COURS

Recherche textuelle : algorithme de Boyer-Moore

Tle

Programme de terminale générale, enseignement de spécialité — BO spécial n°8 du 25 juillet 2019
Relue le 14 novembre 2026

Chercher un mot dans un livre, un gène dans une séquence d'ADN, un virus dans un fichier : il faut trouver un motif dans un texte, le plus vite possible. L'algorithme de Boyer-Moore évite de tester chaque position une à une.

1 L'algorithme naïf

RÈGLE

On place le motif à chaque position du texte et on compare caractère par caractère. Pour un texte de longueur n et un motif de longueur m, le nombre de comparaisons peut atteindre environ **n × m**.

2 L'idée de Boyer-Moore

DÉFINITION

Version simplifiée de Boyer-Moore : on compare le motif au texte **de droite à gauche** ; en cas d'échec, on **décale** le motif grâce à une **table de sauts** calculée à l'avance.

DÉFINITION

Pour chaque caractère du motif (sauf le dernier), distance entre sa dernière occurrence et la fin du motif ; un caractère absent du motif permet de sauter de toute la longueur m.

3 Efficacité

RÈGLE

Plus le motif est long et l'alphabet varié, plus les sauts sont grands : on lit bien moins de caractères que l'algorithme naïf. C'est le principe utilisé par la fonction « Rechercher » des éditeurs de texte.

EXEMPLE

Calcule la table de sauts du motif « NSI ».

1. Longueur m = 3.
2. N est à 2 positions de la fin : saut 2 ; S à 1 position : saut 1.
3. Tout autre caractère : saut 3.

La table de sauts indique de combien décaler le motif selon le caractère du texte aligné avec la fin du motif.

LE PIÈGE QUI COÛTE LE PLUS DE POINTS

Comparer de gauche à droite dans Boyer-Moore. — L'algorithme compare depuis la fin du motif : c'est ce qui permet les grands sauts.

Le réflexe : Toujours partir du dernier caractère du motif.

APPLIQUER BOYER-MOORE-HORSPOOL

1. Calculer la table de sauts du motif.
2. Aligner le motif au début du texte.
3. Comparer de droite à gauche.
4. Décaler selon la table en cas d'échec, jusqu'à la fin du texte.

À VÉRIFIER

- Ai-je défini le vocabulaire ?
- Mon code est-il testé sur un exemple ?
- Ai-je traité les cas limites ?
- Ai-je justifié la complexité ou le résultat ?

À RETENIR

- Algorithme **naïf** : jusqu'à n × m comparaisons.
- **Boyer-Moore-Horspool** : comparaison de droite à gauche, **table de sauts**.
- Grands sauts quand le motif est long : bien plus rapide.