

Mise à jour incrémentale de tableau des suffixes en cours de recodage

Matthias Gallé¹, Pierre Peterlongo¹ et François Coste¹

Symbiose, IRISA, INRIA, CNRS, Université Rennes 1

Résumé

De part sa légèreté, sa puissance et sa simplicité, le tableau des suffixes est une structure de données de plus en plus utilisée dans l'algorithmique du texte, en particulier pour les problèmes nécessitant une indexation.

Cependant, l'un des points faibles de cette structure réside dans le fait que sa mise à jour, suite à une modification du texte indexé, n'est pas aisée. En particulier, elle se prête assez mal à la suppression ou l'ajout de portions de texte.

Motivés par des problématiques d'inférence grammaticale, nous cherchons dans ce travail une solution pour mettre à jour le tableau des suffixes suite à la substitution dans le texte indexé de **toutes les occurrences** d'un mot par un nouveau caractère unique.

De ce travail en cours d'aboutissement, nous présenterons les principaux axes algorithmiques ainsi que les premiers résultats, qui sont encourageants.