# Grands Réseaux d'Interaction
# TP 2 Calculs de distance et de cardialité

Fabien de Montgolfier
`fm@irif.fr`

4 février 2022

**à rendre pour le 11 fevrier 23h59**

**Les règles générales restent valables, cf feuille TP1**

## 1 Définitions

On travaille toujours sur des graphes **orientés** : la structure de donnée reste la même. Pour information, les TPs suivants seront tous sur des graphes non-orientés.

On rappelle que le graphe $G$ possède $n$ sommets, dont les numéros vont de 0 jusqu'à $n-1$ (qui est le numéro max d'un sommet dans le fichier de Stanford).

Pour la **cardialité**, voir le cours n° 2

Les sommets **accessibles** depuis $x$ sont les sommets $y$ tels qu'il y a un chemin de $x$ vers $y$. Attention, cela n'implique pas que $x$ est accessible depuis $y$ !

La **distance** de $x$ à $y$, notée $dist(x, y)$, est la longueur d'un plus court chemin de $x$ à $y$. Si $y$ n'est pas accessible depuis $x$, $dist(x, y)$ n'est pas définie (**pas de distance** $\infty$)

L'**excentricité** de $x$ notée $ex(x)$ est le maximum, pour tout $y$ accessible depuis $x$, de $dist(x, y)$. **x est accessible depuis x** donc $ex(x) \geq 0$ est toujours définie.

Le **diamètre** du graphe $G$ notée $diam(G)$ est l'excentricité maximum d'un sommet de $G$

L'**excentricité moyenne** du graphe $G$ notée $\overline{ex}(G)$ est est $\frac{1}{n}\sum_{x\in G} ex(x)$

La **proximité** de $x$ notée $prox(x)$ est la moyenne, pour tout $y$ accessible depuis $x$, de $dist(x, y)$. $x$ est à distance 0 de lui-même donc un sommet sans voisin a proximité 0

La **proximité moyenne** du graphe est $\overline{prox}(G) = \frac{1}{n}\sum_{x\in G} prox(x)$

Soit $Ac$ l'ensemble des $k$ couples $(x, y)$ de sommets de $G$ tels que $y$ est accessible depuis $x$. $k = |Ac|$ varie de $n$ pour un graphe sans arc (car $Ac$ contient toujours tous les couples $(x, x)$) à $n^2$ pour un graphe fortement connexe (où $Ac$ contient tous les couples possibles). La **distance moyenne** du graphe notée $\overline{dist}(G)$ est

$$\overline{dist}(G) = \frac{\sum_{(x,y)\in Ac} dist(x, y)}{k}$$

Notez que si le graphe est fortement connexe, alors $\overline{prox}(G) = \overline{dist}(G)$

## 2 Travail demandé

Il est demandé de faire un programme qui, étant donnés 3 ou 4 paramètres, dans l'ordre :

1. un nom de fichier de graphe (fichier texte au format de Stanford)
2. son nombre de ligne (cela permet de gagner du temps de lecture) On n'a pas à le vérifier : le comportement du programme n'est pas défini si ce n'est pas le nombre de lignes tel que donné par `wc -l` du premier paramètre.
3. une commande qui peut être la chaîne `1seul` ou `exact` ou `approx`
4. Enfin le quatrième paramètre est :
   — si la commande est `1seul`, alors un numéro $v$ de sommet
   — si la commande est `exact`, alors rien (il n'y a pas de quatrième paramètre)
   — si la commande est `approx`, alors un nombre $s$ de sommets ($0 < s \leq n$)

charge le graphe en mémoire, puis affiche sur la sortie standard

— Soit un message d'erreur en une ligne contenant le mot "ERREUR" en majuscule, plus d'autres informations utiles

— Soit, si la commande est `coeur`, la cardialité du graphe (le nombre $k$ tel qu'il a un $k$-cœur non vide et un $(k+1)$-cœur vide). Affichez un seul nombre, rien d'autre svp !

— Soit, si la commande est `1seul`, les 4 nombres suivants, un par ligne (donc 4 lignes) :
  1. la cardialité de $v$ (le nombre $k$ tel que $v$ est dans le $k$-cœur mais pas dans le $(k+1)$-cœur)
  2. le nombre de sommets accessibles depuis $v$
  3. l'excentricité de $v$
  4. la proximité de $v$

— Soit, si la commande est `exact`, les 4 nombres suivants, un par ligne (donc 4 lignes) :
  1. le diamètre de $G$
  2. l'excentricité moyenne de $G$
  3. la proximité moyenne de $G$
  4. la distance moyenne de $G$

— Enfin, si la commande est `approx`, on doit afficher les 4 mêmes valeurs que pour exact, mais en faisant une **approximation** afin de gagner en vitesse. Pour cela, on tire $s$ sommets **différents** au hasard (ce n'est pas la même chose que de tirer $s$ fois un sommet au hasard !) et on affiche
  1. pour le diamètre approché : la plus grande excentricité des $s$ sommets
  2. pour l'excentricité approchée : la moyenne des excentricités de ces $s$ sommets
  3. pour proximité approchée : la moyenne des proximités de ces $s$ sommets
  4. enfin pour la distance moyenne approchée, on utilise pour $Ac$ l'ensemble des couples $(x, y)$ où $x$ a été tiré au sort et $y$ est accessible depuis $x$

## 3 Exemples

Voyez le fichier `resultats.txt`. Le fichier `exemple.txt` correspond au graphe de l'exemple de cœur orienté dans le cours, et `exemple2.txt` rajoute deux arcs pour le rendre connexe.