Grands Réseaux d'Interaction
TP8 - Attachement préférentiel

Les trois paramètres d, n_0 et n ci-dessous font référence à ceux du modèle de Barabási-Albert dans le sujet de ce TP8.

I) le modèle Barabási–Albert

Graphe initial avec tous les sommets interconnectés

On utilise le modèle d'Erdős–Rényi avec comme argument n_0 et 1.

Ajouter les autres sommets n - n_0

Tout d'abord, on partitionne tous les sommets du graphe selon leur degré de telle sorte à obtenir directement un sommet à partir d'un nombre : Une ArrayList<Vertex> (en Java) semble le mieux adapté. En respect des contraintes du modèle, plus un sommet déjà existant a de degrés, plus il sera présent dans cette liste.

Pour chaque nouveau sommet (nommons a) à ajouter, on marque initialement tous les sommets déjà existants comme éligibles, et on initialise un entier eligibleDegrees = [taille de la partition], puis on répète un total de d fois :

- On choisit aléatoirement un nombre n entre 0 et eligibleDegrees.
- On obtient à partir de n et de la partition le sommet x déjà existant correspondant. Puis
  - Si x est inéligible, on réessaie avec un autre entier tiré au hasard.
  - Si x est éligible, on le marque inéligible, on décrémente de eligibleDegrees le degré de x (impérativement avant la liaison !), et on ajoute x à une liste temporaire helper.

Finalement, on lie tous les sommets de helper à a, et on met à jour la partition : on ajoute deux nouvelles entrées pour chaque liaison de a : une pour a, une pour le x concerné (puisque ce x a désormais un degré de plus). Soit un total de 2*d nouvelles entrées.

Cette partition permet notamment, lors chaque nouvelle liaison, d'éviter d'itérer sur les degrés des sommets éligibles pour savoir lequel il faut lier.

II) Analyse des graphes Barabási–Albert

Avant tout, on remarque que :

- Un graphe standard généré est toujours connexe : le graphe initial est connexe, et chaque nouveau sommet est ajouté au précédent graphe.
- Un graphe orienté généré n'est jamais fortement connexe : le dernier sommet généré n'a aucun sommet entrant.

Exercice 2 : k-cœur

Soit k le nombre de voisins minimal nécessaire pour faire partie du k-coeur correspondant, n_{Core} le nombre de sommets d'un k-coeur, et m_{Core} le nombre total d'arêtes d'un k-coeur.

Graphe non orienté :

Le *nombre d'arêtes de chaque sommet* (en exceptant les sommets initiaux qui en ont au moins plus, sachant que $d < n_0$) est déterminé par le paramètre d lors de la génération du graphe, donc :

- Il existe un nombre d de k-coeurs avec k <= d équivalents au graphe entier, puisque chaque sommet a au moins d arêtes. **Pour k <= d : n_{Core}=n et m_{Core}=** $\sum_{i=1}^{n0} i$ + **$(n-n_0)*d$**. m_{Core}, soit l'ensemble des arêtes du graphe,

  additionne en particulier les arêtes initiales et les arêtes des nouveaux sommets.
- Il existe un nombre n_0-d-1 de k-coeurs, tous composés des sommets initiaux et de leurs arêtes. **Pour k > d et k < n_0, $n_{Core} = n_0$, m_{Core} =** $\sum_{i=1}^{n0} i$.

  Avec $d < n_0$, les sommets ajoutés de n0 à n ne sont en effet plus pris en compte puisque leur nombre d'arêtes avec les sommets initiaux est inférieur à k. Pour d=n0, voir cas décrit précédemment.
- Pour tout graphe non orienté selon le modèle de Barabsi-Albert, il existe donc n0-1 k-coeurs (en excluant k=0).

Par exemple, avec d = 2, n_0 = 7, n = 50, on a :

```
k=1 : n=50, m=107 ; density=2.14
k=2 : n=50, m=107 ; density=2.14
k=3 : n=7, m=21 ; density=3.0
k=4 : n=7, m=21 ; density=3.0
k=5 : n=7, m=21 ; density=3.0
k=6 : n=7, m=21 ; density=3.0
```

Graphe orienté :

Comme le k-coeur d'un graphe orienté prend en considération les sommets sortants d'un sommet, et que ce sont ces mêmes sommets sortants qui sont créés lors de la génération, on a ici les **mêmes propriétés de k-coeur** qu'avec les graphes non orientés. Les résultats obtenus vont de paire avec cette affirmation.

Exercice 3 : diamètre et distance moyenne

On peut commencer en affirmant qu'un graphe généré selon le modèle de Barabási-Albert avec d=0 est **le même** qu'un autre généré selon le modèle d'Erdős–Rényi avec $\mathbf{n_{ER}=n_0}$ et $\mathbf{p_{ER}=1}$. Il y a ici au moins $n_{Er}=n_0$ sommets qui sont à distance 1 l'un de l'autre, puisque tous interconnectés.

N'étant pas certain que mon programme calcule un groupe de distance (Dijkstra) en O(logn), je ne me prononce pas quant à la comparaison de complexité.

Exercice 4 : distribution des degrés

On considère du TP5 que pour une power law, "si $\gamma = 3$ plus de 80% des sommets ont degré 1". Pour obtenir un résultat similaire avec le modèle de Barabási-Albert, il faudrait donc **d=1** pour que les nouveaux sommets soient de degré 1, et s'assurer qu'ils représentent plus 80% des sommets, soit

$$\mathrm{A} = n - n0 \geq 80\% \times n$$

$$\mathrm{A} = 20\% \times n \geq n0$$

$$\mathrm{A} = n \geq 5n0$$

Ainsi, on aurait un nombre n de sommets initiaux, et au moins 4 fois ($5*n_0$) plus de sommets à ajouter ensuite. Exceptions : les sommets ajoutés peuvent toujours être liés à d'autres sommets ajoutés après, et n'auraient plus degré 1. Pour compenser ces exceptions, on peut penser à mettre n beaucoup plus grand. De même, elles seraient moins susceptibles de se produire avec n_0 très grand, augmentant les probabilités d'être liés à la place des nouveaux sommets.