Graphes et lemme des croisements

Eva Philippe

Parimaths, samedi 18 mars 2018

On supposera connues quelques notions sur les graphes (ici on regarde des graphes simples : non-orientés, sans arêtes doubles ni boucles). L'intérêt de cette séance est de montrer des exemples de problèmes où passer par des graphes et certaines de leurs propriété donne des solutions très élégantes.

On s'intéressera à des résultats asymptotiques, d'où une utilisation fréquente de la notation $O(f(n))$, ce qui signifie que la quantité (qui dépend de n) qu'on regarde est bornée par une constante multipliée par $f(n)$. Par exemple, si on a n points dans le plan, le nombre de paires que l'on peut former est exactement $\frac{n(n-1)}{2}$ mais il nous suffira de savoir que c'est un $O(n^2)$.

Quelques problèmes d'introduction

Commençons par trois problèmes qui serviront de fil rouge pour cette séance. Pour commencer, vous pouvez vous familiariser avec ces énoncés et chercher des configurations qui peuvent donner certaines bornes.

Problème 1. On considère n points dans le plan. Quel est le nombre maximal de paires de points qui sont à distance exactement 1 ?

Problème 2. On considère encore n points, mais cette fois qui sont restreints dans un disque de rayon 1. Quel est le nombre minimal de paires de points qui sont à distance inférieure ou égale à $\sqrt{2}$?

Problème 3. On a n nombres et on veut déterminer le plus grand d'entre eux en au plus k étapes. De combien de calculateurs en parallèle (qui peuvent faire chacun une comparaison à chaque étape) a-t-on besoin ?

On pourra commencer par $k = 1$, $k = 2$ et conjecturer une formule générale de la forme $O(n^\alpha)$.

Formule d'Euler et lemme de croisements

Sauf mention contraire, on considère un graphe G qui possède n sommets et m arêtes.

Exercice 1 (Formule d'Euler)

Rappeler la formule d'Euler pour les graphes planaires connexes, ainsi qu'une démonstration possible.

Exercice 2 En déduire une minoration du nombre d'arêtes m en fonction du nombre de sommets n pour les graphes planaires.

Exercice 3 En déduire une majoration du nombre de croisements t dans le cas général (non planaire) en fonction de n et m.

Exercice 4 (Lemme des croisements) Montrer le lemme des croisements : Si $m \geq 4n$, alors le nombre de croisements t est supérieur ou égal à $\frac{m^3}{64n^2}$.

Indication : Une idée, très astucieuse, est de choisir un sous-graphe induit H (c'est-à-dire qu'on choisit un sous-ensemble de sommets et les arêtes dans H correspondent exactement aux arêtes dans G) au hasard et estimer le nombre de croisements de ce sous-graphe.

Exercice 5 (Retour au Problème 1) Pour donner une majoration dans le cas du problème 1, associer judicieusement un graphe à une configuration et utiliser le lemme des croisements pour obtenir une inégalité.

Théorème de Turán

Définition 4. On dit qu'un sous-ensemble de sommets du graphe G forme un *ensemble indépendant* si aucune arête de G ne relie deux points de ce sous-ensemble.

Théorème 5. (Théorème de Turán, 1941)

La taille (cardinal) maximale d'un ensemble indépendant de G, notée $\alpha(G)$, vérifie :

$$\alpha(G) \geq \frac{n^2}{2m+n}.$$

Exercice 6 Voici la stratégie qu'on va suivre pour démontrer le théorème de Turán : on va choisir un ensemble indépendant aléatoirement et minorer l'espérance de sa taille.

On choisit un ordre aléatoire sur les sommets : $v_1, \ldots, v_n$. On dit qu'un sommet v_i est libre si tous ses voisins v_j vérifient $j < i$.

1. Montrer que I, l'ensemble des sommets libres ainsi définis, est un ensemble indépendant.
2. Pour un sommet v quelconque, calculer la probabilité que v soit dans I.
3. En déduire l'espérance de la taille de I et en donner une minoration.

Exercice 7 (Retour au Problème 2) Pour donner une majoration dans le cas du problème 2, associer astucieusement un graphe G à une configuration et montrer que $\alpha(G)$ est borné. En déduire une minoration du nombre de paires de points à distance $\leq \sqrt{2}$.

Exercice 8 (Retour au Problème 3) Montrer par récurrence que le nombre minimal de processeurs est au moins de l'ordre $O(n^{1+\frac{1}{2^k-1}})$. Pour l'hérédité on pourra considérer le graphe des comparaisons effectuées à la première étape et apliquer le théorème de Turán.

Exercice bonus

Exercice 9 Montrer que K_5 (le graphe complet à 5 sommets) et $K_{3,3}$ (le graphe biparti complet entre deux ensembles de trois sommets) ne sont pas planaires.

Le fait de ne pas contenir de subdivision de K_5 et $K_{3,3}$ est en fait une caractérisation complète des graphes planaires, mais c'est difficile à montrer (théorème de Kuratowski).

Solutions

<u>*Solution de l'exercice 1*</u> Pour un graphe planaire, on peut définir la notion de *faces* : les plus grands cycles qui ne contiennent pas d'arête si on dessine le graphe sans croisement, auxquels on ajoute la face extérieure. On note f le nombre de faces. La formule d'Euler relie n le nombre de sommets, m le nombre d'arêtes et f le nombre de faces de la manière suivante :

$$n - m + f = 2.$$

On peut commencer par démontrer cette formule pour les arbres (graphes connexes sans cycles). Or dans un graphe à n sommets, $m = n - 1$ (on peut par exemple enraciner l'arbre en un sommet, et alors chaque sommet est relié par une arête à son père sauf la racine) et $f = 1$. Donc la formule d'Euler est vérifiée.

Ensuite, en partant d'un graphe planaire connexe quelconque, si $f \geq 1$ il y a un cycle et on peut enlever une arête de ce cycle en gardant un graphe planaire connexe et en diminuant de 1 le nombre de faces. Après $f - 1$ telles étapes, le graphe obtenu a n sommets, $m - (f - 1)$ arêtes et $f - (f - 1) = 1$ face. C'est donc un arbre (il est toujours connexe) et vérifie la formule d'Euler, d'où $n - (m - (f - 1)) + (f - (f - 1)) = 2$, *ie* $n - m + f = 2$.

<u>*Solution de l'exercice 2*</u> Il est naturel de chercher une majoration du nombre d'arêtes pour un graphe planaire car pour un nombre de sommets fixé, on sent bien que si on rajoute des arêtes on risque de ne plus pouvoir éviter les croisements et perdre la planarité.

L'idée astucieuse ici est de faire un double-comptage sur les arêtes. Si on additionne les nombres d'arêtes constituant chaque face, on obtient $2m$ car chaque arête est comptée deux fois. De plus, comme il n'y a pas d'arête double, chaque face est constituée de au moins trois arêtes. D'où

$$2m \geq 3f.$$

En intégrant cette inégalité à la formule d'Euler il vient :

$$m \leq 3n - 6.$$

<u>*Solution de l'exercice 3*</u> On considère un graphe G quelconque, dont le dessin dans le plan possède un certain nombre t de croisements. On peut remplacer chaque croisement par un sommet dont partent 4 arêtes, au lieu de 2 arêtes qui se croisent. Se faisant, on obtient un nouveau graphe, planaire cette fois, avec $n + t$ sommets et $m + 2t$ sommets. D'après la question précédente on doit avoir

$$m + 2t \leq 3(n + t) - 6$$

d'où

$$t \geq m - 3n + 6.$$

<u>*Solution de l'exercice 4*</u> Ici , on regarde un graphe qui a déjà un grand nombre d'arêtes, donc son nombre de croisements est plus que linéaire en m.

Ce problème a été originnellement posé dans le cas des graphes bipartis par le mathématicien hongrois Pál Turán pour optimiser le transport de briques entre les fours et les entrepôts dans une usine (les chariots étaient plus difficiles à pousser lorsque les rails se croisaient).

On choisit un sous-graphe induit H de manière aléatoire en sélectionnant chaque sommet indépendamment avec probabilité p (où $p \in [0, 1]$ sera spécifié par la suite). Le nombre moyen de sommets de H est pn et son nombre moyen d'arêtes est p^2m (chaque arête a une probabilité p^2, qui correspond à la probabilité de sélectionner ses deux extrémités, d'être dans H). De plus, un croisement de H correspond au choix de quatre sommets dans G donc le nombre moyen de croisements dans H est p^4t. On déduit de la question précédente :

$$p^4t \geq p^2m - 3pn$$

d'où

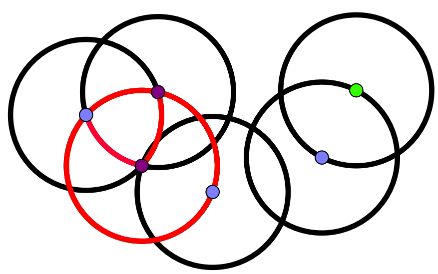
$$t \geq \frac{m}{p^2} - \frac{3n}{p^3}.$$

Comme on a supposé $m > 4n$, on peut poser $p = \frac{4n}{m}$ (ou alors en dérivant le terme de droite en p, on trouve que sa plus grande valeur est atteinte en $p = \frac{9n}{2m}$), ce qui donne

$$t \geq \frac{m^3}{64n^2}.$$

La constante $\frac{1}{64}$ peut être améliorée, actuellement la meilleure constante connue est $\frac{1}{29}$, d'après Wikipédia (`https://en.wikipedia.org/wiki/Crossing_number_inequality`).

Solution de l'exercice 5 Pour chacun de nos n points dans le plan, traçons autour un cercle de rayon 1. Ensuite, on considère le graphe G obtenu en conservant les portions de cercle qui se trouvent entre deux points : les tracés en rouge sur le schéma ci-dessous. Lorsqu'il n'y a que deux points sur un même cercle (par exemple les deux points violets à gauche), on ne conserve qu'un seul arc de cercle, et on ne tient pas compte des cercles où il y a un seul point (comme le point en vert à droite).



Notons u le nombre u de paires de points à distance 1. Chacune de ces paires orientées (pour pouvoir regarder le premier sommet comme un centre et le deuxième comme un point du cercle) va correspondre à 1 arête de G sauf dans au plus n cas (ceux où il n'y a qu'un ou deux points sur le cercle). On en déduit $2u \leq m - n$.

Or entre deux cercles il y a au plus deux croisements et il a n cercles, donc le nombre de croisements dans G est majoré : $t \leq 2\binom{n}{2} = n(n-1) \leq n^2$.

On en déduit

$$m \leq Cn^{\frac{4}{3}},$$

où C est une constante, ce qui majore aussi u par $O(n^{\frac{4}{3}})$.

Je doute que cette majoration soit optimale, à la main la configuration en réseau hexagonal qui semble la plus efficace donne $O(n)$.

<u>*Solution de l'exercice 6*</u> 1. Deux sommets adjacents v_i et v_j ne peuvent pas être dans I car on a soir $i < j$ soit $j < i$.

2. Un sommet v_i est dans I si tous ses voisins sont étiquetés par un j inférieur à i. Pour un ordre possible sur tous les sommets, on s'intéresse donc seulement à l'ordre induit sur v et ses $deg(v)$ voisins (où $deg(v)$ est le *degré* de v, c'est-à-dire justement son nombre de voisinis). Il a au total $(deg(v) + 1)!$ tels ordres induits, et parmi eux il y en a $deg(v)!$ pour lesquels v a l'étiquette maximale. La probabilité que v soit dans I est donc donnée par

$$\mathbb{P}(v \in I) = \frac{deg(v)!}{(deg(v) + 1)!} = \frac{1}{deg(v) + 1}.$$

3. En sommant sur ces probabilités (qui sont aussi les espérances des indicatrices "$v \in I$"), l'espérance sur le cardinal de I est :

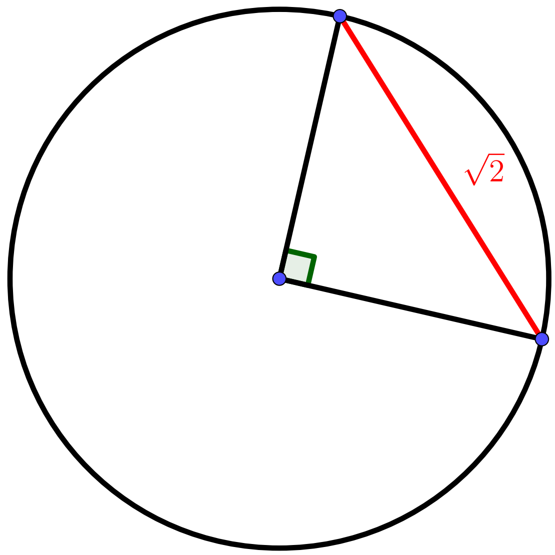
$$\mathbb{E}(|I|) = \sum_{v \text{ sommet de G}} \frac{1}{deg(v) + 1}.$$

De plus, la somme $\sum_{v \text{ sommet de G}}(deg(v)+1)$ est fixée (par double-comptage sur les arêtes, on établit l'identité classique $\sum_{v \text{ sommet de G}} deg(v) = 2m$). Donc la somme ci-dessus est minimale lorsque tous ses termes sont égaux (on peut s'en convaincre par récurrence en commençant par montrer que $\frac{1}{a+b} + \frac{1}{a-b} \geq \frac{2}{a}$). Ce cas correspond à $deg(v) = \frac{2m}{n}$ pour tout sommet v. D'où

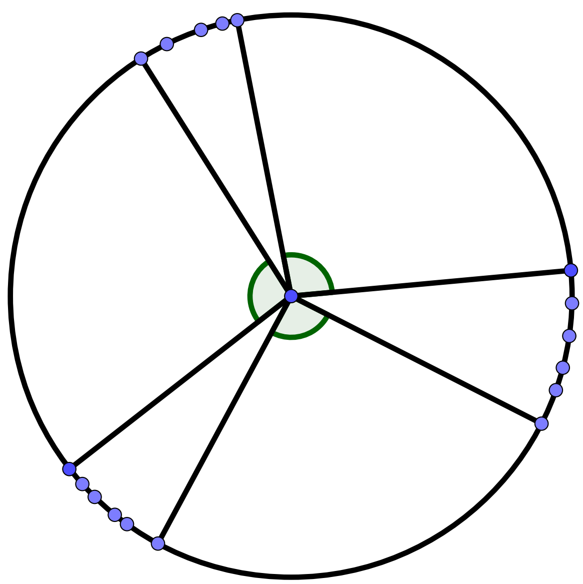
$$\mathbb{E}(|I|) \geq \frac{n^2}{2m + n}.$$

On en conclue qu'il existe un ensemble indépendant de taille minorée par les conditions du théorème de Turán.

<u>*Solution de l'exercice 7*</u> Établissons déjà une stratégie pour essayer de maximiser le nombre de points à distance $> \sqrt{2}$. On se rend rapidement compte qu'il vaut mieux placer les points sur le bord du cercle, et le théorème de Pythagore indique que si l'angle au centre entre deux points est supérieur à un angle droit, alors il sont à distance supérieure à $\sqrt{2}$ (voir schéma ci-dessous).



Une idée consiste alors à répartir les points par paquets de taille $\frac{n}{3}$ en dehors de trois portions d'angle strictement supérieur à l'angle droit (voir deuxième schéma).



Le nombre de paires de points à distance $\leq \sqrt{2}$ est alors de l'ordre de $\frac{n^2}{6} - \frac{n}{2}$.

La question est maintenant de s'assurer que ce sont les bonnes bornes, que l'on ne pourra pas trouver de configuration avec moins de telles paires de points.

Pour cela on construit le graphe G dont les sommets sont nos points et les arêtes relient deux points qui sont à distance $\leq \sqrt{2}$. Dans le disque unité, il ne peut pas y avoir de carré de côté de longueur $> \sqrt{2}$ donc parmi quatre points il y en a au moins 2 qui sont à distance $\leq \sqrt{2}$, *ie* dans G on a $\alpha(G) \leq 3$.

D'où en utilisant le théorème de Turán

$$m \geq \frac{n^2}{6} - \frac{n}{2},$$

ce qui montre que la construction précédente est bien optimale.