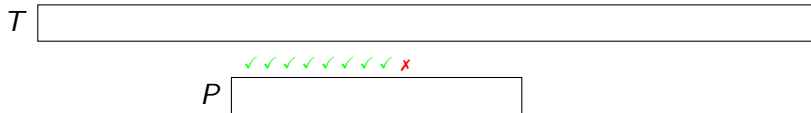


Qu'est-ce qu'un bon décalage ?

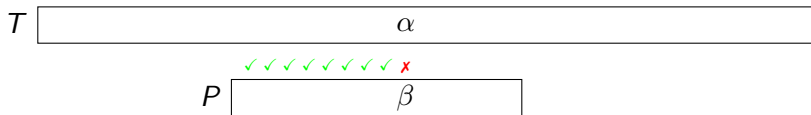
T

P

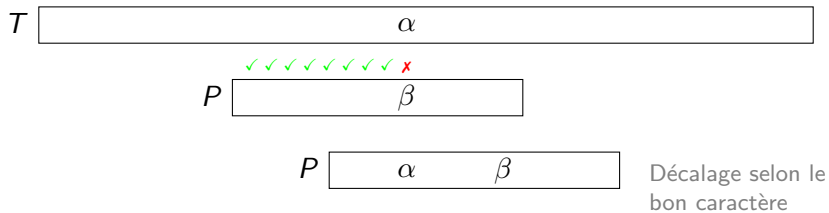
Qu'est-ce qu'un bon décalage ?



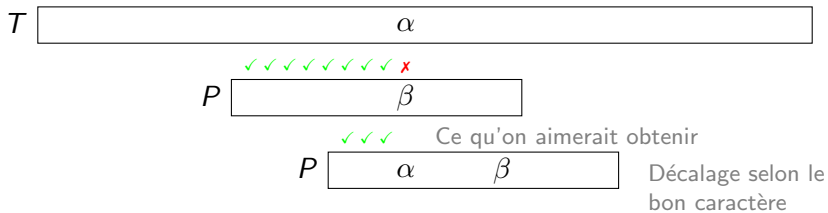
Qu'est-ce qu'un bon décalage ?



Qu'est-ce qu'un bon décalage ?



Qu'est-ce qu'un bon décalage ?



✓✓✓✓✓✓✓✓✓✓x

P β

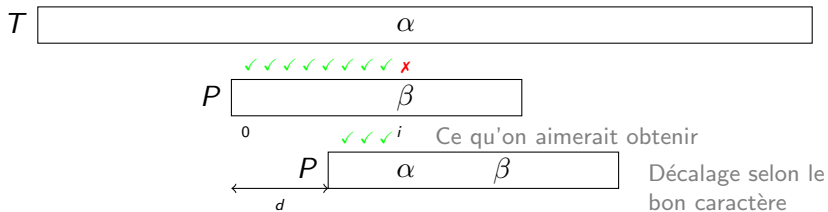
✓✓✓

P α β

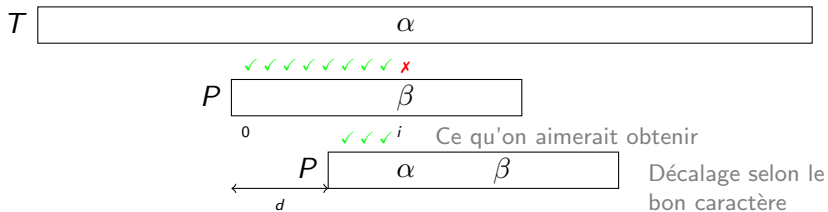
Ce qu'on aimerait obtenir

Décalage selon le bon caractère

Qu'est-ce qu'un bon décalage ?

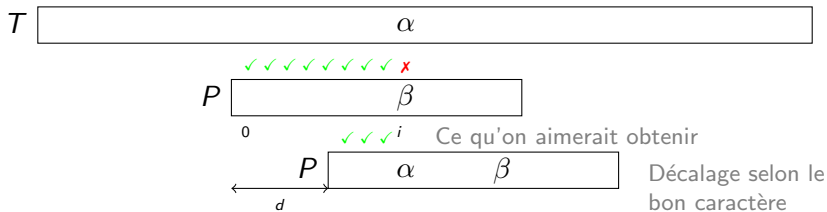


Qu'est-ce qu'un bon décalage ?



Meilleur décalage pour P , après un échec à la position i : plus petite valeur de d telle que $P[0 \dots i - d - 1] = P[d \dots i - 1]$ et $P[i - d] \neq P[i]$

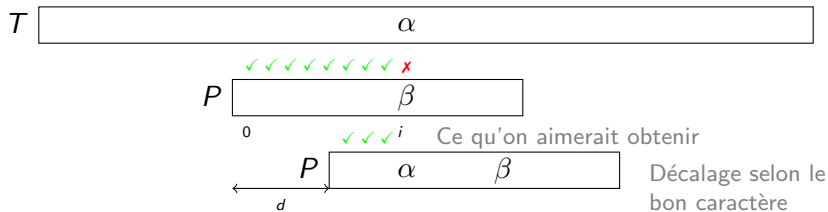
Qu'est-ce qu'un bon décalage ?



Meilleur décalage pour P , après un échec à la position i : plus petite valeur de d telle que $P[0 \dots i - d - 1] = P[d \dots i - 1]$ et $P[i - d] \neq P[i]$

Autrement dit il faut trouver le plus long bord de $P[0 \dots i - 1]$ qui soit suivi par un symbole différent au niveau du préfixe et du suffixe

Qu'est-ce qu'un bon décalage ?



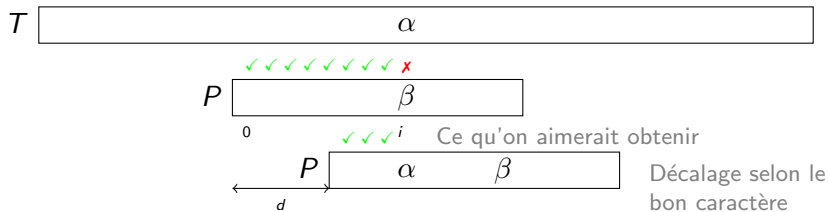
Meilleur décalage pour P , après un échec à la position i : plus petite valeur de d telle que $P[0 \dots i - d - 1] = P[d \dots i - 1]$ et $P[i - d] \neq P[i]$

$P = \text{agagc}$

i 0 1 2 3 4

plus long bord
décalage

Qu'est-ce qu'un bon décalage ?



Meilleur décalage pour P , après un échec à la position i : plus petite valeur de d telle que $P[0 \dots i - d - 1] = P[d \dots i - 1]$ et $P[i - d] \neq P[i]$

$P = \text{agagc}$

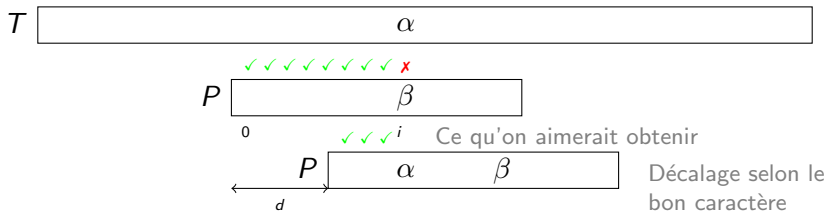
$i = 0$, plus long bord de $P[0 \dots 0 - 1]$ suivis de 2 caractères différents?

a g a g c

i 0 1 2 3 4

plus long bord
décalage

Qu'est-ce qu'un bon décalage ?



Meilleur décalage pour P , après un échec à la position i : plus petite valeur de d telle que $P[0 \dots i - d - 1] = P[d \dots i - 1]$ et $P[i - d] \neq P[i]$

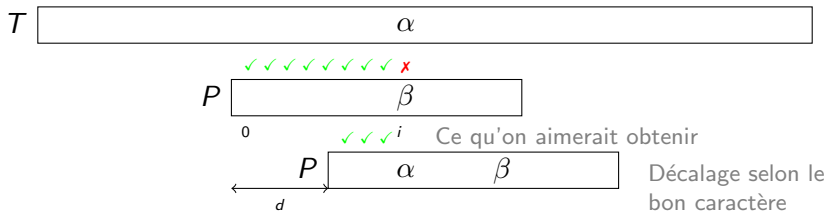
$P = \text{agagc}$

$i = 0$, plus long bord de $P[0 \dots 0 - 1]$ suivis de 2 caractères différents?

a g a g c

i	0	1	2	3	4
plus long bord	—				
décalage	1				

Qu'est-ce qu'un bon décalage ?



Meilleur décalage pour P , après un échec à la position i : plus petite valeur de d telle que $P[0 \dots i - d - 1] = P[d \dots i - 1]$ et $P[i - d] \neq P[i]$

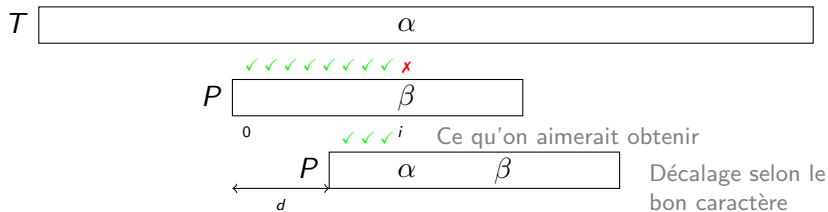
$P = \text{agagc}$

$i = 1$, plus long bord de $P[0 \dots 1 - 1]$ suivis de 2 caractères différents?

a g a g c

i	0	1	2	3	4
plus long bord	—				
décalage	1				

Qu'est-ce qu'un bon décalage ?



Meilleur décalage pour P , après un échec à la position i : plus petite valeur de d telle que $P[0 \dots i - d - 1] = P[d \dots i - 1]$ et $P[i - d] \neq P[i]$

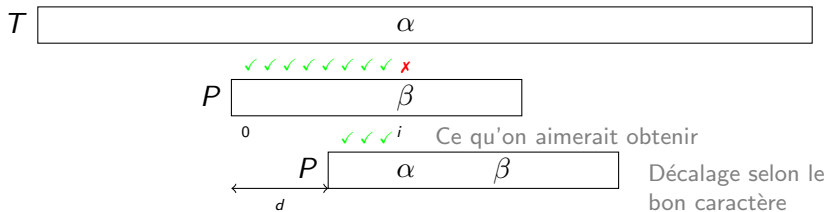
$P = \text{agagc}$

$i = 1$, plus long bord de $P[0 \dots 1 - 1]$ suivis de 2 caractères différents?

a g a g c

i	0	1	2	3	4
plus long bord	-	0			
décalage	1	1			

Qu'est-ce qu'un bon décalage ?



Meilleur décalage pour P , après un échec à la position i : plus petite valeur de d telle que $P[0 \dots i - d - 1] = P[d \dots i - 1]$ et $P[i - d] \neq P[i]$

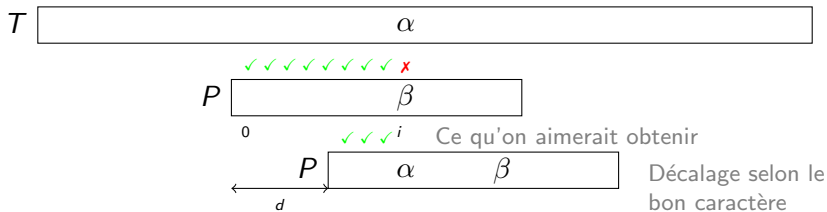
$P = \text{agagc}$

$i = 2$, plus long bord de $P[0 \dots 2 - 1]$ suivis de 2 caractères différents?

a g a g c

i	0	1	2	3	4
plus long bord	-	0			
décalage	1	1			

Qu'est-ce qu'un bon décalage ?



Meilleur décalage pour P , après un échec à la position i : plus petite valeur de d telle que $P[0 \dots i - d - 1] = P[d \dots i - 1]$ et $P[i - d] \neq P[i]$

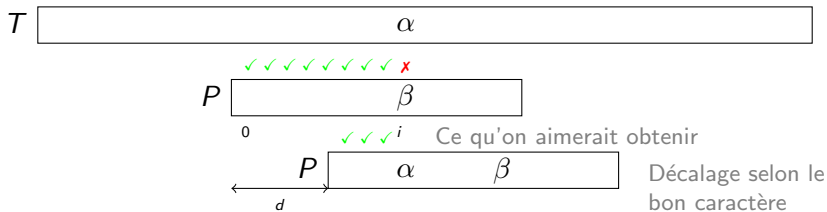
$P = \text{agagc}$

$i = 2$, plus long bord de $P[0 \dots 2 - 1]$ suivis de 2 caractères différents?

a g a g c

i	0	1	2	3	4
plus long bord	–	0	–		
décalage	1	1	3		

Qu'est-ce qu'un bon décalage ?



Meilleur décalage pour P , après un échec à la position i : plus petite valeur de d telle que $P[0 \dots i - d - 1] = P[d \dots i - 1]$ et $P[i - d] \neq P[i]$

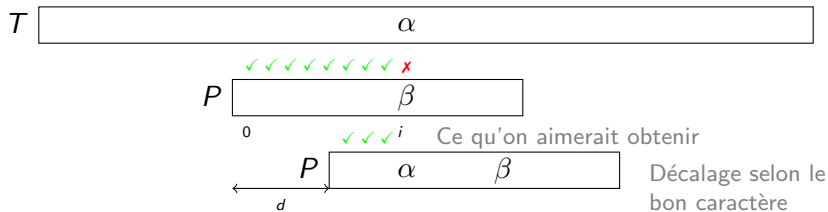
$P = \text{agagc}$

$i = 3$, plus long bord de $P[0 \dots 3 - 1]$ suivis de 2 caractères différents?

a g a g c

i	0	1	2	3	4
plus long bord	-	0	-		
décalage	1	1	3		

Qu'est-ce qu'un bon décalage ?



Meilleur décalage pour P , après un échec à la position i : plus petite valeur de d telle que $P[0 \dots i - d - 1] = P[d \dots i - 1]$ et $P[i - d] \neq P[i]$

$P = \text{agagc}$

$i = 3$, plus long bord de $P[0 \dots 3 - 1]$ suivis de 2 caractères différents?

a g a g c

i	0	1	2	3	4
plus long bord	-	0	-	0	
décalage	1	1	3	3	

Exemple avec l'algorithme de Knuth, Morris et Pratt (1977)

$P = \text{agagc}$

i 0 1 2 3 4

d 1 1 3 3 2

$T = \text{a g a a g a g a g c}$

Exemple avec l'algorithme de Knuth, Morris et Pratt (1977)

$P = agagc$

i	0	1	2	3	4
d	1	1	3	3	2

$T =$	a	g	a	a	g	a	g	a	g	c
	✓	✓	✓	✗						
	a	g	a	g	c					

Exemple avec l'algorithme de Knuth, Morris et Pratt (1977)

$P = agagc$

i	0	1	2	3	4
d	1	1	3	3	2

$T =$

a	g	a	a	g	a	g	a	g	c
✓ a	✓ g	✓ a	✗ g	c					
			✓ a	✓ g	✓ a	✓ g	✗ c		

Exemple avec l'algorithme de Knuth, Morris et Pratt (1977)

$P = agagc$

i	0	1	2	3	4
d	1	1	3	3	2

$T =$

a	g	a	a	g	a	g	a	g	c
✓ a	✓ g	✓ a	✗ g	c					
			✓ a	✓ g	✓ a	✓ g	✗ c		
					a	g	✓ a	✓ g	✓ c

Exemple avec l'algorithme de Knuth, Morris et Pratt (1977)

$P = agagc$

i	0	1	2	3	4
d	1	1	3	3	2

$T =$

a	g	a	a	g	a	g	a	g	c
✓ a	✓ g	✓ a	✗ g	c					
		✓ a	✓ g	✓ a	✓ g	✗ c			
				a	g	✓ a	✓ g	✓ c	

Avec cet algorithme, on effectue **au plus** $2 \times |T| - 1$ comparaisons