Les graphes: un outil mathématique pour le 21ème siècle

Alain Valette
Université de Neuchâtel

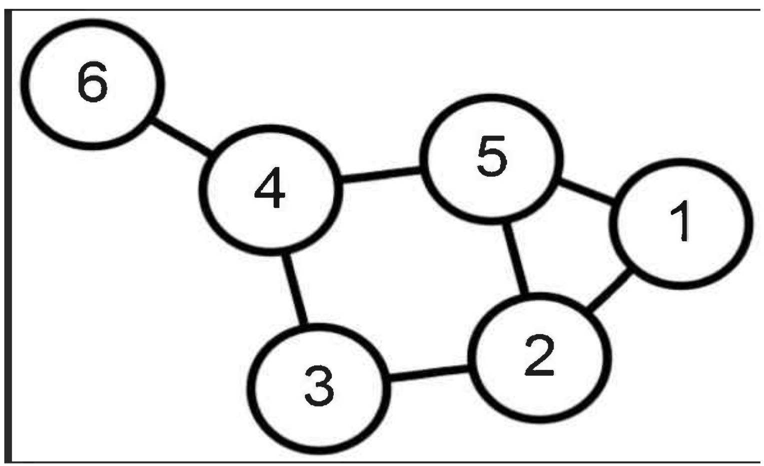
U3A, 11 janvier 2019

1 Graphes

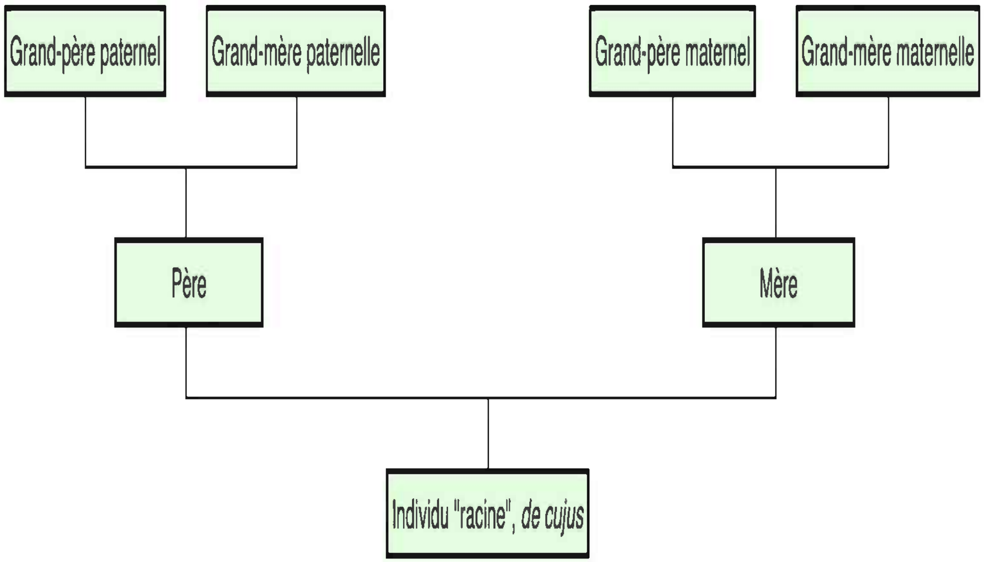
1.1 Un graphe, c'est quoi?

Dans la version la plus simple, un graphe c'est:

- des points, appelés **sommets** ou **noeuds**, qui représentent des choses ou des personnes;
- des **arêtes** ou **arcs**, qui représentent des relations, des interactions, des transitions... entre les sommets.



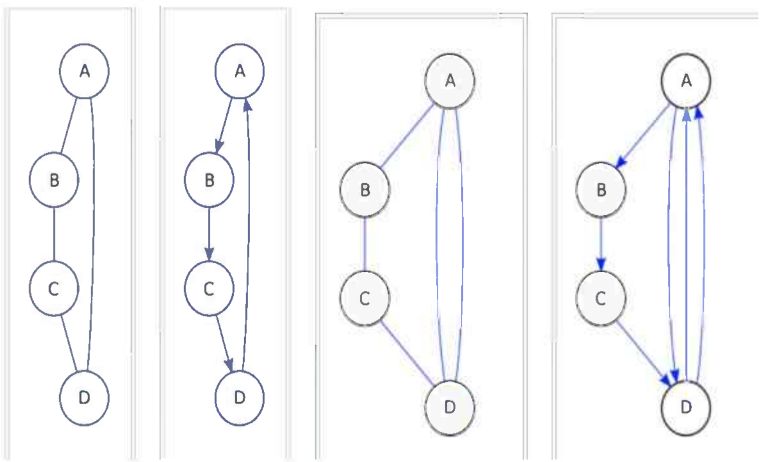
Un usage ancien des graphes: les arbres généalogiques.



Usage plus récent: plan du métro

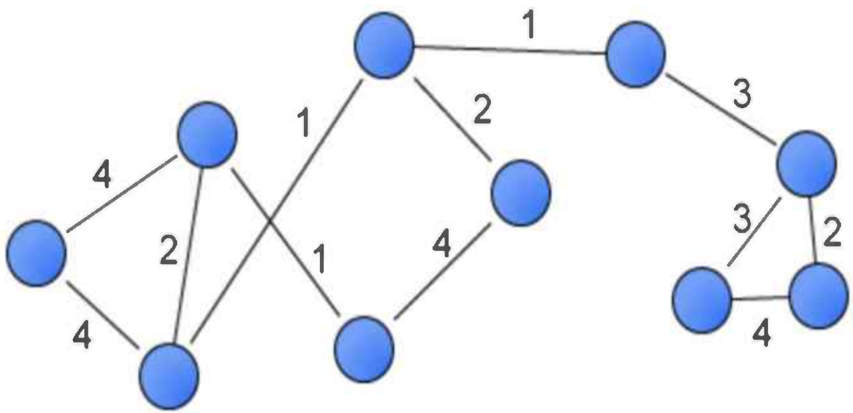
Un des avantages de la notion de graphe est qu'elle admet des variantes utiles:

- arêtes orientées pour indiquer le sens d'une interaction, la temporalité,...
- arêtes multiples s'il y a plusieurs interactions.



De gauche à droite:
graphe simple, graphe orienté simple, multi-graphe, multi-graphe orienté.

On peut aussi marquer des nombres sur les arêtes, qui peuvent représenter des distances (cartes routières), des temps de parcours, des intensités (circuits électriques), des probabilités de transition... On parle alors de **graphe pondéré**:

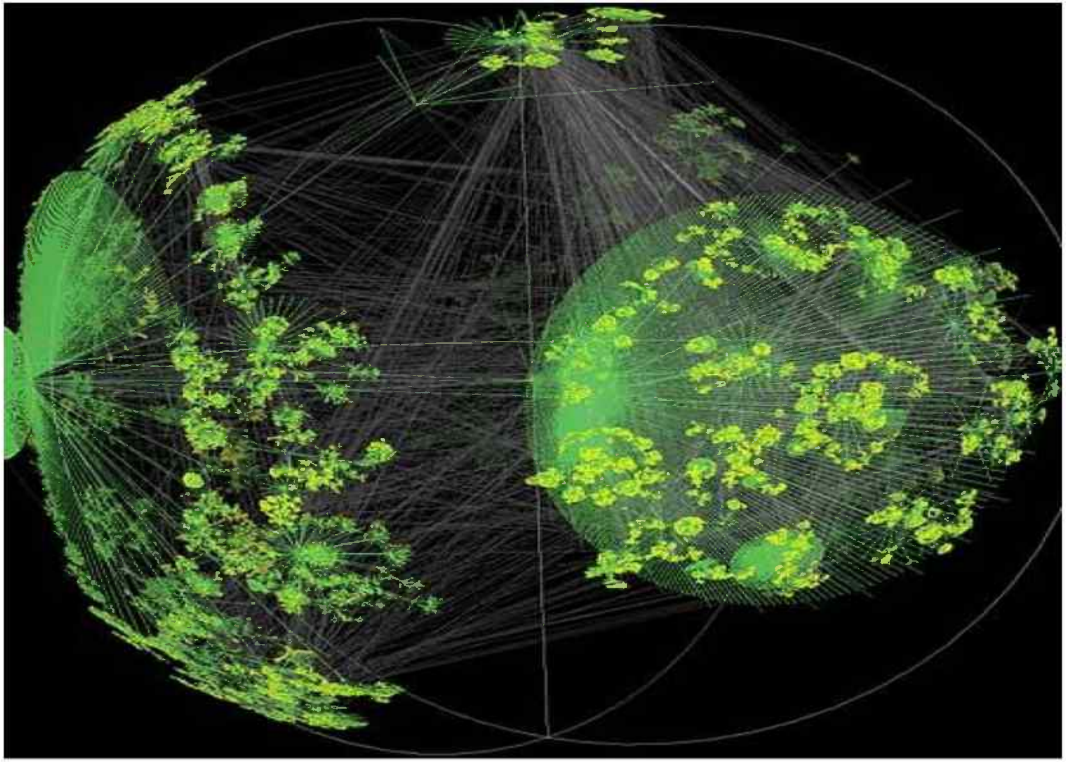


1.2 Applications

Les graphes ont de nombreuses applications dans tous les domaines liés à la notion de réseau (réseau social, réseau informatique, télécommunications, etc.) et dans bien d'autres domaines (par exemple génétique) tant le concept de graphe est général.

1.2.1 En sciences:

- Informatique (réseaux de communication, organisation des données, organisation de calculs complexes...). Un exemple de graphe qu'on aimerait comprendre est le graphe Internet: sommets: les sites Web (estimation: 10^{12} noeuds), arêtes: liens hyper-textes (= graphe orienté).



- Physique (p.ex. réseaux électriques)
- Chimie (modélisation des liaisons chimiques)
- Biologie (modélisation des routes de migration, des canaux de transferts de substances données...)

1.2.2 En sciences humaines

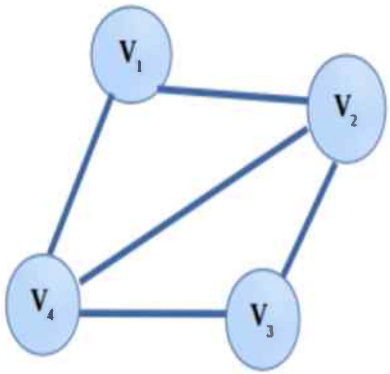
- Linguistique: La syntaxe se prête à une analyse par arborescence, modélisée par un graphe hiérarchique.
- Sociologie: Propagation des rumeurs; analyse et développement des réseaux sociaux comme Facebook (estimation: 10^9 utilisateurs)

Application: tester la théorie des *6 degrés de séparation.*

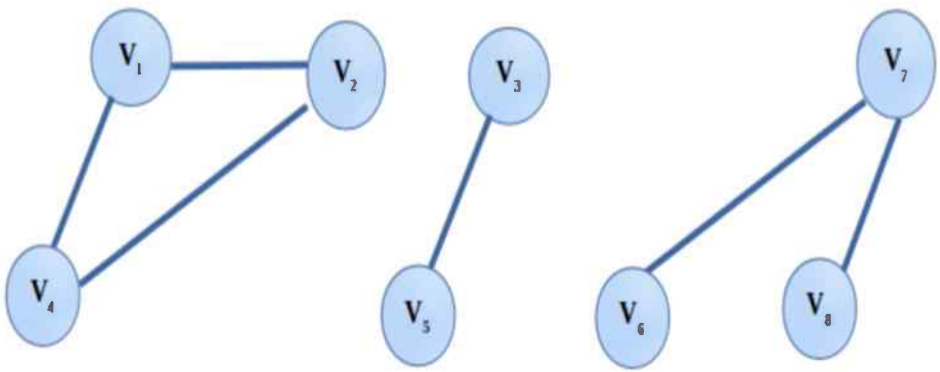
Frigyes KARINTHY, 1929: *"Un jeu fascinant naquit de cette discussion. L'un de nous suggéra de préparer l'expérience suivante afin de prouver que la population de la planète est plus proche maintenant qu'elle ne l'a jamais été dans le passé. Nous devrions sélectionner n'importe quelle personne du 1,5 milliard d'habitants de la planète, n'importe qui, n'importe où. Il paraît que, n'utilisant pas plus de cinq individus, l'un d'entre eux étant une connaissance personnelle, il pourrait contacter les individus choisis en n'utilisant rien d'autre que le réseau des connaissances personnelles."*

1.3 Chemins et distances dans un graphe non orienté

- Deux sommets sont **adjacents** s'ils sont reliés par au moins une arête.
- Un **chemin d'arêtes** entre deux sommets a, b est une suite finie de sommets commençant en a et finissant en b, où chaque sommet est adjacent au suivant.
- La **longueur** d'un chemin d'arêtes est le nombre d'arêtes traversées par le chemin.
- La **distance** entre deux sommets d'un graphe est la longueur d'un chemin d'arêtes minimal joignant ces deux sommets.
- Un graphe est **connexe** si deux sommets quelconques sont reliés par un chemin d'arêtes.



Fig(i):
Connected
Graph



Fig(ii):
Unconnected Graph

Exemples:

- Le **graphe des collaborations**: les sommets sont les mathématicien(ne)s, on trace une arête entre deux personnes si elles ont au moins un article en commun. On obtient la **distance entre collaborateurs**:

AMERICAN MATHEMATICAL SOCIETY
MATHSCINET
MATHEMATICAL REVIEWS

Recherche dans MSC | Distance entre collaborateurs | Revues actuelles | Publications actuelles

MR Collaboration Distance = 3

Cédric Villani	coauthored with	Joel Louis Lebowitz	MR3247073
Joel Louis Lebowitz	coauthored with	Peter Gabriel Bergmann	MR0074332
Peter Gabriel Bergmann	coauthored with	Albert Einstein	MR1503432

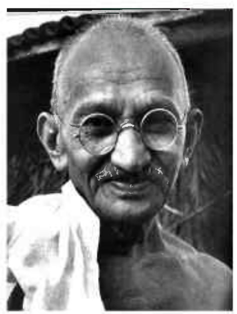
Change First Author | Change Second Author | New Search

MR Collaboration Distance = 5

Alain Valette	coauthored with	Assaf Naor
Assaf Naor	coauthored with	Rina Panigrahy
Rina Panigrahy	coauthored with	Kunal Talwar
Kunal Talwar	coauthored with	Christos H. Papadimitriou
Christos H. Papadimitriou	coauthored with	William H. Gates

- Le **graphe des poignées de mains**: les sommets sont les êtres humains, on trace une arête entre deux personnes chaque fois qu'elles ont échangé au moins une poignée de mains. On obtient la *distance des poignées de main*

Distance 2:   

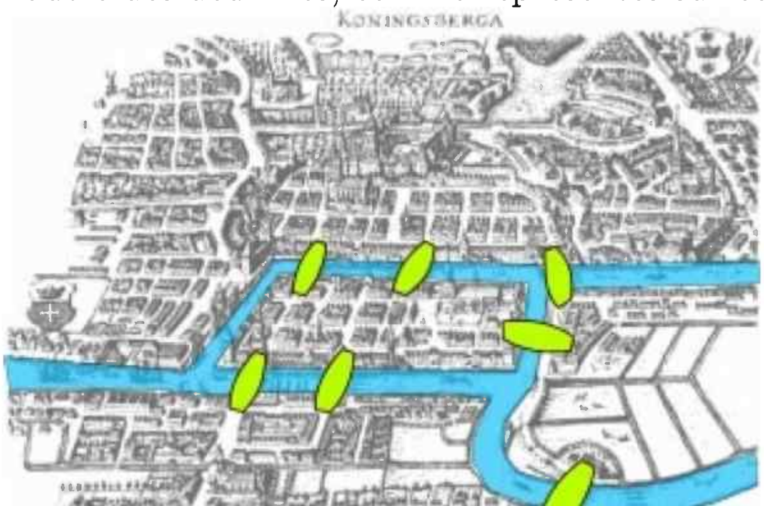
Distance 3:    

Les 6 degrés de séparation prédisent que, dans un réseau social, la distance entre deux personnes est ≤ 6. En 2015: la distance moyenne entre deux utilisateurs de Facebook est 4,74.

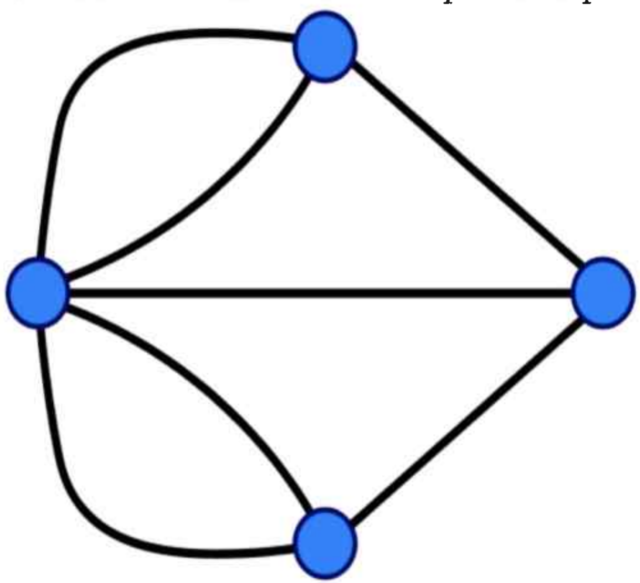
1.4 Des graphes depuis quand?

1.4.1 Euler

Le problème des sept ponts de Königsberg est connu pour être à l'origine de la théorie des graphes. Résolu par Leonhard Euler en 1736, ce problème se présente comme suit : la ville de Königsberg (aujourd'hui Kaliningrad) est construite autour de deux îles situées sur la Pregel et reliées entre elles par un pont. Six autres ponts relient les rives de la rivière à l'une ou l'autre des deux îles, comme représentés sur ce plan:



Le problème consiste à déterminer s'il existe ou non une promenade dans les rues de Königsberg permettant, à partir d'un point de départ au choix, de passer une et une seule fois par chaque pont, et de revenir à son point de départ, étant entendu qu'on ne peut traverser la Pregel qu'en passant sur les ponts.
Euler a l'idée de modéliser la question par un graphe.



Le problème devient: *peut-on tracer ce graphe sans lever le crayon et sans repasser aucune des arêtes, en revenant au point de départ?*
Appelons **degré** d'un sommet le nombre de voisins de ce sommet. Euler observe que, si on peut tracer un graphe d'un trait, sans repasser sur aucune arête, et en revenant au point de départ, alors tout sommet est de degré pair: en effet, chaque fois qu'on arrive à un sommet, on doit aussi en repartir. Or le graphe ci-dessus a tous ses sommets de degré impair, on a démontré que le problème des ponts de Königsberg est impossible.
Les choses ne vont pas mieux si on autorise que le point d'arrivée n'est pas le point de départ. Dans ce cas, le degré du sommet de départ est impair (on le quitte mais on n'y revient pas), le degré du sommet d'arrivée est impair (on y vient mais on n'en repart pas), et le degré de tout autre sommet est pair (on ne fait que passer). Donc: dans ce cas il y a exactement *deux* sommets de degré impair, ce qui n'est pas le cas dans le graphe ci-dessus.

Appelons **chemin eulérien** dans un graphe un chemin qui passe une et une seule fois par chaque arête.

Theorème 1.1. *(Euler)*

1. *Un graphe admet un chemin eulérien fermé (c-à-d. qui revient à son point de départ) si et seulement si tout sommet du graphe est de degré pair.*

2. *Un graphe admet un chemin eulérien non fermé si et seulement si il y a exactement deux sommets de degré impair.*

Remarque 1.2. *Dans un graphe, il y a toujours un nombre pair de sommets de degré impair. En effet, en sommant tous les degrés, on obtient le double du nombre d'arêtes (chaque arête étant comptée deux fois). Donc, dans la somme il y a un nombre pair de termes impairs, pour que la somme soit un nombre pair.*
Application: dans une soirée, les convives serrent la main à certains autres. Il y a toujours un nombre pair de convives qui a serré un nombre impair de mains!

Après Euler, le premier progrès notable est dû à G. Kirchhoff, qui les utilise en 1845 pour les lois des courants électriques.

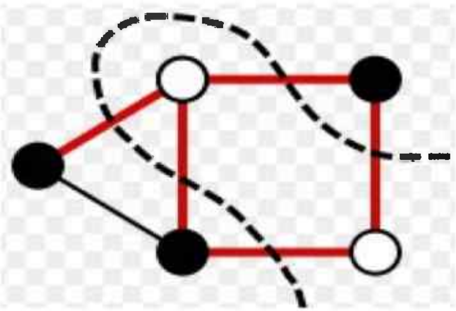
Le terme *"graphe"* est introduit en 1878 par J.J. Sylvester dans un article dans *Nature*, où il fait le parallèle avec les diagrammes moléculaires:

"[...] Every invariant and co-variant thus becomes expressible by a graph precisely identical with a Kekuléan diagram or chemicograph. [...] I give a rule for the geometrical multiplication of graphs, i.e. for constructing a graph to the product of in- or co-variants whose separate graphs are given. [...]"

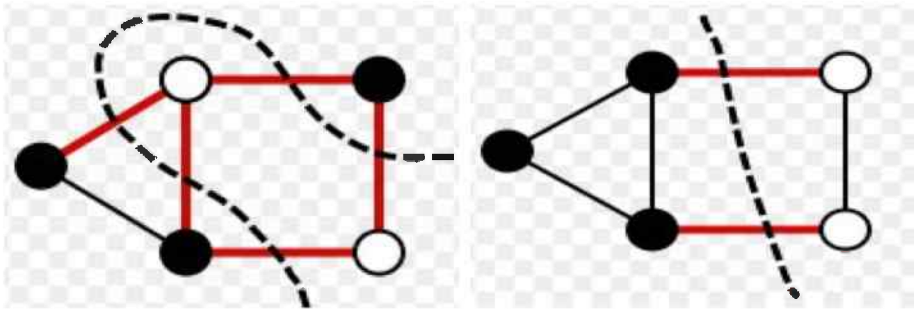
2 Algorithmes de graphe

2.1 Coupures

Une *coupure* dans un graphe X est une partition de l'ensemble des sommets en 2 parties S, S^c (où S^c est le complémentaire de S):



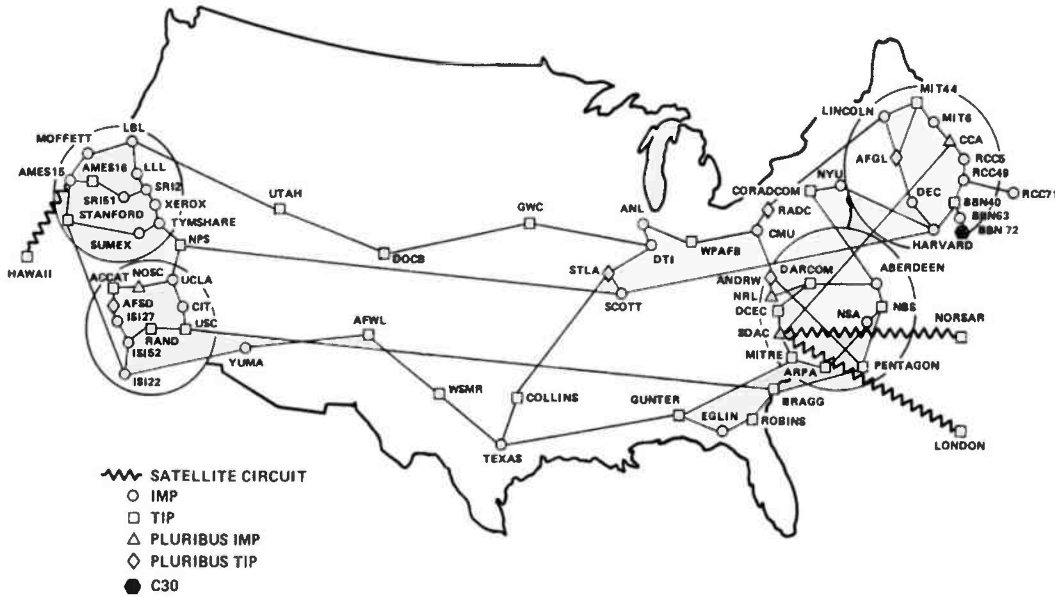
On note $E(S, S^c)$ l'ensemble des arêtes reliant S à S^c; en enlevant les arêtes de $E(S, S^c)$, on déconnecte le graphe en au moins 2 morceaux.



La 1ère coupure n'est pas minimale; la 2nde est minimale.

ARPANET a été le premier réseau d'ordinateurs à transfert de paquets sur le territoire des USA, développé par DARPA (Defense Advanced Research Projects Agency) dans les années 1960. Selon la légende, Arpanet devait permettre aux réseaux de communication militaires de continuer à fonctionner malgré une attaque nucléaire massive de la part de l'Union soviétique: si un des noeuds est détruit, les données empruntent d'autres chemins et d'autres noeuds pour atteindre les destinataires.

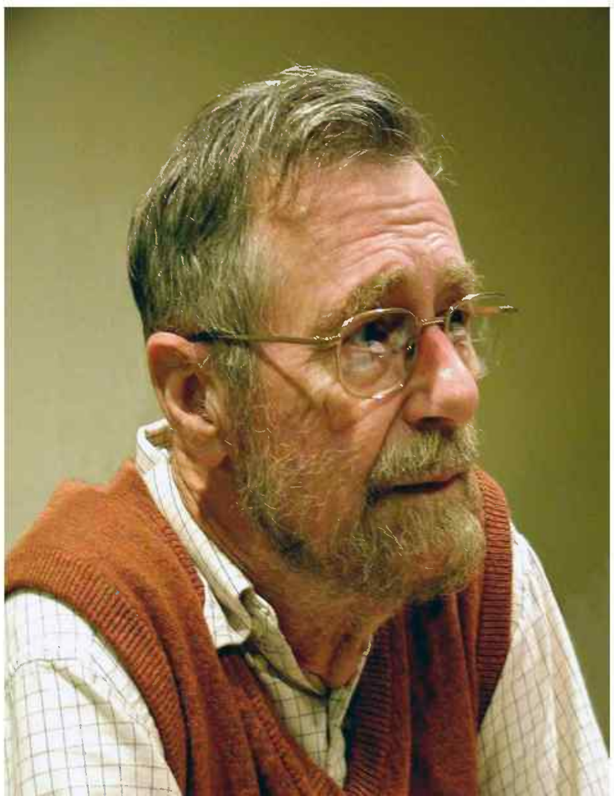
ARPANET GEOGRAPHIC MAP, OCTOBER 1980



(NOTE: THIS MAP DOES NOT SHOW ARPA'S EXPERIMENTAL SATELLITE CONNECTIONS)
NAMES SHOWN ARE IMP NAMES, NOT (NECESSARILY) HOST NAMES

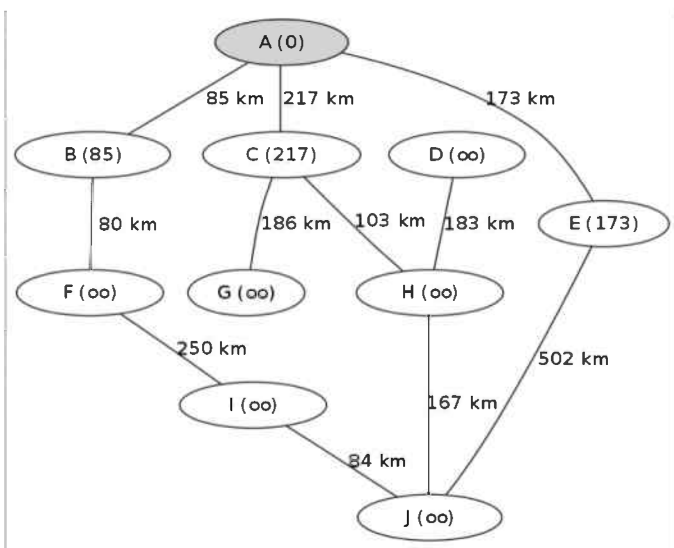
2.2 L'algorithme de Dijkstra (1959)

Edsger W. DIJKSTRA, né à Rotterdam le 11 mai 1930 et mort à Nuenen le 6 août 2002.



L'algorithme de Dijkstra sert à résoudre le problème du plus court chemin. Il permet, par exemple, de déterminer un plus court chemin pour se rendre d'une ville à une autre connaissant le réseau routier d'une région. Plus précisément, il calcule des plus courts chemins à partir d'une source dans un graphe pondéré par des nombres positifs.
Il contourne astucieusement *l'explosion combinatoire*: en effet, si chaque sommet a en moyenne d voisins, au départ de la source on a d chemins de longueur 1, d^2 chemins de longueur 2, d^3 chemins de longueur 3, etc... (croissance exponentielle du nombre de chemins).
L'algorithme prend en entrée un graphe pondéré par des nombres positifs (distances, ou temps de parcours) et un sommet-source. Il s'agit de construire progressivement un sous-graphe dans lequel les différents sommets sont classés par ordre croissant de leur distance minimale au sommet-source. La distance correspond à la somme des poids des arêtes empruntées.

L'exemple ci-dessous montre les étapes successives pour trouver le chemin le plus court dans un graphe. Les noeuds symbolisent des villes identifiées par une lettre et les arêtes indiquent la distance entre ces villes. On cherche à déterminer le plus court trajet pour aller de la ville A à la ville J.



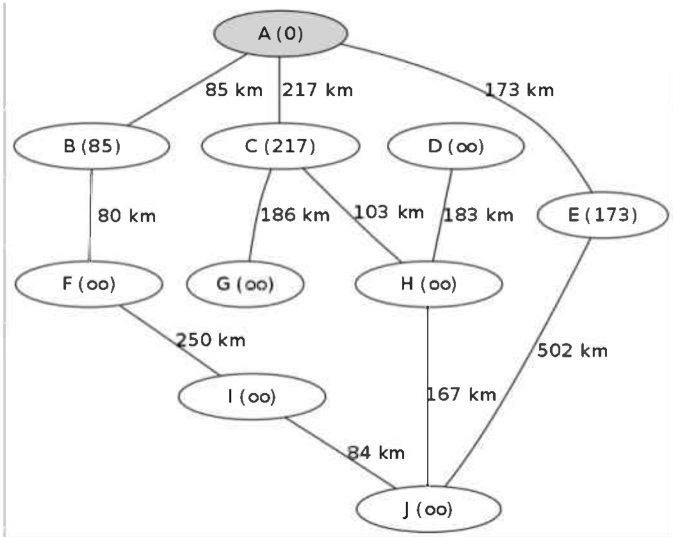
Pour un sommet X du graphe, notons $d(X)$ la distance minimale de A à X, c-à-d. la plus petite somme des poids sur un chemin d'arêtes de A à X. Au départ, on ne connaît pas la valeur définitive de $d(X)$, on se contentera d'une valeur provisoire $\delta(X) \geq d(X)$, donnée par le meilleur chemin actuellement connu. Par approximations successives, on fait évoluer $\delta(X)$ vers la valeur définitive $d(X)$.

Notons Ω l'ensemble des sommets du graphe pour lesquels on sait que $d(X)$ a atteint sa valeur définitive. Chaque itération (=étape) de l'algorithme consiste en deux pas:

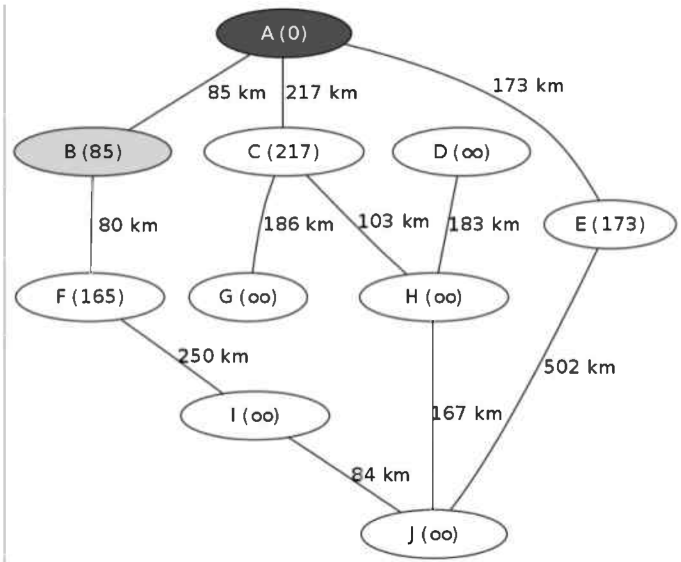
- Déterminer la coupure définie par Ω, c-à-d. les arêtes avec une extrémité dans Ω et l'autre en dehors de Ω; pour les sommets X ainsi déterminés hors de Ω, mettre à jour la fonction δ en posant: $\delta(X)$ = distance de X à A en n'utilisant que les arêtes de Ω et celles dans la coupure.

- Choisir un de ces sommets Y hors de Ω pour lequel $\delta(Y)$ est minimal. Pour ce Y, on a $\delta(Y) = d(Y)$ et on remplace Ω par $\Omega \cup \{Y\}$.

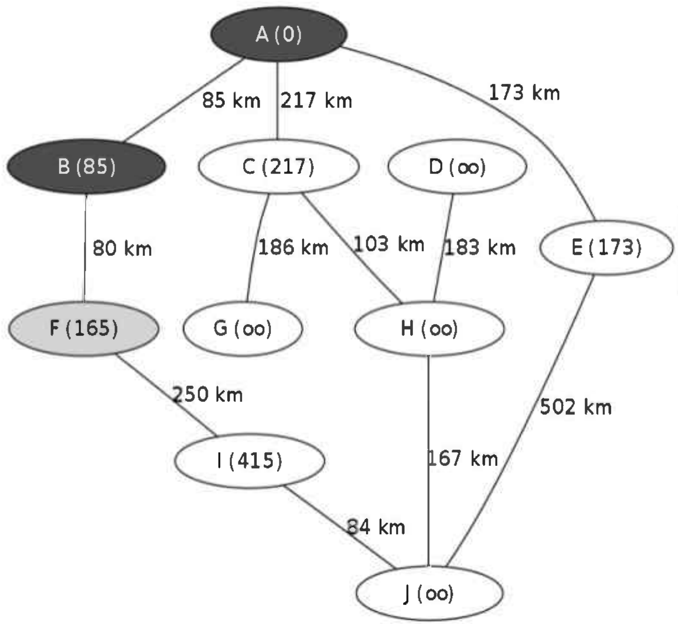
On initialise avec $\Omega = \{A\}, \delta(A) = d(A) = 0$ et $\delta(X) = \infty$ pour $X \neq A$.
Dans notre exemple, en neuf étapes on trouve le chemin minimal de A à J: il passe par C et H et mesure 487 km.



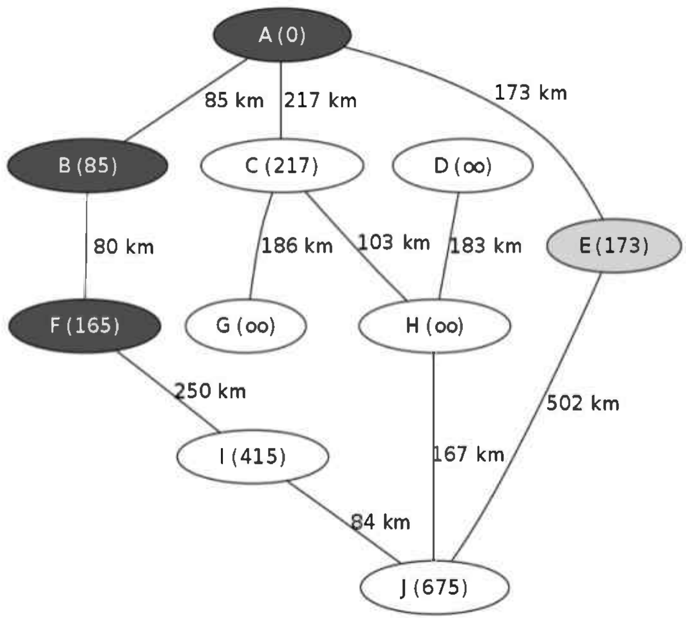
Étape 1 : on choisit la ville A. On met à jour les villes voisines de A qui sont B, C, et E. Leurs distances deviennent respectivement 85, 217, 173, tandis que les autres villes restent à une distance infinie.



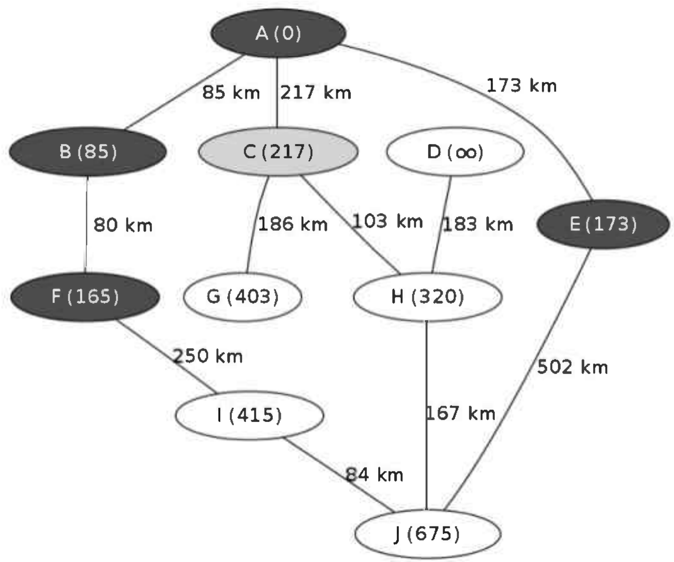
Étape 2 : on choisit la ville B. En effet, c'est la ville hors du sous-graphe qui est à la distance minimale (85). On met à jour le seul voisin (F). Sa distance devient 85+80 = 165.



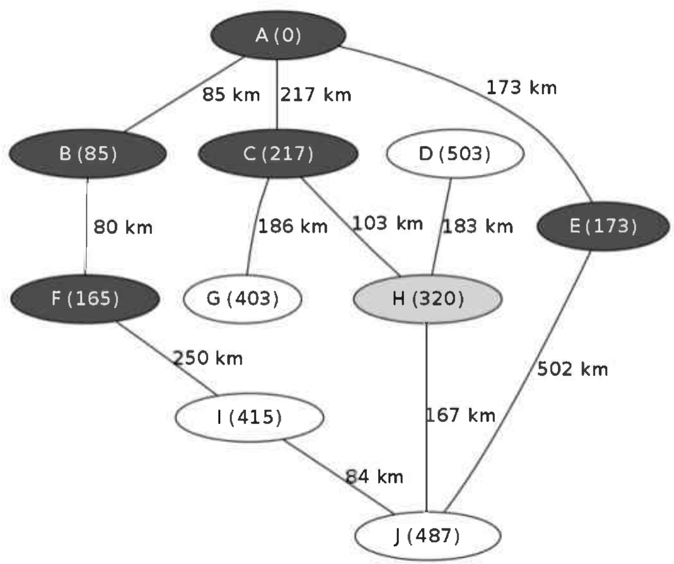
Étape 3 : on choisit F. On met à jour le voisin I (415).



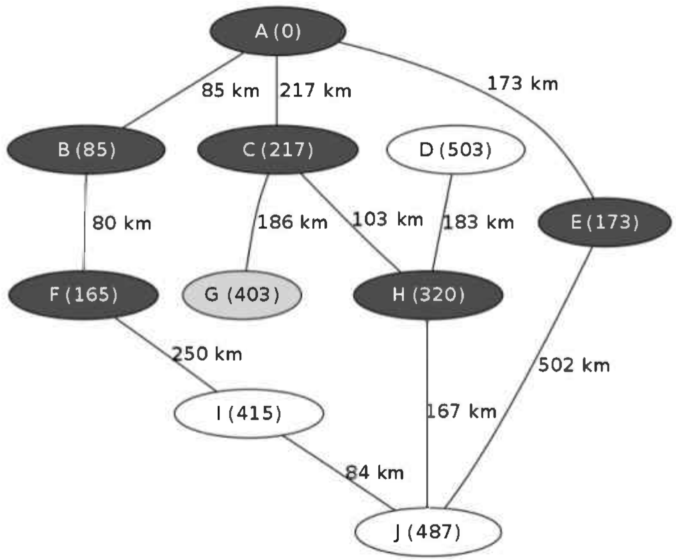
Étape 4 : on choisit E. On met à jour le voisin J (675).



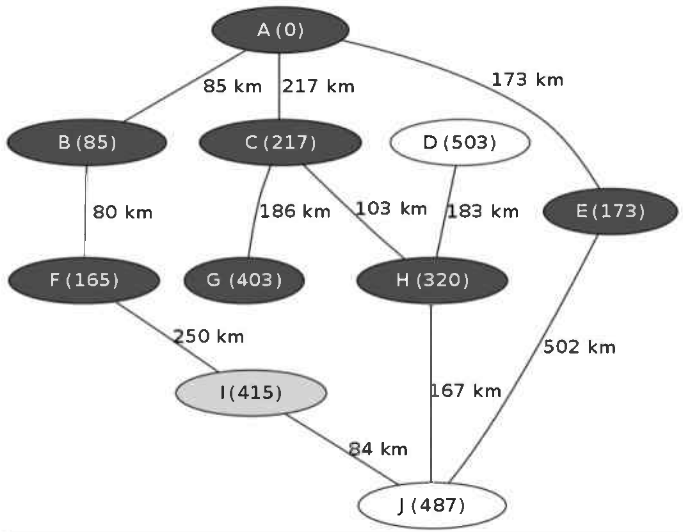
Étape 5 : la distance la plus courte en dehors du sous-graphe est maintenant celle de la ville C. On choisit donc C. On met à jour la ville G (403) et la ville H (320).



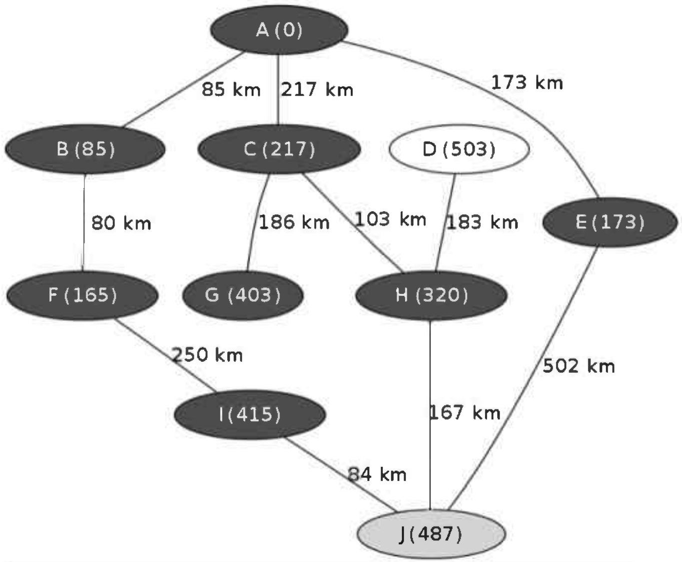
Étape 6 : la distance la plus courte en dehors du sous-graphe est maintenant celle de la ville H (320). On choisit donc H. On met à jour la ville D (503) et la ville J (487< 675).



Étape 7 : la distance la plus courte suivante est celle de la ville G. On choisit G. La mise à jour ne change aucune autre distance.



Étape 8 : la distance la plus courte suivante est celle de la ville I. La distance de la ville voisine J n'est pas modifiée car la distance existante est inférieure à celle que l'on obtiendrait en passant par I (415 + 84 > 487).



Étape 9 : la ville dont la distance est la plus courte est J (487). On choisit J et on l'ajoute au sous-graphe. On s'arrête puisque la ville d'arrivée est maintenant dans le sous-graphe.

Rappelons un ordre de grandeur en pratique:

L'algorithme de Dijkstra est de complexité polynomiale. Plus précisément, pour n noeuds et a arcs, le nombre d'itérations est de l'ordre de $(n + a)\log n$.

Le logarithme:

"Rappel": $\log n$ est le nombre de chiffres de n. C'est la place que tient n dans une mémoire; ce qui est intéressant est que $\log n$ est BEAUCOUP plus petit que n.

Conclusion: L'algorithme de Dijkstra est un de ceux qui font que le monde moderne fonctionne!

Merci pour votre attention!