

Titre: Comptage de répétitions avec gaps dans les mots d'un langage régulier

Encadrant: Philippe DUCHON

Équipe: Combinatoire et Algorithmique

On s'intéresse à l'énumération en moyenne, dans des mots aléatoires, de motifs caractérisés par la répétition d'un sous-motif à une distance qui est bornée par la longueur du sous-motif. Plus précisément, pour un paramètre $\alpha \geq 1$, un α -gapped repeat (respectivement, α -gapped palindrom) dans un mot w est un facteur $u.v.u$ (respectivement, $u.v.\bar{u}$, où \bar{u} désigne le miroir de u) de w , tel que l'on ait $|uv| \leq \alpha|u|$.

Un travail en cours de Duchon, Nicaud et Pivoteau montre que, quand les mots w autorisés sont tous les mots sur un alphabet à $k \geq 2$ lettres, la série génératrice suivante est rationnelle pour tout $\alpha \in \mathbb{Q}$: $S_\alpha(t; x, y, z)$, où

- t compte la longueur des mots
- x compte le nombre de α -gapped repeats (resp. α -gapped palindroms)
- y compte la longueur totale des facteurs u des motifs
- z compte la longueur totale des facteurs v des motifs

À partir de l'expression de cette série génératrice, on obtient une estimation asymptotique des valeurs moyennes des différents paramètres pour des mots aléatoires uniformes de longueur n . Chacun des paramètres est d'ordre $c.n$, pour une constante c qui dépend explicitement du paramètre α .

L'objectif de ce stage de recherche sera d'étendre cette étude au cas où les mots sont pris dans un langage régulier quelconque, et de confirmer (ou d'infirmier) l'intuition selon laquelle les mêmes résultats qualitatifs restent valables: série génératrice rationnelle et asymptotique linéaire, au moins dans le cas des α -gapped repeats. Le point clé devrait être l'obtention de séries rationnelles; les théorèmes généraux de combinatoire analytique devraient fournir à moindre effort le comportement asymptotique.