Théorie des graphes

Chapitre 1 : Définitions et concepts de base

Joseph Gergaud & Géraldine Morin

5 janvier 2021

INP TOULOUSE ENSEEIHT

- Cette UE est divisée en deux parties :

  Théorie des Graphes,

  UE Théorie des Automates et des Langages.
- La partie *Graphes* est composée de 6 CTDs, 5 TP-Projet (en JULIA).
- Évaluation sur 1 Examen (70% de la note) et un BE (30% de la note)
- Il est rappelé qu'un taux supérieur ou égal 30% d'absence dans une UE n'autorise pas à passer le rattrapage.

Définition 1.1.1 – Graphe

Un **graphe** fini $G = (V, E)$ est défini par un ensemble fini, non vide, appelés **sommets** (en anglais *vertex/vertices*)

$$V = \{v_0, v_1, \ldots, v_{n-1}\}$$

et d'un ensemble fini d'arêtes (en anglais edges)

$$E = \{e_0, e_1, \ldots, e_{m-1}\}.$$

Chaque arête est définie un couple de sommets $\{v_i, v_j\}$ [a] appelés **extrémités** de e.

a. Pour les arêtes la notation n'est pas la notation ensembliste, l'arête $\{v_i, v_i\}$ est une boucle et n'est pas l'ensemble $\{v_i\}$.

Définition 1.1.2 – Vocabulaire

- Si $n = \#V$ on dit que $G = (V, E)$ est un graphe d'**ordre** n.
- Si $e = \{v_i, v_j\}$, on dit que l'arête e est **incidente** au sommet v_i (et au sommet v_j aussi d'ailleurs.)
- Si $e = \{v_i, v_j\}$, on dit que v_i et v_j sont **adjacents**.
- Deux arêtes sont **adjacentes** si et seulement si elles ont un sommet commun.

Exemple 1.1.1. Exemples $G = (V, E)$ est-il bien défini pour

1) $V = \{v_1, v_2, v_3, v_4, v_5\}$ et $E = \emptyset$.

▷ **Solution**

YES

2) $V = \{1, 2\}$ et $E = \{\{1, 3\}, \{1, 4\}, \{1, 5\}, \{2, 3\}\}$.

▷ **Solution**

NO

3) $V = \{1, 2, 3\}$ et $E = \{\{1, 1\}, \{1, 2\}, \{2, 3\}\}$.

▷ **Solution**

YES, les boucles sont OK.

4) $V = \{1, 2, 3\}$ et $E = \{\{1, 2\}, \{1, 3\}, \{1, 3\}\}$.

▷ **Solution**

YES, C'est un multigraphe.

Définition 1.1.3

- $e = \{v_i, v_i\}$ est une **boucle** ;
- un graphe dans lequel il peut exister plusieurs arêtes entre deux noeuds est appelé **multigraphe** ;
- un graphe est dit **simple** si

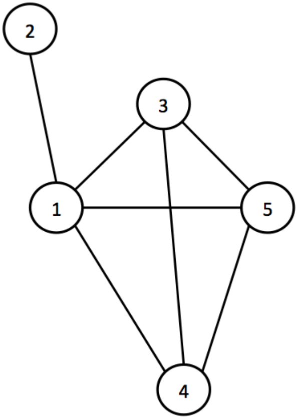
  il n'a pas de boucle,

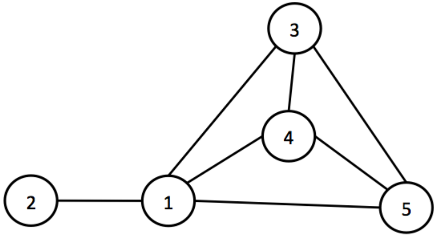
  il a au plus une arête entre deux sommets.

Dans le suite, nous considérons presque toujours des graphes simples (sinon, on le signale !).

Il existe une infinité de manières de représenter graphiquement un graphe, c'est à dire de dessiner un graphe dans le plan.
Exemple 1.2.1. $V = \{1, 2, 3, 4, 5\}$ et
$E = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{1, 5\}, \{3, 4\}, \{3, 5\}, \{4, 5\}\}$.
Dessinez ce graphe de plusieurs façons différentes.

> **Solution**





Définition 1.2.1 – Degré

Le **degré** d'un sommet v_i est le nombre d'arêtes incidentes en v_i. On le note $\delta(v_i)$. Une boucle participe pour 2.

Définition 1.2.2 – Graphe simple complet

Un graphe simple est dit **complet** si tout sommet est adjacent à tout autre.

Exercice 1.2.2. Quel est le nombre d'arêtes du graphe simple complet d'ordre n ?

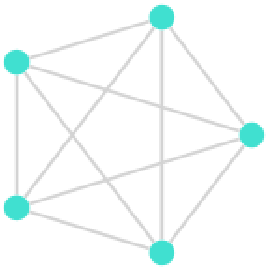
> **Solution**

K_n a $\frac{n(n-1)}{2}$ arêtes.
À calculer en additionnant le nombre d'arêtes partant de chaque sommet, en les "épluchant" au fur et à mesure, i.e., de $n-2$ à 1.

Remarque 1.2.1. Le graphe simple complet d'ordre n est noté K_n.

Exercice 1.2.3. Dessiner K_5.

> **Solution**



***Lemme 1.2.1** (lemme des poignées de main). La somme des degrés de tous les sommets est égale à deux fois le nombre d'arêtes.*

> **Solution**

Chaque arête est comptée 2 fois.

Corollaire 1.2.3

Le nombre de sommets de degré impair est pair.

à justifier.

> **Solution**

$$2m = \sum_k \delta(v_k) = \sum_{\substack{v_k \texttt{ sommet} \\ \texttt{de degré pair}}} \delta(v_k) + \sum_{\substack{v_k \texttt{ sommet} \\ \texttt{de degré impair}}} \delta(v_k).$$

Le premier terme est une somme de termes pairs, c'est donc un nombre pair. Par suite le deuxième terme est aussi pair ; c'est une somme de nombre impair. Il faut donc avoir un nombre de termes pair dans cette somme.

Exercice 1.2.4. Montrer que s'il y a n personnes dans une salle avec $n > 1$, au moins 2 d'entre elles ont le même nombre de connaissances dans la salle.

Solution

On considère le graphe dont les sommets sont les personnes et les arêtes modélisent le fait que 2 personnes se connaissent. On enlève toutes les personnes qui ne se connaissent pas. Si ce nombre de personnes est supérieur ou égale à 2 c'est fini. Sinon, si on enlève une personne, alors $n > 2$. Il reste donc maintenant à montrer le résultat pour un graphe où les degrés de tous les sommets sont supérieurs ou égal à 1. Supposons que pour tout $i \neq j$ on ait $\delta(v_i) \neq \delta(v_j)$ et "ordonnons" les sommet par ordre croissant sur le degré : $\delta(v_0) < \delta(v_1) < \cdots < \delta(v_{n-1})$. Comme $\delta(v_0) \geq 1$ on a $\delta(v_i) \geq i+1$ et en particulier $\delta(v_{n-1}) \geq n$; ce qui est impossible.

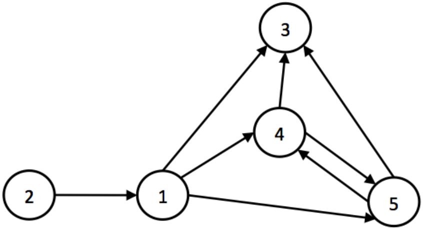
Définition 1.2.4 – Graphe régulier

Un graphe est dit **régulier** si tous ses sommets sont de même degré. Si ce degré est k, on dit que G est k-**régulier**.

Définition 1.3.1 – Graphe orienté

Un graphe $G = (V, E)$ est dit **orienté** si chaque arête est une paire de sommets (donc ordonnée), $e \in E, e = (i, j)$. On appelle les éléments de E des arcs.

Exemple 1.3.1.



Définition 1.3.2 – Vocabulaire

- i est appelé l'**origine** de $e = (i, j)$;
- j est appelée l'**arrivée** de e.
- le **degré sortant** d'un sommet v $\delta^+(v)$ est le nombres d'arcs d'origine v ;
- le **degré entrant** d'un sommet v $\delta^-(v)$ est le nombre d'arcs qui arrivent en v ;
- le **degré total** $\delta(v) = \delta^+(v) + \delta^-(v)$.

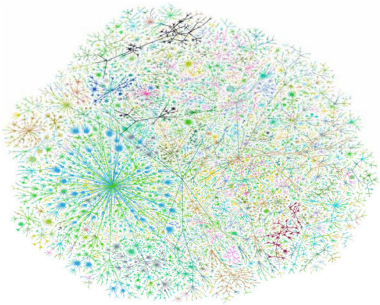
Remarque 1.3.1. on a toujours $\sum_{v \in V} \delta(v) = 2\#E$.

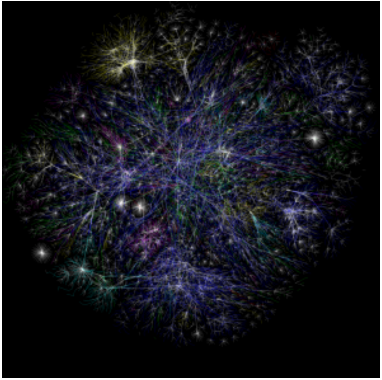
Exercice 1.3.2. Donner des exemples de problèmes que l'on peut modéliser par un graphe.

> **Solution**

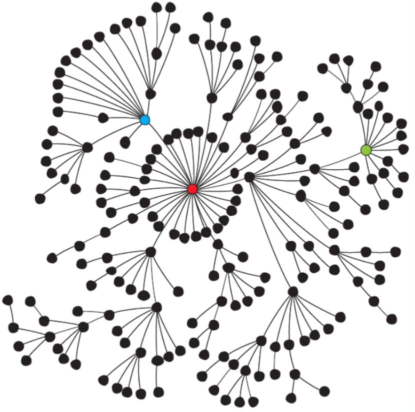
Exemples

- Réseaux de diffusion d'un virus
- Exemples réseaux : À gauche réseaux de neurones, à droite réseau internet



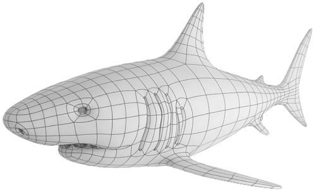


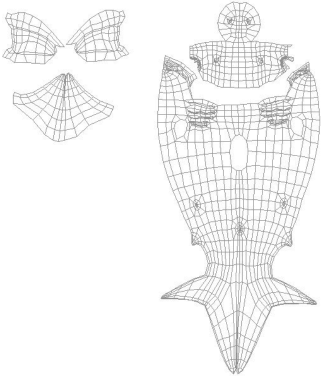
On peut aller de n'importe quel point du réseau à un autre par un chemin assez court. Cela correspond à cette idée populaire qui dit que l'on peut trouver un lien entre 2 personnes prises au hasard dans un pays en moins de 6 connexions. Certains nœuds ont un statut particulier : ce sont des "hub"



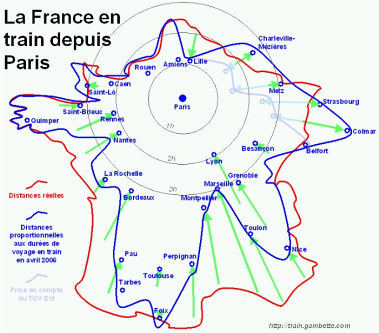
Exemple : maillages

Maillage en 3D : les maillages 3D sont des graphes dont les sommets sont positionnés en 3D (plongement géométrique). Le graphe représente les relations de voisinage entre les sommets (image de droite), appelée la topologie. (image du maillage *Shark* de CGStudio).



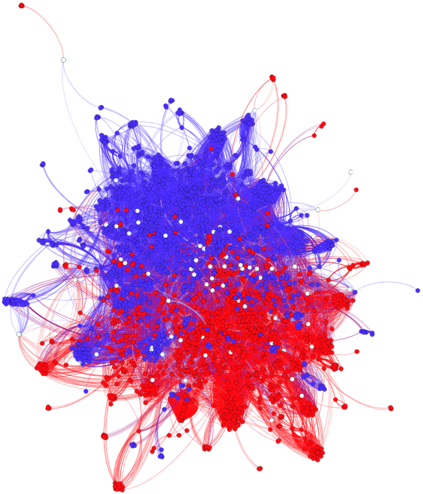


- Calculs de plus courts chemins (cartes de distance par rapport aux temps de transport, (image crée par Philippe Gambette http://philippe.gambette.free.fr/Train/)



- coloration : identification de données indépendantes, par exemple pour effectuer des calculs en parallèle.

identification de *clusters*, c'est à dire, des groupes fortements connectés. Une exemple avec une coloration des *Hashtag* de *Twitter* utilisés majoritairement par des démocrates ou des républicains (images de Thimoty Renner).



Définition 1.4.1 – Sous-graphe

Un **sous-graphe** de $G = (V, E)$ engendré par un sous-ensemble de sommets $V' \subset V$ est $G' = (V', E')$ où E' représente toutes les arêtes de E ayant leur deux extrémités dans V'.

Exemple 1.4.1. Si G est le graphe des routes de France, celui représentant les routes d'Occitanie est un sous-graphe.

Définition 1.4.2 – Graphe partiel

Un **graphe partiel** de $G = (V, E)$ engendré par $E' \subset E$ est le graphe $G' = (V, E')$.

Exemple 1.4.2. Si G est le graphe des routes de France, donner un exemple de graphe partiel.

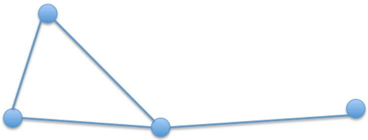
> **Solution**

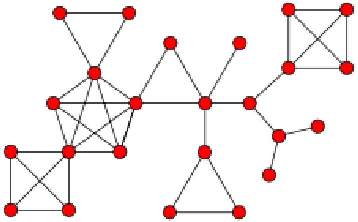
Celui représentant les autoroutes de France est un graphe partiel.

Définition 1.4.3 – Clique

Une **clique** est l'ensemble des sommets d'un sous-graphe complet.

Exemple 1.4.3. Ces graphes contiennent-ils des cliques ?





(image de droite : source Wikipedia)

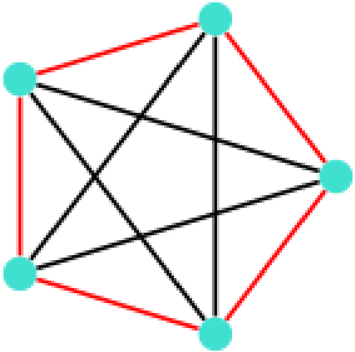
Exercice 1.4.4. Montrez que dans un groupe formé de six personnes, il y en a nécessairement trois qui se connaissent mutuellement ou trois qui ne se connaissent pas (on suppose que si A connaît B, B connaît également A).
Montrez que cela n'est plus nécessairement vrai dans un groupe de cinq personnes.

Solution

Construisons un graphe dont les sommets représentent les six personnes ; deux sommets sont reliés par une arête noire lorsque les personnes se connaissent et rouge dans le cas contraire. Il s'agit de prouver que ce graphe contient une clique K_3 dont les arêtes sont de même couleur. Si l'on ne tient pas compte de la couleur des arêtes, on obtient le graphe complet K_6. De chaque sommet partent cinq arêtes, et au moins trois d'entre elles sont de même couleur (noire ou rouge). Considérons la clique K_4 composée des sommets 1, 2, 3 et 4. Supposons, par exemple, que les arêtes $\{1, 2\}$, $\{1, 3\}$ et $\{1, 4\}$ soient noires. Aucune des arêtes entre $\{2, 3, 4\}$ ne peut être noire, sinon, on obtient un K_3 noir, mais si elles sont toutes les trois rouges, on obtient un K_3 rouge.

> **Solution**

Contre-exemple pour K_5 : on prend les arêtes 'ne se connaissent'/rouge sur la couronne.



- Il existe de nombreuses façons de stocker un graphe
- La complexité des algorithmes dépend de la représentation machine d'un graphe
- On prendra soin pour les représentations proposées de regarder la complexité à la fois en place et en temps (lors d'accès) des représentations proposées

Définition 1.5.1 – Matrice d'incidence

La matrice d'incidence sommets-arcs d'un graphe orienté $G = (V, E)$ d'ordre n ayant m arcs est une matrice $A = (a_{iu})$, $i = 1 \dots n$, $u = 1 \dots m$ à coefficients entiers 0, $+1$, -1, telle que chaque colonne correspond à un arc de E et chaque ligne correspond à un sommet de V. Si $e_k = (v_i, v_j)$ alors la colonne k a tous ses termes nuls sauf $a_{ik} = +1$ et $a_{jk} = -1$.

Exemple 1.5.1. Coder le graphe orienté donné à l' exemple 1.3.1.

> **Solution**

$$A = \begin{pmatrix} -1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & 1 & 1 \\ 0 & 0 & 1 & -1 & 0 & 1 & -1 & 0 \end{pmatrix}$$

Exercice 1.5.2. Trouver comment exprimer le degré entrant et sortant en un sommet.

> **Solution**

les degrés entrants (resp. sortants) sont le nombre de -1 (resp. $+1$) sur la ligne correspondant au sommet

C'est l'équivalent de la matrice sommet-arcs pour un graphe non orienté. Dans ce cas, on met $+1$ partout.

Exemple 1.5.3. Codage du graphe de l'exemple précédent en considérant le graphe non orienté.

Solution

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

Cela correspond en fait à une matrice d'incidence 'sommet-sommet'.

Définition 1.5.2 – Matrice d'adjacence

La matrice d'adjacence d'un graphe simple orienté $G = (V, E)$ est une matrice A à coefficients booléens telle que $A = (a_{ij})$ avec $a_{ij} = 1$ si $(i, j) \in E$.

Remarque 1.5.1. Si le graphe est non orienté, pour chaque arête $\{i, j\}$, $a_{ij} = a_{ji} = 1$; la matrice A est donc symétrique.

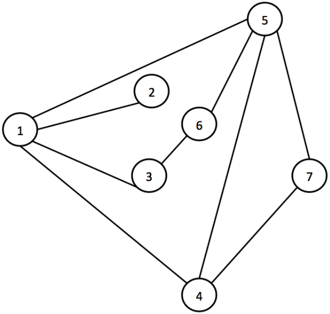
Exercice 1.5.4. Comparez la taille de stockage d'un graphe en utilisant ces différentes matrices. On considère les cas où la matrice est stockée entièrement

Solution

- Matrice d'incidence : nm.
- Matrice d'adjacence : n^2.

On utilise deux tableaux T_V de dimension $n+1$, et T_E de dimension m (cas orienté) ou $2m$ (cas non orienté). Les sommets voisins du sommet i sont stockés entre les indices $T_V(i)$ et $T_V(i+1)-1$ du tableau T_E.
Alternativement, le tableau T_V peut contenir des listes chainées d'indice des sommets voisins.

Exercice 1.5.5. Donner pour le graphe suivant les vecteurs T_V et T_E :



> **Solution**

Les indices des tableaux commencent à 0
$T_V = (0, 4, 5, 7, 10, 14, 16, 18)$, $T_E = (2, 3, 4, 5, 1, 1, 6, 1, 5, 7, 1, 4, 6, 7, 3, 5, 4, 5)$.

Exercice 1.5.6. Calculer le coût de stockage pour une représentation matricielle ou vectorielle. Calculer le temps nécessaire pour savoir si l'arête (i, j) est dans le graphe.

> **Solution**

- Coût de stockage pour un graphe non orienté : $n + 1 + 2m$.
- Temps d'accès

  matrice d'incidence : $O(m)$;

  matrice d'adjacence : $O(1)$;

  codage vectoriel : $O(min(n, m))$ ou $O(\log(min(n, m)))$ si on trie les voisins, ou mieux $O(1 + \max_{i=1...n} \delta(i))$.

Remarque 1.5.2. Codage vectoriel à partir de la matrice d'incidence : on peut représenter, à partir de deux tableaux T_1 et T_2 de dimension m, $T_1(i)$ et $T_2(i)$ donnant les extrémités de la i-ème arête.

Définition 1.6.1 – Graphe pondéré

Un graphe est **pondéré** si à chaque arc, on associe un poids réel positif (ou coût).

Exercice 1.6.1. Adapter les représentations ci-dessus pour prendre en compte des graphes pondérés (chaque arête a un poids positif).

> **Solution**

Réponse : pour les matrices d'incidence, ou peut multiplier une colonne par la valeur du poids de l'arête, pour les matrices d'adjacence, on peut faire apparaitre le poids dans la matrice. Pour la représentation vectorielle, on peut avoir un vecteur poids de taille m associé à T_E.