Channel Coding: Zero-error case

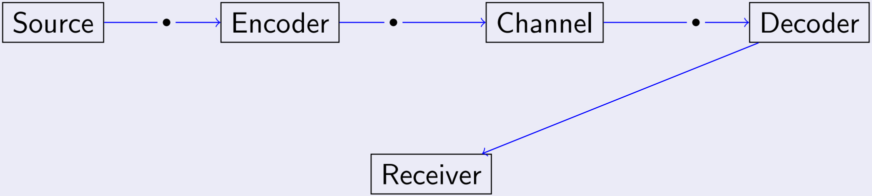
Sander Bet & Ismani Nieuweboer

January 2015

Channel Coding: Zero-error case

Channel Definition

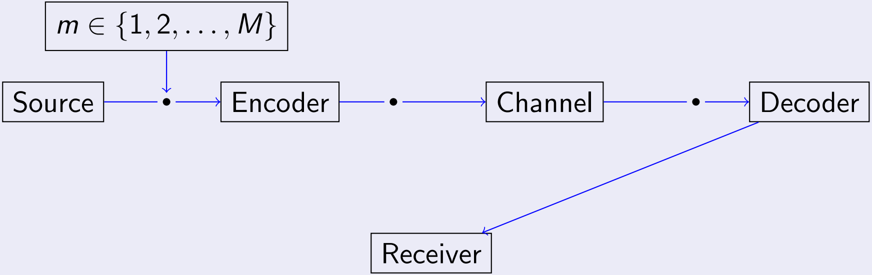
What is a Channel?



Channel Coding: Zero-error case

Channel Definition

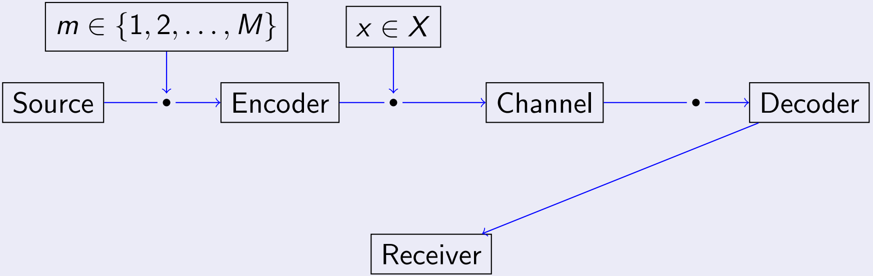
What is a Channel?



Channel Coding: Zero-error case

Channel Definition

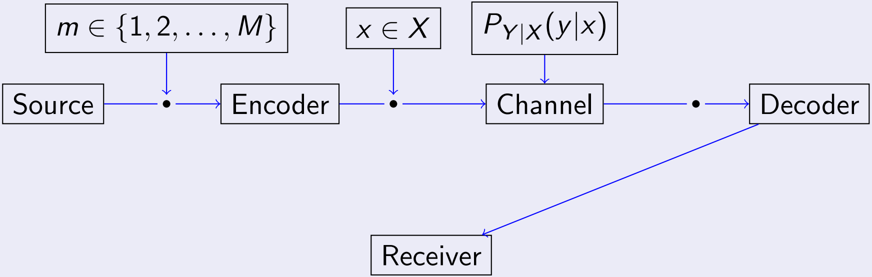
What is a Channel?



Channel Coding: Zero-error case

Channel Definition

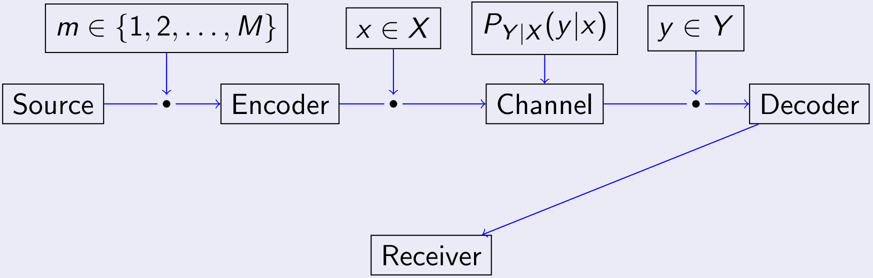
What is a Channel?



Channel Coding: Zero-error case

Channel Definition

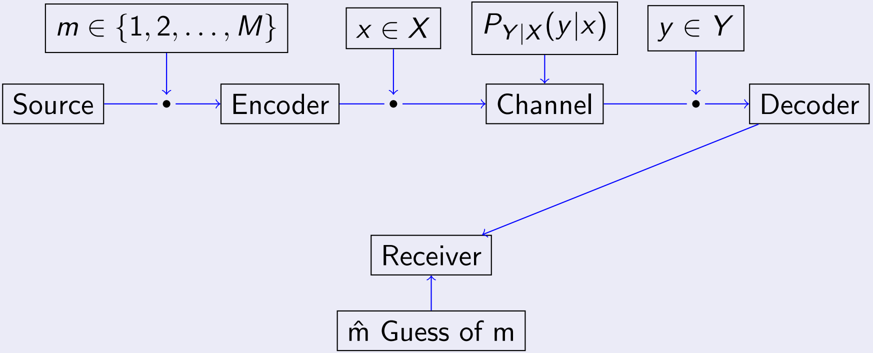
What is a Channel?



Channel Coding: Zero-error case

Channel Definition

What is a Channel?



Channel Coding: Zero-error case

Channel Definition

Definition: Discrete Channel

A discrete channel is denoted by $(\mathcal{X}, P_{Y|X}(y|x), \mathcal{Y})$. Where $\mathcal{X}$ is a finite non-empty input set, $\mathcal{Y}$ a finite output set. And $P_{Y|X}(y|x)$ is a conditional probability distribution that satisfies the following properties;

Channel Coding: Zero-error case

Channel Definition

Definition: Discrete Channel

A discrete channel is denoted by $(\mathcal{X}, P_{Y|X}(y|x), \mathcal{Y})$. Where $\mathcal{X}$ is a finite non-empty input set, $\mathcal{Y}$ a finite output set. And $P_{Y|X}(y|x)$ is a conditional probability distribution that satisfies the following properties;

$$P_{Y|X}(y|x) \geq 0 : \forall x \in \mathcal{X}, \forall y \in \mathcal{Y}$$

$$\sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) = 1 : \forall x \in \mathcal{X}$$

Channel Coding: Zero-error case

Channel Definition

Definition: Discrete Channel

A discrete channel is denoted by $(\mathcal{X}, P_{Y|X}(y|x), \mathcal{Y})$. Where $\mathcal{X}$ is a finite non-empty input set, $\mathcal{Y}$ a finite output set. And $P_{Y|X}(y|x)$ is a conditional probability distribution that satisfies the following properties;

$$P_{Y|X}(y|x) \geq 0 : \forall x \in \mathcal{X}, \forall y \in \mathcal{Y}$$

$$\sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) = 1 : \forall x \in \mathcal{X}$$

Definition: Memory-less Channel

A memory-less channel is a channel the probability distribution $P_{Y|X}(y|x)$ is independent of previous channel inputs and outputs.

Channel Coding: Zero-error case

Channel Definition

Example

$(\mathcal{X} = \{0, 1\},$
$P_{Y|X}(0|0) = p, P_{Y|X}(1|1) = p, P_{Y|X}(1|0) = 1 - p, P_{Y|X}(0|1) = 1 - p,$
$\mathcal{Y} = \{0, 1\})$

Channel Coding: Zero-error case

Channel Definition

Example

$(\mathcal{X} = \{0,1\},$
$P_{Y|X}(0|0) = p, P_{Y|X}(1|1) = p, P_{Y|X}(1|0) = 1-p, P_{Y|X}(0|1) = 1-p,$
$\mathcal{Y} = \{0,1\})$

0

1

Channel Coding: Zero-error case
Channel Definition

Example

$(\mathcal{X} = \{0,1\},$
$P_{Y|X}(0|0) = p, P_{Y|X}(1|1) = p, P_{Y|X}(1|0) = 1-p, P_{Y|X}(0|1) = 1-p,$
$\mathcal{Y} = \{0,1\})$

0 0

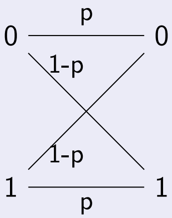
1 1

Channel Coding: Zero-error case

Channel Definition

Example

$(\mathcal{X} = \{0, 1\},$
$P_{Y|X}(0|0) = p, P_{Y|X}(1|1) = p, P_{Y|X}(1|0) = 1 - p, P_{Y|X}(0|1) = 1 - p,$
$\mathcal{Y} = \{0, 1\})$



Channel Coding: Zero-error case

Channel Definition

Multiple uses of a memory-less channel

n uses of the memory-less channel $(\mathcal{X}, P_{Y|X}(y|x), \mathcal{Y})$ corresponds to the memory-less channel $(\mathcal{X}^n, P_{Y^n|X^n}(y^n|x^n), \mathcal{Y}^n)$

Channel Coding: Zero-error case

Channel Definition

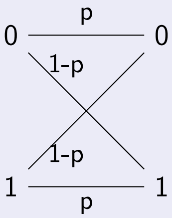
Multiple uses of a memory-less channel

n uses of the memory-less channel $(\mathcal{X}, P_{Y|X}(y|x), \mathcal{Y})$ corresponds to the memory-less channel $(\mathcal{X}^n, P_{Y^n|X^n}(y^n|x^n), \mathcal{Y}^n)$, where $P_{Y^n|X^n}(y^n|x^n) = \prod_i^n P_{Y|X}(y_i|x_i)$.

Channel Coding: Zero-error case

Channel Definition

Example

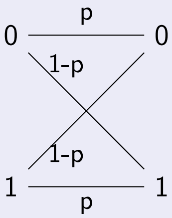


Your input code is $x^2 = 11$. What is the probability that $y^2 = 11$?

Channel Coding: Zero-error case

Channel Definition

Example



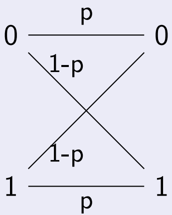
Your input code is $x^2 = 11$. What is the probability that $y^2 = 11$?

$P_{Y^2|X^2}(11|11) = P_{Y|X}(1|1) \cdot P_{Y|X}(1|1) = p^2$

Channel Coding: Zero-error case

Channel Definition

Example



Your input code is $x^2 = 11$. What is the probability that $y^2 = 11$?

$P_{Y^2|X^2}(11|11) = P_{Y|X}(1|1) \cdot P_{Y|X}(1|1) = p^2$

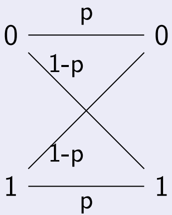
$P_{Y^2|X^2}(01|11) = P_{Y^2|X^2}(10|11) = P_{Y|X}(0|1) \cdot P_{Y|X}(1|1) = (1-p) \cdot p$

$P_{Y^2|X^2}(00|11) = P_{Y|X}(0|1) \cdot P_{Y|X}(0|1) = (1-p)^2$

Channel Coding: Zero-error case

Channel Definition

Example



Your input code is $x^2 = 11$. What is the probability that $y^2 = 11$?

$P_{Y^2|X^2}(11|11) = P_{Y|X}(1|1) \cdot P_{Y|X}(1|1) = p^2$

$P_{Y^2|X^2}(01|11) = P_{Y^2|X^2}(10|11) = P_{Y|X}(0|1) \cdot P_{Y|X}(1|1) = (1-p) \cdot p$

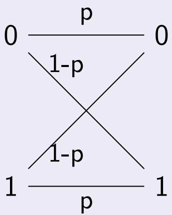
$P_{Y^2|X^2}(00|11) = P_{Y|X}(0|1) \cdot P_{Y|X}(0|1) = (1-p)^2$

Is this a channel?

Channel Coding: Zero-error case

Channel Definition

Example



Your input code is $x^2 = 11$. What is the probability that $y^2 = 11$?
$P_{Y^2|X^2}(11|11) = P_{Y|X}(1|1) \cdot P_{Y|X}(1|1) = p^2$
$P_{Y^2|X^2}(01|11) = P_{Y^2|X^2}(10|11) = P_{Y|X}(0|1) \cdot P_{Y|X}(1|1) = (1-p) \cdot p$
$P_{Y^2|X^2}(00|11) = P_{Y|X}(0|1) \cdot P_{Y|X}(0|1) = (1-p)^2$
Is this a channel? Yes, all the probabilities are positive and
$\sum_{y \in \mathcal{Y}} P_{Y^2|X^2}(y|11) = p^2 + 2 \cdot p(1-p) + (1-p)^2 = (p + (1-p))^2 = 1$

Channel Coding: Zero-error case

Code Definition

Definition: (M, n)-code

A (M, n)-code for channel $(\mathcal{X}, P_{Y|X}(y|x), \mathcal{Y})$ with;

Channel Coding: Zero-error case

Code Definition

Definition: (M, n)-code

A (M, n)-code for channel $(\mathcal{X}, P_{Y|X}(y|x), \mathcal{Y})$ with;

A message index set $\{1, 2, \ldots, M\}$

Channel Coding: Zero-error case

Code Definition

Definition: (M, n)-code

A (M, n)-code for channel $(\mathcal{X}, P_{Y|X}(y|x), \mathcal{Y})$ with;

- A message index set $\{1, 2, \ldots, M\}$
- An encoding function e: $\{1, 2, \ldots, M\} \rightarrow \mathcal{X}^n$

Channel Coding: Zero-error case

Code Definition

Definition: (M, n)-code

A (M, n)-code for channel $(\mathcal{X}, P_{Y|X}(y|x), \mathcal{Y})$ with;

A message index set $\{1, 2, \ldots, M\}$

An encoding function e: $\{1, 2, \ldots, M\} \rightarrow \mathcal{X}^n$

A decoding function d: $\mathcal{Y}^n \rightarrow \{1, 2, \ldots, M\}$

Definition: Transmission Rate

The transmission rate R of a (M, n)-code is $R = \frac{\log M}{n}$.

Zero-error Problem

Given a channel how many bits of information can we send through it without any errors?

Channel Coding: Zero-error case

Zero-error Problem

Zero-error Problem

Given a channel how many bits of information can we send through it without any errors? In other words what is the maximal transmission rate of the channel?

Channel Coding: Zero-error case

Zero-error Problem

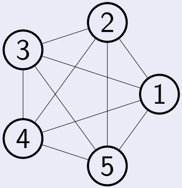
Zero-error Problem

Given a channel how many bits of information can we send through it without any errors? In other words what is the maximal transmission rate of the channel? We use graph theory to clarify the problem.

Graph theory

Graph definition

A graph G is a set of vertices $V(G)$ and a set of edges $E(G)$. Example (full graph with 5 vertices):



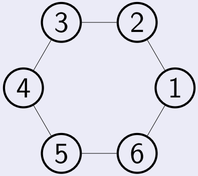
$V = \{1, 2, 3, 4, 5\}$ and $E = \{12, 13, 14, \ldots, 45\}$

Graph theory

Independent set

Definition

For a graph G, an independent set is a set of vertices $I \subset V(G)$ such that no edge $e \in E(G)$ contains two vertices from I



I can be $\{1\}, \{2\}, \{1, 2\}, \{1, 3, 5\}$, etc.

Graph theory

Independent set

Definition

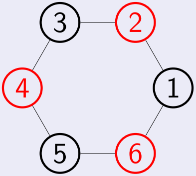
The independence number $\alpha(G)$ is the cardinality of the maximum independent set.

Graph theory

Independent set

Definition

The independence number $\alpha(G)$ is the cardinality of the maximum independent set.

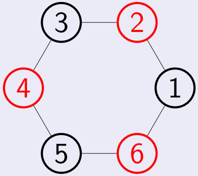


Graph theory

Independent set

Definition

The independence number $\alpha(G)$ is the cardinality of the maximum independent set.



$\alpha(G) = |\{1, 3, 5\}| = |\{2, 4, 6\}| = 3$

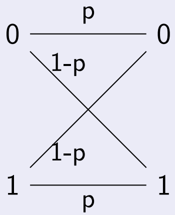
Confusability graph

Definition

Given a discrete channel $(\mathcal{X}, P_{Y|X}(y|x), \mathcal{Y})$, the confusability graph G is defined by $V(G) = \mathcal{X}$ and
$E(G) = \{vw : \exists y : P_{Y|X}(y, v) \neq 0 \wedge P_{Y|X}(y|w) \neq 0\}$ i.e. vertices are connected when they can get confused with each other

Confusability graph

Example



Zero-error codes

Given a (discrete memoryless) channel, how much information can you perfectly send through it?

Zero-error codes

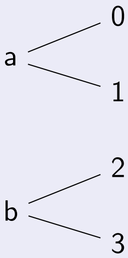
Given a (discrete memoryless) channel, how much information can you perfectly send through it?

When using the channel once, the independence number $\alpha(G)$ of the confusability graph G tells you the maximum rate: $R = \log \alpha(G)$.

Zero-error codes

An ideal situation

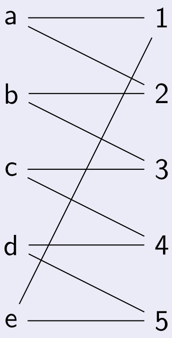
If there is no overlapping output, the maximum independent set is $\mathcal{X}$ itself:



Zero-error codes

Noisy typewriter

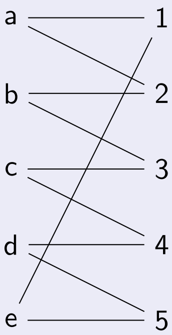
A more interesting case:



Zero-error codes

Noisy typewriter

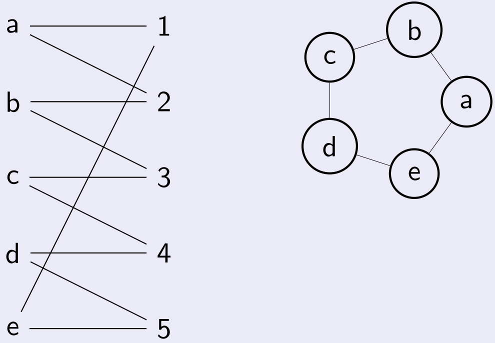
A more interesting case:



What is the confusability graph of this code?

Zero-error codes

Noisy typewriter



Zero-error codes

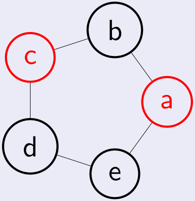
Noisy typewriter

If the channel is used once ($n = 1$) the independence number is $\alpha(G) = 2$:

Zero-error codes

Noisy typewriter

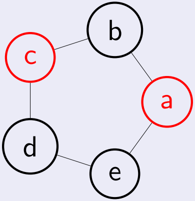
If the channel is used once ($n = 1$) the independence number is $\alpha(G) = 2$:



Zero-error codes

Noisy typewriter

If the channel is used once ($n = 1$) the independence number is $\alpha(G) = 2$:

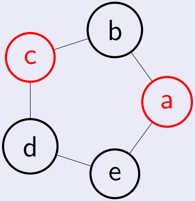


Index set: $\{1, 2\}$

Zero-error codes

Noisy typewriter

If the channel is used once ($n = 1$) the independence number is $\alpha(G) = 2$:

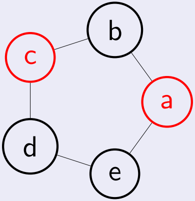


Index set: $\{1, 2\}$ Encoding function: $e(1) = a$ and $e(2) = c$

Zero-error codes

Noisy typewriter

If the channel is used once ($n = 1$) the independence number is $\alpha(G) = 2$:

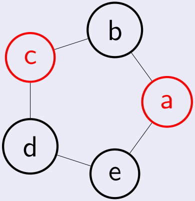


Index set: $\{1, 2\}$ Encoding function: $e(1) = a$ and $e(2) = c$ Decoding function: $d(1) = d(2) = a$ and $d(3) = d(4) = c$

Zero-error codes

Noisy typewriter

If the channel is used once ($n = 1$) the independence number is $\alpha(G) = 2$:



Index set: $\{1, 2\}$ Encoding function: $e(1) = a$ and $e(2) = c$ Decoding function: $d(1) = d(2) = a$ and $d(3) = d(4) = c$ So rate $R = \frac{\log M}{n} = \frac{\log 2}{1} = 1$ bit.

Zero-error codes

Noisy typewriter

What if $n = 2$?

Zero-error codes

Noisy typewriter

What if $n = 2$?
We can do just as well using $\{aa, ac, ca, cc\}$; still no overlap.
Then again $R = \frac{\log M}{n} = \frac{\log 4}{2} = 1$.

Zero-error codes

Noisy typewriter

What if $n = 2$?
We can do just as well using $\{aa, ac, ca, cc\}$; still no overlap.
Then again $R = \frac{\log M}{n} = \frac{\log 4}{2} = 1$.
Claim: There exists a code with index set $M = 5$.

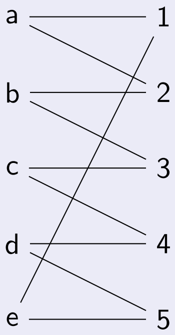
Zero-error codes

Noisy typewriter

What if $n = 2$?
We can do just as well using $\{aa, ac, ca, cc\}$; still no overlap.
Then again $R = \frac{\log M}{n} = \frac{\log 4}{2} = 1$.
Claim: There exists a code with index set $M = 5$. $\{aa, bc, ce, db, ed\}$

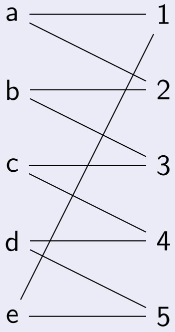
Zero-error codes

Noisy typewriter



Zero-error codes

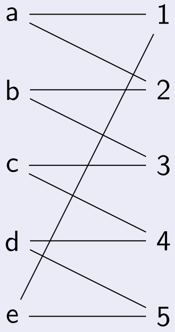
Noisy typewriter



$$aa \to \{11, 12, 21, 22\}$$

Zero-error codes

Noisy typewriter

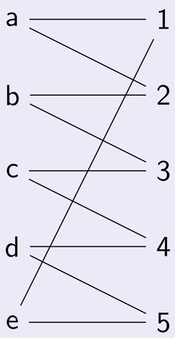


$$aa \rightarrow \{11, 12, 21, 22\}$$

$$bc \rightarrow \{23, 24, 33, 34\}$$

Zero-error codes

Noisy typewriter



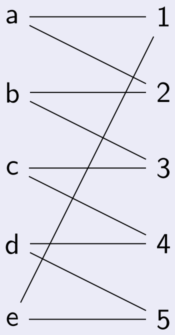
$$aa \rightarrow \{11, 12, 21, 22\}$$

$$bc \rightarrow \{23, 24, 33, 34\}$$

$$ce \rightarrow \{35, 31, 45, 41\}$$

Zero-error codes

Noisy typewriter



$$aa \to \{11, 12, 21, 22\}$$

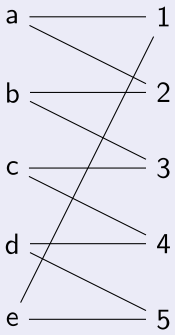
$$bc \to \{23, 24, 33, 34\}$$

$$ce \to \{35, 31, 45, 41\}$$

$$db \to \{42, 43, 52, 53\}$$

Zero-error codes

Noisy typewriter



$aa \to \{11, 12, 21, 22\}$

$bc \to \{23, 24, 33, 34\}$

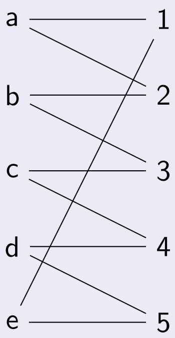
$ce \to \{35, 31, 45, 41\}$

$db \to \{42, 43, 52, 53\}$

$ed \to \{54, 55, 14, 15\}$

Zero-error codes

Noisy typewriter



$$aa \to \{11, 12, 21, 22\}$$

$$bc \to \{23, 24, 33, 34\}$$

$$ce \to \{35, 31, 45, 41\}$$

$$db \to \{42, 43, 52, 53\}$$

$$ed \to \{54, 55, 14, 15\}$$

$$\implies R = \frac{\log 5}{2} > \frac{\log 4}{2} = 1$$

Zero-error codes

Noisy typewriter

$R = \frac{\log 5}{2}$ turns out to be the upper bound for this channel.

Zero-error codes

Noisy typewriter

$R = \frac{\log 5}{2}$ turns out to be the upper bound for this channel.
Proved by Shannon (1956) and Lovasz (1979).

Zero-error codes

Noisy typewriter

$R = \frac{\log 5}{2}$ turns out to be the upper bound for this channel.
Proved by Shannon (1956) and Lovasz (1979).
What happens for bigger graphs? No one knows...

Extra

Multiple channel confusability

We saw an increase in transmission rate, when we used the channel multiple times.

Extra

Multiple channel confusability

We saw an increase in transmission rate, when we used the channel multiple times.It could be helpful to make a graph of the confusability of multiple uses of the channel.

Extra

Multiple channel confusability

We saw an increase in transmission rate, when we used the channel multiple times.It could be helpful to make a graph of the confusability of multiple uses of the channel. The messages are confusable when all the uses are confusable.

Extra

Multiple channel confusability

We saw an increase in transmission rate, when we used the channel multiple times.It could be helpful to make a graph of the confusability of multiple uses of the channel. The messages are confusable when all the uses are confusable.

Reference

Lecture content of Information Theory given by Christian Schaffner (course at University of Amsterdam):
`http://homepages.cwi.nl/~schaffne/courses/inftheory/2014/`
(Blackboard photos from 24 and 26 November)
Many thanks to Christian for letting us use his lecture content!