S2-MINEUR - IA Programmation

Chapitre 1 - La théorie des graphes

MULAPI TITA Ketsia

Année académique 2024-2025

Plan du cours

Ce cours, d'une durée totale de 28 heures, est structuré en trois parties. Chacune d'elles est sanctionnée par un travail pratique et, l'ensemble du cours est également évalué par une présentation de projet : individuelle pour les Data Engineers, et en binôme pour les Data Scientists. Cette présentation finale aura lieu lors de la dernière séance.

Les chapitres :

1. Les notions de la théorie des graphes
2. La recherche opérationnelle : algorithmes
3. Graph Neural Network

1. Notions fondamentales des graphes

Graphe, sommet, arc, arête, boucle

Un **graphe** $G = (X, u)$ est composé de :

- X : ensemble des **sommets** (noeud)
- u : ensemble des **arcs** orientés ou **arêtes** non orientées, avec $u \subseteq X \times X$

G peut alors être un graphe orienté ou non.

Types d'arcs :

- Arc : (x, y) avec orientation de x vers y
- Arête : $(x, y) = (y, x)$ (non orientée ou à double sens)
- Boucle : (x, x) , c'est (une arête ou un arc qui relie un sommet à lui-même)

Illustrations graphe, sommet, arc, arête, boucle

Liaison

Figure 1: arc et arête

Boucle

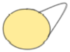


Figure 3: boucle

Sommet

Figure 2: sommet / noeud

Graphe G = (X=2,u=1)

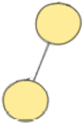
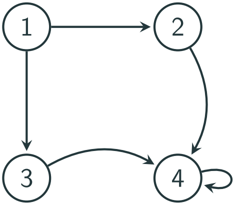


Figure 4: Graphe

Représentation sagittale

Cette façon de représenter les choses s'appelle **représentation sagittale**!

- Sommets : représentés par des points
- Arcs : segments fléchés entre sommets
- Elle n'est pas unique $\rightarrow$ on parle de graphe **topologique**



Ordre d’un graphe

L’ordre d’un graphe $G = (X, u)$ est le cardinal de l’ensemble X , noté $n = |X|$, c’est le **nombre de sommets** du graphe.

Quand on prend le graphe de la page précédente, on dit qu’il est d’ordre 4 car il possède 4 nodes.

Notion des arcs de mêmes formes et des p-graphe

Un graphe est un p-**graphe** si *au plus* p arcs de même forme peuvent exister

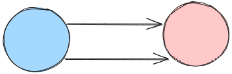
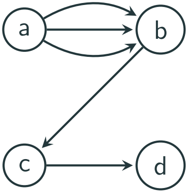


Figure 5: 2-graphe

Cas particulier : un **1-graphe** ne contient qu'un seul arc de chaque forme (x, y)

Un autre exemple illustratif

Graphe G **:** 3-graphe d'orde 4 avec arcs redondants



2. Notions complémentaires

- y est un **successeur** $(\Gamma^+(x))$ de x si $(x, y) \in u$
- x est un **prédécesseur** $(\Gamma^-(y))$ de y si $(x, y) \in u$

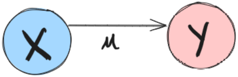


Figure 6: y successeur de x, et x prédécesseur de y

Voisins et isolement

- **Le voisinage :** y est un **voisin** de x ssi $y \in$ au voisinage de x noté :

$$\Gamma(x) = \Gamma^{+}(x) \cup \Gamma^{-}(x)$$

- Un sommet est **isolé** ssi $\Gamma(x) = \emptyset$

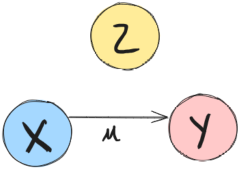


Figure 7: Z est un sommet isolé du graphe G

3. Degré, adjacence et incidence

Degré : demi-degré intérieur et extérieur

- $d^{+}(x)$: demi-degré sortant

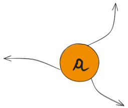


Figure 8: Demi-degré exterieur

- $d^{-}(x)$: demi-degré entrant

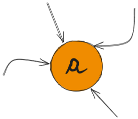


Figure 9: Demi-degré interieur

- $d(x) = d^{+}(x) + d^{-}(x)$: degré total

Un graphe est régulier si tous les sommets ont le même degré.

Adjacence et incidence

- **Sommets adjacents :** 2 sommets daans G sont adjacents si ils sont reliés directement et,
- **Arcs adjacents :** 2 arcs u et v sont adjacents si ils ont une extrémité commune

Incidence (sommet/arc)

Soit le graphe G=(X,U) et A $\in$ X

- $\omega^+(A)$ **:** C'est l'ensemble d'arcs ayant leur extrémité initiale dans A terminale dans U-A
- $\omega^-(A)$ **:** C'est l'ensemble des arcs ayant leur extrémité terminale dans A et initiale dans U-A
- le cocycle de G dans A est $\omega(A) = \omega^+(A) + \omega^-(A)$ **:** c'est l'ensemble des arcs (arêtes) incidents à A $\in$ X
- Incidence numérique :
  1. $m_G^+(x, y)$ **(arcs de x vers y) :** nombre d'arcs de G, ayant comme extrémité initiale x et y en terminale
  2. $m_G^-(x, y)$ **(arcs de y vers x) :** nombre d'arcs de G, ayant comme extrémité terminale x et y en initiale

4. Types particuliers de graphes

Un **graphe simple** ne possède ni boucle, ni arcs multiples et, en non orienté, chaque paire de sommets est reliée par au plus une arête.

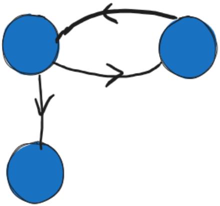


Figure 10: G est un graphe simple

Graphe symétrique

- Pour tout arc (x, y), l'arc inverse (y, x) appartient également au graphe.
- Cela signifie que la relation est **symétrique** :
  $\forall x, y, (x, y) \in u \Rightarrow (y, x) \in u$.

Figure 11: G est un graphe symétrique

Graphe anti-symétrique

Un graphe est **anti-symétrique** si
$\forall x, y, \in G, x \neq y \ et \ (x, y) \in u \Rightarrow (y, x) \in u$.
Il peut contenir des boucles (orientés ou non), mais pas de couples d'arcs opposés.

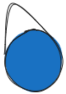


Figure 12: G est un graphe anti-symétrique

Sous-graphe

Lorsque nous avons parlé **d'incidence**, A était un sous-graphe de G, car un **sous-graphe** est obtenu en prenant un sous-ensemble des sommets et des arcs. Il conserve la structure du graphe d'origine sur les éléments sélectionnés.

Soit A $=\{c, d, e\}$ et G $= \{a, b, c, d, e\}$.

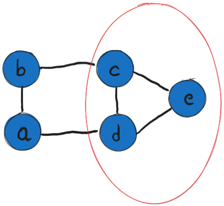


Figure 13: Avec G=(X,U), si A $\in$ X, alors, A est un sous graphe de G noté G_A

Un **graphe partiel** est un cas particulier de sous-graphe car, il contient tous les sommets du graphe d'origine mais seulement une partie de ses arcs.

Si G=(X,U) et V $\in$ U, alors $G^{'}$=(X,V).

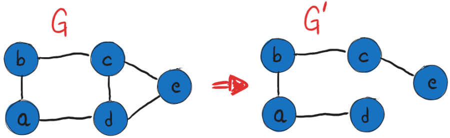


Figure 14: $G^{'}$ est un sous graphe de G

Graphe biparti

2 règles à retenir :

- Un graphe est **biparti** si son ensemble de sommets peut être partitionné en 2 classes : $X = X_1 \cup X_2$.
- Aucun sommet de X_1 n'est adjacent à un autre de X_1, et idem pour X_2.

Nous allons illustrer cela avec un exercice intéressant que j'aime bien et qui vous attend.

Graphe complet (n-clique)

- Un **graphe complet** est un graphe simple dans lequel chaque paire de sommets est reliée par une arête (on dit qu'ils sont 2-à-2 adjacents ou égaux).
- S'il contient n sommets, on parle de n-**clique**.
- Il possède $\dfrac{n(n-1)}{2}$ arêtes.

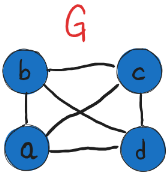


Figure 15: G est un graphe complet

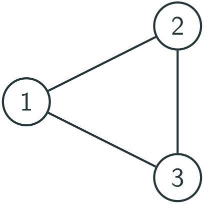
5. Matrices associées aux graphes

Matrice d'incidence sommets-arêtes

- Graphe **non orienté**
- Matrice $A = (a_{ij})$ telle que :

$$a_{ij} = \begin{cases} 1 & \text{si le sommet } i \text{ est incident à l'arête } j \\ 0 & \text{sinon} \end{cases}$$

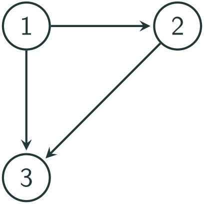
- Chaque colonne représente une arête (non orientée), chaque ligne un sommet.



Matrice d'incidence sommets-arcs

- Graphe **orienté**
- Matrice $A = (a_{ij})$ telle que :

$$a_{ij} = \begin{cases} 1 & \text{si le sommet } i \text{ est extrémité initiale de l'arc } j \\ -1 & \text{si le sommet } i \text{ est extrémité terminale de l'arc } j \\ 0 & \text{sinon} \end{cases}$$

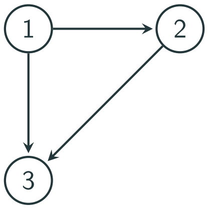


Matrice d'adjacence (sommets-sommets)

- Pour tout graphe (orienté ou non)
- Matrice $A = (a_{ij})$ telle que :

$$a_{ij} = \begin{cases} 1 & \text{si } (i,j) \in u \\ 0 & \text{sinon} \end{cases}$$

- Représente les connexions directes entre sommets.



6. Connexité et forte connexité

Chaînes et cycles (simples et élémentaires)

- Une **chaîne** est une suite d'arêtes reliant une suite de sommets.

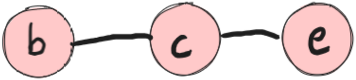


Figure 16: Une chaîne de longueur q = 2

- Un **cycle** est une chaîne fermée (le premier et le dernier sommet sont identiques).

Figure 17: Un cycle

Chaînes et cycles (simples et élémentaires)

- Une **chaîne simple** (aucune arête répétée) : c'est une chaine dans laquelle, en la parcourant, on ne rencontre pas deux fois la même arêtes (donc pas 2 fois égales).
- Une **chaîne élémentaire** (aucune sommet répété) : c'est une chaine dans laquelle, en la parcourant, on ne rencontre pas deux fois le même sommet (on dit qu'il est élémentaire et simple).

Toutes chaînes élémentaires est simple, mais l'inverse est fausse.

- Un **chemin** est une suite de sommets reliés par des arcs orientés.

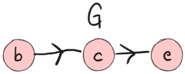


Figure 18: Une chaîne de longueur 2, noté $L_2 = \{(b, c), (c, e)\}$

- Un **circuit** est un chemin fermé (retour au point de départ).

Figure 19: Un cycle

- Ces notions sont centrales pour la **forte connexité**.

Graphes connexes et composantes connexes

- Un graphe est **connexe** s'il existe un chemin (ne jamais tenir compte des orientaitons, donc considerer une chaine) entre chaque paire de sommets...autrement, un graphe est connexe si à partir d'un sommet, on peut atteindre tous les sommets du graphe via une chaine.
- Une **composante connexe** est un sous-graphe connexe maximal.
- Le nombre de composantes connexes est appelé **nombre de connexité K**.

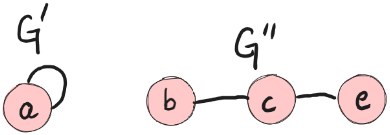


Figure 20: Soit G $= \{G^{'},\ G^{"}\}$, est-ce que G est connexe ? Quelle est la K-connexité de G ?

Isthme et point d'articulation

- Un **isthme** est une arête dont la suppression augmente le nombre de composantes connexes.
- Un **point d'articulation** est un sommet dont la suppression déconnecte le graphe et augmente le nombre de connexité.

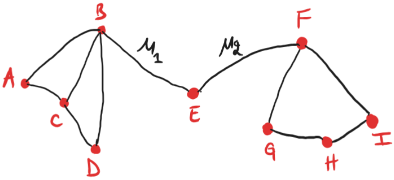


Figure 21: Ce graphe est connexe ? Trouvez un potentiel point d'articulation et un isthme !

Théorèmes et définitions associées (tips)

- Tout graphe connexe a au moins $n-1$ arêtes.
- Si un graphe non orienté a plus de $\frac{(n-1)(n-2)}{2}$ arêtes, alors il est connexe.
- Un graphe sans cycle avec $n-1$ arêtes est un **arbre** (on y reviendra).

Graphe fortement connexe

- Un graphe est **fortement connexe** s'il existe un chemin orienté entre chaque couple de sommets x et y (donc ce sont des vérifications à effectuer).
- Ses **composantes fortement connexes** sont les sous-graphes fortement connexes maximaux.

7. Arbres et arborescences

Définition d'un arbre

- Un **arbre** est un graphe **connexe sans cycle**.
- Il possède $n - 1$ arêtes si le graphe a n sommets.
- Il existe un chemin unique entre chaque paire de sommets ou encore, c'est une chaîne qui n'est pas fermée..

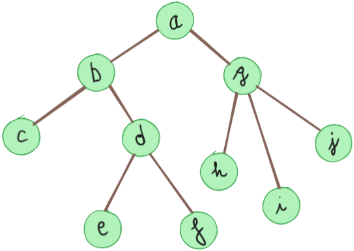


Figure 22: Un arbre

Forêt, branches et cordes

- Une **forêt** est une collection disjointe d'arbres.
- Les **branches** sont les arêtes de l'arbre.
- Une **corde** est une arête ajoutée à un arbre qui crée un cycle.

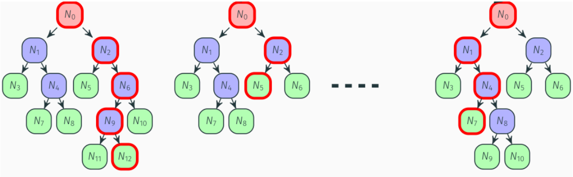


Figure 23: Une forêt

Propriétés caractéristiques des arbres

- Connexe, sans cycle, $n - 1$ arêtes.
- Toute arête est un **pont** (sa suppression déconnecte l'arbre).
- Un arbre est **minimement connexe** : aucune arête n'est superflue.

Définition d'une arborescence

- Une **arborescence** est un **arbre orienté** ayant une **racine** r.
- Il existe un **chemin orienté unique** de r vers tout autre sommet.
- Les sommets sans successeurs sont appelés **feuilles**.

Graphes quasi-fortement connexes

- Un graphe est **quasi-fortement connexe** s'il existe un sommet r tel que tout sommet est accessible depuis r.
- Ce type de graphe peut être représenté par une arborescence enracinée en r.

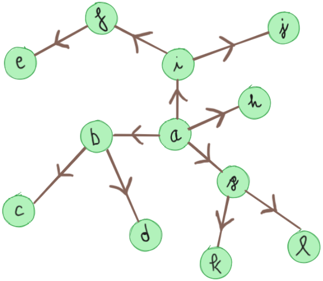


Figure 24: Oui ou non ce graphe est une arborescence ? si oui quel est le sommet racine ?

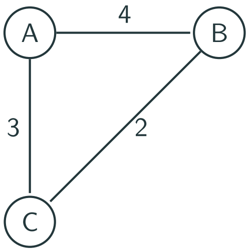
Propriétés des arborescences

- Une arborescence avec n sommets possède $n - 1$ arcs.
- Tous les sommets sauf un (la racine) ont un degré entrant égal à 1.
- Une seule racine avec degré entrant nul.

8. Graphes pondérés, réseaux, réseaux de transport

Graphes pondérés : définition et exemples

- Un **graphe pondéré** est un graphe $G = (X, u, C)$ où chaque arête (arc) (i, j) est associé à un poids $C_{ij} \in \mathbb{R}$.
- Le poids peut représenter un coût, une distance, un temps, une pénalité, une probabilité, etc.



Matrice de pondération

- La **matrice de pondération** $C = (C_{ij})$ contient les poids des arcs du graphe.
- Dans un graphe orienté et pondéré, si $(i, j) \notin u$, on peut poser $C_{ij} = +\infty$ ou une constante $M \gg 0$.

$$C = \begin{pmatrix} 0 & 4 & 3 \\ 4 & 0 & 2 \\ 3 & 2 & 0 \end{pmatrix}$$

Définition formelle d’un réseau

- Un **réseau** est un graphe orienté pondéré sans boucle.
- Il est représenté par $R = (X, u, C)$ où C donne les capacités ou coûts.
- Les arcs sont généralement utilisés pour représenter des flux ou du transport.

Réseau de transport : caractéristiques

- Un **réseau de transport** est un réseau sans boucle ni cycle de coût négatif.
- Utilisé dans les modèles de circulation, distribution, logistique, etc.
- Les pondérations suivent une logique de faisabilité : $C_{ij} = M \gg 0$ si $(i,j) \notin u$.

Bon à savoir : Un réseau informatique est un réseau de transport.

Travaux pratiques

1. Théorique (à la main) : résoudre un bon nombre d'exercices
2. Pratique (sur machine) : utiliser la librairie NetworkX sur NODG.