

Activité : Moteur de recherche/Page Rank

Seconde

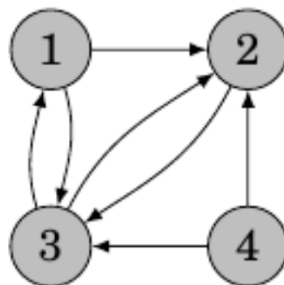
5 Janvier 2026

Le succès du moteur de recherche Google vient de sa capacité à classifier les pages web par popularité :

1. Une page web est populaire si beaucoup d'autres pages web renvoient vers celle-ci (via les liens hypertexte) ;
2. Une page web est d'autant plus populaire qu'elle est citée par d'autres pages populaires.

Pour établir un score de popularité, le moteur de recherche va d'abord prendre la liste de tous les sites web (qui font figurer les mots-clés de la recherche), puis en prendre un au hasard. Ensuite, elle va regarder tous les liens de cette page web (vers les autres pages de la liste) et en visiter un au hasard. Ce processus continue alors, en notant à chaque fois le site visité. Le site le plus visité sera le plus populaire, et inversement.

Exercice 1. Pour modéliser le fonctionnement de Google (ou plutôt de son algorithme PageRank), on utilise un graphe :

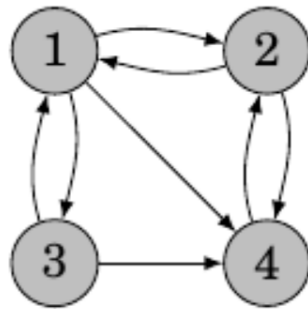


Les sommets représentent les pages web, et les flèches représentent les liens hypertextes. Par exemple, il y a un lien hypertexte dans la page 1 pointant vers la page 2.

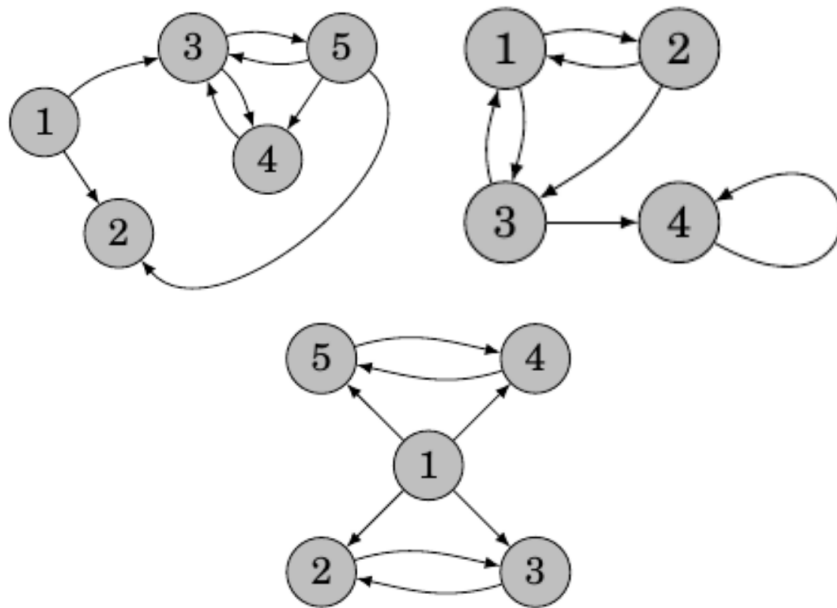
Munissez-vous d'un dé (ou rendez-vous sur random.org). Vous commencez sur la page 1. Ensuite, choisissez au hasard une des flèches issue de la page courante, et visitez-la. Continuez cette procédure trente fois, en comptabilisant chacune des pages visitées.

- a) Quelle page est la plus populaire ?
- b) Comparez votre réponse avec vos camarades.

Exercice 2. Recommencer l'exercice sur le graphe suivant :



Exercice 3. Justifier ce qui pose problème avec les trois graphes suivant. Comment y remédier ?



L'idée des créateurs de Google va être d'ajouter de la « téléportation » : à chaque étape, on décide au hasard si on fait une visite normale, ou si on se « téléporte » sur une des pages au hasard (n'importe laquelle, même s'il n'y pas de lien hypertexte). La téléportation arrive dans 15% des cas.

Exercice 4. Essayez vous-même.