

Conservation d'aléa dans les structures de données

Proposition de sujet de mémoire de recherche

- **Encadrant** : Philippe Duchon (duchon@labri.fr)
- **Équipe** : Combinatoire et Algorithmique
- **Thématique et mots-clés** : algorithmes probabilistes, structures de données randomisées

Conservation d'aléa et structures randomisées

Dans le domaine des structures de données, un grand classique est l'arbre binaire de recherche. L'analyse de la forme d'un arbre binaire de recherche aléatoire (résultant d'une séquence d'insertions dans un ordre aléatoire) révèle que ces arbres sont en général de hauteur logarithmique, ce qui garantit de bonnes complexités aux opérations de mise à jour. Toutefois, les arbres binaires de recherche sont assez sensibles à cet ordre des opérations, et il est bien connu qu'insérer des clés dans un ordre croissant résulte en des arbres fortement déséquilibrés.

Il existe à ce problème des solutions déterministes garantissant que les arbres restent assez équilibrés, mais également des solutions dans lesquelles les opérations de mise à jour sont randomisées (elles font appel à des "tirages aléatoires"), et où l'on garantit que l'arbre résultant d'une séquence d'opérations aura (en loi) la même forme qu'un arbre résultant seulement d'une séquence d'insertions dans un ordre aléatoire uniforme. De telles propriétés de *conservation d'aléa* sont souhaitables, d'une part car elles simplifient l'étude des performances des structures de données (il n'y a pas de risque de "dérive" par rapport à une situation idéale après une séquence mal choisie d'opérations), mais également car elles permettent de transposer automatiquement des résultats connus (et abondants) sur les structures aléatoires.

Une formalisation de cette notion de préservation d'aléa a été donnée par Knuth [2], en termes de propriétés que doivent satisfaire les opérations de mise à jour pour obtenir divers degrés de préservation d'aléa.

Sujet de recherche

L'objectif de ce travail de recherche est d'appliquer le formalisme de Knuth à différents exemples de structures de données probabilistes : les treaps aléatoires d'Aragon et Seidel [1], les arbres binaires de recherche randomisés de Martinez et Roura [3], les skip lists aléatoires de Pugh [4]. Une extension à des structures pouvant exploiter d'autres informations que l'ordre relatif des clés est également envisageable.

Ce sujet pourra éventuellement déboucher sur une thèse dont le sujet serait dans la continuité de celui-ci, appliquée cette fois aux graphes aléatoires dynamiques, avec en ligne de mire la maintenance de réseaux logiques.

Références

- [1] Aragon, C.R., and Seidel, R.G. 1989. *Randomized Search Trees*. In *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. IEEE, New York, pp 540–545.
- [2] Knuth, D. *Deletions that preserve randomness*. IEEE Transactions on Software Engineering **SE-3** (1977), pp 351–359.
- [3] Martínez, C., and Roura, S. 1998. *Randomized binary search trees*. J. Assoc. Comput. Mach., 45(2), pp 288–323, 1998.
- [4] Pugh, W. 1990. *Skip lists : a probabilistic alternative to balanced trees*. Commun. ACM 33, 6 (June), pp 668–676.