



Algorithme Boyer-Moore

par Leif Henriksen et Damien Duchon



Historique

Il s'agit d'un algorithme de recherche de sous-chaîne.

Il a été trouvé par *Robert S. Boyer* et *J Strother Moore* en 1977.



Comparaison avec l'algorithme naïf

	Algorithme Naïf	Algorithme Boyer-Moore
Complexité* (pire)	$O(n*m)$	$O(n+m)$
Benchmark**	63s	25s

* Avec n la taille de la chaîne où l'on recherche et m la taille de la sous-chaîne.

** Pour un même jeu de donnée



Fonctionnement

Là où l'algorithme naïf commence par comparer le premier caractère de la sous-chaîne, celui de Boyer-Moore commence par comparer le dernier caractère. Ce qui lui permet d'avancer beaucoup plus vite dans l'analyse si la sous-chaîne n'est pas présente.

La deuxième heuristique de Boyer-Moore est de créer une table des suffixes pour avancer.

Exemple :

chaîne où l'on cherche : ACGTACCTCAGC

sous-chaîne à trouver : TAC

ACG	TAC	CTCAGC	ACGT	AC	CTCAGC	ACGT	AC	CTCAGC	ACGT	AC	CTCAGC	ACGT	AC	CTCAGC	ACGT	AC	CTCAGC
TAC			TAC			TAC			TAC			TAC			TAC		

résultat : 1 occurrence

Merci !

