

Rice-Poisson-Mellin-Newton-Laplace

Brigitte Vallée
CNRS et Université de Caen
brigitte.vallee@unicaen.fr

Beaucoup d'algorithmes travaillent avec des entrées qui sont des ensembles finis de données : par exemple, les données sont des mots pour les algorithmes du texte ou des points pour les algorithmes géométriques. Le cardinal N de l'ensemble joue un rôle principal, et c'est souvent la taille de l'entrée. On s'intéresse au comportement de l'algorithme quand N devient grand.

Double cadre probabiliste. On adopte en général le cadre probabiliste suivant. Chaque donnée x (mot ou point) est produite avec une distribution, et l'ensemble \mathcal{X} des données est donc un espace probabilisé. On suppose toujours que les données sont tirées indépendamment, et l'ensemble \mathcal{X}^N des entrées de taille N est muni de la probabilité produit.

L'espace des entrées de l'algorithme est ainsi l'ensemble \mathcal{X}^* des suites finies d'éléments de \mathcal{X} , et il y a deux modèles probabilistes principaux.

- (a) Le modèle de Bernoulli \mathcal{B}_n , où le cardinal N est fixé égal à n (puis tend dans un second temps vers ∞);
- (b) Le modèle de Poisson \mathcal{P}_z , où le cardinal N est lui-même une variable aléatoire qui suit une loi de Poisson de paramètre z ,

$$\mathbb{P}[N = n] = e^{-z} \frac{z^n}{n!},$$

le paramètre z ayant vocation à tendre aussi vers ∞ .

Transferts entre les deux modèles. Le modèle de Bernoulli est plus naturel en algorithmique, mais le modèle de Poisson jouit de propriétés probabilistes très intéressantes, notamment d'indépendance. Donc, on procède très souvent comme suit. On considère une variable $R : \mathcal{X}^* \rightarrow \mathbb{N}$ qui décrit le comportement de l'algorithme sur l'entrée $\mathbf{x} \in \mathcal{X}^*$ (par exemple, longueur de cheminement du trie ou du dst construit sur \mathbf{x} , nombre de sommets de l'enveloppe convexe construite sur \mathbf{x}). Et on veut en faire l'analyse dans le modèle \mathcal{B}_n . On commence l'analyse dans le modèle \mathcal{P}_z , puis on cherche à transférer l'analyse dans le modèle \mathcal{B}_n .

Comment opérer ce transfert d'abord dans l'analyse en moyenne du coût R ? Il y a deux principaux chemins qui comparent les espérances B_n de R dans le modèle \mathcal{B}_n et les espérances $P(z)$ de R dans le modèle \mathcal{P}_z .

(DP) *dépoissonisation*. On voit z comme une variable complexe, et on analyse l'asymptotique de $P(z)$ dans des cônes. Sous certaines conditions (DP), et en utilisant notamment la transformée de Mellin, on peut relier l'asymptotique de B_n et celle de la valeur $P(n)$,

(Ri) *méthode de Rice*. On travaille avec les coefficients P_n de la série $P(z)$. Sous certaines conditions (Ri) [existence d'un relèvement "apprivoisé" $\pi(z)$ de la suite $n \mapsto P_n$], alors on peut relier l'asymptotique de B_n et le comportement singulier de $\pi(z)$.

Comparaison des transferts. Dans la communauté AofA, il y a les tenants de la dépoissonisation et ceux de la méthode de Rice, qui ne communiquent pas trop entre eux. Cet exposé cherchera à comparer ces méthodes du point de vue de leur élégance, des outils employés et de leur généralité. Il

- (a) décrira ces deux méthodes;
- (b) expliquera le cycle Rice-Poisson-Mellin-Newton qui les relie en profondeur;
- (c) introduira un cadre assez général, lié à la transformée de Laplace, où les deux méthodes semblent démontrer une efficacité comparable.