Communicate Without Errors

Zifan Hua

November 23, 2022

Table Of Contents

Table of Contents

Introduction

- Channel Capacity
- Zero Error Rate
- Shannon Capacity
- C_n
- $\sqrt{5} \leq \Theta(C_5) \leq 5/2$
- $\Theta(G) = \sqrt{5}$ by Lovász László

Table of Contents

Channel

Definition (channel)

A channel has a sender and a receiver.

Channel

Definition (channel)

A channel has a sender and a receiver.
A message is a finite sequence of characters. The sender sends a message to the receiver.

Channel

Definition (channel)

A channel has a sender and a receiver.
A message is a finite sequence of characters. The sender sends a message to the receiver.

The receiver receives message and decode it.

Channel

Definition (channel)

A channel has a sender and a receiver.
A message is a finite sequence of characters. The sender sends a message to the receiver.

The receiver receives message and decode it.
However, in the procedure of send and receive, the channel may introduce some errors. For instance, here, character b is decoded into c.

Channel

Definition (confusable)

Given two characters a, b. If a and b have chance to be decoded into a same character say c, we say a and b are confusable.
For a two messages of length n, say $a_1 a_2 \ldots a_n$, $b_1 b_2 \ldots b_n$ is confusable if and only if a_i and b_i are confusable for every i.

Channel

Definition (confusable)

Given two characters a, b. If a and b have chance to be decoded into a same character say c, we say a and b are confusable.
For a two messages of length n, say $a_1a_2\dots a_n$, $b_1b_2\dots b_n$ is confusable if and only if a_i and b_i are confusable for every i.

Definition (rate of channel)

The rate of channel actually represent how many distinct character can be send per unit time.
Given a channel that could send r distinct characters per unit time. And send message with length n, so the number of distinct messages the channel can send is r^n.

Channel

Definition (confusable)

Given two characters a, b. If a and b have chance to be decoded into a same character say c, we say a and b are confusable.
For a two messages of length n, say $a_1a_2\dots a_n$, $b_1b_2\dots b_n$ is confusable if and only if a_i and b_i are confusable for every i.

Definition (rate of channel)

The rate of channel actually represent how many distinct character can be send per unit time.
Given a channel that could send r distinct characters per unit time. And send message with length n, so the number of distinct messages the channel can send is r^n.
Conversely, given a channel that could send m distinct messages with length n. And could be able to send only one character per unit time. The rate of channel is $\sqrt[n]{m}$.
Here, we do not care about how much characters. We only care about the number of all messages we can send.

Channel

Definition (zero error rate)

Given a channel which could transfer messages of length n.
We want to find the maximum set of messages M that no two of them is confusable.

$$\text{zero error rate} = \max_{M} \sqrt[n]{|M|} \tag{1}$$

Channel

If given a set of characters S, and some of the characters could be confusable.
We want to find

$$\sup_n \{\text{zero error rate of channel with length } n \text{ and } S \text{ as characters}\} \tag{2}$$

And we call this the shannon Capacity and denoted by $\Theta(S)$.
Clearly, shannon Capacity is actually a function of the set of characters. So, we want a more abstract way to represent the characters.

Graph Representation Of Characters

Definition (graph)

A graph G is a set of vertices with a set of edges connecting pairs of vertices.

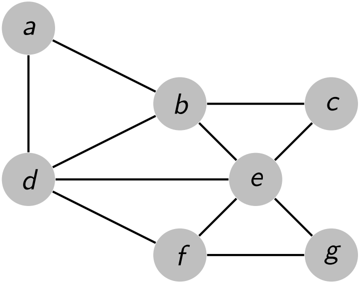


Figure: An example of a graph.

Graph Representation Of Characters

Definition

Given a channel that sending $\{1, 2, \ldots, n\}$ as characters. And some characters i and j could be confused with each other.
Then the graph representation of the characters is the graph with vertices $\{1, 2, \ldots, n\}$ and edges (i, j) if and only if i and j could be confused with each other.

Accordingly, there is the corresponding way that using graphs to represent a message zero error rate and Shannon Capacity.

Product Graph

Definition (graph product)

The product of two graphs can be considered as send a pair of characters (x, y) as one message. So, we have a channel that send messages of length 2.

Product Graph

Definition (graph product)

The product of two graphs can be considered as send a pair of characters (x, y) as one message. So, we have a channel that send messages of length 2.
Recall that before, 2 messages is confusable means that they can be decoded into the same message, which means every characters the two channel use need to be confusable.
Given two graph G and H. The graph product $G \times H$ is the graph with vertices $V(G \times H) = V(G) \times V(H)$ in which (x, y) is adjacent to (x', y') in $G \times H$ if and only if x is adjacent to x' in G and y is adjacent to y' in H.

Product Graph

Definition (graph product)

The product of two graphs can be considered as send a pair of characters (x, y) as one message. So, we have a channel that send messages of length 2.
Recall that before, 2 messages is confusable means that they can be decoded into the same message, which means every characters the two channel use need to be confusable.
Given two graph G and H. The graph product $G \times H$ is the graph with vertices $V(G \times H) = V(G) \times V(H)$ in which (x, y) is adjacent to (x', y') in $G \times H$ if and only if x is adjacent to x' in G and y is adjacent to y' in H.
A graph G product itself for n times will always be denoted by G^n. Which means we could use G^n to represent messages of a channel with length n.

$\alpha(G)$

Definition ($\alpha(G)$)

The $\alpha(G)$ represent the maximum number of characters or message that could not be confused with each other in graph G.
If G represent a set of messages, $\alpha(G)$ is just the zero error rate we have defined before.

$\alpha(G)$

Definition ($\alpha(G)$)

The $\alpha(G)$ represent the maximum number of characters or message that could not be confused with each other in graph G. If G represent a set of messages, $\alpha(G)$ is just the zero error rate we have defined before.

Given a graph G. Given a subgraph H of G, such that every vertex of H is not connected in G. Then $\alpha(G)$ is the maximum number of vertices of H.

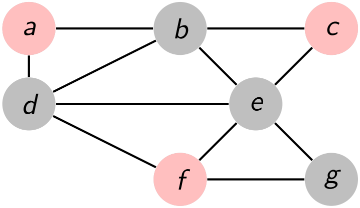


Figure: Example of $\alpha(G)$. Here $\alpha(G) = 3$.

Shannon Capacity

Definition (Shannon Capacity)

Recall that Shannon Capacity is

$$\sup_n \{\text{zero error rate of channel with length } n \text{ and } S \text{ as characters}\} \tag{3}$$

Use the graph G to represent the set of characters, the Shannon capacity $\Theta(G)$ is defined by

$$\Theta(G) = \sup_n \sqrt[n]{\alpha(G^n)} \tag{4}$$

Table of Contents

To get the final result, we still need some more tools.

Properties of $\alpha(G)$

Lemma

$\alpha(G)\alpha(H) \leq \alpha(G \times H)$

Properties of $\alpha(G)$

Lemma

$\alpha(G)\alpha(H) \leq \alpha(G \times H)$

Proof.

Given graph G and H. Let G' and H' be subgraph of G and H such that no vertex of G' or H' is adjacent in G or H, respectively. Then $G' \times H'$ is a subgraph of $G \times H$ such that no vertex of $G' \times H'$ is adjacent in $G \times H$. $\square$

Orthonormal Representation

Here we have the third way to defined a set of characters.

Definition (Orthonormal Representation)

Given a graph G with vertices $1, 2, \ldots, n$, the orthonormal representation of G is a set of unit vectors $\{v_1, v_2, \ldots, v_n\}$ such that v_i and v_j are orthogonal if and only if i and j are not adjacent in G.

Orthonormal Representation

Here we have the third way to defined a set of characters.

Definition (Orthonormal Representation)

Given a graph G with vertices $1, 2, \ldots, n$, the orthonormal representation of G is a set of unit vectors $\{v_1, v_2, \ldots, v_n\}$ such that v_i and v_j are orthogonal if and only if i and j are not adjacent in G.

This existence of the orthonormal representation can be proved by induction.

Tensor Product

Definition (tensor product)

Given two vectors $v = (v_1, \ldots, v_n)$ and $w = (w_1, \ldots, w_n)$, the tensor product $v \circ w$ is defined by

$$v \circ w = (v_1 w_1, \ldots, v_1 w_n, v_2 w_1, \ldots, v_2 w_n, \ldots, v_n w_1, \ldots, v_n w_n) \quad (5)$$

Tensor Product

Definition (tensor product)

Given two vectors $v = (v_1, \ldots, v_n)$ and $w = (w_1, \ldots, w_n)$, the tensor product $v \circ w$ is defined by

$$v \circ w = (v_1 w_1, \ldots, v_1 w_n, v_2 w_1, \ldots, v_2 w_n, \ldots, v_n w_1, \ldots, v_n w_n) \tag{5}$$

Lemma

The inner product of tensor products can be computed by,

$$\langle v \circ w, v' \circ w' \rangle = \langle v, v' \rangle \langle w, w' \rangle \tag{6}$$

Product of Orthonormal Representation

This is the product of graph in the sense of orthonormal representation.

Lemma

Given a graph G with vertices $1, 2, \ldots, n$, and a graph H with vertices $1, 2, \ldots, m$. Then vectors $\{v_i \circ w_j\}$ is an orthonormal representation of $G \times H$.

Theta Function

Given a graph G, the $\theta(G)$ is defined by

$$\theta(G) = \inf_{\{v_1, v_2, \ldots, v_n\}, c} \max_i \frac{1}{\langle c, v_i \rangle^2} \tag{7}$$

where $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal representation of G, and c is any unit vector does not orthogonal to v_i.

Theta Function

Given a graph G, the $\theta(G)$ is defined by

$$\theta(G) = \inf_{\{v_1, v_2, \dots, v_n\}, c} \max_i \frac{1}{\langle c, v_i \rangle^2} \tag{7}$$

where $\{v_1, v_2, \dots, v_n\}$ is an orthonormal representation of G, and c is any unit vector does not orthogonal to v_i.

Lemma

There always exist such an c and orthonormal representation $\{v_1, v_2, \dots, v_n\}$ *such that*

$$\theta(G) = \max_i \frac{1}{\langle c, v_i \rangle^2} \tag{8}$$

This could be proved by proving the set of all possible cases of $\{v_1, v_2, \dots, v_n, c\}$ is compact. And the function $\max \frac{1}{\langle c, v_i \rangle^2}$ is continuous.

Theta Function

Lemma

Given graph G and H, then

$$\theta(G \times H) \leq \theta(G)\theta(H) \tag{9}$$

Theta Function

Lemma

Given graph G and H, then

$$\theta(G \times H) \leq \theta(G)\theta(H) \tag{9}$$

Proof.

Let $\{v_1, v_2, \ldots, v_n\}$ and $\{w_1, w_2, \ldots, w_m\}$ be orthonormal representation of G and H and c_v and c_w such that

$$\max\left\{\frac{1}{\langle c_v, v_i\rangle^2} : i = 1, 2, \ldots, n\right\} = \theta(G) \tag{10}$$

and

$$\max\left\{\frac{1}{\langle c_w, w_i\rangle^2} : i = 1, 2, \ldots, m\right\} = \theta(H) \tag{11}$$

$\square$

Proof.

Then

$$\theta(G \times H) \leq \max\left\{\frac{1}{\langle c_v \circ c_w, v_i \circ w_j\rangle^2}\right\} \tag{12}$$

$$= \max\left\{\frac{1}{\langle c_v, v_i\rangle^2 \langle c_w, w_j\rangle^2}\right\} \tag{13}$$

$$= \max\left\{\frac{1}{\langle c_v, v_i\rangle^2}\right\} \max\left\{\frac{1}{\langle c_w, w_j\rangle^2}\right\} \tag{14}$$

$$= \theta(G)\theta(H) \tag{15}$$

$$\tag{16}$$

$\square$

This lemma relates the $\theta(G)$ and $\alpha(G)$.

Lemma

$$\theta(G) \geq \alpha(G) \tag{17}$$

This lemma relates the $\theta(G)$ and $\alpha(G)$.

Lemma

$$\theta(G) \geq \alpha(G) \tag{17}$$

Proof.

Let $\{1, 2, \ldots, k\}$ be the set of vertices of G such that every point is not adjacency in G. And $k = \alpha(G)$

Let $\{v_1, v_2, \ldots, v_n\}$ be an orthonormal representation of G and c such that

$$\max\left\{\frac{1}{\langle c, v_i\rangle^2}\right\} = \theta(G) \tag{18}$$

Then

$$1 \quad = \quad c^2 \geq \sum_{i=1}^{k} \langle c, v_i\rangle^2 \geq \frac{k}{\theta(G)} \tag{19}$$

$\square$

Theorem

Given a graph G, then

$$\theta(G) \geq \Theta(G) \tag{20}$$

Theorem

Given a graph G, then

$$\theta(G) \geq \Theta(G) \tag{20}$$

Proof.

$$\begin{aligned}
\Theta(G) &= \lim_n \sqrt[n]{\alpha(G^n)} && (21)\\
&\leq \lim_n \sqrt[n]{\theta(G^n)} && (22)\\
&\leq \lim_n \sqrt[n]{\theta(G)^n} && (23)\\
&= \theta(G) && (24)
\end{aligned}$$

$\square$

The final result we want to prove today,

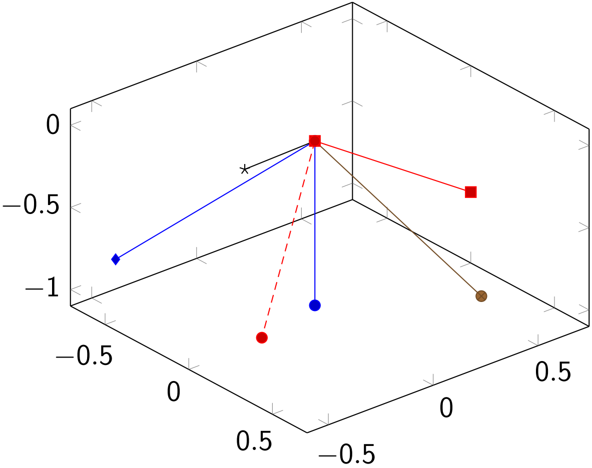
$$\Theta(C_5) \leq \theta(C_5) \leq \sqrt{5} \tag{25}$$

Theorem

$$\theta(C_n) \leq \sqrt{5} \tag{26}$$

Proof.

Consider an umbrella that has a handle and 5 ribs that all have unit length. And also its handle is a unit vector.



□

Proof.

Here, angles between two consecutive ribs are same.
Let v be a rib. And let w be one of the rib that have the largest angle with v. Then, we let the angle between v and w be $\pi/2$.
Then, the 5 ribs of such an umbrella form an orthonormal representation of C_5.
Let d be the vector represent the handle, and $v_1, v_2, \ldots, v_5$ be the 5 ribs. And let γ be the angle between the handle and any rib.
So, by some calculation, we get

$$\begin{aligned}\theta(C_5) &\leq \max \frac{1}{< d, v_i >^2} && (27)\\ &= \left(\frac{1}{\cos(\gamma)}\right)^2 && (28)\\ &= \sqrt{5} && (29)\end{aligned}$$

□

Table of Contents

Conclusion

- We have proved that $Theta(C_5) = \sqrt{5}$
- We could actually proved that $\Theta(C_n)$ is equal to $n\frac{\cos(\pi/n)}{1+\cos(\pi/n)}$.

Open Questions

- Although $\Theta(C_n)$ is equal to $n\frac{\cos(\pi/n)}{1+\cos(\pi/n)}$, but we still don't know the exact value of $\Theta(C_n)$. Even for $n = 7$
- Is there any good lower bound for $\Theta(C_n)$?
- Is there any patterns for n such that $\Theta(c_n)$ is hard to compute?

Discussion

In the real world cases, we always have some kind of relay between the sender and receiver. So the new channel is kind of composite of two channel. Can we compute Shannon Capacity these channels independently and then combine them together to get the Shannon Capacity of the new Channel?

Table of Contents

Reference

- On the Shannon Capacity of a Graph by Laszlo Lovasz
- The zero error capacity of a noisy channel by Claude Shannon