Découverte assistée par ordinateur en théorie des graphes

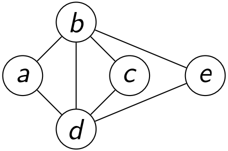
Pierre Hauweele

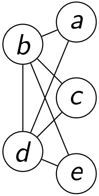
UMONS

25 mars 2016

Théorie des graphes

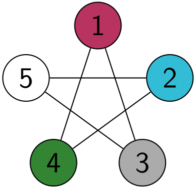
Soit un graphe $G = (V, E)$ où V est l'ensemble de ses sommets (ou nœuds), et E ses arêtes.





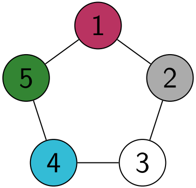
Isomorphisme de graphes — C_5

Deux graphes $G_1 = (V_1, E_1)$ et $G_2 = (V_2, E_2)$ sont isomorphes s'il existe une bijection $f : V_1 \to V_2$ telle que $(u, v) \in E_1 \iff (f(u), f(v)) \in E_2$.



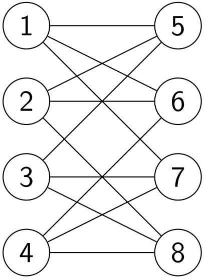
$$\begin{array}{|cccc|} \hline & & & \\ 0 & & & \\ 1 & 0 & & \\ 1 & 1 & 0 & \\ 0 & 1 & 1 & 0 \\ \hline \end{array}$$

$$\simeq$$

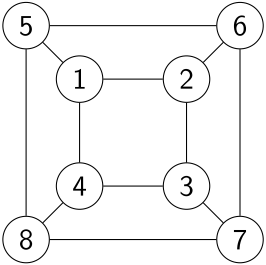


$$\begin{array}{|cccc|} \hline & & & \\ 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 1 & 0 & 0 & 1 \\ \hline \end{array}$$

Isomorphisme de graphes — cube



$$\simeq$$



Forme canonique d'un graphe

La relation d'isomorphisme forme une relation d'équivalence. On associe à un graphe G un représentant de sa classe d'équivalence. On l'appelle la forme canonique de G :

$$\text{Canon} : \mathcal{G} \to \mathcal{G}/\simeq .$$

En pratique, Canon(G) est le plus petit graphe de la classe d'équivalence de G dans l'ordre lexicographique de la représentation par matrice d'adjacence.

$$\text{Canon}(C_5) = "DqK".$$

$$\text{Canon}(cube) = "G?@ipo".$$

Invariant de graphe

Un invariant de graphe est une fonction sur les graphes qui reste constante sur les classes d'équivalence.
Par exemple :

- le nombre de sommets ;
- le nombre d'arêtes ;
- la planarité ;
- la taille de clique maximale.

Théorie extrémale des graphes

La théorie extrémale des graphes s'attache à étudier les graphes extrêmes pour certains invariants sous certaines contraintes.

Par exemple : "Parmi les graphes à n sommets et ne possédant pas de r-clique, lesquels ont un nombre maximal d'arêtes ?"

Théorème (Turán, 1941)

Parmi les graphes à n sommets et ne contenant pas de $(r+1)$-clique, $T(n,r)$ maximise le nombre d'arête, et il est l'unique.
$T(n,r)$ est défini comme l'unique graphe r-parti complet équilibré.

$T(8,3) =$

Théorie extrémale des graphes

La théorie extrémale des graphes s'attache à étudier les graphes extrêmes pour certains invariants sous certaines contraintes.
Par exemple : “Parmi les graphes à *n* sommets et ne possédant pas de *r*-clique, lesquels ont un nombre maximal d'arêtes ?”

Théorème (Turán, 1941)

Parmi les graphes à n sommets et ne contenant pas de $(r+1)$-clique, $T(n,r)$ maximise le nombre d'arête, et il est l'unique.
$T(n,r)$ est défini comme l'unique graphe r-parti complet équilibré.

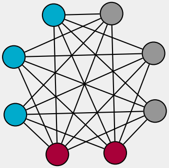
$T(8,3) =$

Théorie extrémale des graphes

La théorie extrémale des graphes s'attache à étudier les graphes extrêmes pour certains invariants sous certaines contraintes.
Par exemple : "Parmi les graphes à n sommets et ne possédant pas de r-clique, lesquels ont un nombre maximal d'arêtes ?"

Théorème (Turán, 1941)

Parmi les graphes à n sommets et ne contenant pas de $(r+1)$-clique, $T(n,r)$ maximise le nombre d'arête, et il est l'unique.
$T(n,r)$ est défini comme l'unique graphe r-parti complet équilibré.

$T(8,3) =$ 

Colorations

Définition

Une coloration est une assignation de couleurs aux sommets de G

Colorations

Définition

Une coloration propre est une assignation de couleurs aux sommets de G telle que deux sommets adjacents ont des couleurs différentes.

Coloration = partition des sommets

Bien sûr, les couleurs elles-mêmes importent peu.

En pratique, une coloration propre est utile puisque c'est une partition des sommets en différentes classes.

Coloration = partition des sommets

Bien sûr, les couleurs elles-mêmes importent peu.

En pratique, une coloration propre est utile puisque c'est une partition des sommets en différentes classes.

Question

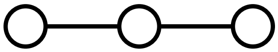
> *Étant donné un graphe, combien de partitions différentes de ses sommets sont possibles ? En d'autres mots, quel est le nombre total de colorations propres différentes d'un graphe ?*

Exemple : quel est le nombre total de colorations propres différentes du chemin P_3 ?

- Aucune coloration utilisant exactement 1 couleur
- Une coloration propre utilisant exactement 2 couleurs :

- Une coloration propre utilisant exactement 3 couleurs :

Exemple : quel est le nombre total de colorations propres différentes du chemin P_3 ?

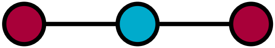


- Aucune coloration utilisant exactement 1 couleur
- Une coloration propre utilisant exactement 2 couleurs :

- Une coloration propre utilisant exactement 3 couleurs :

Exemple : quel est le nombre total de colorations propres différentes du chemin P_3 ?

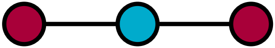
- Aucune coloration utilisant exactement 1 couleur
- Une coloration propre utilisant exactement 2 couleurs :



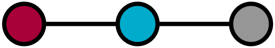
- Une coloration propre utilisant exactement 3 couleurs :

Exemple : quel est le nombre total de colorations propres différentes du chemin P_3 ?

- Aucune coloration utilisant exactement 1 couleur
- Une coloration propre utilisant exactement 2 couleurs :



- Une coloration propre utilisant exactement 3 couleurs :



Donc, le nombre de colorations non-équivalentes est de 2 pour P_3.

Nombre de colorations non-équivalentes

On note $\mathrm{P}(G, k)$ le nombre de colorations non-équivalentes de G qui utilisent *exactement* k couleurs.

$$\mathrm{P}(\mathrm{P}_3, 2) = 1$$

$$\mathrm{P}(\mathrm{P}_3, 3) = 1$$

Nombre total de colorations non-équivalentes

Définition

Le nombre total de colorations non-équivalentes $\mathcal{P}(G)$ d'un graph G est

$$\mathcal{P}(G) = \sum_{k=0}^{n} \mathrm{P}(G,k) = \sum_{k=\chi(G)}^{n} \mathrm{P}(G,k),$$

où $\chi(G)$ est le nombre chromatique de G.

Exemple : $\mathcal{P}(\mathrm{P}_3) = \mathrm{P}(\mathrm{P}_3, 2) + \mathrm{P}(\mathrm{P}_3, 3) = 1 + 1 = 2$

$\mathcal{P}(G)$ est également connu sous le nom de nombre de Bell d'un graphe [Duncan & Peele, 2009].

Le problème NumCol-NumEdges

Les valeurs de $\mathcal{P}(G)$ sont connues pour certaines classes particulières de graphes (p. ex., cycles, r-arbre, etc.).

Nous sommes intéressés par les valeurs extrémales de $\mathcal{P}(G)$

- lorsque l'on fixe seulement le nombre de sommets, K_n et $\overline{K_n}$ ont trivialement les valeurs extrémales pour $\mathcal{P}(G)$
- que se passe-t-il si on fixe d'autres paramètres, comme le nombre d'arêtes ?

Le problème NumCol-NumEdges

Les valeurs de $\mathcal{P}(G)$ sont connues pour certaines classes particulières de graphes (p. ex., cycles, r-arbre, etc.).

Nous sommes intéressés par les valeurs extrémales de $\mathcal{P}(G)$

- lorsque l'on fixe seulement le nombre de sommets, K_n et $\overline{K_n}$ ont trivialement les valeurs extrémales pour $\mathcal{P}(G)$
- que se passe-t-il si on fixe d'autres paramètres, comme le nombre d'arêtes ?

Le problème NumCol-NumEdges

Les valeurs de $\mathcal{P}(G)$ sont connues pour certaines classes particulières de graphes (p. ex., cycles, r-arbre, etc.).

Nous sommes intéressés par les valeurs extrémales de $\mathcal{P}(G)$

- lorsque l'on fixe seulement le nombre de sommets, K_n et $\overline{K_n}$ ont trivialement les valeurs extrémales pour $\mathcal{P}(G)$
- que se passe-t-il si on fixe d'autres paramètres, comme le nombre d'arêtes ?

Comment aborder le problème ?

À la main.

En utilisant l'ordinateur :

- AutoGraphiX : optimisation par approche heuristique (VNS),
- Digenes : optimisation par approche heuristique (génétique),
- Graph, GrInvIn, . . .,
- GraPHedron/Transproof/PHOEG : approche exacte.

Comment aborder le problème ?

À la main.

En utilisant l'ordinateur :

- AutoGraphiX : optimisation par approche heuristique (VNS),
- Digenes : optimisation par approche heuristique (génétique),
- Graph, GrInvIn, . . .,
- GraPHedron/Transproof/PHOEG : approche exacte.

Comment aborder le problème ?

Découverte assistée par ordinateur

À la main.

En utilisant l'ordinateur :

- AutoGraphiX : optimisation par approche heuristique (VNS),
- Digenes : optimisation par approche heuristique (génétique),
- Graph, GrInvIn, . . .,
- GraPHedron/Transproof/PHOEG : approche exacte.

Comment aborder le problème ?

Découverte assistée par ordinateur

À la main.

En utilisant l'ordinateur :

- AutoGraphiX : optimisation par approche heuristique (VNS),
- Digenes : optimisation par approche heuristique (génétique),
- Graph, GrInvIn, . . .,
- GraPHedron/Transproof/PHOEG : approche exacte.

Base de données d'invariants

Graphs
signature
A?
A_
B?
BG
BW
Bw
C?
C@
CB
CF
C‘

NumVertices	
signature	val
A?	2
A_	2
B?	3
BG	3
BW	3
Bw	3
C?	4
C@	4
CB	4
CF	4
C‘	4

NumEdges	
signature	val
A?	0
A_	1
B?	0
BG	1
BW	2
Bw	3
C?	0
C@	1
CB	2
CF	3
C‘	2

NumCol	
signature	val
A?	2
A_	1
B?	5
BG	3
BW	2
Bw	1
C?	15
C@	10
CB	7
CF	5
C‘	7

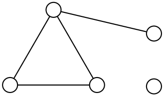
- PostgreSQL DB.
- 12 005 168 graphes non isomorphes de taille 10.

Base de données d'invariants — Requête

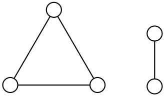
```
WITH tmp AS (
  SELECT n.val AS n, m.val AS m,
    P.signature, P.val AS num_col,
    rank() OVER (
      PARTITION BY n.val, m.val
      ORDER BY P.val
    ) AS pos
  FROM NumVertices n
  INNER JOIN NumEdges m USING(signature)
  INNER JOIN NumCol P USING(signature)
)
SELECT n, m, signature AS sig, num_col AS P
FROM tmp
WHERE pos = 1
ORDER BY n, m, p;
```

```
 n | m | sig | P
---+---+-----+----
 2 | 0 | A?  |  2
 2 | 1 | A_  |  1
 3 | 0 | B?  |  5
 3 | 1 | BG  |  3
 3 | 2 | BW  |  2
 3 | 3 | Bw  |  1
 4 | 0 | C?  | 15
 4 | 1 | C@  | 10
 4 | 2 | C`  |  7
 4 | 2 | CB  |  7
 4 | 3 | CJ  |  4
 4 | 4 | CN  |  3
 4 | 5 | C^  |  2
 4 | 6 | C~  |  1
 5 | 0 | D?? | 52
               [...]
```

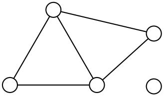
Quelques graphes extrémaux



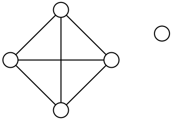
$(n, m) = (5, 4)$



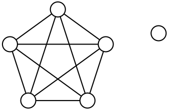
$(n, m) = (5, 4)$



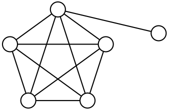
$(n, m) = (5, 5)$



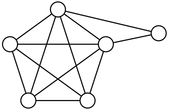
$(n, m) = (5, 6)$



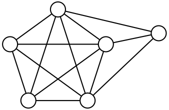
$(n, m) = (6, 10)$



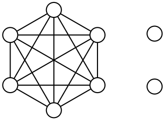
$(n, m) = (6, 11)$



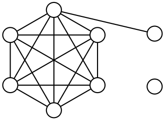
$(n, m) = (6, 12)$



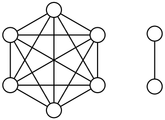
$(n, m) = (6, 13)$



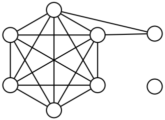
$(n, m) = (8, 15)$



$(n, m) = (8, 16)$



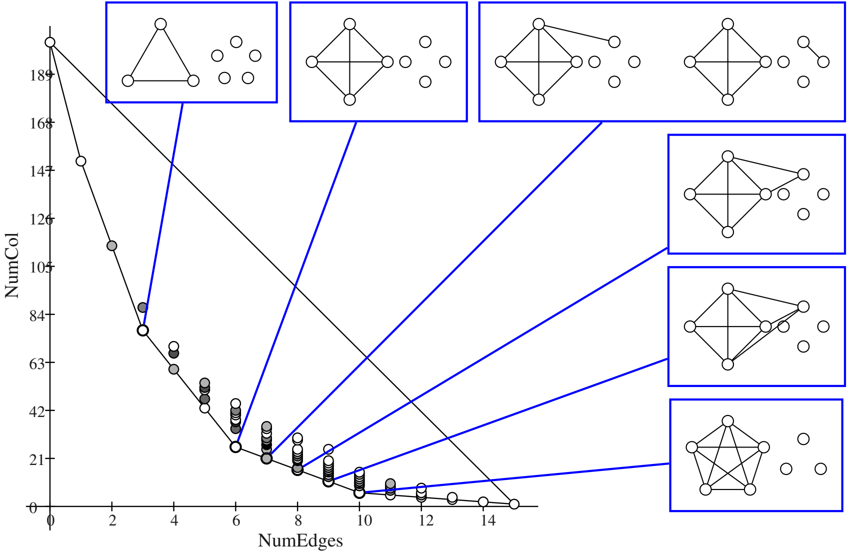
$(n, m) = (8, 16)$



$(n, m) = (8, 17)$

Polytope de graphes, dans l'espace d'invariants

Pour les graphes à 8 sommets



Polytope de graphes, dans l'espace d'invariants — Requête

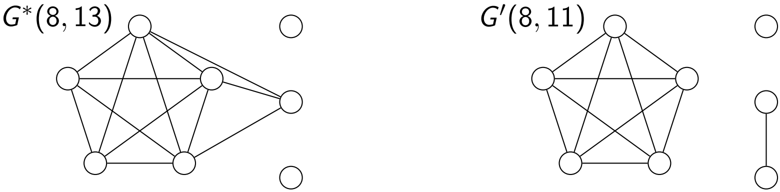
```
CREATE TEMPORARY TABLE tmp AS (
  SELECT NumEdges.val AS x, NumCol.val AS y, COUNT(*) AS n
  FROM NumVertices
  INNER JOIN NumEdges USING(signature)
  INNER JOIN NumCol USING(signature)
  WHERE NumVertices.val = 8
  GROUP BY x, y
);

SELECT ST_AsText(ST_ConvexHull(
  ST_Collect(ST_Point(x, y)))) FROM tmp;
------------------------------------------------------------------------------
 POLYGON((36 1,28 9,21 65,15 365,10 1540,6 4736,3 10481,1 17007,0
21147,36 1))
```

Les graphes extrêmes(?)

Étant donné n et m, soit t_k le plus grand nombre triangulaire tel que $t_k \leq m$. On appelle $r_m = m - t_k$ le reste.

On définit $G^*(n, m)$ comme l'unique graphe formé par $K_{k+1} \bigcup \overline{K}_{n-k-1}$, où un (s'il y en a) sommet de $\overline{K}_{n-k-1}$ est connecté à r_m sommets de la clique.

Si $r_m = 1$, et $n - k - 1 \geq 2$, on définit $G'(n, m)$ comme $K_{k+1} \bigcup \overline{K}_{n-k-1}$, où deux sommets de $K_{k+1} \bigcup \overline{K}_{n-k-1}$ sont connectés.



Conjecture

Conjecture

Soit G un graphe à n sommets et m arêtes. Alors,

$$\mathcal{P}(G) \geq \mathcal{P}(G^*(n,m)),$$

avec égalité ssi G est isomorphe à $G^*(n,m)$, sauf dans le cas où $r_m = 1$ et $n - k - 1 \geq 2$, où il est isomorphe à $G^*(n,m)$ ou $G'(n,m)$.

Ensuite

- Génération de conjecture automatique.
- Classification des graphes (arbre de décision).
- Preuve assistée par ordinateur (cfr. présentation de Gauvain Devillez)

Eccentric Connectivity Index

Définition

Soit $G = (V, E)$ un graphe.
Le degré d'un nœud v est défini comme le nombre de sommets adjacents à v dans G.
L'eccentricité d'un nœud $v \in$ est la distance maximale entre v et n'importe quel autre nœud $u \in V$. On la note $e(v)$.
L'invariant Eccentric Connectivity Index, noté ξ^c, est défini comme

$$\xi^c(G) = \sum_{v \in V} d(v)e(v).$$