(Z=0) = p \times \frac{1}{2} = \frac{p}{2}$$

$P(Y=1)$

$$X + Z = 1$$

these are two ways $X=0, Z=1$ or $X=1, Z=0$

$P(Y=2)$ is possible $X+Z=2$
so $X=1, Z=1$

also $X+Z=1$ if $X=1, Z=0$
 $(1-p) \times \frac{1}{2}$
so $P(Y=1) = (1-p) \times \frac{1}{2} + p \times \frac{1}{2} = \frac{1}{2}$

Therefore

$$C = H(Y) - 1 = \frac{3}{2} - 1 = \frac{1}{2}$$

Hence

$$C = \frac{1}{2} \text{ bit}$$

for

$$a = 1.$$

Case 4: $a = -1$

Now

$$Y = X + Z, \quad Z \in \{0, -1\}.$$

Possible outputs:

$$X = 0 : Y \in \{0, -1\},$$

$$X = 1 : Y \in \{1, 0\}.$$

Again one output overlaps, producing ambiguity exactly as in the previous case. By symmetry, the capacity is identical:

$y =$ $x + \overset{\checkmark}{z} = -1$ $C = \frac{1}{2} \text{ bit}$ ✓

for

$$a = -1.$$

Final Answer

$$C(a) = \begin{cases} 1 \text{ bit}, & a \neq \pm 1, \\ \frac{1}{2} \text{ bit}, & a = \pm 1. \end{cases}$$

Thus the channel loses capacity only when the additive noise causes overlap between output symbols.

2. **Channels with memory have higher capacity.** Consider a binary symmetric channel with $Y_i = X_i \oplus Z_i$, where \oplus is mod 2 addition, and $X_i, Y_i \in \{0, 1\}$.

Suppose that $\{Z_i\}$ has constant marginal probabilities $\Pr\{Z_i = 1\} = p = 1 - \Pr\{Z_i = 0\}$, but that Z_1, Z_2, \dots, Z_n are not necessarily independent. Assume that Z^n is independent of the input X^n . Let $C = 1 - H(p, 1-p)$. Show that $\max_{p(x_1, x_2, \dots, x_n)} I(X_1, X_2, \dots, X_n; Y_1, \dots, Y_n) \geq nC$.

Solution

Given

$$Y_i = X_i \oplus Z_i,$$

with

$$P(Z_i = 1) = p, \quad P(Z_i = 0) = 1 - p,$$

and Z^n independent of X^n .

Define

$$C = 1 - H(p, 1 - p),$$

where

$$H(p, 1 - p) = -p \log_2 p - (1 - p) \log_2 (1 - p).$$

We have

$$I(X^n; Y^n) = H(Y^n) - H(Y^n | X^n).$$

Since

$$Y^n = X^n \oplus Z^n,$$

it follows that

$$H(Y^n | X^n) = H(Z^n).$$

Choose X^n uniformly distributed over $\{0, 1\}^n$. Then Y^n is also uniform, hence

$$H(Y^n) = n.$$

Therefore,

$$I(X^n; Y^n) = n - H(Z^n).$$

Using subadditivity of entropy,

$$H(Z^n) \leq \sum_{i=1}^n H(Z_i).$$

Since each Z_i has Bernoulli(p) distribution,

$$H(Z_i) = H(p, 1 - p).$$

Hence

$$H(Z^n) \leq nH(p, 1 - p).$$

Thus

$$I(X^n; Y^n) \geq n - nH(p, 1 - p).$$

$$\Rightarrow \boxed{I(X_1, \dots, X_n; Y_1, \dots, Y_n) \geq nC}$$

6. **The Z channel.** The Z-channel has binary input and output alphabets and transition probabilities $p(y|x)$ given by the following matrix:

$$Q = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \end{bmatrix} \quad x, y \in \{0, 1\}$$

Find the capacity of the Z-channel and the maximizing input probability distribution.

Capacity of the Z-Channel

The Z-channel has transition matrix

$$Q = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

where

$$P(Y = 0|X = 0) = 1, \quad P(Y = 1|X = 0) = 0,$$

and

$$P(Y = 0|X = 1) = \frac{1}{2}, \quad P(Y = 1|X = 1) = \frac{1}{2}.$$

Let

$$x = P(X = 1).$$

Then

$$P(X = 0) = 1 - x.$$

We compute the conditional entropy $H(Y|X)$.

Using the definition,

$$H(Y|X) = \sum_{x \in \{0,1\}} P(X = x) H(Y|X = x).$$

Hence,

$$H(Y|X) = P(X = 0)H(Y|X = 0) + P(X = 1)H(Y|X = 1).$$

Case $X = 0$

When $X = 0$, the output is deterministic because

$$P(Y = 0|X = 0) = 1.$$

Therefore,

$$H(Y|X = 0) = 0.$$

Case $X = 1$

When $X = 1$,

$$P(Y = 0|X = 1) = \frac{1}{2}, \quad P(Y = 1|X = 1) = \frac{1}{2}.$$

Thus the output is uniformly distributed and

$$H(Y|X = 1) = 1.$$

Substituting these values,

$$H(Y|X) = (1 - x) \cdot 0 + x \cdot 1.$$

Therefore,

$$H(Y|X) = x.$$

Next we compute the entropy of Y .

We first determine $P(Y = 1)$:

$$P(Y = 1) = P(Y = 1|X = 0)P(X = 0) + P(Y = 1|X = 1)P(X = 1).$$

Substituting the probabilities,

$$P(Y = 1) = 0 \cdot (1 - x) + \frac{1}{2}x.$$

Hence,

$$P(Y = 1) = \frac{x}{2}.$$

Therefore,

$$P(Y = 0) = 1 - \frac{x}{2}.$$

Thus,

$$H(Y) = H\left(\frac{x}{2}\right),$$

where the binary entropy function is

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

The mutual information is

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

Hence,

$$I(X; Y) = H\left(\frac{x}{2}\right) - x.$$

To maximize the mutual information, differentiate:

$$\frac{d}{dx}I(X; Y) = \frac{1}{2} \log_2\left(\frac{1-x/2}{x/2}\right) - 1.$$

Setting the derivative equal to zero,

$$\frac{1}{2} \log_2\left(\frac{1-x/2}{x/2}\right) = 1.$$

Multiplying by 2,

$$\log_2\left(\frac{1-x/2}{x/2}\right) = 2.$$

Exponentiating,

$$\frac{1-x/2}{x/2} = 4.$$

Solving for x ,

$$1 - \frac{x}{2} = 2x,$$

$$1 = \frac{5}{2}x,$$

$$x = \frac{2}{5}.$$

Therefore the maximizing input distribution is

$$P(X = 1) = \frac{2}{5}, \quad P(X = 0) = \frac{3}{5}.$$

Finally, the capacity is

$$C = H\left(\frac{1}{5}\right) - \frac{2}{5}.$$

Numerically,

$$H\left(\frac{1}{5}\right) \approx 0.722,$$

so

$$C \approx 0.722 - 0.4 = 0.322.$$

Thus the capacity of the Z-channel is

$$\boxed{C \approx 0.322 \text{ bits.}}$$