Arrangement linéaire des graphes d'intervalles

Johanne Cohen[1] Fedor Fomin[2] Pinar Heggernes[2]
Dieter Kratsch[3] Gregory Kucherov[4]

[1]LORIA/CNRS, Nancy, France.

[2]Department of Informatics, University of Bergen, Norway.

[3]LITA, Université de Metz, France.

[4]LIFL/CNRS, Lille, France

Mulhouse, 30 Novembre, 07

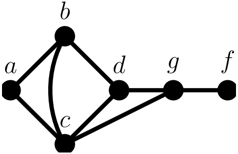
Plan

Plan

Definition

Un graphe G est un couple (V, E) où :

- V est un ensemble dont les éléments sont les sommets ;
- E est un ensemble de paires (non ordonnées) de sommets, appelées arêtes



Les graphes d'intervalles

Definition

Un graphe est un graphe d'intervalles si il existe une bijection entre les sommets du graphe et un ensemble de segments d'une droite tel que deux sommets u et v sont reliés par une arête si et seulement si les segments correspondant à u et v se superposent.

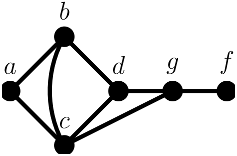


FIG.: Un graphe d'intervalles G

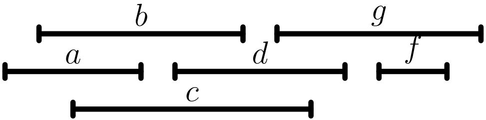


FIG.: Sa représentation graphique

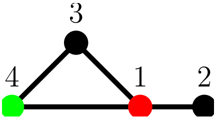
Definition : OLA

Definition

Le problème *arrangement linéaire* est de trouver une numérotation $\mathcal{L}$ des sommets du graphe $G = (V, E)$ tel que $\sum_{\{u,v\} \in E} |\mathcal{L}(u) - \mathcal{L}(v)|$ soit minimal.

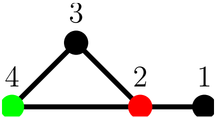
Definition

Le *poids* de L de G est $\mathcal{W}(G, L) = \sum_{(u,v) \in E} |L(u) - L(v)|$. Notons $\mathcal{W}(G) = \min_L \mathcal{W}(G, L)$. Une telle numérotation est un arrangement linéaire optimal (OLA).



numérotation L_1

$\mathcal{W}(G, L_1) = 1 + 3 + 2 + 1$



numérotation L_2

$\mathcal{W}(G, L_2) = 1 + 1 + 2 + 1$

Dans la littérature,

- Calculer le OLA est NP-complet (pour les graphes complets [Garey, Johnson 1979], pour les graphes bipartis. [Even, Shiloach 1975]).

- Mais il reste polynomial pour les arbres [Goldberg, Klipker 1976], pour les hypercubes [Diaz and al 2002].

- Le meilleur algorithme d'approximation (connu) a un rapport d'approximation $O(\log(n))$ [Rao, Richa 1998].

Numérotation et graphes d'interval

- La *bandwidth* ($b(G, L) = \max_{(u,v)\in E} |L(u) - L(v)|$) minimiser est polynomial pour les graphes d'intervalles [Kleitman, Vohra1990].

- Les graphes d'intervalles sont souvent utilisés en bio-informatique (fragments ADN, prédiction de structure de gènes

- L'arrangement linéaire de graphe d'intervalles modélise un chemin moléculaire [Farach-Colton and al 2004].

Nos résultats

- la complexité de ce problème dans un graphe d'intervalles est NP-complet.
- il existe une 2-approximation pour les graphes d'intervalles.

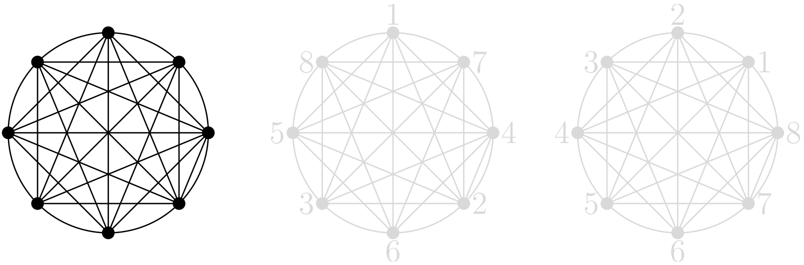
Plan

OLA pour les graphes complets.

Lemma

Soit K_n un graphe complet de n vertices. Then

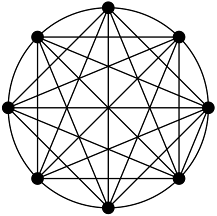
$$\mathcal{W}(K_n) = \frac{(n-1)n(n+1)}{6}.$$

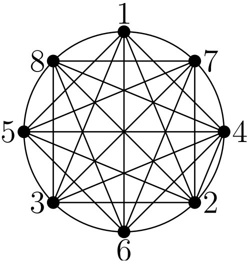


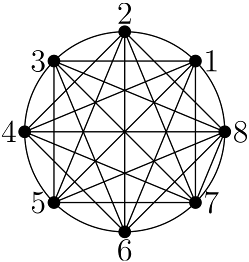
Lemma

Soit K_n un graphe complet de n vertices. Then

$$\mathcal{W}(K_n) = \frac{(n-1)n(n+1)}{6}.$$







OLA pour les étoiles.

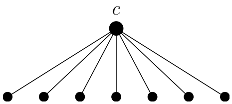
Lemma

Soit S_α une étoile ayant c comme centre et ayant α feuilles. Alors,

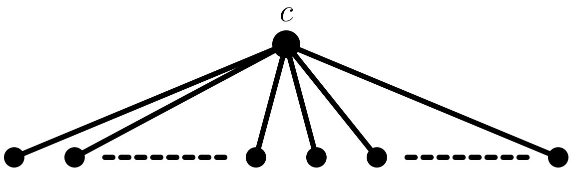
1. *une numérotation L est un arragement linéaire si et seulement si L place c au milieu*

$$\mathcal{W}(S_\alpha) = \left\lceil \frac{\alpha}{2} \right\rceil \left(\left\lfloor \frac{\alpha}{2} \right\rfloor + 1 \right)$$

2. $\mathcal{W}(S_\alpha) \leq \mathcal{W}(S_\alpha, L) \leq 2\mathcal{W}(S_\alpha)$ *pour n'importe quelle numérotation L*

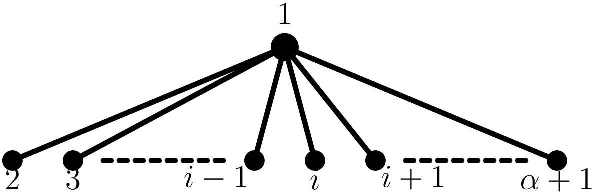


Soit S_α une étoile ayant c comme centre et ayant α feuilles (avec α pair).



Preuve du lemme

Soit S_α une étoile ayant c comme centre et ayant α feuilles (avec α pair).

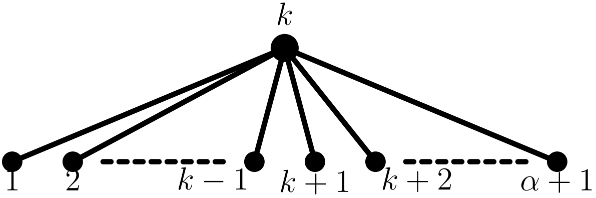


Cas où $L(c) = 1$:

- $\mathcal{W}(S_\alpha, L) = \sum_{i=1}^{\alpha}(i+1-1) = \sum_{i=1}^{\alpha} i = \frac{\alpha(\alpha+1)}{2}$

Preuve du lemme

Soit S_α une étoile ayant c comme centre et ayant α feuilles (avec α pair).



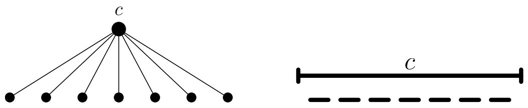
Cas où $L(c) = 1$: lepire cas $(\mathcal{W}(S_\alpha, L) = \frac{\alpha(\alpha+1)}{2})$.
Cas où $L(c) = k$:

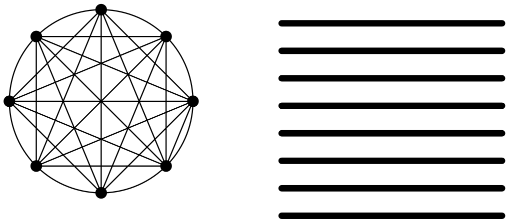
- $\mathcal{W}(S_\alpha, L) = \sum_{i=1}^{k-1} i + \sum_{i=1}^{\alpha+1-k} i$.

Cas où e $L(c) = \alpha/2 + 1$: lemeilleur cas $(\mathcal{W}(S_\alpha, L) = \frac{\alpha}{2}(\frac{\alpha}{2} + 1))$.

Remarque

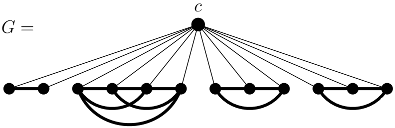
Les étoiles et les graphes sont des graphes intervalles.





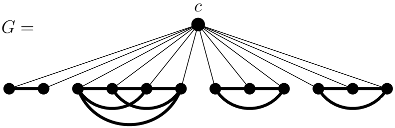
Partition des arêtes.

$G =$



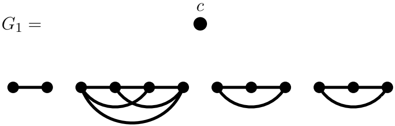
Partition des arêtes.

$G =$

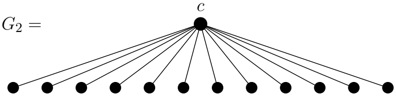


peut être considéré comme 2 graphes : G_1 et G_2

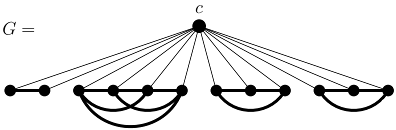
$G_1 =$



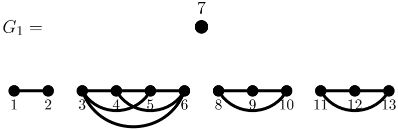
$G_2 =$



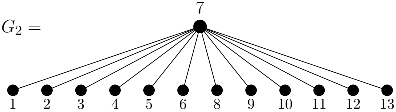
Partition des arêtes.



$$\mathcal{W}(G, L) = \mathcal{W}(G_1, L) + \mathcal{W}(G_2, L)$$



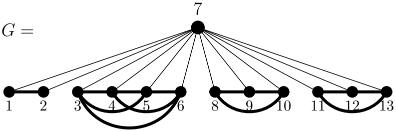
$$\mathcal{W}(G_1, L)$$



$$\mathcal{W}(G_2, L)$$

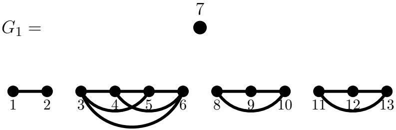
Partition des arêtes.

$G =$



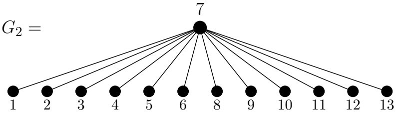
$$\mathcal{W}(G) \geq \mathcal{W}(G_1) + \mathcal{W}(G_2)$$

$G_1 =$



$$\mathcal{W}(G_1)$$

$G_2 =$



$$\mathcal{W}(G_2)$$

Quelques resultats :

Lemma

Soit $G = (V, E)$ *un graphe,* $E = E_1 \cup E_2$ *et* $E_1 \cap E_2 = \emptyset$*. Alors,* $\mathcal{W}(G) \geq \mathcal{W}(G_1) + \mathcal{W}(G_2)$*, où* $G_1 = (V, E_1)$ *et* $G_2 = (V, E_2)$*.*

Quelques resultats :

Lemma

Soit $G = (V, E)$ *un graphe,* $E = E_1 \cup E_2$ *et* $E_1 \cap E_2 = \emptyset$*. Alors,* $\mathcal{W}(G) \geq \mathcal{W}(G_1) + \mathcal{W}(G_2)$*, où* $G_1 = (V, E_1)$ *et* $G_2 = (V, E_2)$*.*

Plus généralement,

Corollary

Soit $G = (V, E)$, $V = V_1 \cup \cdots \cup V_n$, et $E_1, \cdots, E_n$ forment une partition de E. Alors $\mathcal{W}(G) \geq \mathcal{W}(G_1) + \ldots + \mathcal{W}(G_n)$, où $G_i = (V_i, E_i)$, $1 \leq i \leq n$.

Plan

Résultat

Theorem

le problème de décider si pour un graphe d'interval, Leproblem of deciding, for an interval graph $G = (E, V)$ and a constant K, whether $\mathcal{W}(G) \leq K$ est NP-complet.

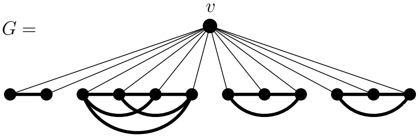
- Le problème appartient à NP.
- La preuve se base sur une réduction du problème 3-Partition :
  Instance : A un ens. de $3m$ entiers $\{a_1, \ldots, a_{3m}\}$, un entier $B \in Z^+$ tq $\sum_{i=1}^{3m} a_i = mB$.
  Question : A Peut-il être partitionné en m parties $A_1, A_2, \ldots, A_m$ tq $\forall i, 1 \leq i \leq m$, $\sum_{a \in A_i} a = B$?

Sketch of proof :

Considérons f la transformation de *PARTITION* en *OLA*

$$f(A,B) = (G,k) \text{ avec } \begin{cases} G \leftarrow \cup_{i=1}^{|A|} K_{a_i} \cup S_{2B} \text{ ayant } v \text{ cô centre} \\ k \leftarrow \mathcal{W}(S_{2B}) + \sum_{i=1}^{|A|} \mathcal{W}(K_{a_i}) \end{cases}$$

$A = \{2,4,3,3\} \quad \rightarrow$
$B = 6$

$G =$ 

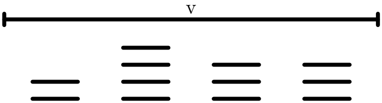
Sketch of proof :

Considérons f la transformation de *PARTITION* en *OLA*

$$f(A,B) = (G,k) \text{ avec } \begin{cases} G \leftarrow \cup_{i=1}^{|A|} K_{a_i} \cup S_{2B} \text{ ayant } v \text{ cô centre} \\ k \leftarrow \mathcal{W}(S_{2B}) + \sum_{i=1}^{|A|} \mathcal{W}(K_{a_i}) \end{cases}$$

G est un graphe d'intervalles.

$A = \{2,4,3,3\} \quad \rightarrow$
$B = 6$

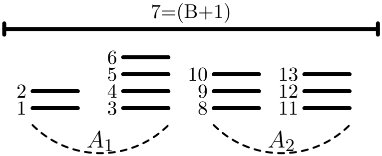


Sketch of proof :

Considérons f la transformation de *PARTITION* en *OLA*

$$f(A,B) = (G,k) \text{ avec } \begin{cases} G \leftarrow \cup_{i=1}^{|A|} K_{a_i} \cup S_{2B} \text{ ayant } v \text{ cô centre} \\ k \leftarrow \mathcal{W}(S_{2B}) + \sum_{i=1}^{|A|} \mathcal{W}(K_{a_i}) \end{cases}$$

$A = \{2,4,3,3\} \quad \rightarrow$
$B = 6$
$A_1 = \{2,4\}$
$A_2 = \{3,3\}$



Sketch of proof :

Considérons f la transformation de *PARTITION* en *OLA*

$$f(A,B) = (G,k) \text{ avec } \begin{cases} G \leftarrow \cup_{i=1}^{|A|} K_{a_i} \cup S_{2B} \text{ ayant } v \text{ cô centre} \\ k \leftarrow \mathcal{W}(S_{2B}) + \sum_{i=1}^{|A|} \mathcal{W}(K_{a_i}) \end{cases}$$

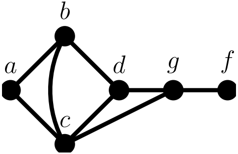
A peut être partitionné en 2 parties A_1, A_2 tq $\sum_{a \in A_1} a = \sum_{a \in A_2} a = B$ $\Leftrightarrow$ il existe une numérotation L tq $\mathcal{W}(G,L) \leq k$

Une transformation de $3 - PARTITION$ vers *OLA* est basée sur la même idée.

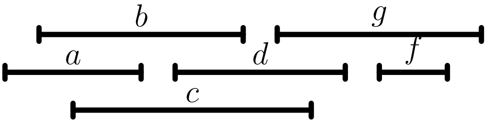
Plan

la numérotation sommet reo

- Considérons la numérotation sommet (*reo*) de G par rapport à $\mathcal{I}$ telle que l'ordre des sommets est défini par l'extrémité droite des intervalles.



un graphe d'intervalles G

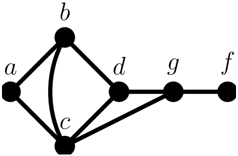


Une représentation graphique

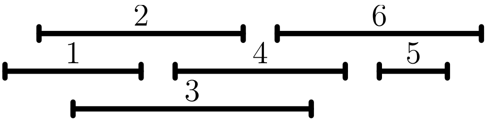
- Quel est son poids ?

la numérotation sommet reo

- Considérons la numérotation sommet (*reo*) de G par rapport à $\mathcal{I}$ telle que l'ordre des sommets est défini par l'extrémité droite des intervalles.



un graphe d'intervalles G

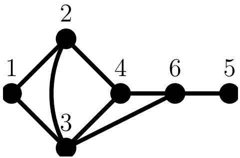


Une représentation graphique

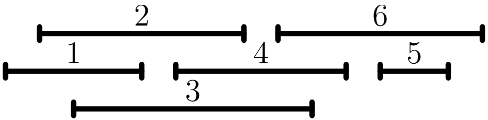
- Quel est son poids ?

la numérotation sommet reo

- Considérons la numérotation sommet (*reo*) de G par rapport à $\mathcal{I}$ telle que l'ordre des sommets est défini par l'extrémité droite des intervalles.



un graphe d'intervalles G

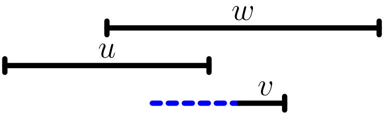


Une représentation graphique

- Quel est son poids ?

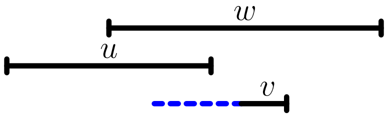
Remarque

Dans la numérotation sommet (*reo*), pour chaque paire de sommets adjacent u et v tel que $reo(u) < reo(w)$, chaque sommet v tel que $reo(u) < reo(v) < reo(w)$ est adjacent à v.



Remarque

Dans la numérotation sommet (*reo*), pour chaque paire de sommets adjacent u et v tel que $reo(u) < reo(w)$, chaque sommet v tel que $reo(u) < reo(v) < reo(w)$ est adjacent à v.

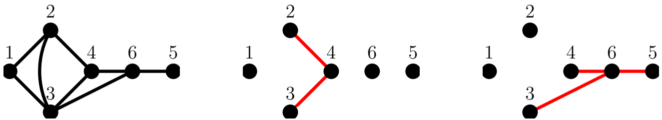


Theorem

Soit $G = (V, E)$ un graphe d'intervalles, et soit $\mathcal{I}$ un de ses représentation graphique. Alors, $\mathcal{W}(G, reo) \leq 2\mathcal{W}(G)$.

Preuve par induction : (1/2)

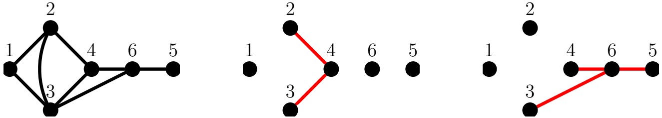
- Partition de E : $E_i = \{(u,v) \mid reo(u) = i \wedge reo(v) < i\} \cap E(G)$



G et reo $\quad$ E_4 $\quad$ E_6

Preuve par induction : (1/2)

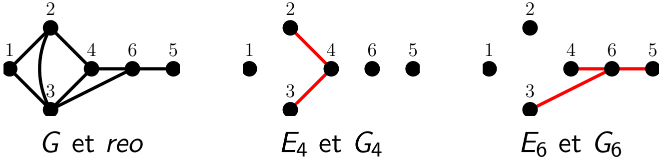
- Partition de E : $E_i = \{(u, v) \mid reo(u) = i \wedge reo(v) < i\} \cap E(G)$
- $G_i = (V, \cup_{\ell=1}^{i} E_\ell)$



G et reo $\quad$ E_4 et G_4 $\quad$ E_6 et G_6

Preuve par induction : (1/2)

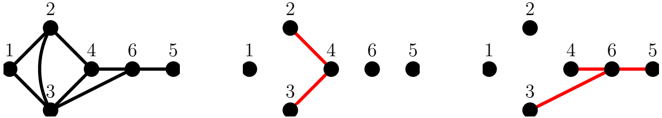
- Partition de E : $E_i = \{(u,v) \mid reo(u) = i \wedge reo(v) < i\} \cap E(G)$
- $G_i = (V, \cup_{\ell=1}^{i} E_\ell)$



- Hypothèse de récurrence sur i : $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$.

Preuve par induction : (1/2)

- Partition de E : $E_i = \{(u,v) \mid reo(u) = i \wedge reo(v) < i\} \cap E(G)$
- $G_i = (V, \cup_{\ell=1}^{i} E_\ell)$

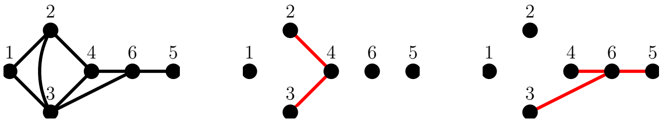


G et reo — E_4 et G_4 — E_6 et G_6

- Hypothèse de récurrence sur i : $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$.

Preuve par induction : (1/2)

- Partition de E : $E_i = \{(u,v) \mid reo(u) = i \wedge reo(v) < i\} \cap E(G)$
- $G_i = (V, \cup_{\ell=1}^{i} E_\ell)$



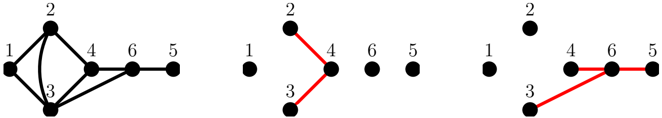
G et reo — E_4 et G_4 — E_6 et G_6

- Hypothèse de récurrence sur i : $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$.
- Pour $i = 1$: $\mathcal{W}(G_1, reo) = \mathcal{W}(G_1)$ car aucune arête
- Pour i : Supposons que $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$. Et pour $i+1$?
- $\mathcal{W}(G_{i+1}, reo) = \mathcal{W}(G_i, reo) + \sum_{e=(u,v)\in E_{i+1}} |L(u) - L(v)|$

Preuve par induction : (2/2)

- Hypothèse de récurrence sur i : $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$.
- Pour $i = 1$: $\mathcal{W}(G_1, reo) = \mathcal{W}(G_1)$ car aucune arête
- Pour i : Supposons que $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$. Et pour $i + 1$?
- $\mathcal{W}(G_{i+1}, reo) = \mathcal{W}(G_i, reo) + \sum_{e=(u,v)\in E_{i+1}} |L(u) - L(v)|$
- (V, E_{i+1}) est une étoile S composé de numéros consécutifs :

$$\mathcal{W}(S) \leq \sum_{(u,v)\in E_{i+1}} |L(u) - L(v)| \leq 2\mathcal{W}(S)$$

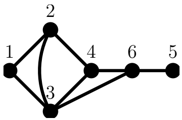
- 

G et reo E_4 E_6

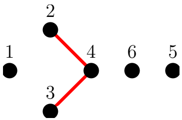
Preuve par induction : (2/2)

- Hypothèse de récurrence sur i : $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$.
- Pour $i = 1$: $\mathcal{W}(G_1, reo) = \mathcal{W}(G_1)$ car aucune arête
- Pour i : Supposons que $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$. Et pour $i+1$?
- $\mathcal{W}(G_{i+1}, reo) = \mathcal{W}(G_i, reo) + \sum_{e=(u,v)\in E_{i+1}} |L(u) - L(v)|$
- (V, E_{i+1}) est une étoile S composé de numéros consécutifs :

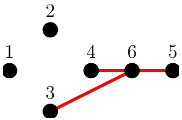
$$\mathcal{W}(S) \leq \sum_{(u,v)\in E_{i+1}} |L(u) - L(v)| \leq 2\mathcal{W}(S)$$

- 

G et reo



E_4 et G_4

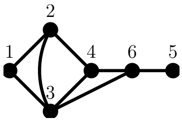


E_6 et G_6

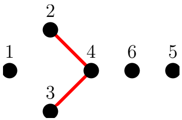
Preuve par induction : (2/2)

- Hypothèse de récurrence sur i : $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$.
- Pour $i = 1$: $\mathcal{W}(G_1, reo) = \mathcal{W}(G_1)$ car aucune arête
- Pour i : Supposons que $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$. Et pour $i + 1$?
- $\mathcal{W}(G_{i+1}, reo) = \mathcal{W}(G_i, reo) + \sum_{e=(u,v)\in E_{i+1}} |L(u) - L(v)|$
- (V, E_{i+1}) est une étoile S composé de numéros consécutifs :

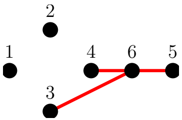
$$\mathcal{W}(S) \leq \sum_{(u,v)\in E_{i+1}} |L(u) - L(v)| \leq 2\mathcal{W}(S)$$

- 

G et reo



E_4 et G_4

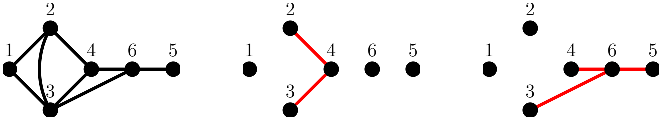


E_6 et G_6

Preuve par induction : (2/2)

- Hypothèse de récurrence sur i : $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$.
- Pour $i = 1$: $\mathcal{W}(G_1, reo) = \mathcal{W}(G_1)$ car aucune arête
- Pour i : Supposons que $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$. Et pour $i+1$?
- $\mathcal{W}(G_{i+1}, reo) = \mathcal{W}(G_i, reo) + \sum_{e=(u,v)\in E_{i+1}} |L(u) - L(v)|$
- (V, E_{i+1}) est une étoile S composé de numéros consécutifs :

$$\mathcal{W}(S) \leq \sum_{(u,v)\in E_{i+1}} |L(u) - L(v)| \leq 2\mathcal{W}(S)$$

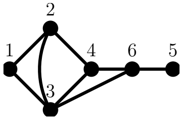
- 

G et reo $\quad$ E_4 et G_4 $\quad$ E_6 et G_6

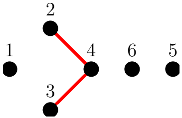
Preuve par induction : (2/2)

- Hypothèse de récurrence sur i : $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$.
- Pour $i = 1$: $\mathcal{W}(G_1, reo) = \mathcal{W}(G_1)$ car aucune arête
- Pour i : Supposons que $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$. Et pour $i + 1$?
- $\mathcal{W}(G_{i+1}, reo) = \mathcal{W}(G_i, reo) + \sum_{e=(u,v)\in E_{i+1}} |L(u) - L(v)|$
- (V, E_{i+1}) est une étoile S composé de numéros consécutifs :

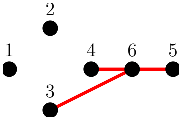
$$\mathcal{W}(S) \leq \sum_{(u,v)\in E_{i+1}} |L(u) - L(v)| \leq 2\mathcal{W}(S)$$

- 

G et reo



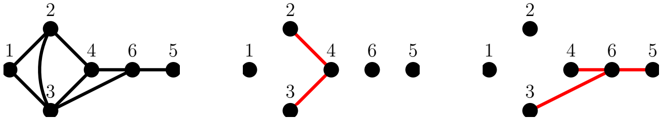
E_4 et G_4



E_6 et G_6

Preuve par induction : (2/2)

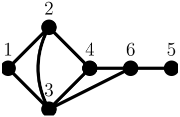
- Hypothèse de récurrence sur i : $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$.
- Pour $i = 1$: $\mathcal{W}(G_1, reo) = \mathcal{W}(G_1)$ car aucune arête
- Pour i : Supposons que $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$. Et pour $i+1$?
- $\mathcal{W}(G_{i+1}, reo) = \mathcal{W}(G_i, reo) + \sum_{e=(u,v)\in E_{i+1}} |L(u) - L(v)|$
- (V, E_{i+1}) est une étoile S composé de numéros consécutifs :
$$\mathcal{W}(S) \leq \sum_{(u,v)\in E_{i+1}} |L(u) - L(v)| \leq 2\mathcal{W}(S)$$
- $\mathcal{W}(S) + \mathcal{W}(G_i, reo) \leq \sum_{(u,v)\in E_{i+1}} |L(u) - L(v)| \leq 2\mathcal{W}(S) + 2\mathcal{W}(G_i, reo)$



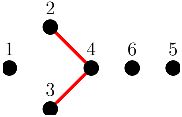
G et reo E_4 et G_4 E_6 et G_6

Preuve par induction : (2/2)

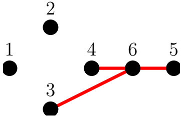
- Hypothèse de récurrence sur i : $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$.
- Pour $i = 1$: $\mathcal{W}(G_1, reo) = \mathcal{W}(G_1)$ car aucune arête
- Pour i : Supposons que $\mathcal{W}(G_i, reo) \leq 2\mathcal{W}(G_i)$. Et pour $i+1$?
- $\mathcal{W}(G_{i+1}, reo) = \mathcal{W}(G_i, reo) + \sum_{e=(u,v)\in E_{i+1}} |L(u) - L(v)|$
- (V, E_{i+1}) est une étoile S composé de numéros consécutifs :
$$\mathcal{W}(S) \leq \sum_{(u,v)\in E_{i+1}} |L(u) - L(v)| \leq 2\mathcal{W}(S)$$
- $\mathcal{W}(G_{i+1}, reo) \leq \sum_{e=(u,v)\in E_{i+1}} |L(u) - L(v)| \leq 2\mathcal{W}(G_{i+1}, reo)$



G et reo



E_4 et G_4



E_6 et G_6

Plan

Conclusion

- Le problème d'arrangement linéaire des graphes d'intervalles est NP-complet.
- Il existe une 2-approximation.
- La complexité d'autres problèmes de numérotation de sommets comme la CUTWIDTH sont des questions ouverts pour les graphes d'intervalles.