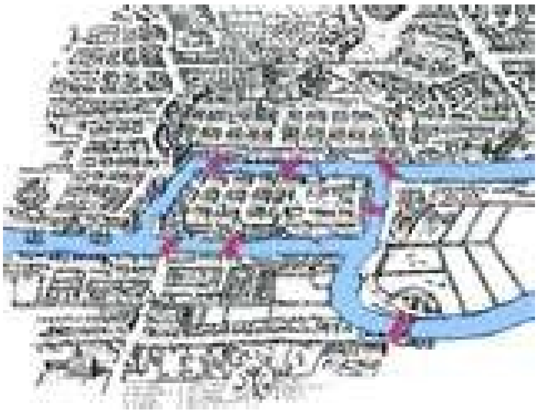
Introduction aux graphes

Philippe Lucaud , Pr. de Mathématiques, Lycée Audiberti Antibes, exercices pour classes de 2nd.

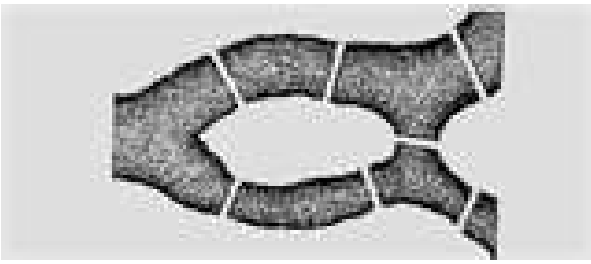
Un peu d'histoire...

En 1735 Karl Ehler propose un problème au mathématicien suisse Leonhard **Euler**. A Königsberg ville de Russie maintenant nommée Kaliningrad, la rivière Pregel possède deux bras à l'Est qui fusionnent en un seul à l'Ouest. On peut alors distinguer deux îles. Sept ponts permettent de traverser la rivière à différents endroits.



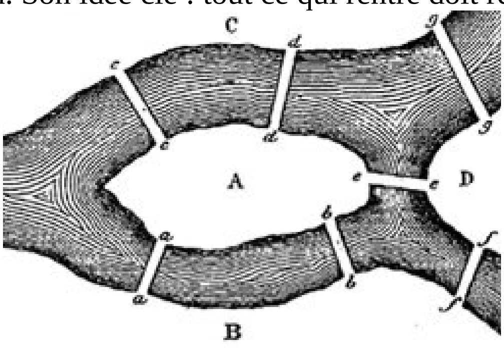
Les habitants de la ville sont partagés. Certains pensent qu'il est possible de faire une promenade en passant une et une seule fois sur chaque pont et en revenant au point de départ. D'autres pensent que cela est impossible.

Euler se saisit du problème qu'il formalise en ne gardant du tracé que ce qui est utile à la réflexion. Il efface le superflu.



Il appelle, d'après les travaux de **Leibnitz**, « *Géométrie de Position*» cette « science qui s'occupe uniquement de la position et des propriétés qui découlent de cette position, indépendamment des grandeurs ».

Il nomme les îles et les berges à l'aide de quatre lettres et il tente de généraliser ce qu'il a découvert de manière empirique, non sans s'être auparavant posé la question de l'utilité de soumettre ce genre de problème à un mathématicien. Son idée clé : tout ce qui rentre doit ressortir...



Pour simplifier le problème il supprime les ponts parallèles et cherche des chemins du type ABD... en prenant soin de ne pas répéter des séries de lettres lorsque le pont est unique (par exemple on ne peut pas trouver AD puis plus loin DA...).

Bien qu'on lui attribue la preuve du théorème découvert et la paternité de la notion de graphe, Euler n'a jamais traduit son problème sous la forme d'un graphe. La preuve de son théorème a été fournie indépendamment de ses travaux en 1873 par Carl **Hierholzer**.

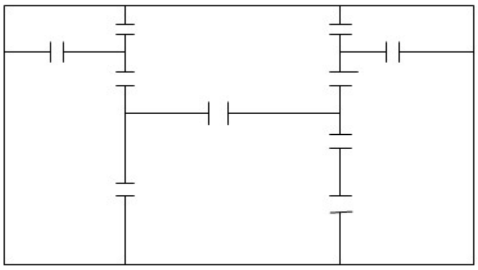
<u>Jouons un peu à Euler...</u>

Voici les plans de quatre musées (tirés de la thèse de Léa **Cartier** : *Le graphe comme outil pour enseigner la preuve et la modélisation*). On peut facilement trouver pour chacun d'entre eux un chemin qui ne passe qu'une seule fois par chacune des salles. Mais tout en visitant ce musée intégralement :

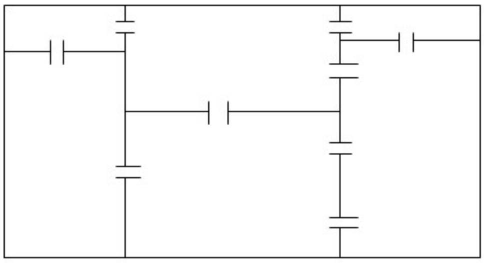
- **Existe t'il un chemin qui ne passe qu'une seule fois par chacune des portes ?**
- **Peut-on trouver un circuit qui ne passe qu'une seule fois par chacune des portes ?**

(Un circuit est un chemin qui revient à son point de départ)

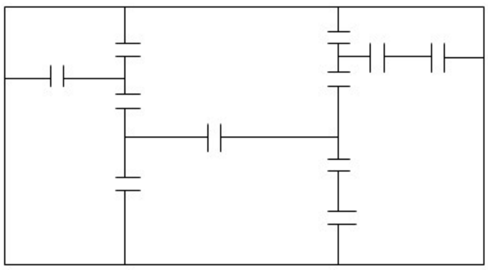
IZIS



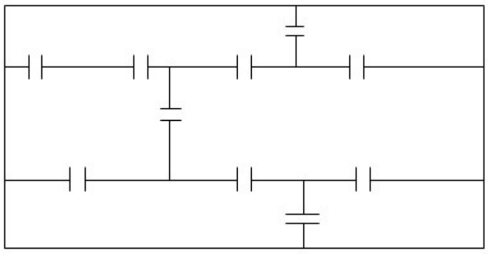
OZ



AZA



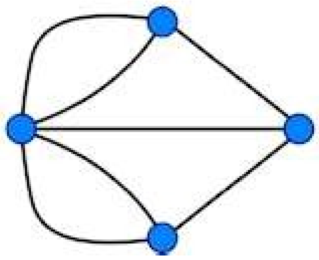
EZE



Comment modéliser cette situation de recherche d'un chemin ?

En 1892 W. W. **Rouse Ball** proposera le *graphe* suivant dans une récréation mathématique :

B



Les points représentent les quatre îles,
et les les traits qui les relient les ponts.

A D

C

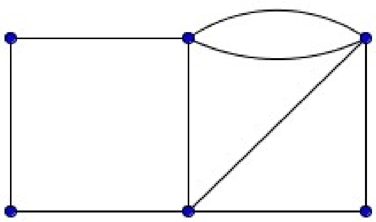
Quelques définitions :

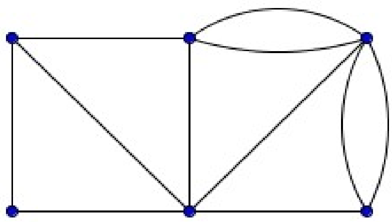
Un **graphe** est un ensemble de points et de lignes reliant certains de ces points.

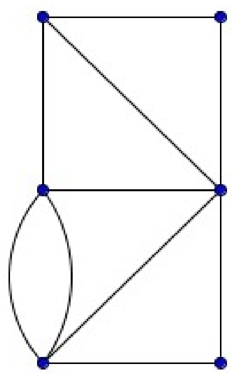
Les points sont appelés les **sommets** du graphe et les lignes sont appelées les **arêtes**.
L'**ordre** du graphe est le nombre de ses sommets.
Le **degré** d'un sommet est le nombre d'arêtes ayant ce sommet comme extrémité.
Un graphe est **connexe** lorsque, pour chaque paire de sommets, il existe au moins une chaîne les reliant.

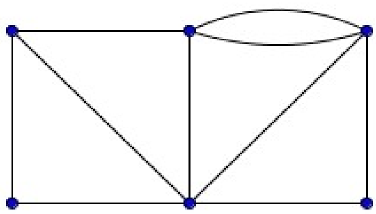
Application à notre problème de chemin dans un musée :

Voici donc les quatre graphes correspondants aux quatre musées proposés. Retrouver et écrire le nom de chacun d'entre eux :









Observer ces graphes et donner pour chacun son ordre et le degré de ses sommets.

Chaîne eulérienne : c'est une chaîne composée de toutes les arêtes du graphe, prises une seule fois.
Cycle eulérien : c'est une chaîne eulérienne dont les sommets coïncident.

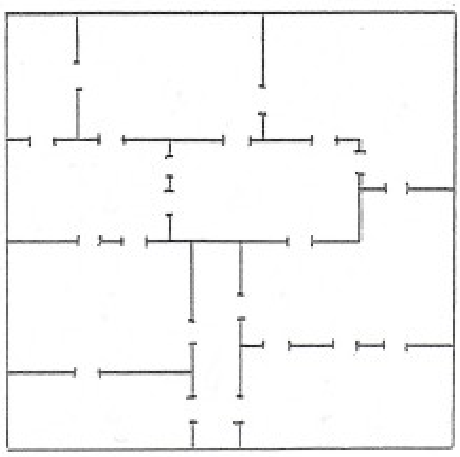
Théorèmes d'existence :

- *Un graphe connexe admet une chaîne eulérienne entre deux sommets si ce sont les deux seuls sommets de degré impair.*
- *Un graphe connexe admet un cycle eulérien si tous ses sommets sont de degré pair.*

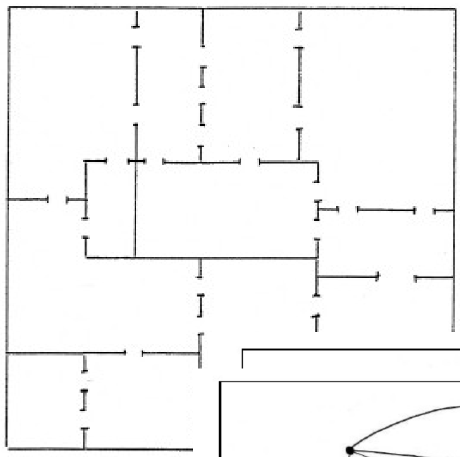
Appliquons ces théorèmes d'existence pour déterminer les graphes de notre exemple qui possèdent :

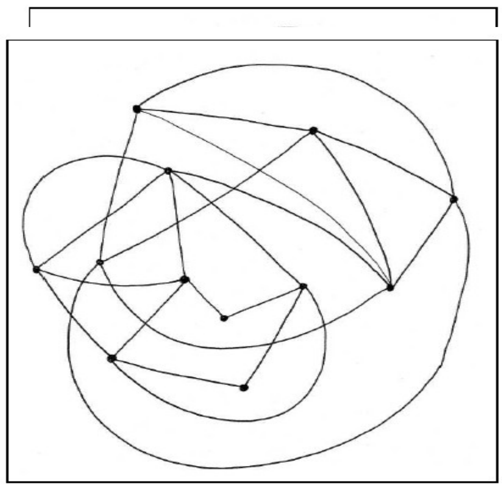
Une Chaîne eulérienne ....
Un Cycle eulérien ....
Ni l'un ni l'autre ....

Avoir modélisé ces musées par des graphes nous a permis de faire une recherche simplifiée de chemins particuliers. Mais, tout aussi important, cette modélisation peut être transposée à d'autres situations (Le voyageur de commerce qui doit rejoindre une ville en passant par toutes les routes pour visiter tout ses clients...)**.**



<u>Pour voir si tout est bien compris...</u> (On se pose les mêmes questions qu'auparavant, mais attention au piège...)

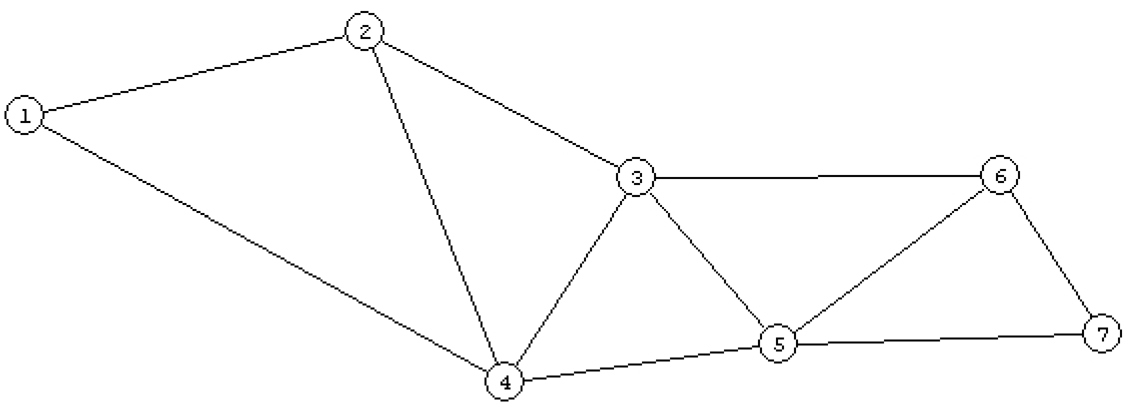




<u>Le plus court chemin...</u>

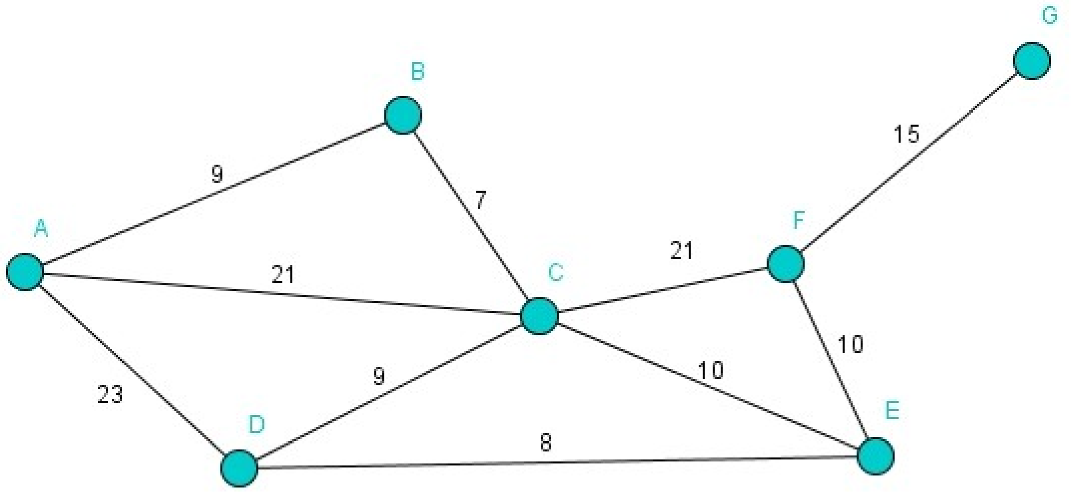
La **distance** entre deux sommets d'un graphe connexe est la longueur de la chaîne la plus courte (ayant le moins d'arêtes) qui relie ces deux sommets .

Exemple : On donne le graphe suivant,



Quelle est la distance entre les sommets 1 et 3 ? ....., 1 et 4 ? ....., 1 et 7 ? ....., 2 et 7 ? .....

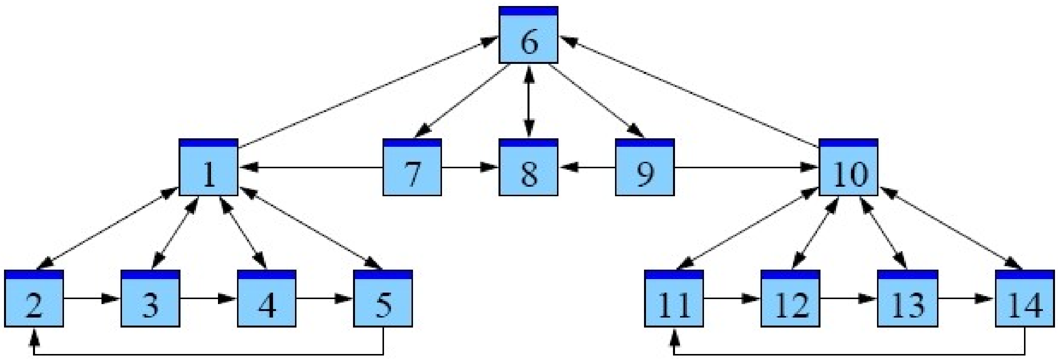
Un **graphe pondéré** est un graphe dont les arêtes sont affectées d'un nombre positif appelé poids de l'arête.



Exemple : Un voyageur de commerce doit partir de la ville A et se rendre à la ville G. Le graphe ci-dessous représente les différentes villes traversées et les routes qui les relient. Les poids des arêtes sont les distances en kilomètre. Qu'elle est le plus court chemin entre A et G ?

<u>Une application originale : comment Google classe-t'il les pages web ?</u>

Un **moteur de recherche** (Google, Bing, Wolfram alpha...) trie par ordre d'importance les résultats d'une requête qui peut être un mot, une phrase... Face à la multitude de textes accessibles sur internet, face aux modifications et aux ajouts continuels, il faut pouvoir aider les utilisateurs à repérer les informations susceptibles de les intéresser. Clarifions le propos. Que peut bien vouloir dire « par ordre d'importance » ? Il est très difficile de donner une définition exacte de ce qu'est l'importance d'une page web. Tout au plus parlerons nous de *l'importance ressentie par les utilisateurs*.



Le web peut être considéré comme un immense graphe dont les sommets sont les pages web et les arêtes (orientées) sont les liens reliant les pages entre elles. Classer efficacement et de manière instantanée les milliards de pages web est mission impossible pour un être humain. Il faut donc automatiser la procédure. On pourra penser avec bon sens qu'une page importante reçoit beaucoup de liens. Réciproquement, si elle est beaucoup citée, c'est qu'elle est importante... Nous allons voir que ce n'est pas si simple...

Intéressons nous à la structure de ce graphe et non au contenu des pages web (qui est bien entendu très important !). Les pages 1 et 10 semblent assez centrales et reçoivent beaucoup de liens (on peut imaginer par exemple une page d'accueil et des annexes...). Dans une moindre mesure, les pages 8 et 6 aussi, mais surtout, la page 6 est pointée par les pages 1 et 10 ce qui signifie sûrement qu'elle contient de l'information essentielle pour tout l'ensemble. Essayons d'établir une mesure de cette importance :

- **Le comptage des liens (comptage naïf) :** On fait la somme des liens reçus par la page considérée. L'importance μ_i de la page P_i est donnée par : $\boldsymbol{\mu_i = \sum 1}$
  ....

- **Le comptage pondéré :** Certaines pages émettent beaucoup de liens. Du coup ceux-ci sont moins spécifiques, et d'une certaine manière, leur poids est plus faible. Si une page P_j émet l_j liens, le poids de chaque lien sera $1/l_j$ et donc l'importance de la page P_i sera donnée par : $\boldsymbol{\mu_i = \sum 1/\mu_j}$
  ....

  (Ces méthodes ne sont pas fiables car on peut artificiellement augmenter l'importance d'une page web, en créant des pages vides qui pointent vers cette page)

- **Le comptage récursif :** On utilise le principe suivant lequel une page paraît importante si beaucoup de pages importantes la citent. Le poids du vote de j vers i est proportionnel au poids de la page émettrice. On obtient alors la formule suivante : $\boldsymbol{\mu_i = \sum \mu_j/l_j}$ qui nous amène à ce que nous voulions, la page P_6 est enfin repérée comme étant la plus importante. Cependant, la mise en œuvre de cette méthode n'est pas au programme de seconde.

  Pour en savoir plus : Michael Eisermann sur Interstices ou sur le site de l'institut Fourier.