Introduction aux graphes, reconnaissance de triangles

N.K. Blanchard

Département d'Informatique, ENS Paris // LIAFA

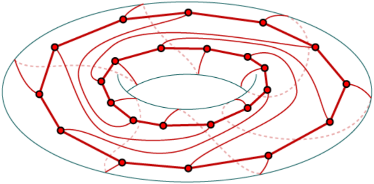
Samedi 9 Mai 2015

Plan de la Conférence

1. Introduction à la théorie des graphes
   - Premières définitions
   - Problèmes passés et actuels
   - Extensions
2. Graphes planaires et Reconnaissance de triangles
   - Caractérisation des graphes planaires
   - Algorithme de reconnaissance de triangles
   - Preuve
   - Pour finir

Les graphes

- Un ensemble V de sommets
- Un ensemble E d'arêtes qui connectent un sommet à un autre
- Le degré d'un sommet est le nombre d'arêtes sortantes
- Un graphe plongé est un graphe dessiné sur une surface tel que les arêtes ne se croisent pas



Types d'objets

- Différentes classes de graphes simples :
  - Lignes, Cycles
  - Complet : chaque paire de sommet est reliée
  - Arbres et Forêts : il n'y a aucun cycle
- Sous-graphes : obtenus avec la suppression de certains sommets et arêtes
- Mineurs : obtenus avec suppression et contraction

Premières propriétés

- Connexe : de tout point pouvoir aller vers tout point
- Biparti : pouvoir séparer le graphe en deux sous-graphes n'ayant aucune arête interne
- Planaire : peut se dessiner sur un plan sans que les arêtes se croisent
- Line graph : chaque arête est transformée en sommet et on relie celles qui ont une extremité en commun

Parcours

- Parcours Eulérien : passer une fois par chaque arête
  Facile, il suffit de regarder si :
  - Le graphe est connexe
  - Le degré de chaque sommet (sauf deux) est pair
- Analogue : parcours Hamiltonien, passant une fois par chaque sommet
  Extrêmement difficile

Problèmes intéressants

- Planarité : cns pour être un graphe planaire (résolu théoriquement en 1930, algorithmiquement en 1974)
- Coloration : colorier un graphe sans que deux sommets adjacents aient la même couleur (nombreux problèmes liés toujours ouverts)
- Décomposition : séparer un graphe en plusieurs sous-graphes
- Caractérisation : cns pour faire partie d'une classe de graphe (graphes toroïdaux par exemple)

Problèmes difficiles

- Hamiltonicité : même pour les graphes planaires de degré maximal 3
- Isomorphisme : deux graphes sont-ils les mêmes à permutation près ?
- Clique-Max : trouver le plus grand sous-graphe complet
- Conjecture de Steinberg : Graphe planaire sans (4 ou 5)-cycle est 3-coloriable
- Couverture par double cycle : Pour chaque graphe sans cycle trouver un ensemble de cycles tel que chaque arêtes soit présente exactement deux fois

Extensions de graphes

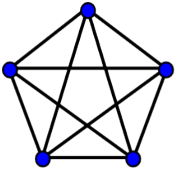
- Orientés : chaque arête a un sens
- Pondérés : on pose des poids sur les arêtes
- Infinis : dénombrables ou non
- Hypergraphes : des arêtes généralisées reliant des ensembles de points
- Dynamiques : les arêtes changent avec le temps
- Automates : graphes où l'on se déplace selon des caractères lus

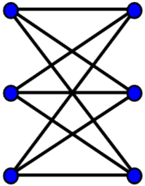
Problèmes qui m'intéressent

- Nombre de séparateurs minimaux
- Clustering en graphes dynamiques
- Cycles en graphes orientés
- Parcours LexBFS dans des graphes pondérés
- Bioinformatique (calcul de repli de molécule)

L'énigme des trois maisons

- Deux groupes de 3 sommets, peut-on connecter chaque membre d'un groupe à tous ceux de l'autre ?
- De manière équivalente : peut-on plonger $K_{3,3}$ dans le plan ?
- En fait on ne peut pas, tout comme on ne peut pas plonger K_5 (graphe complet à 5 sommets)





Caractérisation des graphes planaires

- Sont plongeables dans le plan exactement les graphes qui ne se réduisent pas au deux graphes précédent (Théorème de Kuratowski)
- Se réduire = on obtient le réduit par contraction et suppression d'arêtes

Formule d'Euler

- Sur les graphes connexes :

$$F + V = E + 2$$

- Preuve : par induction sur le nombre de faces (plus de 20 preuves distinctes)
- La constante 2 change sur d'autres surfaces - avec des conditions - exemple, $F + V = E$ sur le tore

Nombre de Triangles

- Chaque arête sépare au plus deux faces triangulaires
- Chaque face est entourée d'au moins trois arêtes
- Donc $3F \leq 2E$, et $F \leq \frac{2}{3}E$
- $E \leq 3V - 6$ avec la formule d'Euler puis $F \leq 2V - 4$

Nombre de Triangles (2)

- On a $F \leq 2V - 4$
- Il y a des triangles qui ne sont pas des faces
- Ils contiennent tous plusieurs faces
- On peut faire une hiérarchie (et un arbre)
- Il y en a au plus V, d'où $T \leq 3V$

Arboricité

- On couvre G par un ensemble de forêts arêtes-disjointes
- On cherche à minimiser le nombre de forêts
- En général

$$a(G) \leq (2m+n)^{1/2}/2$$

et (Nash-Williams 1961)

$$a(G) = \max_{H \subseteq G} \lceil q/(p-1) \rceil$$

- Pour les graphes planaires $a(G) \leq 3$

Entrée / Sortie

- On donne un graphe en entrée (par liste d'adjacence doublement chaînée)
- On cherche à sortir tous les couples (a, b, c) tels qu'on ait un triangle abc dans le graphe

Algorithme (Chiba - Nishizeki)

- Trier les sommets par degré décroissant
- Pour $1 \leq i \leq n-1$
  - Marquer tous les voisins de v_i
  - Pour chaque voisin u :
    - pour chaque voisin w marqué sortir $v_i uw$:
    - enlever la marque de u
  - enlever v_i du graphe
- Fin

Analyse

- On veut montrer que l'algorithme prend $O(a(G) \times m)$ en temps, donc $O(n)$ sur les graphes planaires
- On doit borner le coût du tri. Cela se fait avec un bucket sort.
- On doit montrer que

$$Temps \leq O(m) + O(n) + \sum_{v_i \in V} O\left(d(v_i) + \sum_{u \in N(v_i)} d(u)\right)$$

- On doit aussi montrer que

$$O(m) + O(n) + \sum_{v_i \in V} O\left(d(v_i) + \sum_{u \in N(v_i)} d(u)\right) = O(a(G) \times m)$$

Analyse (suite)

- Le coût est égal à (coût de tri) + pour chaque sommet (visite des voisins de ses voisins)
- Tri : $O(m) + O(n)$
- Pour chaque sommet : $\sum_{v_i \in V}$
- accès à la liste des voisins et visite des voisins de ses voisins : $d(v_i) + \sum_{u \in N(v_i)} d(u)$
- Total égal à

$$Temps \leq O(m) + O(n) + \sum_{v_i \in V} O\left(d(v_i) + \sum_{u \in N(v_i)} d(u)\right)$$

Lemme n°1

$$\sum_{u,v \in E} \min\{d(u), d(v)\} \leq 2 \times a(G) \times m$$

- On choisit un ensemble de forêts, et un sommet sur chaque arbre, et on oriente chaque arête pour partir des racines de l'arbre
- Chaque arête est associée à sa source (donc chaque sommet sauf les racines est associé à exactement une arête)

$$\sum_{(u,v) \in E} \min\{d(u), d(v)\} \leq \sum_{1 \leq i \leq a(G)} \sum_{e \in E(F_i)} d(h(e))$$

$$\leq \sum_{1 \leq i \leq a(G)} \sum_{v \in V} d(v) = 2a(G) \times m$$

Lemme n°2

$$\sum_{v_i \in V} O\left(d(v_i) + \sum_{u \in N(v_i)} d(u)\right) \leq O\left(\sum_{u,v \in E} \min\{d(u), d(v)\}\right)$$

- Comme les $d(v_i)$ sont décroissants on a $d(v_i) \geq d(u)$
- Comme on supprime v_i à la fin de chaque étape on ne considère les arêtes (u, v_i) qu'une seule fois

$$\sum_{v_i \in V}\left(d(v_i) + \sum_{u \in N(v_i)} d(u)\right) = \sum_{v_i \in V} \sum_{u \in N(v_i)} d(u) + \sum_{v_i \in V} d(v_i)$$

- Somme sur les arêtes de $d(u)$, avec $d(u) = \min\{d(u), d(v)\}$

Analyse (fin)

- Total égal à

$$Temps \leq O(m) + O(n) + \sum_{v_i \in V} O\left(d(v_i) + \sum_{u \in N(v_i)} d(u)\right)$$

- Or

$$\sum_{v_i \in V} O\left(d(v_i) + \sum_{u \in N(v_i)} d(u)\right) = O(a(G) \times m)$$

- Dans le cas planaire

$$O(a(G) \times m) = O(m) = O(n)$$

Résumé

- Graphes : de nombreuses applications et plein de questions ouvertes
- La théorie donne des algorithmes très rapides en pratique
- Algorithme : trier par ordre décroissant puis vérifier tous les triangles autour d'un point en ayant les bonnes structures de données

Pour aller plus loin

- Desbons livres d'introduction de Jean-Pierre Petit : Topologicon et Géometricon, disponibles librement sur internet
- De nombreux livres spécialisés et polycopiés accessibles sur les graphes planaires, par exemple celui de Colin de Verdière
- Le papier "Arboricity and subgraph listing algorithms" sur lequel je me base
- Des questions ?

(Les deux images proviennent de wikipédia)