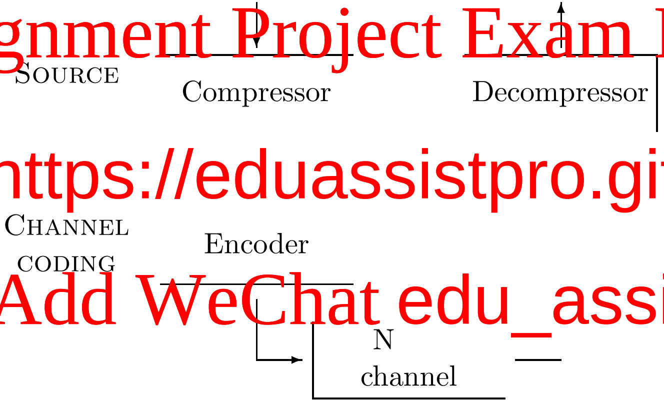
COMP2610/6261 - Information Theory

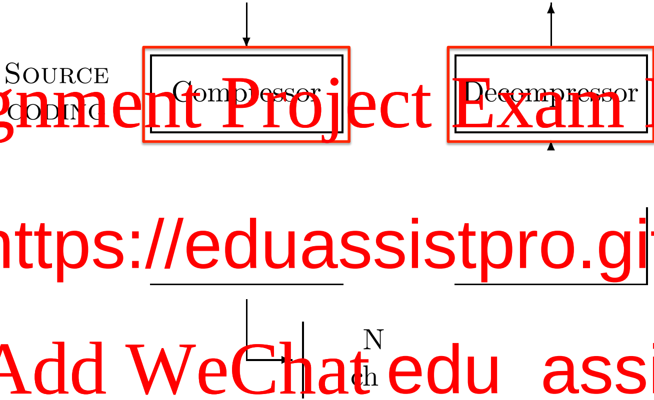
Lecture 17: Noisy Channels

Australian National University

2 October, 2018

The Big Picture



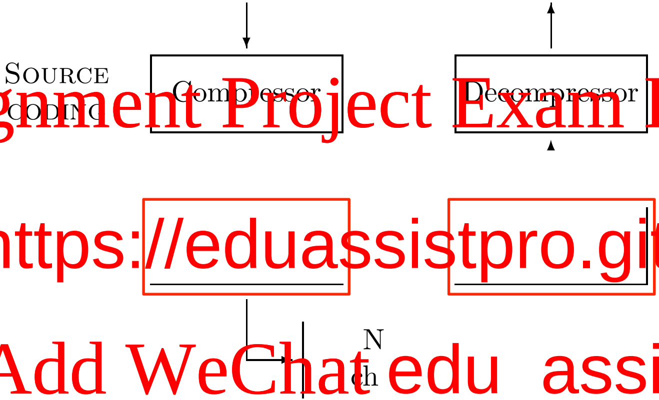


Concept: Expected code length

Theorem: Source coding theorem

Algorithms: { Huffman, Arithmetic } codes

The Big Picture



Concept: Channel capacity

Theorem: Channel coding theorem

Algorithms: Repetition codes, Hamming codes

Communication over Noisy Channels

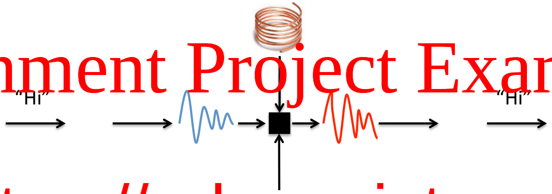
A channel is some medium for transmitting messages

A noisy ch
sender's

The Prob

“The fundamental problem of communication is th
point either exactly or approximately a message se
(Claude Shannon, 1948)

Example: Telephone Network



Source	: Aditya
Encoder	: Telephone handset
Channel	: Analogue telephone line
Decoder	: Telephone handset
Destination	: Mark

Key Questions

How do we model noisy communication abstractly?

What are the theoretical limits of noise correction?

What are the practical approaches to noise correction?

1. Communication over Noisy Channels: Big Picture
2. Nois
3. Exa
4. Probability of Error

Encoder/Decoder Pairs

Suppose we have some set $\mathcal{S} = \{1, 2, \ldots, S\}$ of possible messages

- e.g. codewords from Huffman coding on some ensemble
- Sender and receiver agree on what these are

Encoder/Decoder Pairs

Suppose we have some set $\mathcal{S} = \{1, 2, \ldots, S\}$ of possible messages

- e.g. codewords from Huffman coding on some ensemble

- Sender and receiver agree on what these are

When co
into some i

Encoder/Decoder Pairs

Suppose we have some set $\mathcal{S} = \{1, 2, \ldots, S\}$ of possible messages

- e.g. codewords from Huffman coding on some ensemble

- Sender and receiver agree on what these are

When co
into some i

The receiver then receives some (possibly corrup
output alphabet $\mathcal{Y}$

- Simple case: $\mathcal{X} = \mathcal{Y} = \{0, 1\}$

- The bit the sender transmits may not be what the receiver sees

Encoder/Decoder Pairs

Formally, the sender encodes messages via

$$\mathtt{enc}\colon \mathcal{S} \to \mathcal{X}^N$$

Encoder/Decoder Pairs

Formally, the sender encodes messages via

$$\texttt{enc}\colon \mathcal{S} \to \mathcal{X}^N$$

The receiver then decodes messages via

Encoder/Decoder Pairs

Formally, the sender encodes messages via

$$\texttt{enc}\colon \mathcal{S} \to \mathcal{X}^N$$

The receiver then decodes messages via

Isn't the compressor already "encoding" a message?

- Yes, but we might want to add something for noi

Might have $\mathcal{X} \neq \mathcal{Y}$

- e.g. if we allow a special "erased" symbol

$N > 1$ can be thought of as multiple uses of a channel

- e.g. use a bitstring of length 4 to represent messages

Channels: Informally

A **discrete channel** Q will:

- acc
- pro

Channels: Informally

A **discrete channel** Q will:

- acc
- pro

There is a

- This represents some inherent noise
- Noise could depend on the input

Channels: Formally

A **discrete channel** Q consists of:

- an *input alphabet* $\mathcal{X} = \{a_1, \ldots, a_I\}$

- an *output alphabet* $= b_1, \ldots, b_J$

- *tran*

Channels: Formally

A **discrete channel** Q consists of:

- an *input alphabet* $\mathcal{X} = \{a_1, \ldots, a_I\}$
- an *output alphabet* $= b_1, \ldots, b_J$
- *tran*

The channel Q can be expressed as a matrix

$$Q_{j,i} = P(y = b_j |$$

This represents the probability of observing b_j given that we transmit a_i

Channels: Example

Example: A channel Q with inputs $\mathcal{X} = \{a_1, a_2, a_3\}$, outputs $\mathcal{Y} = \{b_1, b_2\}$, and transition probabilities expressed by the matrix

So $P(b_1|a_1) = 0.8 = P(b_2|a_3)$ and $P(b_1|a_2) = P(b_2|a_2) = 0.5$.

Channels: Example

Example: A channel Q with inputs $\mathcal{X} = \{a_1, a_2, a_3\}$, outputs $\mathcal{Y} = \{b_1, b_2\}$, and transition probabilities expressed by the matrix

So $P(b_1|a_1) = 0.8 = P(b_2|a_3)$ and $P(b_1|a_2) = P(b_2|a_2) = 0.5$.

We arrange the inputs along the columns, and outp

Actual details of alphabet are abstracted away

1. Communication over Noisy Channels: Big Picture

2. Nois

3. Exa

4. Probability of Error

The Binary Noiseless Channel

One of the simplest channels is the **Binary Noiseless Channel**. The received symbol is always equal to the transmitted symbol – there is no probability of error, hence *noiseless*.

Inputs $= 0, 1$; Outputs $= \{0, 1\}$;

The Binary Noiseless Channel

One of the simplest channels is the **Binary Noiseless Channel**. The received symbol is always equal to the transmitted symbol – there is no probability of error, hence *noiseless*.

$$\text{Inputs} = 0, 1 \;;\ \text{Outputs} = \{0, 1\};$$

What was transmitted over the channel if 0000 1111 was received?

$$\xrightarrow{Q} 0000\ 1111$$

The Binary Noiseless Channel

One of the simplest channels is the **Binary Noiseless Channel**. The received symbol is always equal to the transmitted symbol – there is no probability of error, hence *noiseless*.

$$\text{Inputs} = 0, 1 \;; \text{Outputs} = \{0, 1\};$$

What was transmitted over the channel if $0000\,1111$ was received?

$$0000\,1111 \xrightarrow{Q} 0000\,1111$$

The Noisy Non-overlapping Channel

Even if there is some uncertainty about the output given the input, it may still be possible to perfectly infer what was transmitted.

Inputs $\mathcal{X} = \{0, 1\}$; Outputs $= \mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}$; Transition probabilities

.

The Noisy Non-overlapping Channel

Even if there is some uncertainty about the output given the input, it may still be possible to perfectly infer what was transmitted.

Inputs $\mathcal{X} = \{0, 1\}$; Outputs

$= \ \mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}$; Transition probabilities

.

What was transmitted over the channel if `abba ddcd` was received?

$\xrightarrow{Q}$ `abba ddcd`

The Noisy Non-overlapping Channel

Even if there is some uncertainty about the output given the input, it may still be possible to perfectly infer what was transmitted.

Inputs $\mathcal{X} = \{0, 1\}$; Outputs

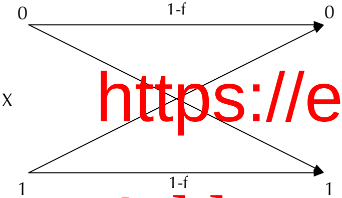
$= \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}\}$; Transition probabilities

What was transmitted over the channel if abba ddcd was received?

$$\color{red}{\mathtt{0000\,1111}} \xrightarrow{Q} \mathtt{abba\,ddcd}$$

The Binary Symmetric Channel

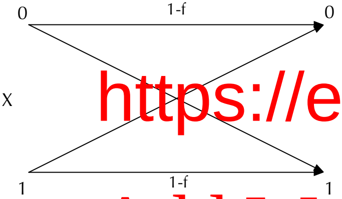
Each symbol sent across a **binary symmetric channel** has a chance of being "flipped" to its counterpart ($0 \rightarrow 1$; $1 \rightarrow 0$)



Inputs $\mathcal{X} = \{0, 1\}$; Outputs $\mathcal{Y} = \{0, 1\}$;
Transition probabilities with $P(\text{flip}) = f$

The Binary Symmetric Channel

Each symbol sent across a **binary symmetric channel** has a chance of being “flipped” to its counterpart ($0 \to 1$; $1 \to 0$)



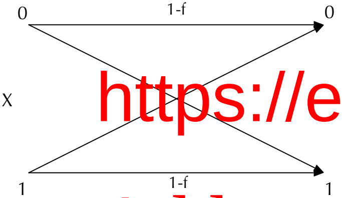
Inputs $\mathcal{X} = \{0, 1\}$; Outputs $\mathcal{Y} = \{0, 1\}$;
Transition probabilities with $P(\text{flip}) = f$

What was most likely transmitted over the channel if s received, assuming $f = 0.1$ and $P(x = 0) = P(x = 1) = 0.5$?

$$\xrightarrow{Q} 0010\,1001$$

The Binary Symmetric Channel

Each symbol sent across a **binary symmetric channel** has a chance of being "flipped" to its counterpart ($0 \rightarrow 1$; $1 \rightarrow 0$)



Inputs $\mathcal{X} = \{0, 1\}$; Outputs $\mathcal{Y} = \{0, 1\}$;
Transition probabilities with $P(\text{flip}) = f$

What was most likely transmitted over the channel if s received, assuming $f = 0.1$ and $P(x = 0) = P(x = 1) = 0.5$?

$$0010\,1001 \xrightarrow{Q} 0010\,1001$$

Inferring the Input

Suppose we know the $P(x)$ — the probability x is transmitted.

Given a pa likely was it that input

Inferring the Input

Suppose we know the $P(x)$ — the probability x is transmitted.

Given a pa likely was it that input

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} = \frac{}{\sum}$$

Inferring the Input: Example

Suppose $P(x = 0) = P(x = 1) = 0.5$. What are the probability that a $x = 0$ was transmitted over a *binary symmetric channel* Q with $f = 0.1$ given that a $y = 0$ was received?

Inferring the Input: Example

Suppose $P(x = 0) = P(x = 1) = 0.5$. What are the probability that a $x = 0$ was transmitted over a *binary symmetric channel* Q with $f = 0.1$ given that a $y = 0$ was received?

$$P(x = 0\ y = 0) = \frac{0.9 \times 0.5}{} = 0.9$$

Similarly

Inferring the Input: Example

Suppose $P(x = 0) = P(x = 1) = 0.5$. What are the probability that a $x = 0$ was transmitted over a *binary symmetric channel* Q with $f = 0.1$ given that a $y = 0$ was received?

$$P(x = 0 \; y = 0) = \frac{0.9 \times 0.5}{} = 0.9$$

Similarly

What if $P(x = 0) = 0.01$?

Inferring the Input: Example

Suppose $P(x=0) = P(x=1) = 0.5$. What are the probability that a $x = 0$ was transmitted over a *binary symmetric channel* Q with $f = 0.1$ given that a $y = 0$ was received?

$$P(x = 0 \; y = 0) = \frac{0.9 \times 0.5}{} = 0.9$$

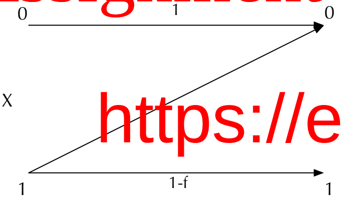
Similarly

What if $P(x=0) = 0.01$?

$$P(x = 0 | y = 0) = \frac{0.9 \times}{0.9 \times 0.01 +}$$

The Z Channel

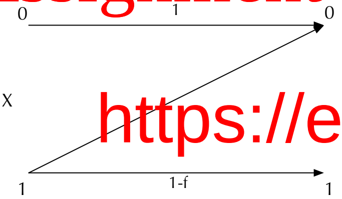
Symbols may be corrupted over the channel asymmetrically.



Inputs $\mathcal{X} = \{0, 1\}$; Outputs $\mathcal{Y} = \{0, 1\}$;
Transition probabilities

The Z Channel

Symbols may be corrupted over the channel asymmetrically.



Inputs $\mathcal{X} = \{0, 1\}$; Outputs $\mathcal{Y} = \{0, 1\}$;
Transition probabilities

Inferring the input: Clearly $P(x = 1|y =$

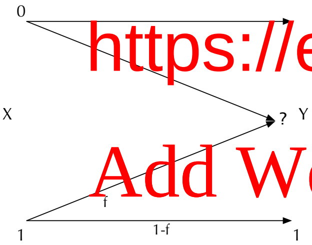
$$P(x = 0|y = 0) = \frac{P(y = 0|x = 0)P(x = 0)}{\sum_{x' \in \mathcal{X}} P(y = 0|x')P(x')} = \frac{P(x = 0)}{P(x = 0) + f\,P(x = 1)}$$

So $P(x = 0|y = 0) \to 1$ as $f \to 0$, and goes to $P(x = 0)$ as $f \to 1$

The Binary Erasure Channel

We can model a channel which “erases” bits by letting one of the output symbols be the symbol ’?’ with associated probability f. The receiver knows which bits are erased.

$\{0, ?, 1\}$;



Example:

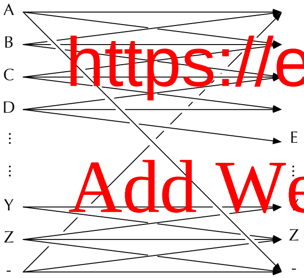
$$0000\,1111 \xrightarrow{Q} 00?0\,?11?$$

The Noisy Typewriter Channel

This channel simulates a noisy "typewriter". Inputs and outputs are 26 letters A through Z plus space. With probability $\frac{1}{3}$, each letter is either: unchanged; changed to the next letter, changed to the previous letter.

The Noisy Typewriter Channel

This channel simulates a noisy “typewriter”. Inputs and outputs are 26 letters A through Z plus space. With probability $\frac{1}{3}$, each letter is either: unchanged; changed to the next letter; changed to the previous letter.

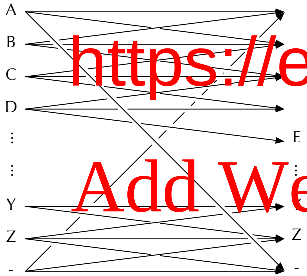


Inputs $= \mathtt{A}, \mathtt{B}, \ldots, \mathtt{Z}, \_\ ;$

$$Q = \begin{bmatrix} & & 0 & 0 & & 0 & \frac{1}{3} \\ & & & & & & \\ & & & & & & \vdots \\ \frac{1}{3} & 0 & 0 & & \cdots & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

The Noisy Typewriter Channel

This channel simulates a noisy “typewriter”. Inputs and outputs are 26 letters A through Z plus space. With probability $\frac{1}{3}$, each letter is either: unchanged; changed to the next letter; changed to the previous letter.



$$\text{Inputs} = \mathtt{A}, \mathtt{B}, \ldots, \mathtt{Z}, \_ \;;$$

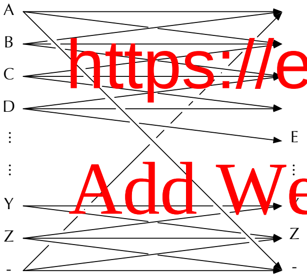
$$Q = \begin{bmatrix} [illegible] & [illegible] & 0 & 0 & & 0 & \frac{1}{3} \\ & & & & & & \vdots \\ \frac{1}{3} & 0 & 0 & & \cdots & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

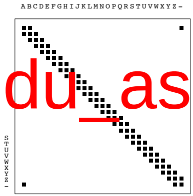
The transition matrix for this channel has a diagonal structure: all of the probability mass is concentrated around the diagonal.

The Noisy Typewriter Channel

This channel simulates a noisy “typewriter”. Inputs and outputs are 26 letters A through Z plus space. With probability $\frac{1}{3}$, each letter is either: unchanged; changed to the next letter, changed to the previous letter.

Inputs $= \mathtt{A}, \mathtt{B}, \ldots, \mathtt{Z}, \_\ ;$





The transition matrix for this channel has a diagonal structure: all of the probability mass is concentrated around the diagonal.

Communicating over Noisy Channels

Suppose we know we have to communicate over some channel Q and we want build an *encoder/decoder* pair to reliably send a message $\mathbf{s}$ over Q.

Communicating over Noisy Channels

Suppose we know we have to communicate over some channel Q and we want build an *encoder/decoder* pair to reliably send a message $\mathbf{s}$ over Q.

Reliability is measured via **probability of err** f
incorrectly decoding $\mathbf{s}_{out}$ given $\mathbf{s}_{in}$ as input:

$$P(\mathbf{s}_{out} \neq \mathbf{s}_{in}) = \sum_{\mathbf{s}} P(\mathbf{s}_{out} \neq \mathbf{s}$$

Probability of Error: Example

Let $\mathcal{S} = \{a, b\}$ and $\mathcal{X} = \mathcal{Y} = \{0, 1\}$

Assume

Conside

$$Q = \begin{bmatrix} 1 - f & f \\ f & 1 - f \end{bmatrix}$$

with $f = 0.1$

Probability of Error: Example

If base probabilities of symbol transmission are $(p_{\mathtt{a}}, p_{\mathtt{b}}) = (0.5, 0.5)$,

$$P(\mathbf{s}_{in} \neq \mathbf{s}_{out}) = P(\mathbf{s}_{in} = \mathtt{a}, \mathbf{s}_{out} = \mathtt{b}) + P(\mathbf{s}_{in} = \mathtt{b}, \mathbf{s}_{out} = \mathtt{a})$$

Probability of Error: Example

If base probabilities of symbol transmission are $(p_{\mathtt{a}}, p_{\mathtt{b}}) = (0.5, 0.5)$,

$$P(\mathbf{s}_{in} \neq \mathbf{s}_{out}) = P(\mathbf{s}_{in} = \mathtt{a}, \mathbf{s}_{out} = \mathtt{b}) + P(\mathbf{s}_{in} = \mathtt{b}, \mathbf{s}_{out} = \mathtt{a})$$

$$= P(\mathbf{s}_{out} = \mathtt{b} | \mathbf{s}_{in} = \mathtt{a}) P(\mathbf{s}_{in} = \mathtt{a}) +$$

Probability of Error: Example

If base probabilities of symbol transmission are $(p_{\mathtt{a}}, p_{\mathtt{b}}) = (0.5, 0.5)$,

$$
\begin{aligned}
P(\mathbf{s}_{in} \neq \mathbf{s}_{out}) &= P(\mathbf{s}_{in} = \mathtt{a}, \mathbf{s}_{out} = \mathtt{b}) + P(\mathbf{s}_{in} = \mathtt{b}, \mathbf{s}_{out} = \mathtt{a}) \\
&= P(\mathbf{s}_{out} = \mathtt{b} \mid \mathbf{s}_{in} = \mathtt{a}) P(\mathbf{s}_{in} = \mathtt{a}) +
\end{aligned}
$$

$\mathtt{b}$

Probability of Error: Example

If base probabilities of symbol transmission are $(p_{\mathtt{a}}, p_{\mathtt{b}}) = (0.5, 0.5)$,

$$P(\mathbf{s}_{in} \neq \mathbf{s}_{out}) = P(\mathbf{s}_{in} = \mathtt{a}, \mathbf{s}_{out} = \mathtt{b}) + P(\mathbf{s}_{in} = \mathtt{b}, \mathbf{s}_{out} = \mathtt{a})$$

$$= P(\mathbf{s}_{out} = \mathtt{b} | \mathbf{s}_{in} = \mathtt{a}) P(\mathbf{s}_{in} = \mathtt{a}) +$$

Probability of Error: Example

If base probabilities of symbol transmission are $(p_{\mathtt{a}}, p_{\mathtt{b}}) = (0.5, 0.5)$,

$$
\begin{aligned}
P(\mathbf{s}_{in} \neq \mathbf{s}_{out}) &= P(\mathbf{s}_{in} = \mathtt{a}, \mathbf{s}_{out} = \mathtt{b}) + P(\mathbf{s}_{in} = \mathtt{b}, \mathbf{s}_{out} = \mathtt{a}) \\
&= P(\mathbf{s}_{out} = \mathtt{b} | \mathbf{s}_{in} = \mathtt{a}) P(\mathbf{s}_{in} = \mathtt{a}) + \\
&= 0.1
\end{aligned}
$$

A Simple Coding Scheme

Suppose $\mathbf{s} \in \{\mathtt{a}, \mathtt{b}\}$ and we encode by $\mathtt{a} \rightarrow \mathtt{000}$ and $\mathtt{b} \rightarrow \mathtt{111}$.
To decode we count the number of 1s and 0s and set all bits to the majority count to determine $\mathbf{s}$

b

A Simple Coding Scheme

Suppose $\mathbf{s} \in \{\mathtt{a}, \mathtt{b}\}$ and we encode by $\mathtt{a} \rightarrow \mathtt{000}$ and $\mathtt{b} \rightarrow \mathtt{111}$.
To decode we count the number of 1s and 0s and set all bits to the majority count to determine $\mathbf{s}$

b

If the chan

$$
\begin{aligned}
P(\mathbf{s}_{in} \neq \mathbf{s}_{out}) &= P(\mathbf{y} \in B | 000)\, p_{\mathtt{a}} \\
&= [f^3 + 3f^2(1-f)] \\
&= f^3 + 3f^2(1-f) = 0.028
\end{aligned}
$$

A Simple Coding Scheme

Suppose $\mathbf{s} \in \{\mathtt{a}, \mathtt{b}\}$ and we encode by $\mathtt{a} \to \mathtt{000}$ and $\mathtt{b} \to \mathtt{111}$.
To decode we count the number of $\mathtt{1}$s and $\mathtt{0}$s and set all bits to the majority count to determine $\mathbf{s}$

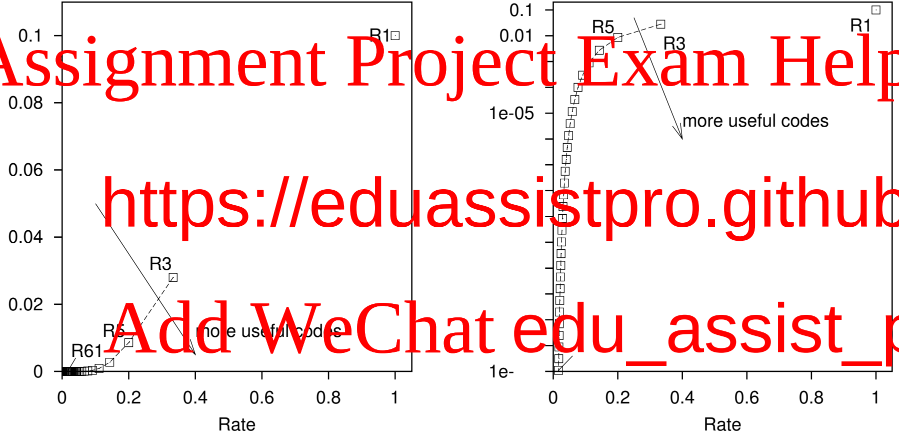
b

If the chan

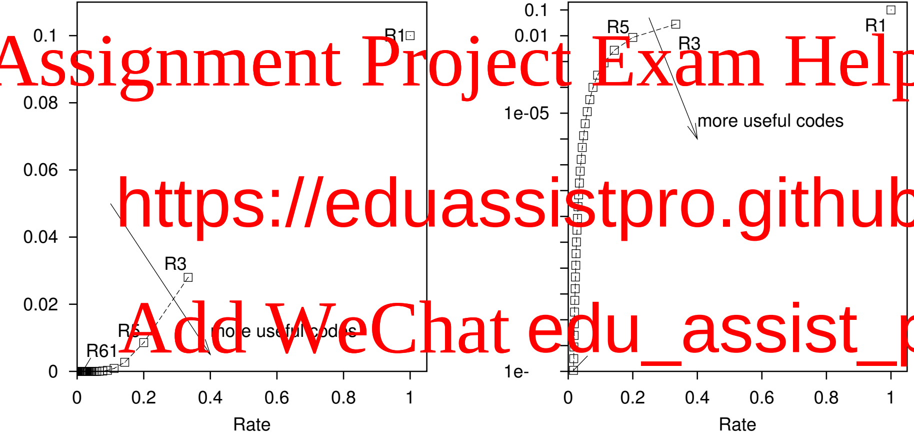
$$
\begin{aligned}
P(\mathbf{s}_{in} \neq \mathbf{s}_{out}) &= P(\mathbf{y} \in B|000)\, p_{\mathtt{a}} \\
&= [f^3 + 3f^2(1-f)] \\
&= f^3 + 3f^2(1-f) = 0.028
\end{aligned}
$$

So the *error* has dropped from 0.1 to 0.028
but so has the *rate*: from 1 symbol/bit to 1/3 symbol/bit.

A Simple Coding Scheme



A Simple Coding Scheme



Can we make the error arbitrarily small without the rate going to zero?

Summary and Reading

Main Points:

- Modelling Noisy Channels
  - Noiseless, Overlap, Symmetric, Z, Erasure
- Sim
  -

Reading:

- MacKay §9.1 - §9.5

- Cover & Thomas §7.1 - §7.3