Communicate Without Errors

Zifan Hua

March 28, 2023

Table Of Contents

Table of Contents

Introduction

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

Table of Contents

Channel

Definition (channel)

Characters is a finite set.

A message is a finite sequence of characters.

A channel has a sender and a receiver. The sender sends a message to the receiver.

Channel

Definition (channel)

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

The receiver receives message and decode it.

Channel

Definition (channel)

Characters is a finite set.
A message is a finite sequence of characters.
A channel has a sender and a receiver. 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.

Confusable

Definition (confusable)

Given two distinct 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 distinct 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** or **same** for every i.

Rate of Channel

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 for n unit time, so the number of distinct messages the channel can send is r^n.

Rate of Channel

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 for n unit time, so the number of distinct messages the channel can send is r^n.

Conversely, given a channel that could send m **distinct messages** in n **unit time**. Then the rate of channel is $\sqrt[n]{m}$.

Zero Error Rate of Channel

Definition (zero error rate)

Given a channel that could send messages in n unit time.
We want to find the maximum set of messages M that **no two of them is confusable**. Which means if we send these messages then, there is no chance of get an error.

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

Zero Error Rate of Channel

Definition

If given a set of characters S, and some of the characters could be confusable.
In the reality, we usually do not fix n, so the thing we really want to find is

$$\sup_n \{\text{zero error rate of channel send messages for } n \text{ unit time}\} \tag{2}$$

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.

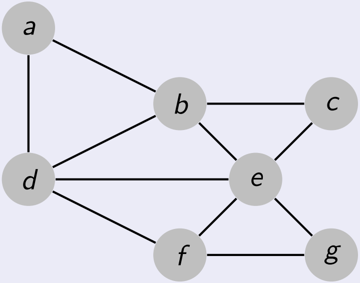


Figure: An example of a graph.

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 is **distinct** and could be **confused with each other**.

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

Product Graph

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 for 2 unit time.

Product Graph

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 for 2 unit time.

Recall that before, 2 distinct 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 or the same.

Definition (graph product)

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' or the **same** in G and y is **adjacent** to y' or the **same** in H.

Product Graph

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 for 2 unit time.

Recall that before, 2 distinct 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 or the same.

Definition (graph product)

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' or the **same** in G and y is **adjacent** to y' or the **same** 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 that send messages for n unit time.

$\alpha(G)$

Definition ($\alpha(G)$)

Given a finite graph G. $\alpha(G)$ is the **maximum** number of vertices such that every two of them is not adjacent in G.
If G represent a set of characters, $\sqrt[n]{\alpha(G^n)}$ is just the zero error rate we have defined before.

$\alpha(G)$

Definition ($\alpha(G)$)

Given a finite graph G. $\alpha(G)$ is the **maximum** number of vertices such that every two of them is not adjacent in G.
If G represent a set of characters, $\sqrt[n]{\alpha(G^n)}$ is just the zero error rate we have defined before.

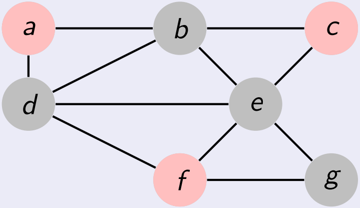


Figure: Example of $\alpha(G)$. Here $\alpha(G) = 3$.

Shannon Capacity

Recall that Shannon Capacity is

$$\sup_n \{\text{zero error rate of channel send messages for } n \text{ unit time}\} \tag{3}$$

Definition (Shannon Capacity)

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

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

Introduction of the Problem

We use C_n to represent a graph look like regular polygons with n vertices.

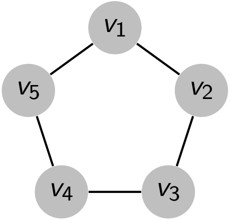


Figure: Example of C_5

Introduction of the Problem

We use C_n to represent a graph look like regular polygons with n vertices.

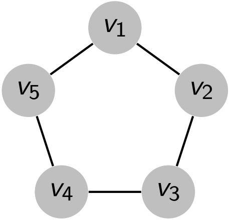


Figure: Example of C_5

Calculate the Shannon Capacity of C_n which seems to be a simple graph, turned out to be a fairly hard problem.
In 1979, L Lovász give the proof that $\Theta(C_5) = \sqrt{5}$ which is about 20 years after Shannon's paper.
And $\Theta(C_7)$ is still unknown.

Table of Contents

Lower Bound of $\Theta(C_5)$

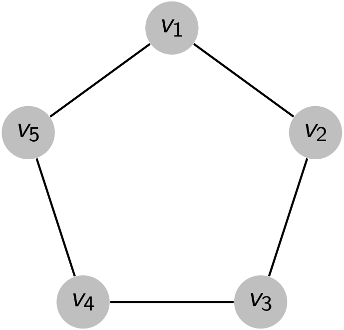


Figure: C_5

The $\alpha(C_5)$ is clearly equal to 2.

Lower Bound of $\Theta(C_5)$

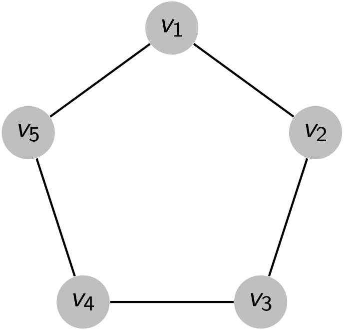


Figure: C_5

The $\alpha(C_5)$ is clearly equal to 2.
We want to show that $\alpha((C_5)^2) \geq 5$.
Thus $\Theta(C_5) = \sup \sqrt[n]{(C_5)^n} \geq \sqrt{\alpha((C_5)^2)} \geq \sqrt{5}$

Lower Bound of $\Theta(C_5)$

Figure: $(C_5)^2$

Lower Bound of $\Theta(C_5)$

It is messy to draw edges in the graph of $(C_5)^2$. So we use the following graph to give some idea of $(C_5)^2$. But it is clear that every point is adjacent to the 8 points around it.

Point v_{ij} represent the vertex (v_i, v_j) in $(C_5)^2$.

The green points are the same with the corresponding red points, we use it just for the convenience of visualization.

We will choose five points v_{11}, v_{23}, v_{35}, v_{42}, v_{54}, and show that they are mutually not adjacent.

Lower Bound of $\Theta(C_5)$

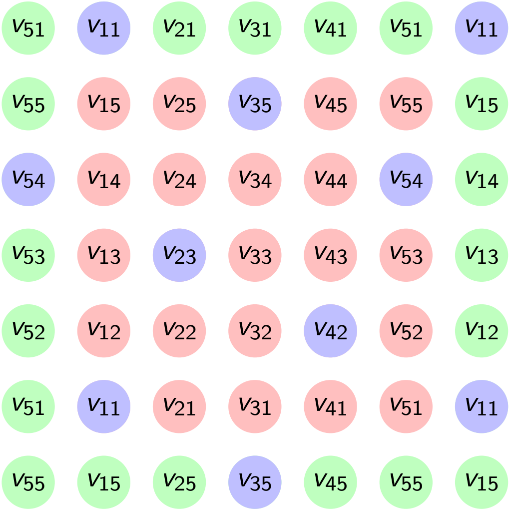


Figure: $(C_5)^2$

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 if i and j are **not adjacent** in G then v_i and v_j are **orthogonal**.

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 if i and j are **not adjacent** in G then v_i and v_j are **orthogonal**.

Tensor Product

Definition (tensor product)

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

$$v \circ w = (v_1 w_1, v_1 w_2, \ldots, v_1 w_n, v_2 w_1, \ldots, v_2 w_n, \ldots, v_n w_1, \ldots, v_n w_n)^T \quad (5)$$

Somehow equal to stack up the columns of vw^T.

Tensor Product

Definition (tensor product)

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

$$v \circ w = (v_1 w_1, v_1 w_2, \ldots, v_1 w_n, v_2 w_1, \ldots, v_2 w_n, \ldots, v_n w_1, \ldots, v_n w_n)^T \quad (5)$$

Somehow equal to stack up the columns of vw^T.

Lemma

The inner product of tensor products can be computed by,

$$\langle v \circ w, v' \circ w' \rangle = \langle v, v' \rangle \langle w, w' \rangle \quad (6)$$

Product of Orthonormal Representation

Lemma

Given a graph G with vertices $1, 2, \ldots, n$, and a graph H with vertices $1, 2, \ldots, m$. Let $\{v_1, v_2, \ldots, v_n\}$ and $\{w_1, w_2, \ldots, w_m\}$ be orthonormal representations of G and H respectively. Then vectors $\{v_i \circ w_j\}$ is an orthonormal representation of $G \times H$. If the vector $v_i w_j$ correspond to the vertex (i, j) in $G \times H$

Product of Orthonormal Representation

Lemma

Given a graph G with vertices $1, 2, \ldots, n$, and a graph H with vertices $1, 2, \ldots, m$. Let $\{v_1, v_2, \ldots, v_n\}$ and $\{w_1, w_2, \ldots, w_m\}$ be orthonormal representations of G and H respectively. Then vectors $\{v_i \circ w_j\}$ is an orthonormal representation of $G \times H$. If the vector $v_i w_j$ correspond to the vertex (i, j) in $G \times H$

Proof.

Let (i, j) and (i', j') be two vertices in $G \times H$.
If (i, j) and (i', j') not adjacent, then at least one of i and i' is not adjacent to j and j' respectively, which means at least one of $v_i, v_{i'}$ and $w_j, w_{j'}$ is orthogonal.

$$\langle v_i \circ w_j, v_{i'} \circ w_{j'} \rangle = \langle v_i, v_{i'} \rangle \langle w_j, w_{j'} \rangle = 0 \tag{7}$$

□

Theta Function

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

$$\theta(G) = \inf_{\{v_1, v_2, \ldots, v_n\}, c} \max_i \frac{1}{\langle c, v_i \rangle^2} \tag{8}$$

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

$$\theta(G) = \inf_{\{v_1, v_2, \ldots, v_n\}, c} \max_i \frac{1}{\langle c, v_i \rangle^2} \tag{8}$$

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.

Lemma

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

$$\theta(G) = \max_i \frac{1}{\langle c, v_i \rangle^2} \tag{9}$$

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

Lemma

Given a set of real finite dimensional unit vectors $\{v_1, v_2, \ldots, v_n\}$*, that is mutually orthogonal. And given an unit vector* c*, then*

$$\sum_{i=1}^{n} \langle c, v_i \rangle \leq 1 \tag{10}$$

Relation of Theta Function and Alpha Function

Lemma

$$\theta(G) \geq \alpha(G) \tag{11}$$

Proof.

Let $\{1, 2, \ldots, k\}$ be the maximum set of vertices of G such that every point is not adjacency in G. Thus $k = \alpha(G)$.
Let $\{v_1, v_2, \ldots, v_n\}$ be an orthonormal representation of G.
Then, $\{v_1, v_2, \ldots, v_k\}$ should be mutually orthogonal. And, for any unit vector c,

$$\sum_{i=1}^{k} \langle c, v_i \rangle \leq 1 \tag{12}$$

$\exists j$ such that

$$\begin{aligned} \langle c, v_j \rangle &\leq \frac{1}{k} \qquad (13) \\ \frac{1}{\langle c, v_j \rangle} &\geq k \qquad (14) \end{aligned}$$

□

Proof.

$$\theta(G) = \inf_{\{v_1, v_2, \ldots, v_n\}, c} \max_i \frac{1}{\langle c, v_i \rangle^2} \tag{15}$$

$$\geq \inf_{\{v_1, v_2, \ldots, v_n\}, c} \frac{1}{\langle c, v_j \rangle^2} \tag{16}$$

$$\geq k \tag{17}$$

$$= \alpha(G) \tag{18}$$

$\square$

Theta Function of Product Graph

Lemma

Given graph G and H, then

$$\theta(G \times H) \leq \theta(G)\theta(H) \tag{19}$$

Theta Function of Product Graph

Lemma

Given graph G and H, then

$$\theta(G \times H) \leq \theta(G)\theta(H) \tag{19}$$

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

$$\max\left\{\frac{1}{\langle c_v, v_i\rangle^2} : i = 1, 2, \ldots, n\right\} = \theta(G) \tag{20}$$

and

$$\max\left\{\frac{1}{\langle c_w, w_i\rangle^2} : i = 1, 2, \ldots, m\right\} = \theta(H) \tag{21}$$

□

Proof.

Then

$$\begin{aligned}
\theta(G \times H) &\leq \max\left\{\frac{1}{\langle c_v \circ c_w, v_i \circ w_j\rangle^2}\right\} && (22)\\
&= \max\left\{\frac{1}{\langle c_v, v_i\rangle^2 \langle c_w, w_j\rangle^2}\right\} && (23)\\
&\leq \max\left\{\frac{1}{\langle c_v, v_i\rangle^2}\right\} \max\left\{\frac{1}{\langle c_w, w_j\rangle^2}\right\} && (24)\\
&= \theta(G)\theta(H) && (25)\\
& && (26)
\end{aligned}$$

$\square$

Relation of Theta Function and Shannon Capacity

Theorem

Given a graph G, then

$$\theta(G) \geq \Theta(G) \tag{27}$$

Relation of Theta Function and Shannon Capacity

Theorem

Given a graph G, then

$$\theta(G) \geq \Theta(G) \tag{27}$$

Proof.

$$\begin{aligned} \Theta(G) &= \sup_n \sqrt[n]{\alpha(G^n)} && (28) \\ &\leq \sup_n \sqrt[n]{\theta(G^n)} && (29) \\ &\leq \sup_n \sqrt[n]{\theta(G)^n} && (30) \\ &= \theta(G) && (31) \end{aligned}$$

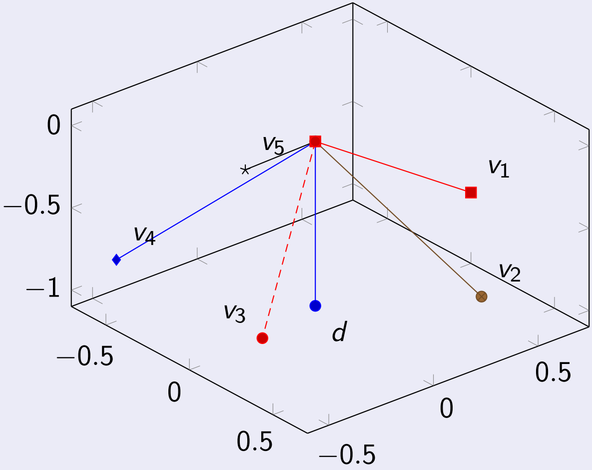
□

Theorem

$$\theta(C_5) \leq \sqrt{5} \tag{32}$$

Proof.

Consider an umbrella that has a handle and 5 ribs that all have unit length. And also its handle is a unit vector. Also, angles between two consecutive ribs are same.



Proof.

Let angle between v_1 and v_3, v_2 and v_4, v_3 and v_5, v_4 and v_1, v_5 and v_2 be $\frac{\pi}{2}$.

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

$$\begin{aligned} \theta(C_5) &= \inf_{\{v_1, v_2 \ldots v_5\}, c} \max_i \frac{1}{\langle c, v_i \rangle} & (33) \\ &\leq \max \frac{1}{< d, v_i >^2} & (34) \\ &= \left(\frac{1}{\cos(\gamma)} \right)^2 & (35) \\ &= \sqrt{5} & (36) \end{aligned}$$

□

Table of Contents

Conclusion

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

Open Questions

- Although $\Theta(C_n)$ is less or equal than $n\frac{\cos(\pi/n)}{1+\cos(\pi/n)}$, 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? Such as, odd, or prime.

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