6. Graphes et arbres

Graphes et informatique

- Modélisation par les graphes :
  - gestion des réseaux
  - routage dans les circuits VLSI
  - autres routages
  - automates
  - …
- Graphes comme structures de données :
  - arbres (binaires)
  - graphes orientés ou non
- Problèmes classiques sur les graphes :
  - recherche de plus courts chemins
  - ordonnancement de tâches
  - problème de recherche opérationnelle
  - problèmes de flots
  - problème de l'arbre recouvrant
  - backtracking
  - …

Graphes

- Un graphe non orienté G = (S,A) est la donnée de 2 ensembles:
  - S est un ensemble de sommets
  - A est un ensemble d'arêtes (s,t) où on confond (s,t) et (t,s)

  $A \subseteq S \times S$

- Un graphe orienté G = (S,A) est la donnée de 2 ensembles :
  - S est un ensemble de sommets
  - A est un ensemble d'arcs (s,t)

  $A \subseteq S \times S$

- Une arête ou un arc (s,s) est appelée une boucle.

- On peut associer à G une valuation v : A → V, V quelconque.

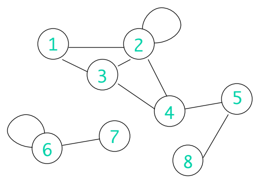
On étudie ici les <u>graphes finis</u> i.e. dont S et A sont deux ensembles <u>finis</u>.
On ne traite pas du tout les multigraphes.

Exemples

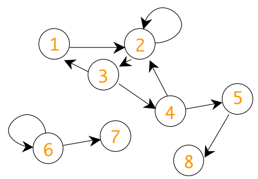
- Graphe G = (S,A) avec S = {1, 2, 3, 4, 5, 6, 7, 8} et
  A = {(1,2), (2,2), (2,3), (3,1), (3,4), (4,2), (4,5), (5,8), (6,6), (6,7)}

  <u>Remarque</u> on parle indifféremment de l'arête (i,j) ou de l'arête (j,i).

Version *non orientée*



Version *orientée*



Orientation

- Un graphe non orienté modélise une relation symétrique.
  Chaque arête « compte pour deux arcs … symétriques ».

- Un graphe orienté modélise une relation non forcément symétrique.

- Si une relation est réflexive, on peut choisir de ne faire figurer aucune boucle afin d'alléger le graphe.

- Un ordre est modélisé par un graphe orienté.
  Habituellement, on ne fait pas figurer les relations que l'on peut obtenir par transitivité, ni celles obtenues par réflexivité.
  Remarque : ce graphe orienté est à distinguer du diagramme de Hasse.

Représentation informatique des graphes

- liste des arêtes ou des arcs
  ((1, 2), (2, 2), (2, 3), (3, 1), (3, 4), (4, 2), (4, 5), (5, 8), (6, 6), (6, 7))

- matrice d'adjacence $M_G = (m_{ij})_{1 \le i,j \le |S|}$ $m_{ij} = 1$ si $(i,j) \in A$, $m_{ij} = 0$ sinon.

0	1	0	0	0	0	0	0
0	1	1	0	0	0	0	0
1	0	0	1	0	0	0	0
0	1	0	0	1	0	0	0
0	0	0	0	0	0	0	1
0	0	0	0	0	1	1	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

  G valué, les coefficients sont $m_{ij} = v(i,j)$ pour tout $(i,j) \in A$, $m_{ij} = 0$ sinon.
- matrice d'incidence $N_G = (n_{ij})_{1 \le i \le |S|; 1 \le j \le |A|}$
  n_{ij} est le nombre de fois où l'arête (ou arc) j est incidente au sommet i.
- liste d'adjacence
  liste de sommets et pour chacun, liste de ses successeurs.
  ((1, (2)), (2, (2, 3)), (3, (1, 4)), (4, (2, 5)), (5, (8)), (6, (6, 7)), (7, ( )), (8, ( )))

Vocabulaire

- Nombre de sommets |S| : ordre du graphe.
- Nombre d'arêtes ou d'arcs |A| : taille du graphe

- Nombre d'arêtes ou d'arcs incidents à s : d(s) degré de s
  Somme des degrés des sommets : d(G) degré du graphe

- Pour une arête (s,t) donnée, s et t en sont les extrémités.
  Pour un arc (s,t) donné, s est l'extrémité initiale et t l'extrémité terminale ; s est alors un prédécesseur de t, t un successeur de s.

- Sous-graphe d'un graphe G : on retire des arêtes/arcs à G.

- Graphe partiel du graphe G : on retire des sommets ainsi que leurs arêtes/arcs incidents.

Vocabulaire (suite)

Graphe non orienté :

- Une chaîne est une suite non vide d'arêtes adjacentes qui relient un sommet s à un sommet t.

- Un cycle est une chaîne telle que
  - la même arête ne figure pas deux fois
  - les deux sommets en bout de chaîne coïncident.

Graphe orienté :

- Un chemin est une suite non vide d'arcs consécutifs qui relient un sommet s à un sommet t.

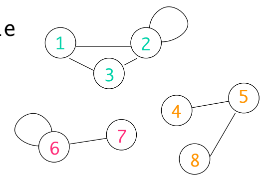
- Un circuit est à un chemin ce qu'un cycle est à une chaîne i.e. un chemin *fermé* ne passant pas 2 fois par le même arc.

Connexité

- Un graphe non orienté est connexe s'il existe une chaîne entre tout couple de sommets distincts.

- On peut partitionner l'ensemble des sommets en p classes $C_1,\dots, C_p$ de telle sorte que les p graphes partiels $G_i = (C_i, A_i)_{1\le i\le p}$ soient connexes.
  Les $(G_i)_{1\le i\le p}$ sont appelées les composantes connexes de G.

Exemple



Ce graphe a 3 composantes connexes.

Connexité et acyclicité

Théorème Tout graphe connexe d'ordre n>0 a au moins n-1 arêtes

Dém. par récurrence généralisée
- si n=1 c'est vrai.
- hyp.réc.gén : on suppose que la propriété est vraie aux ordres k avec 1<k<n.
- soit un graphe G = (S, A) à n sommets, prenons s un sommet de G.
- considérons le graphe partiel G' dont l'ensemble de sommets est S\{s}.
- partitionnons G' en composantes connexes $G_i=(S_i, A_i)_{1\le i\le p}$. G connexe donc
  $p \le d(s)$
  On a $|S| = 1 + \sum_{1\le i\le p}|S_i|$ et $|A|= \sum_{1\le i\le p}|A_i|+ d(s)$.
  Par hyp.réc.gén., pour tout i,

$$|S_i|- 1 \le |A_i|$$
$$\sum_{1\le i\le p}|S_i| - p \le \sum_{1\le i\le p} |A_i|$$
$$|S|-1 -p \le |A|- d(s)$$
$$|S|-1 \le |A|$$

Théorème Tout graphe sans cycle d'ordre n>0 a au plus n-1 arêtes.

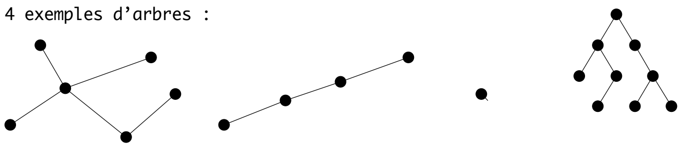
On pourrait montrer ces 2 théorèmes simplement en raisonnant par l'absurde.

Arbre

Par définition, un arbre est un graphe connexe sans cycle.

Théorème Soit un graphe G=(S, A) d'ordre n (n≥2). Les propriétés suivantes sont équivalentes :
1. G est un arbre
2. tout couple de sommets est relié par une chaîne unique
3. G est sans cycle et en ajoutant une arête on en crée un
4. G est connexe et si on supprime une arête il ne l'est plus
5. G est sans cycle et admet n-1 arêtes
6. G est connexe et admet n-1 arêtes

4 exemples d'arbres :



Premières propriétés

Propriété 1 Un arbre d'ordre n avec n>1 a au moins deux feuilles.

démonstration
- on considère une chaîne de longueur maximale et d'extrémités s et t.
- il existe un x tel que l'arête (s, x) existe.
- il n'existe pas d'arête (s, y) pour un autre sommet y de la chaîne car sinon, on aurait un cycle.
- il n'existe pas non plus d'arête (s, z) pour un sommet z hors de la chaîne car sinon, la longueur de la chaîne ne serait pas maximale.

Propriété 2 Un arbre d'ordre n avec n>0 a n-1 arêtes.

dém. récurrence sur l'ordre (on peut à la place utiliser les 2 théorèmes du tr.10)
- un arbre à 1 sommet a 0 arête.
- hyp.réc : supposons qu'un arbre à n sommets (n>1) ait n-1 arêtes.
- considérons un arbre à n+1 sommets. Par la propriété précédente, il a au moins une feuille f.
- le reste de l'arbre reste un arbre et a n sommets donc, par hyp.réc., il a n-1 arêtes.
- en ajoutant l'unique arête reliant f à l'arbre, il a n arêtes.

Démonstration du théorème

- (1⇒ 2)
  G est connexe donc tout couple de sommets est relié par une chaîne. Elle est unique car sinon, on a un cycle.

- (2⇒ 3)
  Si il y avait un cycle, deux sommets ne seraient pas connectés par une chaîne unique. Si on ajoute une arête entre deux sommets qui n'étaient pas reliés, elle forme un cycle avec la chaîne déjà existante entre les deux sommets.

- (3⇒ 1)
  G est sans cycle. Prenons deux sommets dans G. Si il existe une arête entre ces deux sommets, a fortiori il existe une chaîne entre les deux. Supposons qu'il n'existe pas d'arête entre les deux sommets. En ajoutant une arête entre les deux sommets, on crée un cycle. C'est qu'il y avait une déjà une chaîne entre ces deux sommets. Donc G est bien connexe. G est connexe sans cycle, c'est donc un arbre.

Démonstration du théorème (suite)

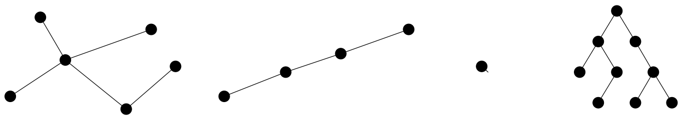
(1⇒ 4)
- G est un arbre donc G est connexe. Si on enlève une arête, on coupe l'unique chaîne qui reliait ses deux extrémités, le graphe n'est plus connexe.
- supposons la propriété 4 vraie et aussi que G a un cycle. En supprimant une arête du cycle, G reste connexe, d'où la contradiction. Donc G n'a pas cycle. Comme il est connexe, c'est un arbre.

(1⇒ 5)
- G est un arbre donc pas de cycle, d'ordre n donc à n-1 arêtes (t.12).
- G est sans cycle et a n-1 arêtes. Si on ajoute une arête, le graphe a n arêtes et donc on a créé un cycle (cf.t.10). La propriété 3 est vraie, elle implique la 1.

(1⇒ 6)
- G est un arbre d'ordre n alors il est connexe et a n-1 arêtes.
- G est connexe et a n-1 arêtes. Si on enlève une arête, le graphe a n-2 arêtes. Il n'est plus connexe (cf.t.10). G vérifie la propriété 4.

Vocabulaire forestier

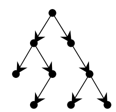
- Une forêt est un graphe dont chaque composante connexe est un arbre. Autrement dit, une forêt est un graphe sans cycle :



- Tout sommet d'un arbre est appelé un nœud.
- On appelle feuille un sommet d'un arbre adjacent à une seule arête.
- On peut distinguer un sommet particulier d'un arbre. On l'appelle racine et l'arbre sera dit enraciné.
- Une chaîne reliant la racine d'un arbre à une feuille est appelée une branche.

Arborescence

- Une arborescence k-aire est un arbre enraciné orienté dont chaque sommet a au plus k successeurs : ses fils.
  Quand k=2, l'arborescence est binaire.
  On parle abusivement mais couramment d'arbres, et plutôt d'arbres dessinés i.e. dont les fils sont ordonnés.



- Un niveau est un ensemble des nœuds qui ont en commun d'être équidistants de la racine. On les numérote à partir de 0.

- La hauteur (ou profondeur) d'un arbre binaire est le nombre d'arcs sur un chemin de longueur maximale.
  Selon le contexte, on peut aussi la définir comme étant le nombre de sommets sur un tel chemin. Les 2 définitions diffèrent de 1.
  (on a utilisé au TD4 la deuxième de ces définitions).

- Un arbre k-aire est complet si tous ses nœuds ont 0 ou k successeurs. Un arbre complet est saturé si, pour une hauteur donnée, il a un nombre maximal de nœuds.

Propriétés des arbres binaires

On considère un arbre binaire <u>saturé</u> à n nœuds (et la première définition pour la hauteur).

1. Le niveau i, $i \geq 0$, contient 2^i nœuds.

2. Le nombre de feuilles est donc égal à 2^h.

3. Le nombre de nœuds est égal à
$$n = 2^{h+1}-1$$

4. La hauteur est égale à
$$h = \log_2(n+1) - 1$$

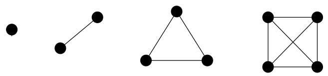
5. Le nombre de feuilles est égal à $(n+1)/2$.

Représentation informatique des arbres

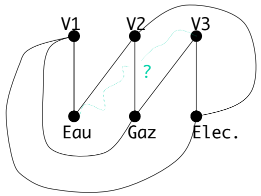
- Représentations générales des graphes
- Représentation séquentielle des arbres binaires quasi-saturés
  - on prend un tableau de n cases ($n \leq 2^{h+1}-1$, h étant la hauteur)
  - la case n°1 contient la racine de l'arbre
  - la valeur du <u>fils gauche</u> du nœud i, $0<i\leq n$, est dans la case 2i
  - la valeur du <u>fils droit</u> du nœud i, $0<i\leq n$, est dans la case 2i+1
- Représentation récursive des arbres binaires
  - un arbre est un pointeur sur un nœud
  - un nœud est une valeur et deux pointeurs
    - un sur l'arbre fils gauche
    - un sur l'arbre fils droit
- Représentation récursive des arbres k-aires
  - un arbre est un pointeur sur un nœud
  - un nœud a une valeur et deux pointeurs
    - un sur l'arbre appelé premier fils
    - un sur la liste des frères

Graphes complets

- Un graphe sans boucle est complet s'il existe une arête entre tout couple de sommets.
  Exemples : K_1 , K_2 , K_3 et K_4



- Problème des 3 villas et des 3 usines…

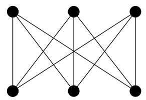


On dit qu'un graphe est planaire si on peut relier toute paire de sommets sans que deux arêtes ne se coupent.

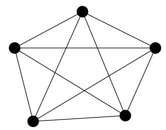
Graphes planaires

Théorème (Kuratowski, 1930)
Un graphe est planaire si et seulement si il ne contient pas de sous-graphe partiel isomorphe* à $K_{3,3}$ ou à K_5.
(admis)
Ces deux graphes ne sont pas planaires :

$K_{3,3}$



K_5



* Deux graphes sont <u>isomorphes</u> si ils sont égaux à un renommage de leurs sommets près.

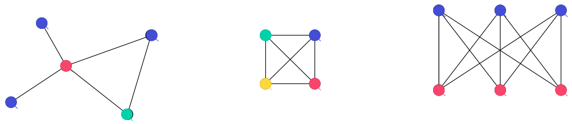
Nombre chromatique

Soit un graphe sans boucle $G=(S,A)$.

- Un coloriage est une application c de S dans un ensemble de couleurs $\{c_1,\dots, c_n\}$ telle que pour tout $(s,t) \in A$,

$$c(s) \neq c(t)$$

- On appelle nombre chromatique d'un graphe sans boucle G le plus petit entier k permettant d'effectuer un coloriage du graphe avec k couleurs.



k=3 k=4 k=2

Coloriabilité des graphes planaires

Théorème Tout graphe planaire sans boucle $G=(S,A)$ admet un coloriage en 5 couleurs.

(admis : la preuve se ferait par récurrence généralisée)

Qu'est ce que la conjecture des 4 couleurs ?...