Capacité de Shannon, théorie des graphes et intution divine

Léo Gagnon

June 19, 2020

Université de Montréal

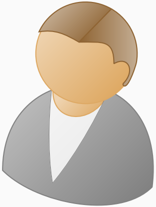
Information et capacité de Shannon sur un canal parfait

Information de Shannon

Alice

$m \in \Sigma$

Bob



$\Sigma = \{\text{oui, non}\}$

$m \in \Sigma$

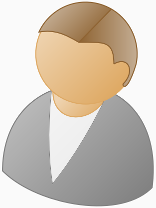
Quel est le contenu **d'information** d'un symbole envoyé par Alice?

Information de Shannon

Alice

$$\xrightarrow{\quad m \in \Sigma \quad}$$

Bob



$\Sigma = \{0, 1\}$

$m \in \Sigma$

Quel est le contenu **d'information** d'un symbole envoyé par Alice?

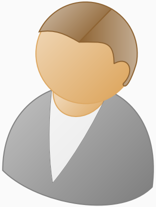
$\log_2(2) = 1$ bit d'information

Information de Shanon

Alice

$$m \in \Sigma$$

Bob



$\Sigma = \{\text{rouge, bleu, vert}\}$

$m \in \Sigma$

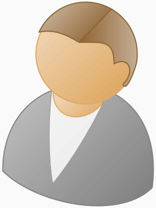
Quel est le contenu **d'information** d'un symbole envoyé par Alice?

Information de Shanon

Alice

$\xrightarrow{m \in \Sigma}$

Bob



$\Sigma = \{\text{rouge, bleu, vert}\}$

$m \in \Sigma$

Quel est le contenu **d'information** d'un symbole envoyé par Alice?

$$\log_2(3) = 1.58 \text{ bit d'information}$$

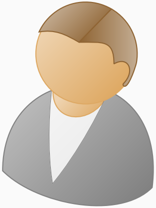
Information de Shannon

Alice

$m \in \Sigma^*$

$\longrightarrow$

Bob



$\Sigma = \{a, b, c, \ldots\}$

$m \in \Sigma^*$

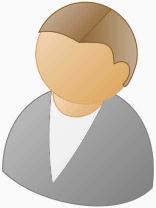
Quel est le contenu **d'information** d'un par symbole du message d'Alice?

Information de Shannon

Alice

$$\xrightarrow{m \in \Sigma^*}$$

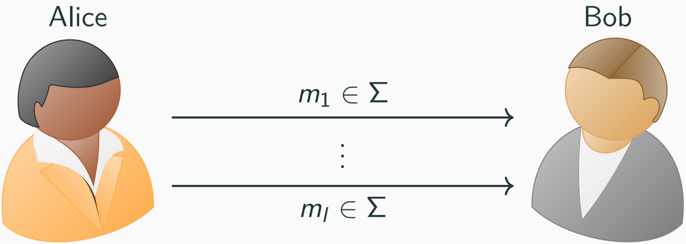
Bob



$\Sigma = \{a, b, c, \ldots\}$

$m \in \Sigma^*$

$$\frac{\log_2(26^k)}{k} = \frac{k \log_2(26)}{k} = \log_2(26)$$



$\Sigma = \{a, 5, \square, , +, \ldots\}$ $\qquad$ $m \in \Sigma^*$

Taux de transmission maximal d'un canal parfait $= \log_2(|\Sigma|)$

Taux de transmission

La **quantité d'information** contenue dans un symbole d'Alice est égale au nombre de bits nécéssaire à l'encodage de tout les messages possibles.

Le **taux de transmission maximal** d'un canal est la quantité maximale d'information qui peux être transmise par symbole sur des mots de longeur arbitraire.

Une borne supérieure pour la **taux de transmission** est $\log_2(|\Sigma|)$.

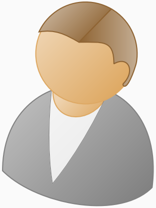
Information et capacité de Shannon sur un canal à confusion

Communication sans erreur sur un canal à confusion

Alice

$$\rightsquigarrow^{\,m \in \Sigma^*}$$

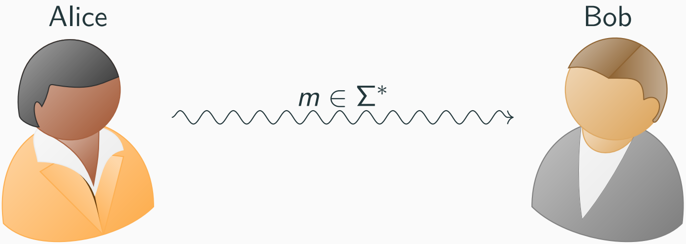
Bob



$\Sigma = \{a, 5, \square, , +, \ldots\}$

$m \in \Sigma^*$

Communication sans erreur sur un canal à confusion



$\Sigma = \{a, 5, \square, , +, \ldots\}$ $\qquad$ $m \in \Sigma^*$

Quel est le **taux de transmission** du canal si on ne tolère pas les erreurs?

Modélisation du canal à confusion

Étant donné un alphabet Σ, un canal de confusion est décrit par un graphe de confusion $G = (\Sigma, E)$ où

$$E = \{uv : u, v \in \Sigma \text{ et } u \text{ peux être confondu avec } v\}$$

Le taux de transmission du canal **si on transmet les symboles un à la fois** est égal au taux de transmission d'un canal parfait avec comme alphabet le stable maximal de G : $\log_2(\alpha(G))$

Transmission par chaînes de symboles

Deux chaînes $u_1 u_2$ et $v_1 v_2$ peuvent être confondues si et seulement si un des trois énoncés suivants est vrai

- $u_1 = v_1$ et u_2 peux être confondu avec v_2
- $u_2 = v_2$ et u_1 peux être confondu avec v_1
- $u_1 \neq v_1$ peuvent être confondus ou $u_2 \neq v_2$ peuvent être confondus

Transmission par chaînes de symboles

Deux chaînes $u_1 u_2$ et $v_1 v_2$ peuvent être confondues si et seulement si un des trois énoncés suivants est vrai

- $u_1 = v_1$ et u_2 peux être confondu avec v_2
- $u_2 = v_2$ et u_1 peux être confondu avec v_1
- $u_1 \neq v_1$ peuvent être confondus ou $u_2 \neq v_2$ peuvent être confondus

Exactement la définition du produit fort de G avec lui-même ($G \boxtimes G$)!!

Modélisation du canal à confusion (bis)

Étant donné un alphabet Σ, un canal de confusion est décrit par un graphe de confusion $G = (\Sigma, E)$ où

$$E = \{uv : u, v \in \Sigma \text{ et } u \text{ peux être confondu avec } v\}$$

Le taux de transmission du canal si Alice envoie k **symboles à la fois** est égal à

$$\frac{\log_2(\alpha(G^k))}{k} \neq \log_2(\alpha(G))$$

où $G^k = G^{k-1} \boxtimes G$.

Communication sans erreur sur un canal à confusion

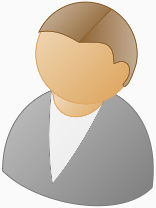
Alice

$m_1 \in \Sigma$

$\vdots$

$m_l \in \Sigma$

Bob

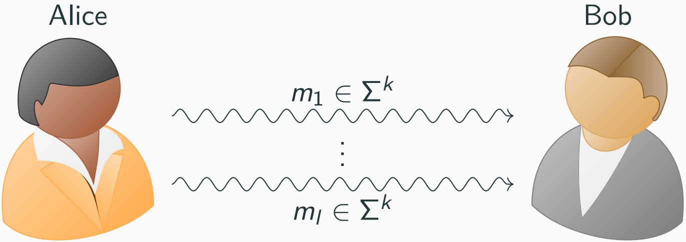


$\Sigma = \{a, 5, \square,, +, \ldots\}$

$m \in \Sigma^*$

Taux de transmission : $\log_2(\alpha(G))$

Communication sans erreur sur un canal à confusion



$\Sigma = \{a, 5, \square,, +, \ldots\}$ $\qquad$ $m \in \Sigma^*$

Taux de transmission : $\frac{\log_2(\alpha(G^k))}{k}$

Reformulation du problème

Soit G un graphe de confusion. Définissons la *capacité de Shannon* du graphe comme

$$\Theta(G) = \sup_{n \geq 1} \sqrt[n]{\alpha(G^n)}$$

Ainsi, le taux de transmission maximal du canal de confusion du graphe G est de

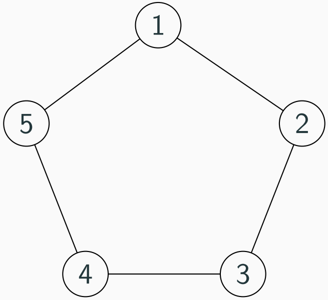
$$\log_2(\Theta(G)) = \sup_{n \geq 1} \frac{\log_2(\alpha(G^n))}{n}$$

.

Exemple : C_5

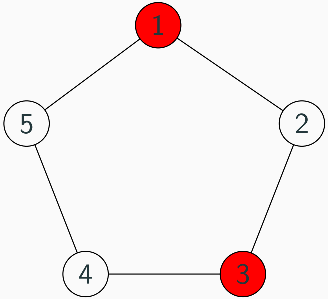
Exemple : C_5

Soit $\Sigma = \mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ et G le graphe suivant



Exemple : C_5

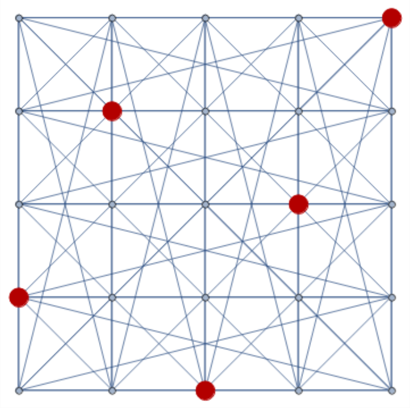
Soit $\Sigma = \mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ et G le graphe suivant



$\alpha(G) = 2$
Taux de transmission $= \log_2(2) = 1$ bit d'information

Exemple : C_5

Soit $\Sigma = \mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ et $G \boxtimes G$ le graphe suivant

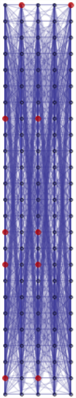


$\alpha(G^2) = 5$

Taux de transmission $= \frac{\log_2(5)}{2} \approx 1.16$ bit d'information!!

Exemple : C_5

Soit $\Sigma = \mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ et $G^2 \boxtimes G$ le graphe suivant

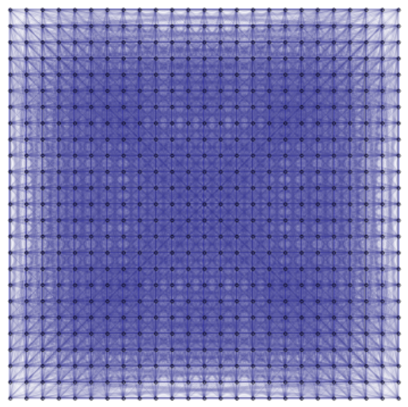


$\alpha(G^3) = 10$

Taux de transmission $= \frac{\log_2(10)}{3} \approx 1.11$ bit d'information...

Exemple : C_5

Soit $\Sigma = \mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ et $G^3 \boxtimes G$ le graphe suivant



$\alpha(G^4) = ?$

Exemple : C_5

On a donc

$$? \geq \Theta(G) \geq \sqrt{5}$$

Comment trouver une borne supérieure

Comment trouver une borne supérieure

Si on trouve une fonction f telle que

- $\alpha(G) \leq f(G)$
- $f(G^n) \leq f(G)^n$

alors

$$\Theta(G) = \sup_{n \geq 1} \sqrt[n]{\alpha(G^n)} \leq \sup_{n \geq 1} \sqrt[n]{f(G^n)} \leq \sup_{n \geq 1} \sqrt[n]{f(G)^n} = f(G)$$

Première borne supérieure

Couverture par clique d'un graphe

Définition

Le nombre de couverture par clique $\overline{\chi}(G)$ d'un graphe G est le nombre minimum de cliques nécéssaires pour couvrir le graphe G.

Remarque : $\overline{\chi}(G) = \chi(\overline{G})$

Première borne

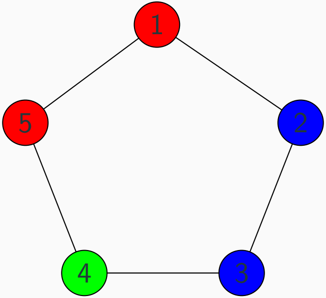
Théorème

$\Theta(G) = \sup_{n\geq 1} \sqrt[n]{\alpha(G^n)} \leq \overline{\chi}(G)$

Preuve:

1. $\alpha(G) \leq \overline{\chi}(G)$ pour tout graphe G
2. $\overline{\chi}(G \boxtimes H) \leq \overline{\chi}(G)\overline{\chi}(H)$ car le produit fort de deux cliques demeure une clique :
   Soit $(u,x),(v,y) \in K_n \boxtimes K_m$. Alors
   $[(u,x),(v,y)] \in E(K_n \boxtimes K_m)$

Exemple : C_5



$$\sqrt{5} = 2.24 \leq \Theta(G) = \sup_{n\geq 1} \sqrt[n]{\alpha(G^n)} \leq \overline{\chi}(G) = 3$$

Resserer la borne avec de la programmation linéaire

Dualité de $\alpha(G)$ et $\overline{\chi}(G)$

Definition

$\alpha(G)$ est la taille maximale de

$$\{u_1, u_2, \ldots, u_k\} \in 2^{V(G)}$$

tel que aucun sommet n'appartient à une même clique dans G.

Definition

$\overline{\chi}(G)$ est la taille minimale de

$$\{H_1, H_2, \ldots, H_s\}, \text{ où } H_i \text{ est une clique dans } G$$

tel que tout sommet appartient à au moins un H_i

Dualité de $\alpha(G)$ et $\overline{\chi}(G)$

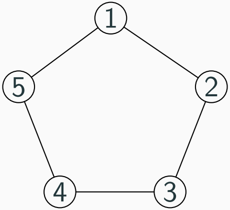
Définition

La matrice d'incidence sommet-clique d'un graphe possédant n sommets et m sous-cliques est une matrice $n \times m$ possédant un 1 à la position (i, j) si et seulement si le sommet i fait partie de la clique j.

Cette matrice servira de matrice de contraintes.

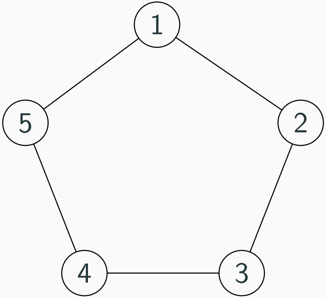
Exemple : C_5

	{1,2}	{2,3}	{3,4}	{4,5}	{5,1}	{1}	{2}	{3}	{4}	{5}
1	1	0	0	0	1	1	0	0	0	0
2	1	1	0	0	0	0	1	0	0	0
3	0	1	1	0	0	0	0	1	0	0
4	0	0	1	1	0	0	0	0	1	0
5	0	0	0	1	1	0	0	0	0	1



Exemple : C_5

	1	2	3	4	5
$\{1,2\}$	1	1	0	0	0
$\{2,3\}$	0	1	1	0	0
$\{3,4\}$	0	0	1	1	0
$\{4,5\}$	0	0	0	1	1
$\{5,1\}$	1	0	0	0	1
$\{1\}$	1	0	0	0	0
$\{2\}$	0	1	0	0	0
$\{3\}$	0	0	1	0	0
$\{4\}$	0	0	0	1	0
$\{5\}$	0	0	0	0	1



Dualité de $\alpha(G)$ et $\overline{\chi}(G)$

Soit A la matrice d'incidence sommet-clique de G.

Definition

$\alpha(G)$ est le maximum de $\mathbf{1}_n^\intercal \mathbf{x}$ avec les contraintes suivantes

1. $A\mathbf{x} \leq \mathbf{1}_m$
2. $\mathbf{x} \geq \mathbf{0}$
3. $\mathbf{x}$ est entier

Definition

$\overline{\chi}(G)$ est le minimum de $\mathbf{1}_m^\intercal \mathbf{y}$ avec les contraintes suivantes :

- $A^\intercal \mathbf{y} \geq \mathbf{1}_n$
- $\mathbf{y} \geq \mathbf{0}$
- $\mathbf{y}$ est entier

Théorème de dualité forte

Théorème

$$\min_{A\mathbf{x}\geq\mathbf{b},\ \mathbf{x}\geq\mathbf{0}} \mathbf{c}^\mathsf{T}\mathbf{x} = \max_{A\mathbf{y}\leq\mathbf{b},\ \mathbf{y}\geq\mathbf{0}} \mathbf{b}^\mathsf{T}\mathbf{y}$$

Dualité de $\alpha_f(G)$ et $\overline{\chi}_f(G)$

Soit A la matrice d'incidence sommet-clique de G.

Definition

$\alpha_f(G)$ est le maximum de $\mathbf{1}_n^\intercal \mathbf{x}$ avec les contraintes suivantes

1. $A\mathbf{x} \leq \mathbf{1}_m$
2. $\mathbf{x} \geq \mathbf{0}$
3. ~~$\mathbf{x}$ est entier~~

Definition

$\overline{\chi}_f(G)$ est le minimum de $\mathbf{1}_m^\intercal \mathbf{y}$ avec les contraintes suivantes :

- $A^\intercal \mathbf{y} \geq \mathbf{1}_n$
- $\mathbf{y} \geq \mathbf{0}$
- ~~$\mathbf{y}$ est entier~~

Resserer la borne

On a donc les inégalités suivantes

1. $\alpha_f(G) = \overline{\chi}_f(G)$
2. $\alpha(G) \leq \alpha_f(G)$
3. $\overline{\chi}(G) \geq \overline{\chi}_f(G)$

$$\implies \alpha(G) \leq \alpha_f(G) = \overline{\chi}_f(G) \leq \overline{\chi}(G)$$
$$\implies \alpha(G) \leq \overline{\chi}_f(G)$$

Resserer la borne

Théorème

$\Theta(G) = \sup_{n\geq 1} \sqrt[n]{\alpha(G^n)} \leq \overline{\chi}_f(G)$

Preuve:

1. $\alpha(G) \leq \overline{\chi}_f(G)$ pour tout graphe G
2. $\overline{\chi}_f(G \boxtimes H) \leq \overline{\chi}_f(G)\overline{\chi}(H)$ car le produit fort de deux cliques demeure une clique

Exemple : C_5

Grâce à Mathematica, on trouve que la solution minimale non-négative à

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{x} \geq \mathbf{1}$$

est $\mathbf{x}^\mathsf{T} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$
avec $\mathbf{1}^\mathsf{T}\mathbf{x} = \frac{5}{2} = 2.5$

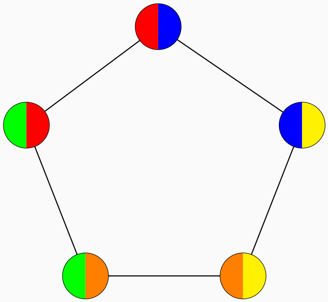
Exemple : C_5

Grâce à Mathematica, on trouve que la solution maximale non-négative à

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{y} \leq \mathbf{1}$$

est $\mathbf{y}^{\intercal} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$
avec $\mathbf{1}^{\intercal}\mathbf{y} = \frac{5}{2} = 2.5$

Exemple : C_5



$$\sqrt{5} = 2.24 \leq \Theta(G) = \sup_{n \geq 1} \sqrt[n]{\alpha(G^n)} \leq \overline{\chi}_f(G) = 2.5$$

Deuxième borne supérieure et intuition divine (ou autre substance illicite)

Algèbre linéaire : rappels et notations

Notation Bra-Ket :

- $\langle x|$ est un vecteur rangée (bra)
- $|y\rangle = \langle y|^\intercal$ est un vecteur colonne (ket)
- $\langle x|y\rangle$ est un produit scalaire (bra-ket)
- $|x\rangle\langle y|$ est un produit externe (ket-bra)

Représentation orthnormale d'un graphe

Définition

Une **représentation orthonormale** de dimension k $T = \{|v_1\rangle, \ldots, |v_n\rangle\}$ de G associe à chaque sommet un vecteur dans $\mathbb{R}^k$ de norme 1 de façon à ce que deux vecteurs associés à des sommets non-adjacent dans G soit orthogonaux dans $\mathbb{R}^k$.

Représentation orthnormale d'un graphe

Définition

Une **représentation parapluie** de dimension k de G est une représentation orthonormale de dimension k de G telle que tout les vecteurs $|v_i\rangle$ on le même angle avec la poignée $|h_T\rangle := \frac{1}{n}(|v_1\rangle + \cdots + |v_n\rangle) \in \mathbb{R}^k$.

On note $\langle v_i|h_T\rangle = \cos(\theta) = \sigma_T$ la *constante* de la représentation parapluie T.

Exemple : C_5

En route vers une borne

Soit $G = (V, E)$ ayant une représentation parapluie T.
Notons

$$\mu(\mathbf{x}) := x_1 |v_1\rangle + \cdots + x_n |v_n\rangle$$

la combinaire aléatoire des vecteurs de T déterminé par la distribution de probabilité discrète $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$.

Lovasz s'intéresse aux valeur possibles de $\|\mu(\mathbf{x})\|^2$, en particulier $\mu_T(G) := \inf_{\mathbf{x}} \|\mu(\mathbf{x})\|^2$

En route vers une borne

Soit U le plus grand stable de G (de taille α). Remarquons qu'en prenant la distribution

$$\mathbf{x}_U \text{ telle que } x_i = \begin{cases} \frac{1}{\alpha} & \text{si } v_i \in U \\ 0 & \text{sinon} \end{cases}$$

on obtient que

$$\begin{aligned} \|\mu(\mathbf{x})\|^2 &= \left(\sum_{|v_i\rangle \in U} \frac{1}{\alpha} \langle v_i| \right) \left(\sum_{|v_j\rangle \in U} \frac{1}{\alpha} |v_j\rangle \right) \\ &= \frac{1}{\alpha^2} \sum_{|v_i\rangle \in U} \left(\langle v_i| \sum_{|v_j\rangle \in U} |v_j\rangle \right) \\ &= \frac{1}{\alpha^2} \sum_{|v_i\rangle \in U} \sum_{|v_j\rangle \in U} \langle v_i|v_j\rangle = \frac{1}{\alpha^2} \alpha = \frac{1}{\alpha} \end{aligned}$$

En route vers une borne

Toute borne inférieure sur $\|\mu(\mathbf{x})\|^2$ pourra donc servir à construire une borne supérieure pour α car

$$M \leq \|\mu(\mathbf{x}_U)\|^2 = \frac{1}{\alpha} \implies \alpha \leq \frac{1}{M}$$

Exemple : C_5

En route vers une borne

Théorème

Soit T une représentation parapluie d'un graphe quelquonque, alors

$$\mu_T(G) = \inf_{\mathbf{x}} \|\mu(\mathbf{x})\|^2 = \||h_T\rangle\|^2$$

En route vers une borne

Théorème

Soit T une représentation parapluie d'un graphe quelconque, alors

$$\mu_T(G) = \inf_{\mathbf{x}} \|\mu(\mathbf{x})\|^2 = \||h_T\rangle\|^2$$

Corollaire

Si T est une représentation parapluie de G, alors $\alpha(G) \leq \frac{1}{\||h_T\rangle\|^2}$ où $|h_T\rangle$ est la poignée de T.

Étendre la borne à $\Theta(G)$

Théorème

Soit G et H deux graphes avec leur représentation parapluie respectivement S et T, alors

$$S \otimes T = \{|v\rangle \otimes |w\rangle : |v\rangle \in S, |w\rangle \in T\}$$

est une représentation parapluie pour $G \boxtimes H$

Étendre la borne à $\Theta(G)$

Preuve

Soient $(u, x), (v, y) \in V(G \boxtimes H)$ deux sommets non-adjacent et S, T les représentations parapluie de G et H. Alors

$$(\langle u| \otimes \langle x|)(|v\rangle \otimes |y\rangle) = \langle u|v\rangle \otimes \langle x|y\rangle = 0$$

où $|a\rangle$ est le vecteur associé au sommet a.

De plus, pour tout $(u, x) \in V(G \boxtimes H)$,

$$(\langle u| \otimes \langle x|)(|h_S\rangle \otimes |h_T\rangle) = \langle u|h_S\rangle \otimes \langle x|h_T\rangle = \sigma_S \sigma_T$$

donc $S \otimes T$ est une représentation parapluie pour $G \boxtimes H$ avec comme poignée $|h_{S \otimes T}\rangle = |h_S\rangle \otimes |h_T\rangle$

Étendre la borne à $\Theta(G)$

Corollaire

$$\begin{aligned}\| |h_{S\otimes T}\rangle \|^2 &= \| |h_S\rangle \otimes |h_T\rangle \|^2 \\ &= (\langle h_S| \otimes \langle h_T|)(|h_S\rangle \otimes |h_T\rangle) \\ &= \langle h_S|h_S\rangle \otimes \langle h_T|h_T\rangle \\ &= \| |h_S\rangle \|^2 \| |h_T\rangle \|^2\end{aligned}$$

Troisième borne

Théorème

Soit G un graphe et T sa représentation parapluie, alors

$$\Theta(G) = \sup_{n \geq 1} \sqrt[n]{\alpha(G^n)} \leq \frac{1}{\left\| |h_T\rangle \right\|^2}$$

$$\alpha(G^n) \leq \frac{1}{\left\| |h_{T^{\otimes n}}\rangle \right\|^2} = \frac{1}{\left\| |h_T\rangle \right\|^{2n}} \implies \sqrt[n]{\alpha(G^n)} \leq \frac{1}{\left\| |h_T\rangle \right\|^2}$$

Exemple : C_5

La longeur de la poignée pour le parapluie de C_5 est de $\frac{1}{\sqrt{5}}$ donc

on a

$$\Theta(G) = \sqrt{5}$$