

# TD\_Boyer Moore

<http://whocouldthat.be/visualizing-string-matching/>

[https://en.wikipedia.org/wiki/Boyer%E2%80%93Moore\\_string-search\\_algorithm#Description\\_3](https://en.wikipedia.org/wiki/Boyer%E2%80%93Moore_string-search_algorithm#Description_3)

Bon suffixe si le motif est A B A B A  
On cherche une liste de taille le motif BS.

**Pour BS [4].** Si on utilise BS [4] c'est que la dernière lettre du motif et du texte ne matche pas.

A B A B **A**

On décale de 1. Améliorable en disant qu'on décale jusqu'à ne plus avoir cette dernière lettre (MC...)

**Pour BS [3].** Si on utilise BS [3] c'est que les dernières lettres matchent MAIS pas les précédentes.

A B A **B A**

Dans ce cas,

- On cherche s'il y a, dans le motif, quelque chose du style **X A** où X est une lettre différente de B. Ici il n'y en a pas ! (c'est le cas simple du suffixe que l'on pourrait retrouver...)

- Sinon, on cherche s'il y a, au début du motif, le sous-motif **A**.

**A B A B A** C'est le cas : On décale de 4 pour faire correspondre les deux A.

Si, on trouve bien A au tout début, il faut considérer que la lettre précédente qui se "trouverait" à la position -1 est toujours différente de B. Le calcul  $i-k+1 = 3-0+1 = 4$  permet de trouver le même décalage de 4.

**Pour BS [2].** Si on utilise BS [2] c'est que les 2 dernières lettres matchent MAIS pas les précédentes.

A B **A B A**

Dans ce cas,

- On cherche s'il y a, dans le motif, quelque chose du style **X B A** où X est une lettre différente de A. Ici il n'y en a pas ! (c'est le cas simple du suffixe que l'on pourrait retrouver...)

- Sinon, on cherche s'il y a, au début du motif, le sous-motif **B A**. NON

- Sinon, on cherche s'il y a, au début du motif, le sous-motif **A**.

**A B A B A** C'est le cas : On décale de 4 pour faire correspondre les deux A.

**Pour BS [1].** Si on utilise BS [1] c'est que les 3 dernières lettres matchent MAIS pas les précédentes.

A **B A B A**

Dans ce cas,

- On cherche s'il y a, dans le motif, quelque chose du style **X A B A** où X est une lettre différente de B. Ici il n'y en a pas ! (c'est le cas simple du suffixe que l'on pourrait retrouver...)

- Sinon, on cherche s'il y a, au début du motif, le sous-motif **A B A**.

**A B A B A** C'est le cas : les deux motifs se chevauchent... On décale de 2 pour les faire correspondre.

Si, on trouve bien ABA au tout début, il faut considérer que la lettre précédente qui se "trouverait" à la position -1 est toujours différente de B. Le calcul  $i-k+1 = 1-0+1 = 2$  permet de trouver le même décalage de 2.

**Pour BS [0].** Si on utilise BS [0] c'est que toutes les lettres, sauf la première, matchent.

**A B A B A**

Dans ce cas,

- On cherche s'il y a, dans le motif, quelque chose du style **X B A B A** où X est une lettre différente de A. Ici il n'y en a pas ! (c'est le cas simple du suffixe que l'on pourrait retrouver...)

- Sinon, on cherche s'il y a, au début du motif, le sous-motif **B A B A**. NON

- Sinon, on cherche s'il y a, au début du motif, le sous-motif **A B A**.

**A B A B A** C'est le cas : les deux motifs se chevauchent... On décale de 2 pour les faire correspondre.

Cela fonctionne (cf. page wikipedia) mais je trouve plus simple de raisonner selon le cours : si le sous-motif n'est pas trouvé, alors on passe automatiquement au cas n°2. On cherche le plus grand préfixe qui est également suffixe, ici "ABA". Il se trouve en position  $p=2$  (slide 19) => décalage de 2. Je suis d'accord que tout cela revient à la fin à aligner les deux sous-motifs ABA.

Donc BS = [2, 2, 4, 4, 1]