

Examen du cours Moteurs de recherche M2 Pro, Université Paris Diderot, Paris 7

Michel Habib et Antoine Meyer

March 6, 2008

1 Questions de cours

1. L'équipe technique d'un moteur de recherche décide de procéder à une catégorisation des pages WEB qu'elle possède dans son gigantesque système de fichiers.
Quelles méthodes peut-elle utiliser ?
2. Comment caractériser une communauté dans un réseau social. Donner une définition la plus opérationnelle (i.e. facilement calculable) possible.
3. Proposer un algorithme de détection de communautés dans un réseau (par exemple le graphe des pages WEB). Justifier votre choix.
4. Dans le cadre de la conception d'un nouveau moteur de recherche généraliste, on se pose la question de l'utilisation de PageRank, pour l'affichage des résultats d'une requête. Comment justifier l'utilisation de cet ordre et comment l'améliorer (le but étant que les réponses soient classées en fonction de leur pertinence) ? (Piste: il est possible d'utiliser un profil d'utilisateur, un profil de requête ...)

2 Exercice

1. Considérons la fonctionnalité suivante proposée par le site Amazon. Lorsqu'un internaute consulte un livre le système lui propose automatiquement des livres voisins. Donner d'autres exemples d'applications de telles fonctionnalités.
2. Comment réaliser une telle fonctionnalité ?
3. Préciser les spécifications de la méthode que vous préconisez.