

Structures d'indexation compressées et dynamiques pour le texte

Mikaël Salson

LITIS, Université de Rouen

8 décembre 2009



Constat

Le volume de données numériques croît, il est indispensable de pouvoir rechercher l'information qui nous intéresse parmi celles-ci.

Constat

Le volume de données numériques croît, il est indispensable de pouvoir rechercher l'information qui nous intéresse parmi celles-ci.

Recherche de données

Nous voulons rechercher toutes les occurrences de *motifs* dans un texte connu à l'avance.

Constat

Le volume de données numériques croît, il est indispensable de pouvoir rechercher l'information qui nous intéresse parmi celles-ci.

Recherche de données

Nous voulons rechercher toutes les occurrences de *motifs* dans un texte connu à l'avance.

Deux possibilités :

- ▶ texte court ou peu de recherches à effectuer : parcours du texte (Knuth, Morris et Pratt, 1977 ; Boyer et Moore, 1977) ;
- ▶ texte long ou beaucoup de recherches à effectuer : indexation du texte.

Indexation non compressée

- ▶ arbre des suffixes (Weiner, 1973);
- ▶ automate des suffixes (Blumer *et al.*, 1983);
- ▶ table des suffixes (Manber et Myers, 1990; Gonnet *et al.*, 1992);
- ▶ vecteur des suffixes (Monostori *et al.*, 2001);
- ▶ arbre contracté des suffixes (Ehrenfeucht *et al.*, 2009).

Indexation compressée

- ▶ LZ-index (Kärkkäinen et Ukkonen, 1998);
- ▶ FM-index (Ferragina et Manzini, 2000);
- ▶ table compressée des suffixes (Grossi et Vitter, 2000);
- ▶ table compacte des suffixes (Mäkinen, 2000).

Problème

À chaque texte correspond un index particulier.
Si le texte est modifié, l'index ne lui correspond plus, il devient caduque.

Problème

À chaque texte correspond un index particulier.
Si le texte est modifié, l'index ne lui correspond plus, il devient caduque.

Objectif

Proposer une structure d'indexation supportant n'importe quel type de mise à jour (insertions, délétions et substitutions).

Indexation non compressée

- ▶ arbre des suffixes (Weiner, 1973);
- ▶ automate des suffixes (Blumer *et al.*, 1983);
- ▶ table des suffixes (Manber et Myers, 1990; Gonnet *et al.*, 1992);
- ▶ vecteur des suffixes (Monostori *et al.*, 2001);
- ▶ arbre contracté des suffixes (Ehrenfeucht *et al.*, 2009).

Indexation compressée

- ▶ LZ-index (Kärkkäinen et Ukkonen, 1998);
- ▶ **FM-index (Ferragina et Manzini, 2000);**
- ▶ table compressée des suffixes (Grossi et Vitter, 2000);
- ▶ table compacte des suffixes (Mäkinen, 2000).

- ▶ Introduit par Ferragina et Manzini (2000);
- ▶ structure d'indexation compressée basée sur la transformée de Burrows-Wheeler;
- ▶ utilise :
 - ▶ la transformée de Burrows-Wheeler;
 - ▶ un échantillon de la table des suffixes.

- ▶ Introduit par Ferragina et Manzini (2000) ;
- ▶ structure d'indexation compressée basée sur la transformée de Burrows-Wheeler ;
- ▶ utilise :
 - ▶ la transformée de Burrows-Wheeler ;
 - ▶ un échantillon de la table des suffixes.

- ▶ Introduit par Ferragina et Manzini (2000) ;
- ▶ structure d'indexation compressée basée sur la transformée de Burrows-Wheeler ;
- ▶ utilise :
 - ▶ la transformée de Burrows-Wheeler ;
 - ▶ un échantillon de la table des suffixes.

- ▶ introduite en 1994 ;
- ▶ sert dans la compression de données (bzip2).

$$T = \begin{array}{cccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ T = & C & G & A & G & A & C & G & A & A & \$ \end{array}$$

$$T = \begin{array}{cccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{array}$$

Permutations circulaires

$$T = \begin{array}{cccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{array}$$

Permutations circulaires

C G A G A C G A A \$

$$T = \begin{array}{cccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{array}$$

Permutations circulaires

CGAGACGAA\$
\$CGAGACGAA

$$T = \begin{array}{cccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{array}$$

Permutations circulaires

CGAGACGAA\$
 \$CGAGACGAA
 A\$CGAGACGA

$$T = \begin{array}{cccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{array}$$

Permutations circulaires

```

CGAGACGAA$
$CGAGACGAA
A$CGAGACGA
AA$CGAGACG
    
```

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{matrix}$$

Permutations circulaires

```

CGAGACGAA$
$CGAGACGAA
A$CGAGACGA
AA$CGAGACG
GAA$CGAGAC
CGAA$CGAGA
ACGAA$CGAG
GACGAA$CGA
AGACGAA$CG
GAGACGAA$C
    
```

$$T = \begin{array}{cccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{array}$$

Permutations circulaires

C G A G A C G A A \$
 \$ C G A G A C G A A
 A \$ C G A G A C G A
 A A \$ C G A G A C G
 G A A \$ C G A G A C
 C G A A \$ C G A G A
 A C G A A \$ C G A G
 G A C G A A \$ C G A
 A G A C G A A \$ C G
 G A G A C G A A \$ C

Permutations triées

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{matrix}$$

Permutations circulaires

Permutations triées

CGAGACGAA\$

\$CGAGACGAA

A\$CGAGACGA

AA\$CGAGACG

GAA\$CGAGAC

CGAA\$CGAGA

ACGAA\$CGAG

GACGAA\$CGA

AGACGAA\$CG

GAGACGAA\$C

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{matrix}$$

Permutations circulaires

C G A G A C G A A \$
 \$ C G A G A C G A A
 A \$ C G A G A C G A
 A A \$ C G A G A C G
 G A A \$ C G A G A C
 C G A A \$ C G A G A
 A C G A A \$ C G A G
 G A C G A A \$ C G A
 A G A C G A A \$ C G
 G A G A C G A A \$ C

Permutations triées

\$ C G A G A C G A A

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{matrix}$$

Permutations circulaires

CGAGACGAA\$

\$CGAGACGAA

A\$CGAGACGA

AA\$CGAGACG

GAA\$CGAGAC

CGAA\$CGAGA

ACGAA\$CGAG

GACGAA\$CGA

AGACGAA\$CG

GAGACGAA\$C

Permutations triées

\$CGAGACGAA

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{matrix}$$

Permutations circulaires

C G A G A C G A A \$
 \$ C G A G A C G A A
 A \$ C G A G A C G A
 A A \$ C G A G A C G
 G A A \$ C G A G A C
 C G A A \$ C G A G A
 A C G A A \$ C G A G
 G A C G A A \$ C G A
 A G A C G A A \$ C G
 G A G A C G A A \$ C

Permutations triées

\$ C G A G A C G A A
 A \$ C G A G A C G A

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ T = & C & G & A & G & A & C & G & A & A & \$ \end{matrix}$$

Permutations circulaires

CGAGACGAA\$

\$CGAGACGAA

A\$CGAGACGA

AA\$CGAGACG

GAA\$CGAGAC

CGAA\$CGAGA

ACGAA\$CGAG

GACGAA\$CGA

AGACGAA\$CG

GAGACGAA\$C

Permutations triées

\$CGAGACGAA

A\$CGAGACGA

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{matrix}$$

Permutations circulaires

C G A G A C G A A \$
 \$ C G A G A C G A A
 A \$ C G A G A C G A
 A A \$ C G A G A C G
 G A A \$ C G A G A C
 C G A A \$ C G A G A
 A C G A A \$ C G A G
 G A C G A A \$ C G A
 A G A C G A A \$ C G
 G A G A C G A A \$ C

Permutations triées

\$ C G A G A C G A A
 A \$ C G A G A C G A
 A A \$ C G A G A C G

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ T = & C & G & A & G & A & C & G & A & A & \$ \end{matrix}$$

Permutations circulaires

C G A G A C G A A \$
 \$ C G A G A C G A A
 A \$ C G A G A C G A
 A A \$ C G A G A C G
 G A A \$ C G A G A C
 C G A A \$ C G A G A
A C G A A \$ C G A G
 G A C G A A \$ C G A
 A G A C G A A \$ C G
 G A G A C G A A \$ C

Permutations triées

\$ C G A G A C G A A
 A \$ C G A G A C G A
 A A \$ C G A G A C G

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ T = & C & G & A & G & A & C & G & A & A & \$ \end{matrix}$$

Permutations circulaires

C G A G A C G A A \$
 \$ C G A G A C G A A
 A \$ C G A G A C G A
 A A \$ C G A G A C G
 G A A \$ C G A G A C
 C G A A \$ C G A G A
 A C G A A \$ C G A G
 G A C G A A \$ C G A
 A G A C G A A \$ C G
 G A G A C G A A \$ C

Permutations triées

\$ C G A G A C G A A
 A \$ C G A G A C G A
 A A \$ C G A G A C G
 A C G A A \$ C G A G

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ T = & C & G & A & G & A & C & G & A & A & \$ \end{matrix}$$

Permutations circulaires

CGAGACGAA\$
 \$CGAGACGAA
 A\$CGAGACGA
 AA\$CGAGACG
 GAA\$CGAGAC
 CGAA\$CGAGA
 ACGAA\$CGAG
 GACGAA\$CGA
 AGACGAA\$CG
 GAGACGAA\$C

Permutations triées

\$CGAGACGAA
 A\$CGAGACGA
 AA\$CGAGACG
 ACGAA\$CGAG
 AGACGAA\$CG
 CGAA\$CGAGA
 CGAGACGAA\$
 GAA\$CGAGAC
 GACGAA\$CGA
 GAGACGAA\$C

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{matrix}$$

Permutations circulaires

CGAGACGAA\$
 \$CGAGACGAA
 A\$CGAGACGA
 AA\$CGAGACG
 GAAS\$CGAGAC
 CGAA\$CGAGA
 ACGAA\$CGAG
 GACGAA\$CGA
 AGACGAA\$CG
 GAGACGAA\$C

Permutations triées

\$CGAGACGAA
 A\$CGAGACGA
 AA\$CGAGACG
 ACGAA\$CGAG
 AGACGAA\$CG
 CGAA\$CGAGA
 CGAGACGAA\$
 GAA\$CGAGAC
 GACGAA\$CGA
 GAGACGAA\$C



$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{matrix}$$

Permutations circulaires

CGAGACGAA\$
 \$CGAGACGAA
 A\$CGAGACGA
 AA\$CGAGACG
 GAAS\$CGAGAC
 CGAA\$CGAGA
 ACGAA\$CGAG
 GACGAA\$CGA
 AGACGAA\$CG
 GAGACGAA\$C

Permutations triées

\$CGAGACGAA
 A\$CGAGACGA
 AA\$CGAGACG
 ACGAA\$CGAG
 AGACGAA\$CG
 CGAA\$CGAGA
 CGAGACGAA\$
 GAA\$CGAGAC
 GACGAA\$CGA
 GAGACGAA\$C

$$TBW(T) = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & A & A & G & G & G & A & \$ & C & A & C \end{matrix}$$

$$T = \begin{array}{cccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{array}$$

F											L
\$	C	G	A	G	A	C	G	A	A		A
A	\$	C	G	A	G	A	C	G	A		A
A	A	\$	C	G	A	G	A	C	G		G
A	C	G	A	A	\$	C	G	A	G		G
A	G	A	C	G	A	A	\$	C	G		G
C	G	A	A	\$	C	G	A	G	A		A
C	G	A	G	A	C	G	A	A	\$		
G	A	A	\$	C	G	A	G	A	C		
G	A	C	G	A	A	\$	C	G	A		
G	A	G	A	C	G	A	A	\$	C		

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \end{matrix}$$

F	L
\$	A
A	A
A	G
A	G
A	G
C	A
C	\$
G	C
G	A
G	C

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix}_1$$

F	L
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix}_1$$

F	L
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1

Propriété

Les lettres identiques sont dans le même ordre dans F et dans L .

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & _1 \end{matrix}$$

F	L
$\$1$	$A4$
$A4$	$A3$
$A3$	$G3$
$A2$	$G2$
$A1$	$G1$
$C2$	$A2$
$C1$	$\$1$
$G3$	$C2$
$G2$	$A1$
$G1$	$C1$

Propriété

Les lettres identiques sont dans le même ordre dans F et dans L .

Fonction LF

Passage d'une permutation circulaire à la précédente.

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix}_1$$

F	L
$\$1$	$A4$
$A4$	$A3$
$A3$ ←	$G3$
$A2$	$G2$
$A1$	$G1$
$C2$	$A2$
$C1$	$\$1$
$G3$	$C2$
$G2$	$A1$
$G1$	$C1$

Propriété

Les lettres identiques sont dans le même ordre dans F et dans L .

Fonction LF

Passage d'une permutation circulaire à la précédente.

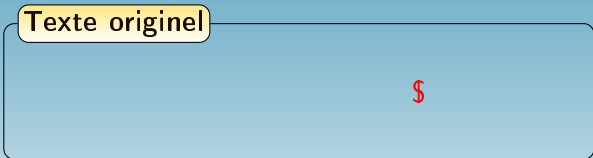
$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1

Texte original

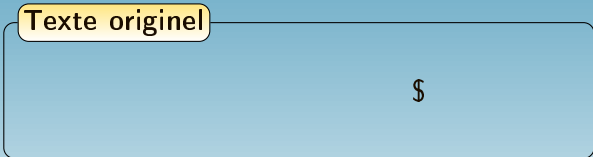
$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1



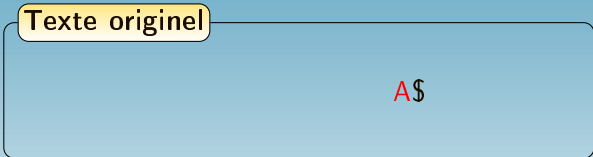
$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1



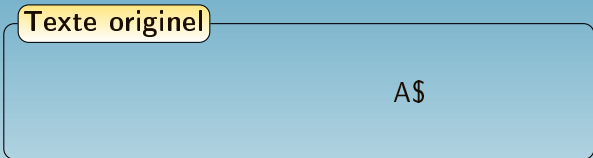
$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \\ & 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 4 & 1 \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1



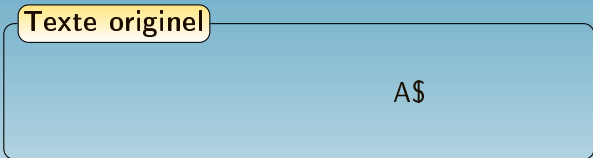
$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>		<i>L</i>
$\$1$	$\swarrow LF$	A_4
A_4		A_3
A_3		G_3
A_2		G_2
A_1		G_1
C_2		A_2
C_1		$\$1$
G_3		C_2
G_2		A_1
G_1		C_1



$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
$A_4 \longrightarrow$	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1



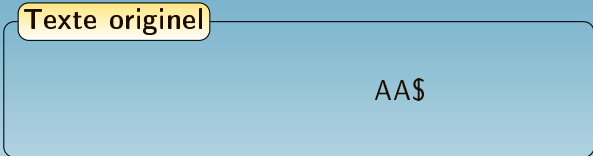
$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \\ & \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1



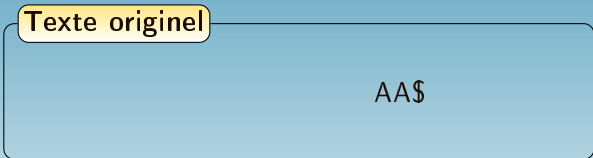
$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \\ & 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 4 & 1 \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1



$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \\ & 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 4 & 1 \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
$A_3 \rightarrow$	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1



$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1



$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix}_1$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1

LF



$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
$G_3 \rightarrow$	C_2
G_2	A_1
G_1	C_1



$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \\ & \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1



$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1



$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & C & G & A & G & A & C & G & A & A & \$ \\ & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
$C_2 \longrightarrow$	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1



$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1



$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
$\$1$	A_4
A_4	A_3
A_3	G_3
A_2	G_2
A_1	G_1
C_2	A_2
C_1	$\$1$
G_3	C_2
G_2	A_1
G_1	C_1

Texte original

$T = CGAGACGAA\$$

Utilisation de la fonction LF

- ▶ parcours du motif P de droite à gauche ;
- ▶ à la i^{e} itération, l'intervalle contenant toutes les occurrences du suffixe de P de longueur i est connu.

Utilisation de la fonction LF

- ▶ parcours du motif P de droite à gauche ;
- ▶ à la i^e itération, l'intervalle contenant toutes les occurrences du suffixe de P de longueur i est connu.

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

Recherchons $P = \overset{0\ 1\ 2}{A\ G\ A}$ dans $T = \overset{0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9}{C\ G\ A\ G\ A\ C\ G\ A\ A\ \$}$

F	L
\$	C G A G A C G A A
A	\$ C G A G A C G A
A	A \$ C G A G A C G
A	C G A A \$ C G A G
A	G A C G A A \$ C G
C	G A A \$ C G A G A
C	G A G A C G A A \$
G	A A \$ C G A G A C
G	A C G A A \$ C G A
G	A G A C G A A \$ C

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

F		L
\$	CGAGACGA	A
A	\$CGAGACGA	A
AA	\$CGAGACG	G
ACGA	\$CGAG	G
AGACGA	\$CG	G
CGAA	\$CGAGA	A
CGAGACGA	A	\$
GAA	\$CGAGAC	C
GACGA	\$CGA	A
GAGACGA	A	\$C

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

F L

\$	C	G	A	G	A	C	G	A	A
A	\$	C	G	A	G	A	C	G	A
A	A	\$	C	G	A	G	A	C	G
A	C	G	A	\$	C	G	A	G	G
A	G	A	C	G	A	\$	C	G	G
C	G	A	A	\$	C	G	A	G	A
C	G	A	G	A	C	G	A	A	\$
G	A	A	\$	C	G	A	G	A	C
G	A	C	G	A	\$	C	G	A	A
G	A	G	A	C	G	A	A	\$	C

Recherchons $P = A^0 G^1 A^2$ dans $T = C^0 G^1 A^2 G^3 A^4 C^5 G^6 A^7 A^8 \9

F										
\$	C	G	A	G	A	C	G	A	A	\$
A	\$	C	G	A	G	A	C	G	A	\$
A	A	\$	C	G	A	G	A	C	G	\$
A	C	G	A	A	\$	C	G	A	G	\$
A	G	A	C	G	A	A	\$	C	G	\$
C	G	A	A	\$	C	G	A	G	A	\$
C	G	A	G	A	C	G	A	A	\$	
G	A	A	\$	C	G	A	G	A	C	
G	A	C	G	A	A	\$	C	G	A	
G	A	G	A	C	G	A	A	\$	C	

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

F L

\$	C	G	A	G	A	C	G	A	A	\$	↑ Zéro G
A	\$	C	G	A	G	A	C	G	A	A	} Trois G
A	A	\$	C	G	A	G	A	C	G	A	
A	C	G	A	\$	C	G	A	G	A	A	
A	G	A	C	G	A	\$	C	G	A	A	
C	G	A	\$	C	G	A	G	A	A	\$	
C	G	A	G	A	C	G	A	A	\$	\$	
G	A	A	\$	C	G	A	G	A	C	C	
G	A	C	G	A	\$	C	G	A	A	C	
G	A	G	A	C	G	A	\$	C	A	C	

Recherchons $P = A \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = C \overset{0}{G} \overset{1}{A} \overset{2}{G} \overset{3}{A} \overset{4}{C} \overset{5}{G} \overset{6}{A} \overset{7}{A} \overset{8}{C} \overset{9}{G}$

F	L																														
\$	CGAGACGA A																														
A	\$CGAGACGA																														
A	A\$CGAGACG																														
A	CGAA\$CGAG																														
A	GACGAAS\$CG																														
C	GAA\$CGAGA																														
Zéro G ↑	C GAGACGAA \$																														
Trois G {	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="padding: 2px;">G</td><td style="padding: 2px;">A</td><td style="padding: 2px;">A</td><td style="padding: 2px;">\$</td><td style="padding: 2px;">C</td><td style="padding: 2px;">G</td><td style="padding: 2px;">A</td><td style="padding: 2px;">G</td><td style="padding: 2px;">A</td><td style="padding: 2px;">C</td></tr> <tr><td style="padding: 2px;">G</td><td style="padding: 2px;">A</td><td style="padding: 2px;">C</td><td style="padding: 2px;">G</td><td style="padding: 2px;">A</td><td style="padding: 2px;">A</td><td style="padding: 2px;">\$</td><td style="padding: 2px;">C</td><td style="padding: 2px;">G</td><td style="padding: 2px;">A</td></tr> <tr><td style="padding: 2px;">G</td><td style="padding: 2px;">A</td><td style="padding: 2px;">G</td><td style="padding: 2px;">A</td><td style="padding: 2px;">C</td><td style="padding: 2px;">G</td><td style="padding: 2px;">A</td><td style="padding: 2px;">A</td><td style="padding: 2px;">\$</td><td style="padding: 2px;">C</td></tr> </table>	G	A	A	\$	C	G	A	G	A	C	G	A	C	G	A	A	\$	C	G	A	G	A	G	A	C	G	A	A	\$	C
G	A	A	\$	C	G	A	G	A	C																						
G	A	C	G	A	A	\$	C	G	A																						
G	A	G	A	C	G	A	A	\$	C																						

Recherchons $P = \overset{0}{A}\overset{1}{G}\overset{2}{A}$ dans $T = \overset{0}{C}\overset{1}{G}\overset{2}{A}\overset{3}{G}\overset{4}{A}\overset{5}{C}\overset{6}{G}\overset{7}{A}\overset{8}{A}\overset{9}{\$}$

$F \quad \uparrow \quad \quad \quad L$

\$	C	G	A	G	A	C	G	A	A	\$
A	\$	C	G	A	G	A	C	G	A	
A	A	\$	C	G	A	G	A	C	G	
A	C	G	A	A	\$	C	G	A	G	
A	G	A	C	G	A	A	\$	C	G	
C	G	A	A	\$	C	G	A	G	A	
C	G	A	G	A	C	G	A	A	\$	
G	A	A	\$	C	G	A	G	A	C	
G	A	C	G	A	A	\$	C	G	A	
G	A	G	A	C	G	A	A	\$	C	

Recherchons $P = \overset{0\ 1\ 2}{A\ G\ A}$ dans $T = \overset{0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9}{C\ G\ A\ G\ A\ C\ G\ A\ A\ \$}$

F		L
\$	C G A G A C G A A	
A	\$ C G A G A C G A	
A	A \$ C G A G A C G	
A	C G A A \$ C G A G	
A	G A C G A A \$ C G	
C	G A A \$ C G A G A	
C	G A G A C G A A	\$↑ Trois A
G	A A \$ C G A G A C	} Un A
G	A C G A A \$ C G A	
G	A G A C G A A \$ C	

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

$F \quad \uparrow \quad L$

	\$	C	G	A	G	A	C	G	A	A	
A	\$	C	G	A	G	A	C	G	A	A	
A	A	\$	C	G	A	G	A	C	G	A	
Trois A	A	C	G	A	A	\$	C	G	A	G	
Un A	{	A	G	A	C	G	A	A	\$	C	G
	C	G	A	A	\$	C	G	A	G	A	
	C	G	A	G	A	C	G	A	A	\$	
	G	A	A	\$	C	G	A	G	A	C	
	G	A	C	G	A	A	\$	C	G	A	
	G	A	G	A	C	G	A	A	\$	C	

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

F L

\$	C	G	A	G	A	C	G	A	A
A	\$	C	G	A	G	A	C	G	A
A	A	\$	C	G	A	G	A	C	G
A	C	G	A	A	\$	C	G	A	G
A	G	A	C	G	A	A	\$	C	G
C	G	A	A	\$	C	G	A	G	A
C	G	A	G	A	C	G	A	A	\$
G	A	A	\$	C	G	A	G	A	C
G	A	C	G	A	A	\$	C	G	A
G	A	G	A	C	G	A	A	\$	C

Il y a une seule occurrence de AGA dans T .

Recherchons $P = \overset{0\ 1\ 2}{A\ G\ A}$ dans $T = \overset{0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9}{C\ G\ A\ G\ A\ C\ G\ A\ A\ \$}$

<i>F</i>	<i>L</i>
\$ C G A G A C G A A	
A \$ C G A G A C G A	
A A \$ C G A G A C G	
A C G A A \$ C G A G	
A G A C G A A \$ C G	
C G A A \$ C G A G A	
C G A G A C G A A \$	
G A A \$ C G A G A C	
G A C G A A \$ C G A	
G A G A C G A A \$ C	

Il y a une seule occurrence de AGA dans T .
 Mais à quelle position ?

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

	<i>F</i>	<i>L</i>	
9	\$	C G A G A C G A A	
8	A	\$ C G A G A C G A	
7	A A	\$ C G A G A C G	
4	A C G A A	\$ C G A G	
2	A G A C G A A	\$ C G	← occurrence
5	C G A A	\$ C G A G A	
0	C G A G A C G A A	\$	
6	G A A	\$ C G A G A C	
3	G A C G A A	\$ C G A	
1	G A G A C G A A	\$ C	

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

	F	L
9	\$ C G A G A C G A A	
8	A \$ C G A G A C G A	
7	A A \$ C G A G A C G	←
4	A C G A A \$ C G A G	
2	A G A C G A A \$ C G	← occurrence
5	C G A A \$ C G A G A	
0	C G A G A C G A A \$	
6	G A A \$ C G A G A C	
3	G A C G A A \$ C G A	
1	G A G A C G A A \$ C	

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

F	L
9	\$ C G A G A C G A A
8	A \$ C G A G A C G A
7	A A \$ C G A G A C G ←
4	A C G A A \$ C G A G
2	A G A C G A A \$ C G ← occurrence
5	C G A A \$ C G A G A
0	C G A G A C G A A \$
6	G A A \$ C G A G A C
3	G A C G A A \$ C G A
1	G A G A C G A A \$ C

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

<i>F</i>	<i>L</i>
9	\$ C G A G A C G A A
8	A \$ C G A G A C G A
7	A A \$ C G A G A C G
4	A C G A A \$ C G A G
2	A G A C G A A \$ C G ← occurrence
5	C G A A \$ C G A G A
0	C G A G A C G A A \$
6	G A A \$ C G A G A C
3	G A C G A A \$ C G A ←
1	G A G A C G A A \$ C

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

<i>F</i>	<i>L</i>
9	\$ C G A G A C G A A
8	A \$ C G A G A C G A
7	A A \$ C G A G A C G
4	A C G A A \$ C G A G
2	A G A C G A A \$ C G ← occurrence
5	C G A A \$ C G A G A
0	C G A G A C G A A \$
6	G A A \$ C G A G A C
3	G A C G A A \$ C G A ←
1	G A G A C G A A \$ C

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

	<i>F</i>	<i>L</i>
9	\$	C G A G A C G A A
8	A	\$ C G A G A C G A
7	A A	\$ C G A G A C G
4	A C G A A	\$ C G A G
2	A G A C G A A	\$ C G ← occurrence
5	C G A A	\$ C G A G A
0	C G A G A C G A A	\$
6	G A A	\$ C G A G A C
3	G A C G A A	\$ C G A
1	G A G A C G A A	\$ C

Table des suffixes ←

← occurrence

Recherchons $P = \overset{0\ 1\ 2}{A\ G\ A}$ dans $T = \overset{0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9}{C\ G\ A\ G\ A\ C\ G\ A\ A\ \$}$

	<i>F</i>	<i>L</i>	
	<hr/>	<hr/>	
	9	\$ C G A G A C G A A	
	8	A \$ C G A G A C G A	
	7	A A \$ C G A G A C G	
	4	A C G A A \$ C G A G	
Table des suffixes ←	2	A G A C G A A \$ C G ← occurrence	
	5	C G A A \$ C G A G A	
↓	0	C G A G A C G A A \$	
Trop d'espace	6	G A A \$ C G A G A C	
	3	G A C G A A \$ C G A	
	1	G A G A C G A A \$ C	

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

	<i>F</i>	<i>L</i>	
	<u> </u>	<u> </u>	
	9	\$ C G A G A C G A A	
	8	A \$ C G A G A C G A	
	7	A A \$ C G A G A C G	
	4	A C G A A \$ C G A G	
Table des suffixes ←	2	A G A C G A A \$ C G ← occurrence	
↓	5	C G A A \$ C G A G A	
Trop d'espace	0	C G A G A C G A A \$	
↓	6	G A A \$ C G A G A C	
Échantillonnage	3	G A C G A A \$ C G A	
	1	G A G A C G A A \$ C	

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

F	L
9	\$ C G A G A C G A A
	A \$ C G A G A C G A
	A A \$ C G A G A C G
	A C G A A \$ C G A G
	A G A C G A A \$ C G ← occurrence
	C G A A \$ C G A G A
0	C G A G A C G A A \$
6	G A A \$ C G A G A C
3	G A C G A A \$ C G A
	G A G A C G A A \$ C

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

	<i>F</i>		<i>L</i>
9	\$	C	A
	A	\$	A
	A	A	G
	A	C	G
	A	G	G ← occurrence
	C	\$	A
0	C	<i>LE</i>	\$
6	G	\$	C
3	G	A	A
	G	A	C

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

	<i>F</i>		<i>L</i>
9	\$	C	A
	A	\$	A
	A	A	G
	A	C	G
	A	G	G
	C	A	G
0	C	\$	\$
6	G	A	C
3	G	A	A
	G	A	C

← occurrence

LF

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

	<i>F</i>	<i>L</i>	
9	\$	C G A G A C G A A	
	A	\$ C G A G A C G A	
	A A	\$ C G A G A C G	
	A C G A A	\$ C G A G	
	A G A C G A A	\$ C G	← occurrence
	C G A A	\$ C G A G A	
0	C	G A G A C G A A	\$
6	G	A A \$ C G A G A C	
3	G	A C G A A \$ C G A	
	G	A G A C G A A \$ C	

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

	<i>F</i>	<i>L</i>	
9	\$	C G A G A C G A A	
	A	\$ C G A G A C G A	
	A A	\$ C G A G A C G	
	A C G A A	\$ C G A G	
	A G A C G A A	\$ C G	← occurrence
	C G A A	\$ C G A G A	
0	C	G A G A C G A A	\$
6	G	A A \$ C G A G A C	
3	G	A C G A A \$ C G A	
1	G	A G A C G A A \$ C	

Recherchons $P = \overset{0}{A} \overset{1}{G} \overset{2}{A}$ dans $T = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{C} \overset{6}{G} \overset{7}{A} \overset{8}{A} \overset{9}{\$}$

	F		L
9	\$	C G A G A C G A A	
	A	\$ C G A G A C G A	
	A	A \$ C G A G A C G	
	A	C G A A \$ C G A G	
2	A	G A C G A A \$ C G	← occurrence
	C	G A A \$ C G A G A	
0	C	G A G A C G A A \$	
6	G	A A \$ C G A G A C	
3	G	A C G A A \$ C G A	
1	G	A G A C G A A \$ C	

Synthèse

- ▶ Permet de compter le nombre d'occurrences d'un motif de longueur m en temps $O(m)$;
- ▶ trouve la position d'une occurrence en temps proportionnel à la distance d'échantillonnage $O(\log^{1+\varepsilon} n)$, $\varepsilon > 0$.

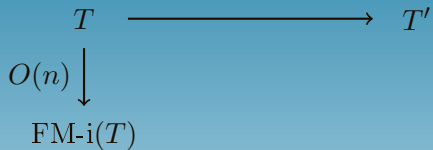
Synthèse

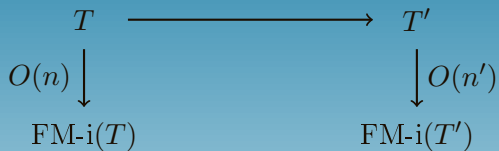
- ▶ Permet de compter le nombre d'occurrences d'un motif de longueur m en temps $O(m)$;
- ▶ trouve la position d'une occurrence en temps proportionnel à la distance d'échantillonnage $O(\log^{1+\varepsilon} n)$, $\varepsilon > 0$.

Limitation

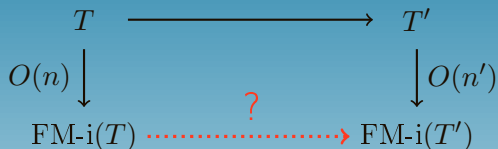
La seule solution existante pour mettre à jour un FM-index ne permet que d'ajouter ou supprimer des textes parmi une collection (Mäkinen et Navarro, 2008).

$$\begin{array}{c}
 T \\
 O(n) \downarrow \\
 \text{FM-i}(T)
 \end{array}$$









Mise à jour du FM-index

- ▶ Mise à jour de la transformée de Burrows-Wheeler.
- ▶ Mise à jour de l'échantillon de la table des suffixes.

$$T = \begin{array}{cccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C_1 & G & A & G & A & C & G & A & A & \$ & C_1 \end{array}$$

<i>F</i>	<i>L</i>
$\$$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A ₂
C ₁	$\$$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁

$$T = \begin{array}{cccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C_1 & G & A & G & A & C & G & A & A & \$ & \end{array} \longrightarrow T' = \begin{array}{cccccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C_1 & G & A & G & A & G & A & G & C & G & A & A & \$ & \end{array}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A ₂
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁

$$T = \begin{array}{cccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C_1 & G & A & G & A & C & G & A & A & \$ & \end{array} \longrightarrow T' = \begin{array}{cccccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C_1 & G & A & G & A & G & C & G & A & A & \$ & \end{array}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A ₂
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).

$$T = \begin{array}{cccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C_1 & G_1 & A_1 & G_2 & A_2 & C_2 & G_3 & A_3 & G_4 & A_4 \\ \hline \end{array} \$ \longrightarrow T' = \begin{array}{ccccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C_1 & G_1 & A_1 & G_2 & A_2 & G_3 & A_3 & G_4 & A_4 & C_1 & \$ \\ \hline \end{array}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A ₂
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).

$$T = \begin{array}{cccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C_1 & G_1 & A_1 & G_2 & A_2 & C_2 & G_3 & A_3 & G_4 & A_4 \\ \hline & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 4 & 1 \end{array} \longrightarrow T' = \begin{array}{ccccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C_1 & G_1 & A_1 & G_2 & A_2 & G_3 & C_2 & G_4 & A_3 & G_4 & A_4 \\ \hline & 1 & 1 & 1 & 2 & 2 & 3 & 3 & 3 & 4 & 1 \end{array}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄ ←
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A ₂
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).

$$T = \begin{array}{cccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{array} \longrightarrow T' = \begin{array}{ccccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{4} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} \end{array}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A ₂
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁

Étapes

1. Ne rien faire ;

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).

$$T = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{matrix} \longrightarrow T' = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{4} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} \end{matrix}$$

<i>F</i>	<i>L</i>
----------	----------

$\$1$	$A4$
$A4$	$A3 \leftarrow$
$A3$	$G3$
$A2$	$G2$
$A1$	$G1$
$C2$	$A2$
$C1$	$\$1$
$G3$	$C2$
$G2$	$A1$
$G1$	$C1$

Étapes

1. Ne rien faire ;

Remarques

- La mise à jour s'effectue de droite à gauche (fonction LF).

$$T = \begin{array}{cccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{array} \longrightarrow T' = \begin{array}{ccccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{4} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{array}$$



<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A ₂
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁

Étapes

1. Ne rien faire ;

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{matrix} \longrightarrow T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{4} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{matrix}$$

<i>F</i>	<i>L</i>
$\$$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃ ←
A ₂	G ₂
A ₁	G ₁
C ₂	A ₂
C ₁	$\$$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁

Étapes

1. Ne rien faire ;

Remarques

- La mise à jour s'effectue de droite à gauche (fonction *LF*).

$$T = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} \end{matrix} \longrightarrow T' = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{4} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} \end{matrix}$$



<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A ₂
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁

Étapes

1. Ne rien faire ;

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).

$$T = \begin{array}{cccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{array} \longrightarrow T' = \begin{array}{ccccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{4} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{array}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A ₂
C ₁	\$ ₁
G ₃	C ₂ ←
G ₂	A ₁
G ₁	C ₁

Étapes

1. Ne rien faire ;

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).

$$T = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ \\ C_1 & G_1 & A_1 & G_2 & A_2 & C_2 & G_3 & A_3 & A_4 & \$_1 \end{matrix} \longrightarrow T' = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ C_1 & G_1 & A_1 & G_2 & A_2 & G_3 & C_2 & G_3 & A_3 & A_4 & \$_1 \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A ₂
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁

Étapes

1. Ne rien faire ;

Remarques

- La mise à jour s'effectue de droite à gauche (fonction LF).

$$T = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} \end{matrix} \longrightarrow T' = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{1} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A ₂ ←
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁

Étapes

1. Ne rien faire ;

Remarques

- La mise à jour s'effectue de droite à gauche (fonction *LF*).

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix} \longrightarrow T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁

Étapes

1. Ne rien faire ;

Remarques

- La mise à jour s'effectue de droite à gauche (fonction *LF*).

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix} \longrightarrow T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁

Étapes

1. Ne rien faire ;
2. substitution ;

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix} \longrightarrow T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁

Étapes

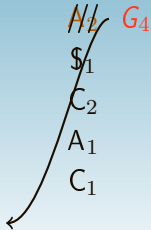
1. Ne rien faire ;
2. substitution ;

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).
- ▶ La position d'insertion est donnée par la fonction *LF*.

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix} \longrightarrow T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \end{matrix}$$

F	L
$\$$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A₂ G ₄
C ₁	$\$$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁



Étapes

1. Ne rien faire ;
2. substitution ;

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction LF).
- ▶ La position d'insertion est donnée par la fonction LF .

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{matrix} \longrightarrow T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{4} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁
G ₄	A ₂

Étapes

1. Ne rien faire ;
2. substitution ;

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).
- ▶ La position d'insertion est donnée par la fonction *LF*.

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix} \longrightarrow T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁
G ₄	A ₂

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).
- ▶ La position d'insertion est donnée par la fonction *LF*.

Étapes

1. Ne rien faire ;
2. substitution ;
3. insertion ;

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix} \longrightarrow T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁
G ₄	A ₂

Étapes

1. Ne rien faire ;
2. substitution ;
3. insertion ;

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).
- ▶ La position d'insertion est donnée par la fonction *LF*.

$$T = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{matrix} \longrightarrow T' = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{4} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁
G ₄	A ₂

Étapes

1. Ne rien faire ;
2. substitution ;
3. insertion ;

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).
- ▶ La position d'insertion est donnée par la fonction *LF*.
- ▶ *LF* donne la nouvelle position de la permutation à déplacer.

$$T = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ \\ \hline C_1 & G_1 & A_1 & G_1 & A_2 & C_2 & G_2 & A_2 & A_3 & \$_1 \end{matrix} \longrightarrow T' = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \hline C_1 & G_1 & A_1 & G_1 & A_2 & G_2 & C_2 & G_2 & A_2 & A_3 & \$_1 \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁
G ₄	A ₂

Étapes

1. Ne rien faire ;
2. substitution ;
3. insertion ;

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).
- ▶ La position d'insertion est donnée par la fonction *LF*.
- ▶ *LF* donne la nouvelle position de la permutation à déplacer.

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix} \longrightarrow T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \end{matrix}$$

F	L
$\$$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₂	G ₂
A ₁	G ₁
C ₂	A₂ G ₄
C ₁	$\$$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁
G ₄	A ₂

Étapes

1. Ne rien faire ;
2. substitution ;
3. insertion ;
4. réordonnement.

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction LF).
- ▶ La position d'insertion est donnée par la fonction LF .
- ▶ LF donne la nouvelle position de la permutation à déplacer.

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix} \longrightarrow T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₁	G ₁
A ₂	G ₂
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁
G ₄	A ₂

Étapes

1. Ne rien faire ;
2. substitution ;
3. insertion ;
4. réordonnement.

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).
- ▶ La position d'insertion est donnée par la fonction *LF*.
- ▶ *LF* donne la nouvelle position de la permutation à déplacer.

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix} \longrightarrow T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₁	G ₁
A ₂	G ₂
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁
G ₄	A ₂

Étapes

1. Ne rien faire ;
2. substitution ;
3. insertion ;
4. réordonnement.

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction LF).
- ▶ La position d'insertion est donnée par la fonction LF .
- ▶ LF donne la nouvelle position de la permutation à déplacer.

$$T = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ \\ \hline C_1 & G_1 & A_1 & G_1 & A_2 & C_2 & G_2 & A_2 & A_3 & \$_1 \end{matrix} \longrightarrow T' = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \hline C_1 & G_1 & A_1 & G_1 & A_2 & G_2 & A_2 & G_3 & A_3 & A_4 & \$_1 \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₁	G ₁
A ₂	G ₂
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₂	A ₁
G ₁	C ₁
G ₄	A ₂

Étapes

1. Ne rien faire ;
2. substitution ;
3. insertion ;
4. réordonnement.

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction LF).
- ▶ La position d'insertion est donnée par la fonction LF .
- ▶ LF donne la nouvelle position de la permutation à déplacer.

$$T = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{4} & \subscript{1} & \end{matrix} \longrightarrow T' = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{4} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₁	G ₁
A ₂	G ₂
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₁	C ₁
G ₂	A ₁
G ₄	A ₂

Étapes

1. Ne rien faire ;
2. substitution ;
3. insertion ;
4. réordonnement.

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction LF).
- ▶ La position d'insertion est donnée par la fonction LF .
- ▶ LF donne la nouvelle position de la permutation à déplacer.

$$T = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ \\ \hline C_1 & G_1 & A_1 & G_1 & A_2 & C_2 & G_2 & A_3 & A_4 & \$_1 \end{matrix} \longrightarrow T' = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \hline C_1 & G_1 & A_1 & G_1 & A_2 & G_4 & C_2 & G_2 & A_3 & A_4 & \$_1 \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₁	G ₁
A ₂	G ₂
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₁	C ₁
G ₂	A ₁
G ₄	A ₂

Étapes

1. Ne rien faire ;
2. substitution ;
3. insertion ;
4. réordonnement.

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction *LF*).
- ▶ La position d'insertion est donnée par la fonction *LF*.
- ▶ *LF* donne la nouvelle position de la permutation à déplacer.

$$T = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \end{matrix} \longrightarrow T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₁	G ₁
A ₂	G ₂
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₁	C ₁
G ₂	A ₁
G ₄	A ₂

Étapes

1. Ne rien faire ;
2. substitution ;
3. insertion ;
4. réordonnement.

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction LF).
- ▶ La position d'insertion est donnée par la fonction LF .
- ▶ LF donne la nouvelle position de la permutation à déplacer.
- ▶ Lorsque la nouvelle position et l'ancienne coïncident, l'algorithme est terminé.

$$T = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ \\ \hline C_1 & G_1 & A_1 & G_1 & A_2 & C_2 & G_2 & A_3 & A_4 & \$_1 \end{matrix} \longrightarrow T' = \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \hline C_1 & G_1 & A_1 & G_1 & A_2 & G_4 & C_2 & G_2 & A_3 & A_4 & \$_1 \end{matrix}$$

<i>F</i>	<i>L</i>
\$ ₁	A ₄
A ₄	A ₃
A ₃	G ₃
A ₁	G ₁
A ₂	G ₂
C ₂	A₂ G ₄
C ₁	\$ ₁
G ₃	C ₂
G ₁	C ₁
G ₂	A ₁
G ₄	A ₂

Étapes

1. Ne rien faire ;
2. substitution ;
3. insertion ;
4. réordonnement.

Remarques

- ▶ La mise à jour s'effectue de droite à gauche (fonction LF).
- ▶ La position d'insertion est donnée par la fonction LF .
- ▶ LF donne la nouvelle position de la permutation à déplacer.
- ▶ Lorsque la nouvelle position et l'ancienne coïncident, l'algorithme est terminé.

Résultat

$$\begin{aligned} \text{TBW}(T) &= A_4 A_3 G_3 G_2 G_1 A_2 \$_1 C_2 A_1 C_1 \\ \text{TBW}(T') &= A_4 A_3 G_3 G_1 G_2 G_4 \$_1 C_2 C_1 A_1 A_2 \end{aligned}$$

Les différentes étapes pour l'insertion d'un facteur S

1. ne rien faire;
2. substituer un élément dans L ;
3. insérer $|S|$ éléments dans L et F , en utilisant la fonction LF ;
4. réordonner un nombre indéterminé de permutations circulaires.



Les différentes étapes pour l'insertion d'un facteur S

1. ne rien faire ;
2. substituer un élément dans L ;
3. insérer $|S|$ éléments dans L et F , en utilisant la fonction LF ;
4. réordonner un nombre indéterminé de permutations circulaires.



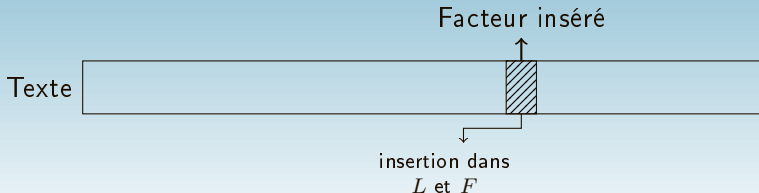
Les différentes étapes pour l'insertion d'un facteur S

1. ne rien faire;
2. substituer un élément dans L ;
3. insérer $|S|$ éléments dans L et F , en utilisant la fonction LF ;
4. réordonner un nombre indéterminé de permutations circulaires.



Les différentes étapes pour l'insertion d'un facteur S

1. ne rien faire;
2. substituer un élément dans L ;
3. insérer $|S|$ éléments dans L et F , en utilisant la fonction LF ;
4. réordonner un nombre indéterminé de permutations circulaires.



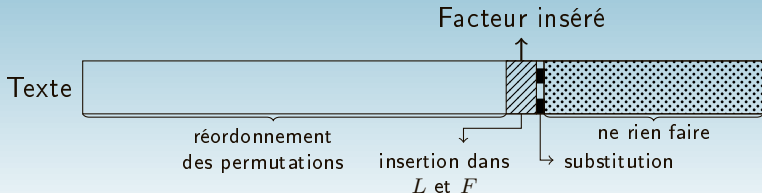
Les différentes étapes pour l'insertion d'un facteur S

1. ne rien faire;
2. substituer un élément dans L ;
3. insérer $|S|$ éléments dans L et F , en utilisant la fonction LF ;
4. réordonner un nombre indéterminé de permutations circulaires.



Les différentes étapes pour l'insertion d'un facteur S

1. ne rien faire ;
2. substituer un élément dans L ;
3. insérer $|S|$ éléments dans L et F , en utilisant la fonction LF ;
4. réordonner un nombre indéterminé de permutations circulaires.



	0	1	2	3	4	5	6	7	8	9	
$T =$	C	G	A	G	A	C	G	A	A	\$	
	0	1	2	3	4	5	6	7	8	9	10
$T' =$	C	G	A	G	A	G	C	G	A	A	\$

$TBW(T) = AAGGGA\$CAC$
 $TBW(T') = AAGGGG\$CCAA$

$$\begin{array}{cccccccccc}
 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
 T = & C & G & A & G & A & C & G & A & A & \$ \\
 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
 T' = & C & G & A & G & A & G & C & G & A & A & \$
 \end{array}$$

$$\begin{array}{l}
 \text{TBW}(T) = \text{AAGGGGA\$CAC} \\
 \text{TBW}(T') = \text{AAGGGG\$CCAA}
 \end{array}$$

$$\begin{array}{l}
 \text{TS}(T) = 9 \ 8 \ 7 \ 4 \ 2 \ 5 \ 0 \ 6 \ 3 \ 1 \\
 \text{TS}(T') = 10 \ 9 \ 8 \ 2 \ 4 \ 6 \ 0 \ 7 \ 1 \ 3 \ 5
 \end{array}$$

Remarque

La transformée de Burrows-Wheeler et la table des suffixes sont conceptuellement très proches.

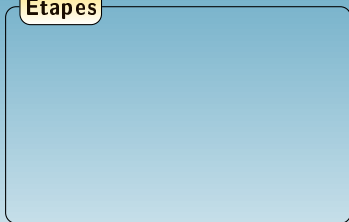
Approche

Nous allons adapter notre algorithme de mise à jour de la transformée de Burrows-Wheeler à la table des suffixes.

$$T = \begin{array}{cccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ C & G & A & G & A & C & G & A & A & \$ & \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{array}$$

TS	F	L
9	$\$$ ₁	A ₄
8	A ₄	A ₃
7	A ₃	G ₃
4	A ₂	G ₂
2	A ₁	G ₁
5	C ₂	A ₂
0	C ₁	$\$$ ₁
6	G ₃	C ₂
3	G ₂	A ₁
1	G ₁	C ₁

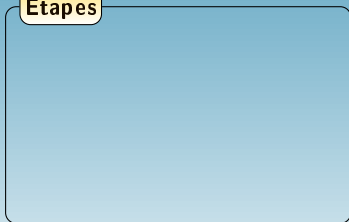
Étapes



$$T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & C & G & A & A & S & & \\ \substack{1 \\ 1} & \substack{1 \\ 1} & \substack{1 \\ 2} & \substack{2 \\ 2} & & & \substack{2 \\ 3} & \substack{3 \\ 3} & \substack{4 \\ 1} & & & \end{matrix}$$

TS	F	L
9	S_1	A_4
8	A_4	A_3
7	A_3	G_3
4	A_2	G_2
2	A_1	G_1
5	C_2	A_2
0	C_1	S_1
6	G_3	C_2
3	G_2	A_1
1	G_1	C_1

Étapes



$$T' = \begin{array}{cccccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & C & G & A & A & \$ & & \\ \hline & 1 & 1 & 1 & 2 & 2 & & 2 & 3 & 3 & 4 & 1 \end{array}$$

TS	F	L
9	$\$$ ₁	A ₄
8	A ₄	A ₃
7	A ₃	G ₃
4	A ₂	G ₂
2	A ₁	G ₁
5	C ₂	A ₂
0	C ₁	$\$$ ₁
6	G ₃	C ₂
3	G ₂	A ₁
1	G ₁	C ₁

Étapes

- ▶ incrémentation ;

$$T' = \begin{array}{cccccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & C & G & A & A & \$ & & \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{array}$$

TS	F	L
10	$\$$ ₁	A ₄
9	A ₄	A ₃
8	A ₃	G ₃
4	A ₂	G ₂
2	A ₁	G ₁
6	C ₂	A ₂
0	C ₁	$\$$ ₁
7	G ₃	C ₂
3	G ₂	A ₁
1	G ₁	C ₁

Étapes

- ▶ incrémentation ;

$$T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \substack{1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 4 \\ 2 \\ 3 \\ 3 \\ 4 \\ 1} \end{matrix}$$

TS	F	L
10	$\$$ ₁	A ₄
9	A ₄	A ₃
8	A ₃	G ₃
4	A ₂	G ₂
2	A ₁	G ₁
6	C ₂	A₂ G ₄
0	C ₁	$\$$ ₁
7	G ₃	C ₂
3	G ₂	A ₁
1	G ₁	C ₁

Étapes

- ▶ incrémentation ;
- ▶ ne rien faire ;

$T' = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{G} \overset{6}{C} \overset{7}{G} \overset{8}{A} \overset{9}{A} \overset{10}{\$}$
 $\quad \quad \quad \underset{1}{1} \underset{1}{1} \underset{1}{1} \underset{2}{2} \underset{2}{2} \underset{4}{4} \underset{2}{2} \underset{3}{3} \underset{3}{3} \underset{4}{4} \underset{1}{1}$

TS	F	L
10	S_1	A_4
9	A_4	A_3
8	A_3	G_3
4	A_2	G_2
2	A_1	G_1
6	C_2	A_2 G_4
0	C_1	S_1
7	G_3	C_2
3	G_2	A_1
1	G_1	C_1

Étapes

- ▶ incrémentation ;
- ▶ ne rien faire ;

$$T' = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{G} \overset{6}{C} \overset{7}{G} \overset{8}{A} \overset{9}{A} \overset{10}{\$}$$

$$T' = \underset{1}{C} \underset{1}{G} \underset{1}{A} \underset{2}{G} \underset{2}{A} \underset{4}{G} \underset{2}{C} \underset{3}{G} \underset{3}{A} \underset{4}{A} \underset{1}{\$}$$

TS	<i>F</i>	<i>L</i>
10	$\$$ ₁	A ₄
9	A ₄	A ₃
8	A ₃	G ₃
4	A ₂	G ₂
2	A ₁	G ₁
6	C ₂	A₂ G ₄
0	C ₁	$\$$ ₁
7	G ₃	C ₂
3	G ₂	A ₁
1	G ₁	C ₁
5	G ₄	A ₂

Étapes

- ▶ incrémentation ;
- ▶ ne rien faire ;
- ▶ insertion ;

$T' = \overset{0}{C} \overset{1}{G} \overset{2}{A} \overset{3}{G} \overset{4}{A} \overset{5}{G} \overset{6}{C} \overset{7}{G} \overset{8}{A} \overset{9}{A} \overset{10}{\$}$
 $\quad \quad \quad \underset{1}{1} \underset{1}{1} \underset{1}{1} \underset{2}{2} \underset{2}{2} \underset{4}{4} \underset{2}{2} \underset{3}{3} \underset{3}{3} \underset{4}{4} \underset{1}{1}$

TS	F	L
10	S_1	A_4
9	A_4	A_3
8	A_3	G_3
4	A_2	G_2
2	A_1	G_1
6	C_2	A_2 G_4
0	C_1	S_1
7	G_3	C_2
3	G_2	A_1
1	G_1	C_1
5	G_4	A_2

Étapes

- ▶ incrémentation ;
- ▶ ne rien faire ;
- ▶ insertion ;

$$T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \substack{1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 4 \\ 2 \\ 3 \\ 3 \\ 4 \\ 1} \end{matrix}$$

TS	F	L
10	$\$$ ₁	A ₄
9	A ₄	A ₃
8	A ₃	G ₃
4	A ₂	G ₂
2	A ₁	G ₁
6	C ₂	A₂ G ₄
0	C ₁	$\$$ ₁
7	G ₃	C ₂
3	G ₂	A ₁
1	G ₁	C ₁
5	G ₄	A ₂

Étapes

- ▶ incrémentation ;
- ▶ ne rien faire ;
- ▶ insertion ;
- ▶ réordonnements.

$$T' = \begin{array}{cccccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{4} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{array}$$

TS	F	L
10	$\$$ ₁	A ₄
9	A ₄	A ₃
8	A ₃	G ₃
2	A ₁	G ₁
4	A ₂	G ₂
6	C ₂	A ₂ G ₄
0	C ₁	$\$$ ₁
7	G ₃	C ₂
3	G ₂	A ₁
1	G ₁	C ₁
5	G ₄	A ₂

Étapes

- ▶ incrémentation ;
- ▶ ne rien faire ;
- ▶ insertion ;
- ▶ réordonnements.

$$T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{4} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} \end{matrix}$$

TS	<i>F</i>	<i>L</i>
10	\$ ₁	A ₄
9	A ₄	A ₃
8	A ₃	G ₃
2	A ₁	G ₁
4	A ₂	G ₂
6	C ₂	A₂ G ₄
0	C ₁	\$ ₁
7	G ₃	C ₂
3	G ₂	A ₁
1	G ₁	C ₁
5	G ₄	A ₂

Étapes

- ▶ incrémentation ;
- ▶ ne rien faire ;
- ▶ insertion ;
- ▶ réordonnements.

$$T' = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{4} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{matrix}$$

TS	F	L
10	$\$1$	$A4$
9	$A4$	$A3$
8	$A3$	$G3$
2	$A1$	$G1$
4	$A2$	$G2$
6	$C2$	$A2$ $G4$
0	$C1$	$\$1$
7	$G3$	$C2$
3	$G2$	$A1$
1	$G1$	$C1$
5	$G4$	$A2$

Étapes

- ▶ incrémentation ;
- ▶ ne rien faire ;
- ▶ insertion ;
- ▶ réordonnements.

$$T' = \begin{array}{cccccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ C & G & A & G & A & G & C & G & A & A & \$ & \\ \subscript{1} & \subscript{1} & \subscript{1} & \subscript{2} & \subscript{2} & \subscript{4} & \subscript{2} & \subscript{3} & \subscript{3} & \subscript{4} & \subscript{1} & \end{array}$$

TS	F	L
10	$\$$ ₁	A ₄
9	A ₄	A ₃
8	A ₃	G ₃
2	A ₁	G ₁
4	A ₂	G ₂
6	C ₂	A ₂ G ₄
0	C ₁	$\$$ ₁
7	G ₃	C ₂
1	G ₁	C ₁
3	G ₂	A ₁
5	G ₄	A ₂

Étapes

- ▶ incrémentation ;
- ▶ ne rien faire ;
- ▶ insertion ;
- ▶ réordonnements.

Question

Combien coûte la mise à jour de la transformée de Burrows-Wheeler pour l'insertion d'un facteur de longueur ℓ en position i dans un texte de longueur n dont l'alphabet est de taille σ ?

Coût des quatre étapes

1. Ne rien faire : $\Theta(0)$.
2. Substituer : $\text{modif}_L(n, \sigma)$.
3. Insérer : $\ell \times (\text{lf}(n', \sigma) + \text{modif}_L(n', \sigma))$.
4. Réordonner : au pire $i \times (\text{lf}(n', \sigma) + \text{modif}_L(n', \sigma))$.

Question

Combien coûte la mise à jour de la transformée de Burrows-Wheeler pour l'insertion d'un facteur de longueur ℓ en position i dans un texte de longueur n dont l'alphabet est de taille σ ?

Coût des quatre étapes

1. Ne rien faire : $\Theta(0)$.
2. Substituer : $modif_L(n, \sigma)$.
3. Insérer : $\ell \times (lf(n', \sigma) + modif_L(n', \sigma))$.
4. Réordonner : au pire $i \times (lf(n', \sigma) + modif_L(n', \sigma))$.

Question

Combien coûte la mise à jour de la transformée de Burrows-Wheeler pour l'insertion d'un facteur de longueur ℓ en position i dans un texte de longueur n dont l'alphabet est de taille σ ?

Coût des quatre étapes

1. Ne rien faire : $\Theta(0)$.
2. Substituer : $modif_L(n, \sigma)$.
3. Insérer : $\ell \times (lf(n', \sigma) + modif_L(n', \sigma))$.
4. Réordonner : au pire $i \times (lf(n', \sigma) + modif_L(n', \sigma))$.

Question

Combien coûte la mise à jour de la transformée de Burrows-Wheeler pour l'insertion d'un facteur de longueur ℓ en position i dans un texte de longueur n dont l'alphabet est de taille σ ?

Coût des quatre étapes

1. Ne rien faire : $\Theta(0)$.
2. Substituer : $\text{modif}_L(n, \sigma)$.
3. Insérer : $\ell \times (\text{lf}(n', \sigma) + \text{modif}_L(n', \sigma))$.
4. Réordonner : au pire $i \times (\text{lf}(n', \sigma) + \text{modif}_L(n', \sigma))$.

Question

Combien coûte la mise à jour de la transformée de Burrows-Wheeler pour l'insertion d'un facteur de longueur ℓ en position i dans un texte de longueur n dont l'alphabet est de taille σ ?

Coût des quatre étapes

1. Ne rien faire : $\Theta(0)$.
2. Substituer : $\text{modif}_L(n, \sigma)$.
3. Insérer : $\ell \times (\text{lf}(n', \sigma) + \text{modif}_L(n', \sigma))$.
4. Réordonner : au pire $i \times (\text{lf}(n', \sigma) + \text{modif}_L(n', \sigma))$.

Question

Combien coûte la mise à jour de la transformée de Burrows-Wheeler pour l'insertion d'un facteur de longueur ℓ en position i dans un texte de longueur n dont l'alphabet est de taille σ ?

Coût des quatre étapes

1. Ne rien faire : $\Theta(0)$.
2. Substituer : $\text{modif}_L(n, \sigma)$.
3. Insérer : $\ell \times (\text{lf}(n', \sigma) + \text{modif}_L(n', \sigma))$.
4. Réordonner : au pire $i \times (\text{lf}(n', \sigma) + \text{modif}_L(n', \sigma))$.

$\text{modif}_L(n, \sigma)$ et $\text{lf}(n, \sigma)$

Stockage compressé de L , modification de L et calcul de LF en temps $O(\log n(1 + \log \sigma / \log \log n))$ (González et Navarro, 2009).

Question

Pourquoi doit-on réordonner des éléments ?

Question

Pourquoi doit-on réordonner des éléments ?

Exemple

ACGAA\$
AGACGAA\$

<i>F</i>		<i>L</i>
	\$	CGAGACGAA
	A	\$CGAGACGA
	AA	\$CGAGACG
	A	CGAA\$CGAG
	A	GACGAA\$CG
	C	GAA\$CGAGA
	C	GAGACGAA\$
	G	AA\$CGAGAC
	G	ACGAA\$CGA
	G	AGACGAA\$C

Question

Pourquoi doit-on réordonner des éléments ?

Exemple

```

A C G A A $
A G A C G A A $
    
```

PLPC

PLPC = A et LPLPC = 1

<i>F</i>	<i>L</i>
\$	C G A G A C G A A
A	\$ C G A G A C G A
A	A \$ C G A G A C G
A	C G A A \$ C G A G
A	G A C G A A \$ C G
C	G A A \$ C G A G A
C	G A G A C G A A \$
G	A A \$ C G A G A C
G	A C G A A \$ C G A
G	A G A C G A A \$ C

Question

Pourquoi doit-on réordonner des éléments ?

Exemple

```

A G C G A A $
A G A G C G A A $
    
```

PLPC

PLPC = A et LPLPC = 1

<i>F</i>	<i>L</i>
\$	C G A G A C G A A
A	\$ C G A G A C G A
A	A \$ C G A G A C G
A	C G A A \$ C G A G
A	G A C G A A \$ C G
C	G A A \$ C G A G A
C	G A G A C G A A \$
G	A A \$ C G A G A C
G	A C G A A \$ C G A
G	A G A C G A A \$ C

Question

Pourquoi doit-on réordonner des éléments ?

Exemple

A G C G A A \$

A G A G C G A A \$

PLPC

PLPC = AG et LPLPC = 2

<i>F</i>	<i>L</i>
\$	C G A G A C G A A
A	\$ C G A G A C G A
A	A \$ C G A G A C G
A	C G A A \$ C G A G
A	G A C G A A \$ C G
C	G A A \$ C G A G A
C	G A G A C G A A \$
G	A A \$ C G A G A C
G	A C G A A \$ C G A
G	A G A C G A A \$ C

Question

Pourquoi doit-on réordonner des éléments ?

Exemple

```

A G C G A A $
A G A G C G A A $
  
```

PLPC

PLPC = AG et LPLPC = 2

<i>F</i>	<i>L</i>
\$	C G A G A C G A A
A	\$ C G A G A C G A
A	A \$ C G A G A C G
A	C G A A \$ C G A G
A	G A C G A A \$ C G
C	G A A \$ C G A G A
C	G A G A C G A A \$
G	A A \$ C G A G A C
G	A C G A A \$ C G A
G	A G A C G A A \$ C

Question

Pourquoi doit-on réordonner des éléments ?

Exemple

G A C G A A \$

G A G A C G A A \$

PLPC

PLPC = GA et LPLPC = 2

<i>F</i>	<i>L</i>
\$	C G A G A C G A A
A	\$ C G A G A C G A
A	A \$ C G A G A C G
A	C G A A \$ C G A G
A	G A C G A A \$ C G
C	G A A \$ C G A G A
C	G A G A C G A A \$
G	A A \$ C G A G A C
G	A C G A A \$ C G A
G	A G A C G A A \$ C

Question

Pourquoi doit-on réordonner des éléments ?

Exemple

```
GA G C G A A $
GA GA G C G A A $
```

PLPC

PLPC = GA et LPLPC = 2

<i>F</i>	<i>L</i>
\$	C G A G A C G A A
A	\$ C G A G A C G A
A	A \$ C G A G A C G
A	C G A A \$ C G A G
A	G A C G A A \$ C G
C	G A A \$ C G A G A
C	G A G A C G A A \$
G	A A \$ C G A G A C
G	A C G A A \$ C G A
G	A G A C G A A \$ C

Question

Pourquoi doit-on réordonner des éléments ?

Exemple

G A **G** C G A A \$

G A G A **G** C G A A \$

PLPC

PLPC = GAG et LPLPC = 3

<i>F</i>	<i>L</i>
\$	C G A G A C G A A
A	\$ C G A G A C G A
A	A \$ C G A G A C G
A	C G A A \$ C G A G
A	G A C G A A \$ C G
C	G A A \$ C G A G A
C	G A G A C G A A \$
G	A A \$ C G A G A C
G	A C G A A \$ C G A
G	A G A C G A A \$ C

Question

Pourquoi doit-on réordonner des éléments ?

Exemple

G A G C G A A \$
 G A G A G C G A A \$

PLPC

PLPC = GAG et LPLPC = 3

<i>F</i>	<i>L</i>
\$	C G A G A C G A A
A	\$ C G A G A C G A
A	A \$ C G A G A C G
A	C G A A \$ C G A G
A	G A C G A A \$ C G
C	G A A \$ C G A G A
C	G A G A C G A A \$
G	A A \$ C G A G A C
G	A C G A A \$ C G A
G	A G A C G A A \$ C

Question

Pourquoi doit-on réordonner des éléments ?

Exemple

G	A	G	C	G	A	A	\$		
G	A	G	A	G	C	G	A		A

PLPC

PLPC = GAG et LPLPC = 3

<i>F</i>		<i>L</i>							
\$	C	G	A	G	A	C	G	A	A
A	\$	C	G	A	G	A	C	G	A
A	A	\$	C	G	A	G	A	C	G
A	C	G	A	A	\$	C	G	A	G
A	G	A	C	G	A	A	\$	C	G
C	G	A	A	\$	C	G	A	G	A
C	G	A	G	A	C	G	A	A	\$
G	A	A	\$	C	G	A	G	A	C
G	A	C	G	A	A	\$	C	G	A
G	A	G	A	C	G	A	A	\$	C

Question

Pourquoi doit-on réordonner des éléments ?

Exemple

A G A C G A A \$
C G A A \$

<i>F</i>	<i>L</i>
\$	C G A G A C G A A
A	\$ C G A G A C G A
A A	\$ C G A G A C G
A C G A A	\$ C G A G
A G A C G A A	\$ C G
C G A A	\$ C G A G A
C G A G A C G A A	\$
G A A	\$ C G A G A C
G A C G A A	\$ C G A
G A G A C G A A	\$ C

Question

Pourquoi doit-on réordonner des éléments ?

Exemple

```

| A G A C G A A $
| C G A A $

```

PLPC

LPLPC = 0

<i>F</i>		<i>L</i>							
\$	C	G	A	G	A	C	G	A	A
A	\$	C	G	A	G	A	C	G	A
A	A	\$	C	G	A	G	A	C	G
A	C	G	A	A	\$	C	G	A	G
A	G	A	C	G	A	A	\$	C	G
C	G	A	A	\$	C	G	A	G	A
C	G	A	G	A	C	G	A	A	\$
G	A	A	\$	C	G	A	G	A	C
G	A	C	G	A	A	\$	C	G	A
G	A	G	A	C	G	A	A	\$	C

Question

Pourquoi doit-on réordonner des éléments ?

Exemple

```

| A G A G C G A A $
| C G A A $
  
```

PLPC

LPLPC = 0

<i>F</i>	<i>L</i>
\$	C G A G A C G A A
A	\$ C G A G A C G A
A	A \$ C G A G A C G
A	C G A A \$ C G A G
A	G A C G A A \$ C G
C	G A A \$ C G A G A
C	G A G A C G A A \$
G	A A \$ C G A G A C
G	A C G A A \$ C G A
G	A G A C G A A \$ C

Question

Pourquoi doit-on réordonner des éléments ?

Question

Peut-on calculer le nombre de réordonnements en moyenne et dans le pire des cas ?

Question

Pourquoi doit-on réordonner des éléments ?

Question

Peut-on calculer le nombre de réordonnements en moyenne et dans le pire des cas ?

Réponse

Si la modification est au-delà de la $LPLPC+1$, il n'y a pas de réordonnements.

⇒ la $LPLPC$ moyenne d'un texte nous donne le nombre moyen de réordonnements.

Question

Combien vaut la LPLPC moyenne d'un texte ?

Question

Combien vaut la LPLPC moyenne d'un texte ?

Réponse

Fayolle et Ward (2005) montrent qu'elle est logarithmique pour des textes issus d'un modèle de Markov d'ordre 1.

Question

Combien vaut la LPLPC moyenne d'un texte ?

Réponse

Fayolle et Ward (2005) montrent qu'elle est logarithmique pour des textes issus d'un modèle de Markov d'ordre 1.

Expérimentations

Calcul sur des textes réels :

- ▶ séquences génomiques les plus répétées (*M. flagellatus*);
- ▶ séquences génomiques d'organismes d'intérêt (homme, souris);
- ▶ textes en langages naturels (ouvrages littéraires, Wikipedia).

Définitions

- L_{max} Valeur maximale parmi les LPLPC.
- L_{moy} Valeur moyenne des LPLPC.
- L_{cent} Dernier centile des LPLPC (valeur maximale parmi les 99 % plus petites).
- R_n L_{moy} divisée par la longueur du texte.

Texte	Longueur	L_{max}	L_{moy}	L_{cent}	R_n
-------	----------	-----------	-----------	------------	-------

Définitions

- L_{max} Valeur maximale parmi les LPLPC.
- L_{moy} Valeur moyenne des LPLPC.
- L_{cent} Dernier centile des LPLPC (valeur maximale parmi les 99 % plus petites).
- R_n L_{moy} divisée par la longueur du texte.

Texte	Longueur	L_{max}	L_{moy}	L_{cent}	R_n
<i>M. flagellatus</i>	2 971 519	143 034	3 452	113 320	$1 \cdot 10^{-3}$

Définitions

- L_{max} Valeur maximale parmi les LPLPC.
- L_{moy} Valeur moyenne des LPLPC.
- L_{cent} Dernier centile des LPLPC (valeur maximale parmi les 99 % plus petites).
- R_n L_{moy} divisée par la longueur du texte.

Texte	Longueur	L_{max}	L_{moy}	L_{cent}	R_n
<i>M. flagellatus</i>	2 971 519	143 034	3 452	113 320	$1 \cdot 10^{-3}$
Homme (chr. 17)	79 601 503	15 692	32	302	$4 \cdot 10^{-7}$

Définitions

- L_{max} Valeur maximale parmi les LPLPC.
- L_{moy} Valeur moyenne des LPLPC.
- L_{cent} Dernier centile des LPLPC (valeur maximale parmi les 99 % plus petites).
- R_n L_{moy} divisée par la longueur du texte.

Texte	Longueur	L_{max}	L_{moy}	L_{cent}	R_n
<i>M. flagellatus</i>	2 971 519	143 034	3 452	113 320	$1 \cdot 10^{-3}$
Homme (chr. 17)	79 601 503	15 692	32	302	$4 \cdot 10^{-7}$
Souris (chr. 1)	191 477 429	7 279	28	345	$1 \cdot 10^{-7}$
Souris (chr. Y)	2 702 555	16 589	192	3 303	$7 \cdot 10^{-5}$

Définitions

- L_{max} Valeur maximale parmi les LPLPC.
- L_{moy} Valeur moyenne des LPLPC.
- L_{cent} Dernier centile des LPLPC (valeur maximale parmi les 99 % plus petites).
- R_n L_{moy} divisée par la longueur du texte.

Texte	Longueur	L_{max}	L_{moy}	L_{cent}	R_n
<i>M. flagellatus</i>	2 971 519	143 034	3 452	113 320	$1 \cdot 10^{-3}$
Homme (chr. 17)	79 601 503	15 692	32	302	$4 \cdot 10^{-7}$
Souris (chr. 1)	191 477 429	7 279	28	345	$1 \cdot 10^{-7}$
Souris (chr. Y)	2 702 555	16 589	192	3 303	$7 \cdot 10^{-5}$
Ouvrages	91 070 340	1 727	12	32	$1 \cdot 10^{-7}$

Définitions

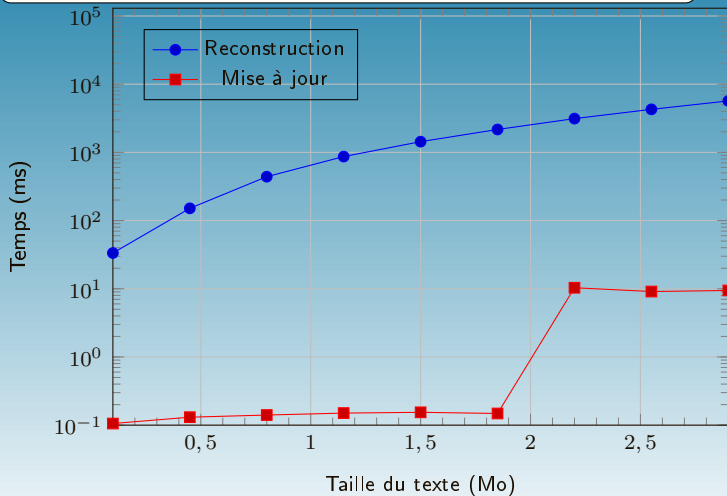
- L_{max} Valeur maximale parmi les LPLPC.
- L_{moy} Valeur moyenne des LPLPC.
- L_{cent} Dernier centile des LPLPC (valeur maximale parmi les 99 % plus petites).
- R_n L_{moy} divisée par la longueur du texte.

Texte	Longueur	L_{max}	L_{moy}	L_{cent}	R_n
<i>M. flagellatus</i>	2 971 519	143 034	3 452	113 320	$1 \cdot 10^{-3}$
Homme (chr. 17)	79 601 503	15 692	32	302	$4 \cdot 10^{-7}$
Souris (chr. 1)	191 477 429	7 279	28	345	$1 \cdot 10^{-7}$
Souris (chr. Y)	2 702 555	16 589	192	3 303	$7 \cdot 10^{-5}$
Ouvrages	91 070 340	1 727	12	32	$1 \cdot 10^{-7}$
Wikipedia (afrikaans)	68 989 658	34 205	66	383	$1 \cdot 10^{-6}$

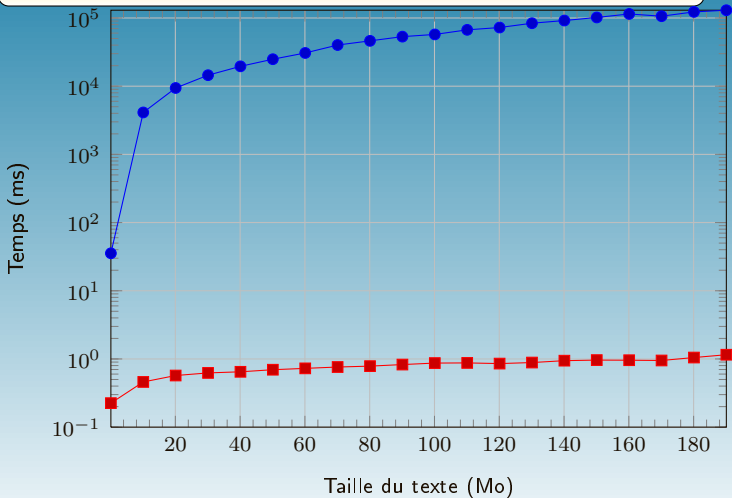
Question

La mise à jour du FM-index d'origine est-elle plus rapide que la construction du FM-index du texte modifié ?

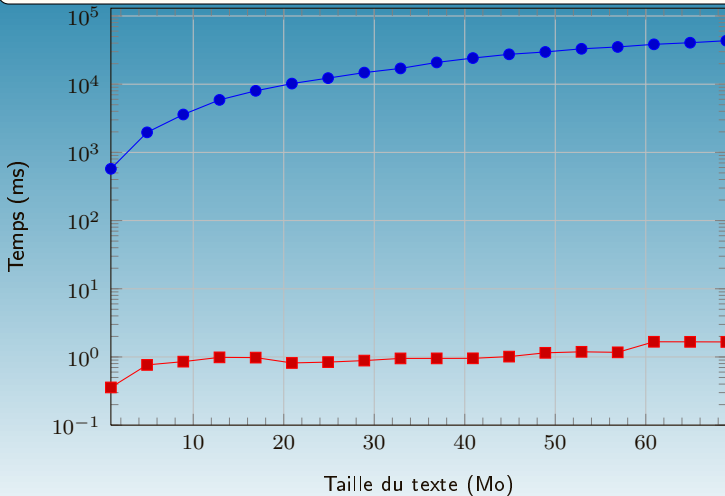
M. flagellatus – insertion d'un facteur de 20 lettres



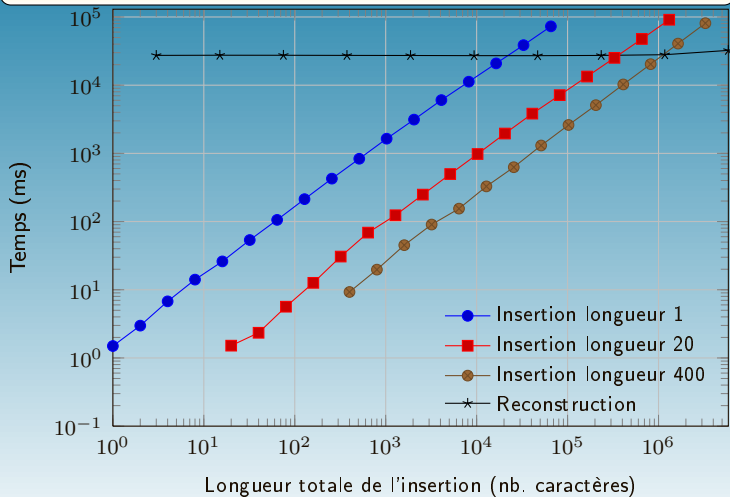
Souris (chr. 1) – insertion d'un facteur de 20 lettres



Wikipedia (afrikaans) – insertion d'un facteur de 20 lettres



Wikipedia, afrikaans (69 Mo) – seuil de reconstruction



Mise à jour d'un FM-index

- ▶ Résultats :
 - ▶ FM-index supportant n'importe quel type de mise à jour ;
 - ▶ algorithme poly-logarithmique en théorie, en moyenne ;
 - ▶ très efficace par rapport à la reconstruction, en pratique.
- ▶ Publications :
 - ▶ *Theoretical Computer Science*, 2009 ;
 - ▶ *Journal of Discrete Algorithms*, 2009 ;
 - ▶ *Journal on Experimental Algorithmics*, en cours de révision.
- ▶ Communications :
 - ▶ *Prague Stringology Conference 2008* ;
 - ▶ *Stringology Research Workshop 2009*, Tel-Aviv, invité ;
 - ▶ *London Stringology Days*, 2009 ;
 - ▶ séminaires (Londres, Caen, Lille, Montpellier).
- ▶ Implantation.

Collaborations

- ▶ LZ factorisation (B. Smyth, A. Al-Hafeedh – Université McMaster, Canada) :
 - ▶ article soumis à la *Data Compression Conference 2010*.
- ▶ Bioinformatique (N. Philippe, É. Rivals – Montpellier) :
 - ▶ localisation des séquences sur le génome de référence en dépit des erreurs de séquence, SNP et autres causes biologiques ;
 - ▶ identification des erreurs de séquences, SNP et autres au sein de chaque séquence.

Structures d'indexation compressées

- ▶ utilisation de notre structure dynamique :
 - ▶ bioinformatique :
 - ▶ *Freie Universität Berlin* (présentation à venir) ;
 - ▶ *Institute for Systems Biology, Seattle* ;
 - ▶ *Center for Bioinformatics and Computational Biology, university of Maryland* ;
 - ▶ traitement automatique des langues (université d'Édimbourg) ;
- ▶ pistes de réflexion :
 - ▶ stockage efficace de l'échantillonnage de la table des suffixes ;
 - ▶ mise à jour efficace de la table PLPC.

Poursuite des collaborations

- ▶ LZ-factorisation :
 - ▶ implantation de la méthode ;
 - ▶ tests exhaustifs à réaliser ;
- ▶ bioinformatique :
 - ▶ rassembler les séquences provenant d'une même région, afin d'améliorer le temps de recherche ;
 - ▶ validation biologique des résultats.

Merci de votre attention.

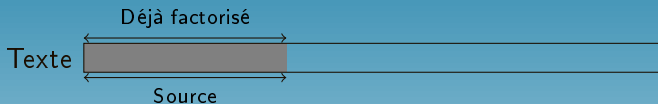
- ▶ LZ factorisation [▶ Go!](#)
- ▶ Bioinformatique [▶ Go!](#)
- ▶ Stockage de TS [▶ Go!](#)
- ▶ Stockage de L [▶ Go!](#)

Problème

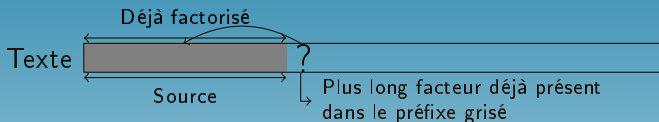
Texte



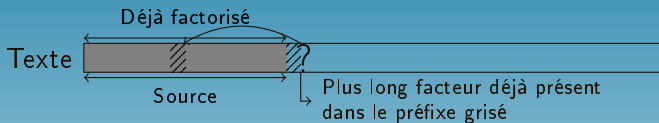
Problème



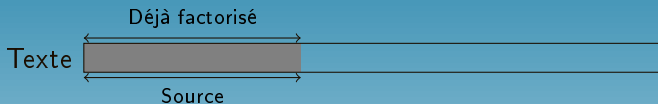
Problème



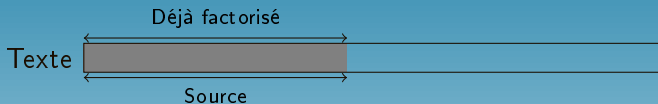
Problème



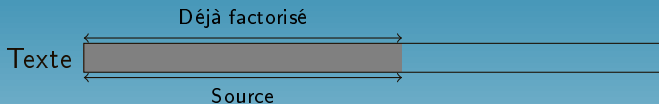
Problème



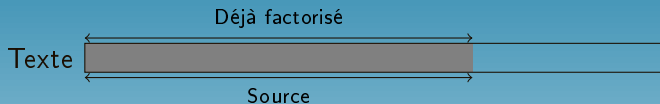
Problème



Problème



Problème



Problème



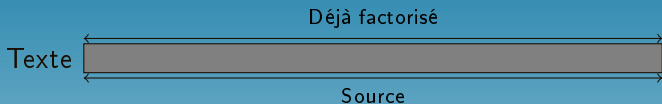
Problème



Problème



Problème

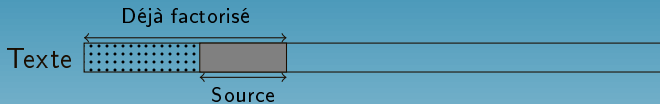


Solution

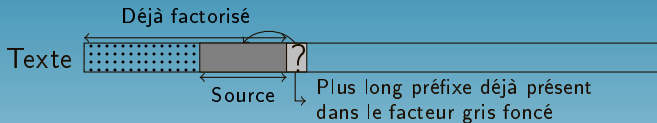
Indexer le préfixe croissant grisé (Okanohara et Sadakane, 2008).

Si on recherche dans un facteur de taille constante

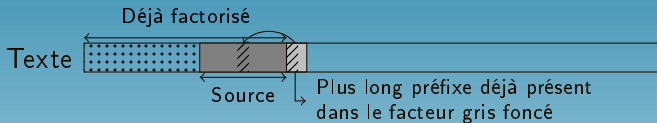
Si on recherche dans un facteur de taille constante



Si on recherche dans un facteur de taille constante



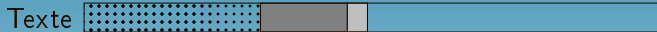
Si on recherche dans un facteur de taille constante



Si on recherche dans un facteur de taille constante



Si on recherche dans un facteur de taille constante



Si on recherche dans un facteur de taille constante



Si on recherche dans un facteur de taille constante



Si on recherche dans un facteur de taille constante



Si on recherche dans un facteur de taille constante



Si on recherche dans un facteur de taille constante



Si on recherche dans un facteur de taille constante



Si on recherche dans un facteur de taille constante

Texte 

Si on recherche dans un facteur de taille constante



Si on recherche dans un facteur de taille constante

Texte 

Solutions actuelles

- ▶ arbre de recherche ;
- ▶ table des « préfixes » incrémentale (Okanohara et Sadakane, 2008) ;
- ▶ reconstruction de la table des suffixes (Ferreira *et al.*, 2009).

Si on recherche dans un facteur de taille constante

Texte 

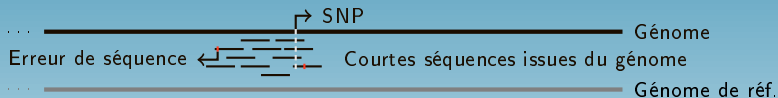
Solutions actuelles

- ▶ arbre de recherche ;
- ▶ table des « préfixes » incrémentale (Okanohara et Sadakane, 2008) ;
- ▶ reconstruction de la table des suffixes (Ferreira *et al.*, 2009).

Idée

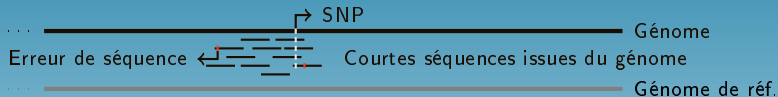
- ▶ Maintenir une liste des modifications à réaliser (suppressions et insertions).
- ▶ Mettre à jour la table des suffixes d'après cette liste des modifications.

Problème



Comment localiser au mieux ces courtes séquences sur le génome de référence ?

Problème

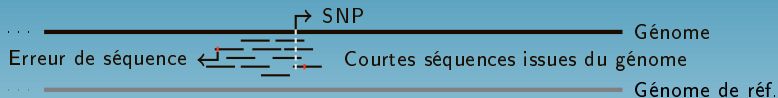


Comment localiser au mieux ces courtes séquences sur le génome de référence ?

Solutions actuelles

- ▶ Bowtie (Langmead *et al.*, 2009) ;
- ▶ BWA (Li et Durbin, 2008) ;
- ▶ BWT-SW (Lam *et al.*, 2009) ;
- ▶ SOAP2 (Li *et al.*, 2009).

Problème

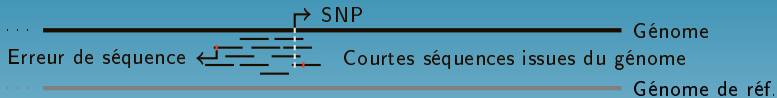


Comment localiser au mieux ces courtes séquences sur le génome de référence ?

Stratégie

Rechercher les facteurs de longueur ℓ de chaque séquence.

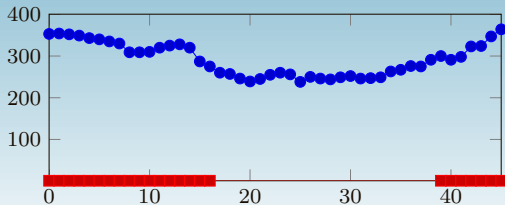
Problème



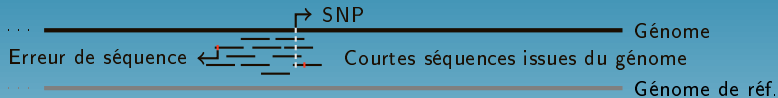
Comment localiser au mieux ces courtes séquences sur le génome de référence ?

Stratégie

Rechercher les facteurs de longueur ℓ de chaque séquence.



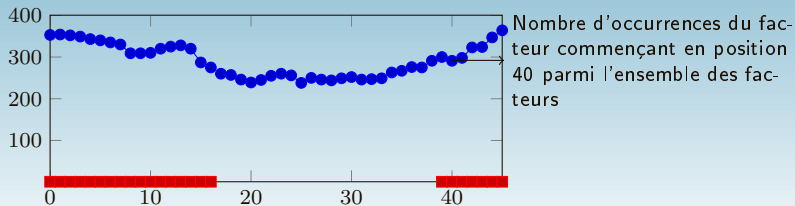
Problème



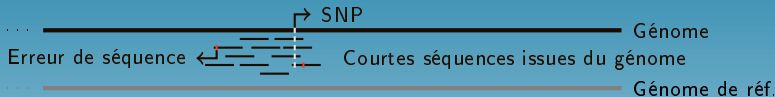
Comment localiser au mieux ces courtes séquences sur le génome de référence ?

Stratégie

Rechercher les facteurs de longueur ℓ de chaque séquence.



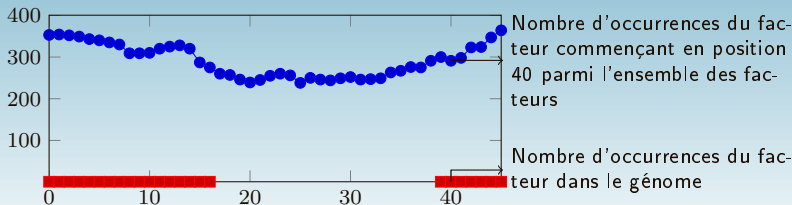
Problème



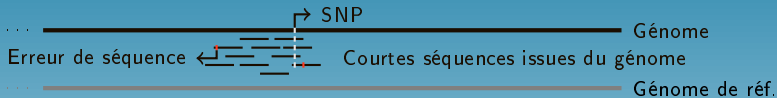
Comment localiser au mieux ces courtes séquences sur le génome de référence ?

Stratégie

Rechercher les facteurs de longueur ℓ de chaque séquence.



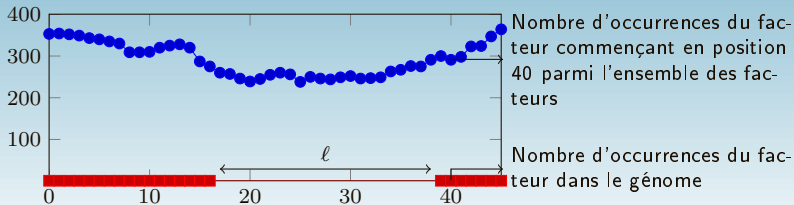
Problème



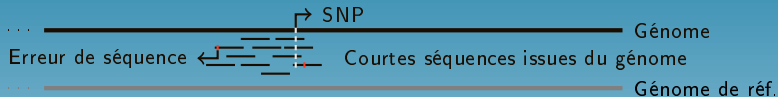
Comment localiser au mieux ces courtes séquences sur le génome de référence ?

Stratégie

Rechercher les facteurs de longueur ℓ de chaque séquence.



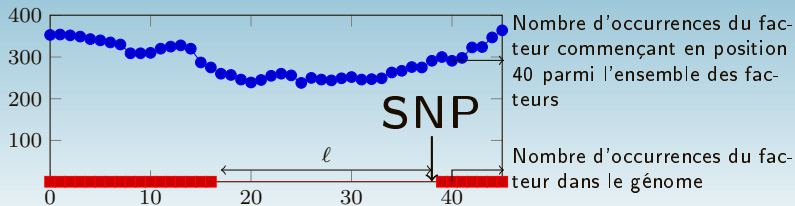
Problème



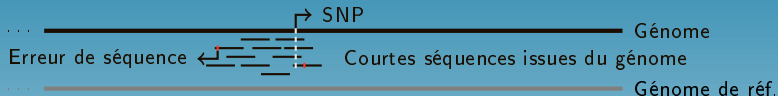
Comment localiser au mieux ces courtes séquences sur le génome de référence ?

Stratégie

Rechercher les facteurs de longueur ℓ de chaque séquence.



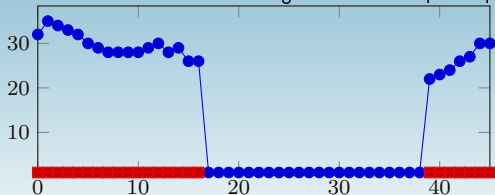
Problème



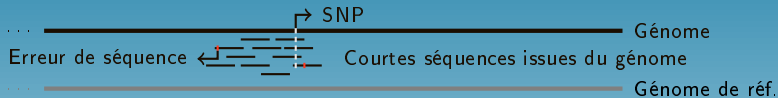
Comment localiser au mieux ces courtes séquences sur le génome de référence ?

Stratégie

Rechercher les facteurs de longueur ℓ de chaque séquence.



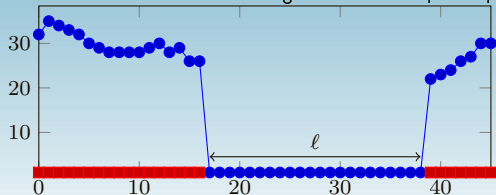
Problème



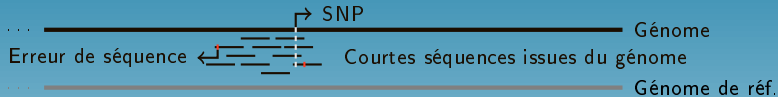
Comment localiser au mieux ces courtes séquences sur le génome de référence ?

Stratégie

Rechercher les facteurs de longueur ℓ de chaque séquence.



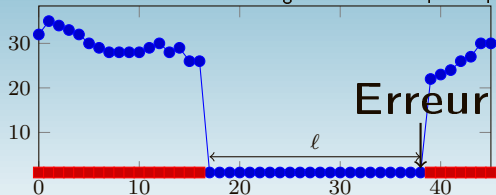
Problème



Comment localiser au mieux ces courtes séquences sur le génome de référence ?

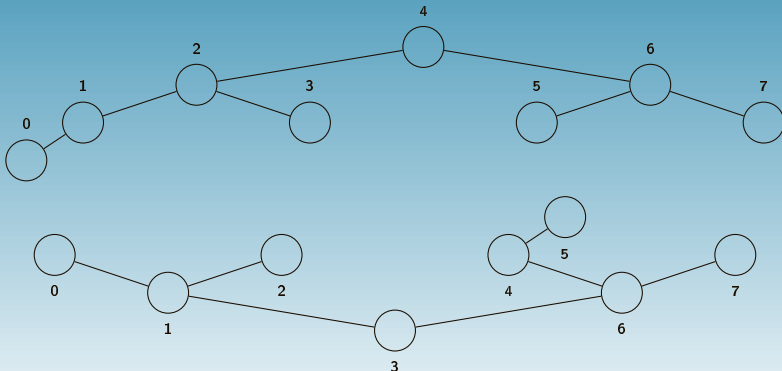
Stratégie

Rechercher les facteurs de longueur ℓ de chaque séquence.



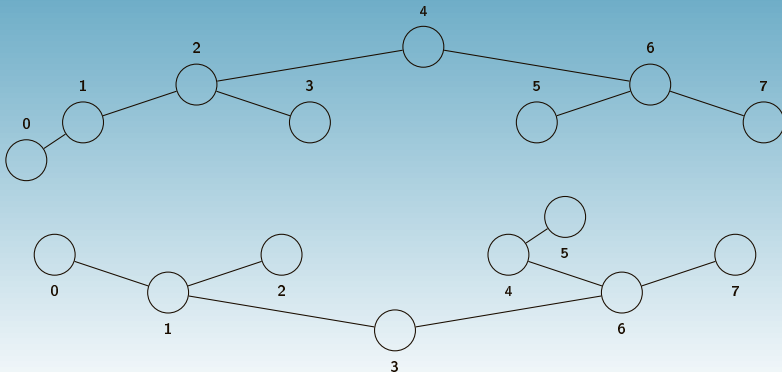
Beaucoup d'incrémentations dans TS.

Beaucoup d'incrémentations dans TS.



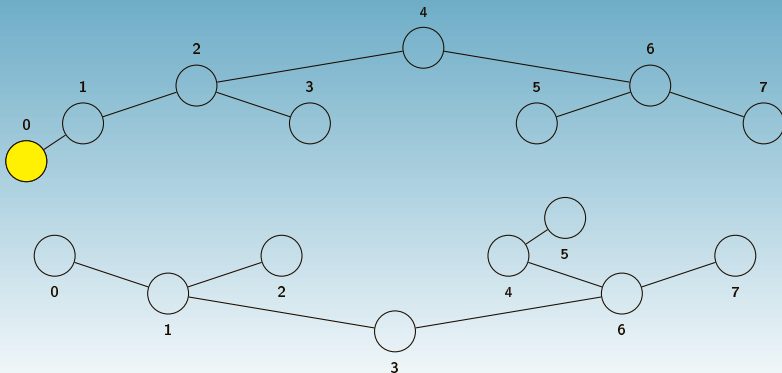
Considérons :

i	0	1	2	3	4	5	6	7
TS	7	4	1	0	5	2	6	3



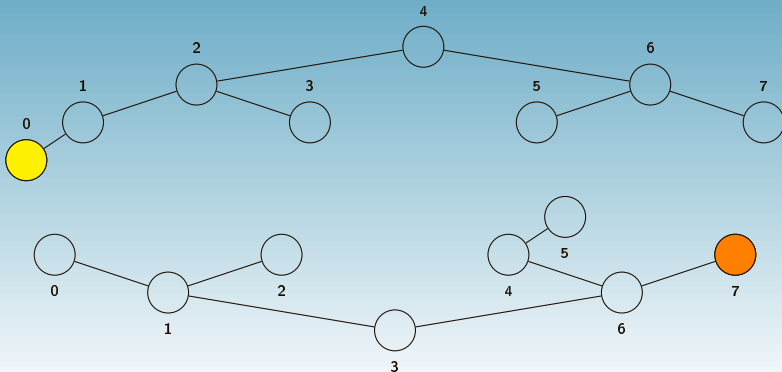
Considérons :

i	0	1	2	3	4	5	6	7
TS	7	4	1	0	5	2	6	3



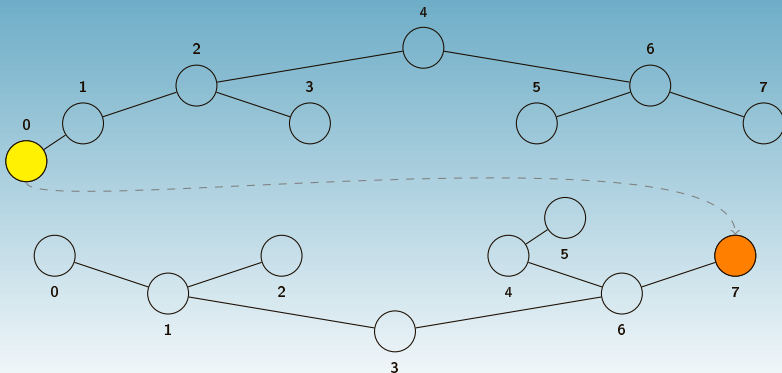
Considérons :

i	0	1	2	3	4	5	6	7
TS	7	4	1	0	5	2	6	3



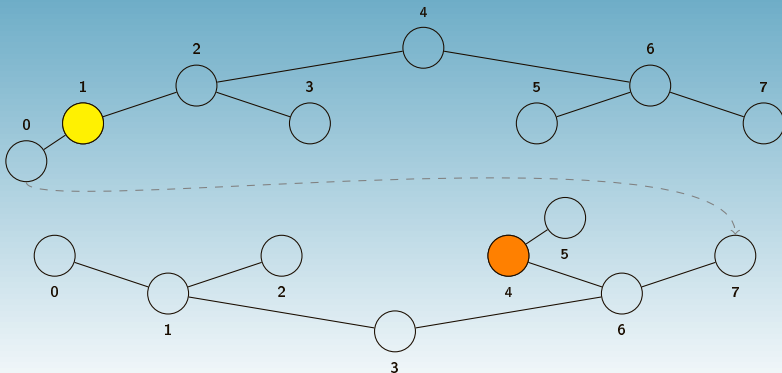
Considérons :

i	0	1	2	3	4	5	6	7
TS	7	4	1	0	5	2	6	3



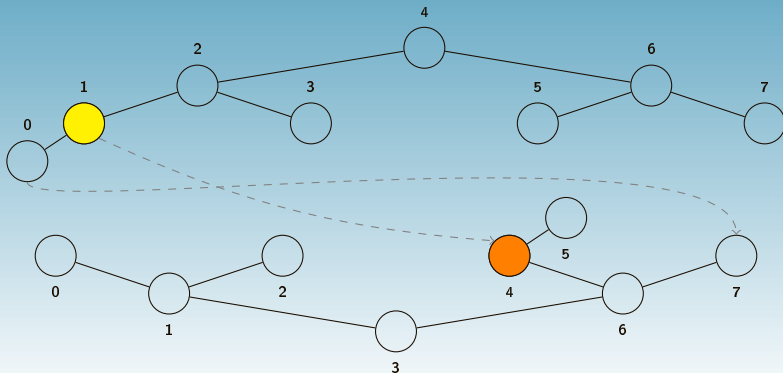
Considérons :

i	0	1	2	3	4	5	6	7
TS	7	4	1	0	5	2	6	3



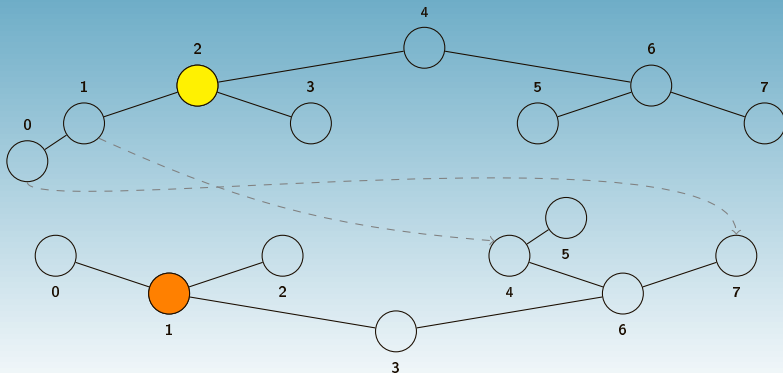
Considérons :

i	0	1	2	3	4	5	6	7
TS	7	4	1	0	5	2	6	3



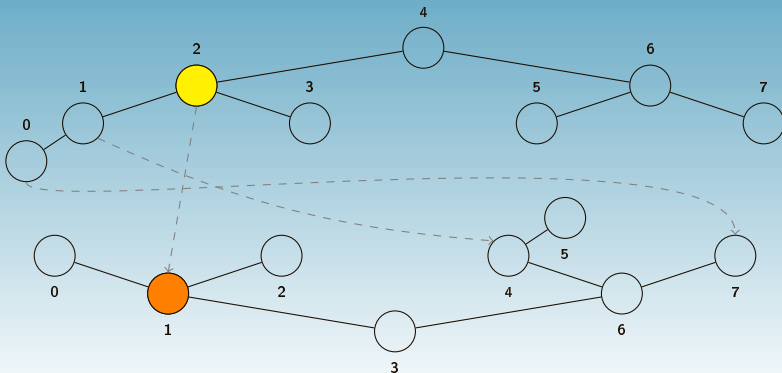
Considérons :

i	0	1	2	3	4	5	6	7
TS	7	4	1	0	5	2	6	3



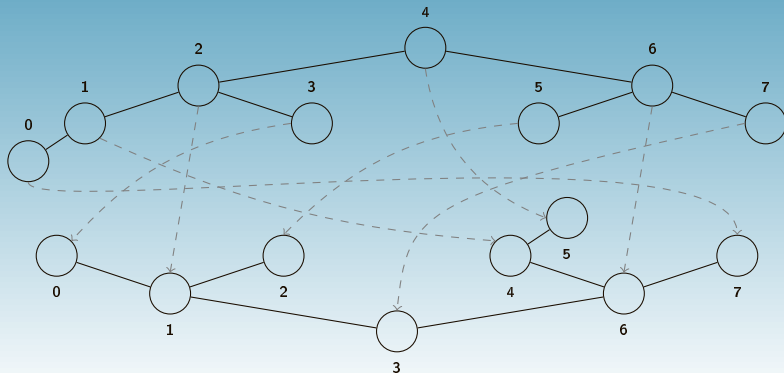
Considérons :

i	0	1	2	3	4	5	6	7
TS	7	4	1	0	5	2	6	3



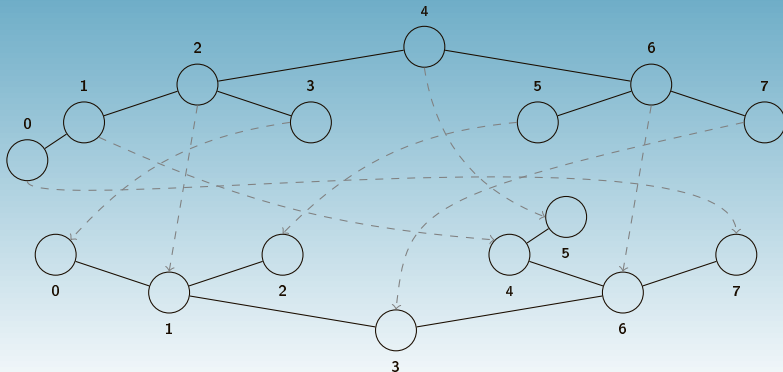
Considérons :

i	0	1	2	3	4	5	6	7
TS	7	4	1	0	5	2	6	3



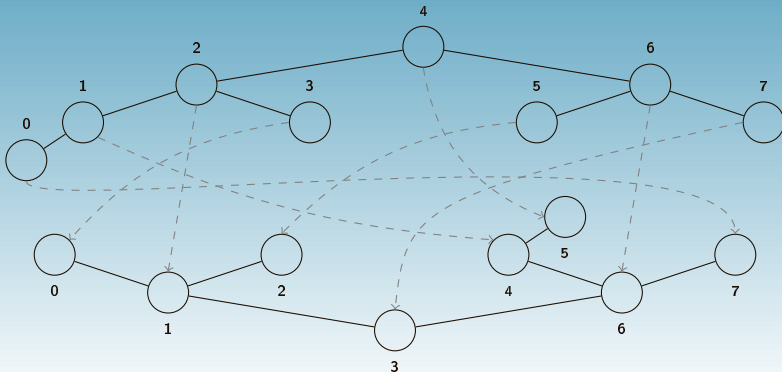
Considérons :

i	0	1	2	3	4	5	6	7	8
TS	7	4	1	3	0	5	2	6	3



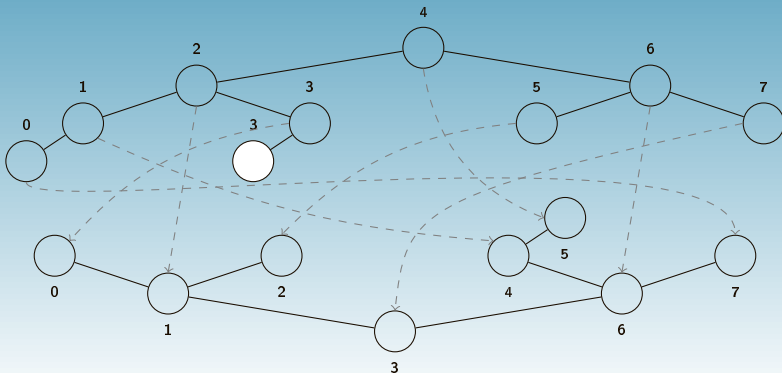
Considérons :

i	0	1	2	3	4	5	6	7	8
TS	8	5	1	3	0	6	2	7	4



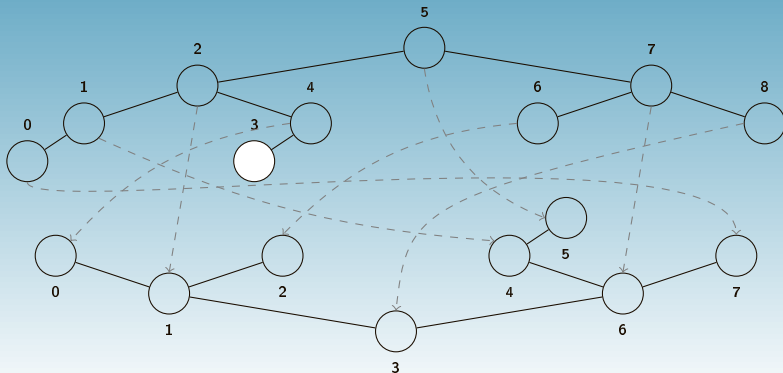
Considérons :

i	0	1	2	3	4	5	6	7	8
TS	8	5	1	3	0	6	2	7	4



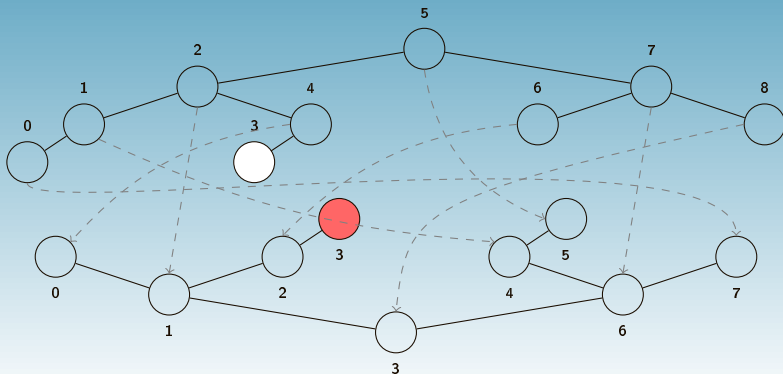
Considérons :

i	0	1	2	3	4	5	6	7	8
TS	8	5	1	3	0	6	2	7	4



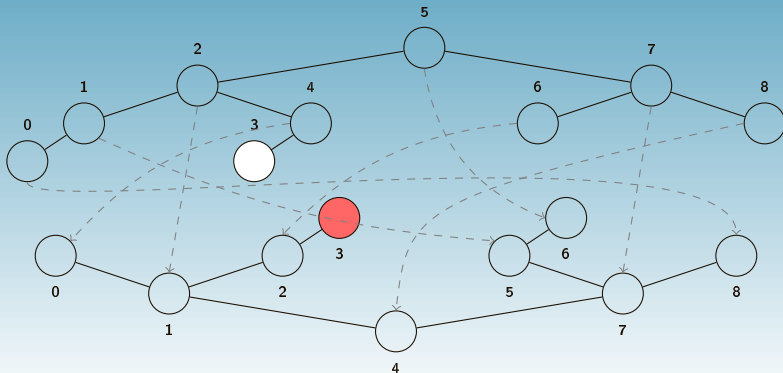
Considérons :

i	0	1	2	3	4	5	6	7	8
TS	8	5	1	3	0	6	2	7	4



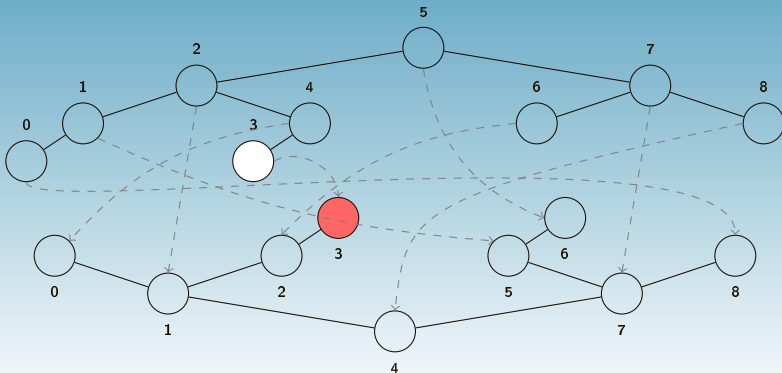
Considérons :

i	0	1	2	3	4	5	6	7	8
TS	8	5	1	3	0	6	2	7	4

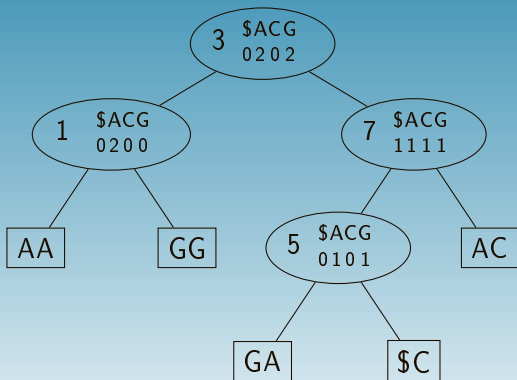


Considérons :

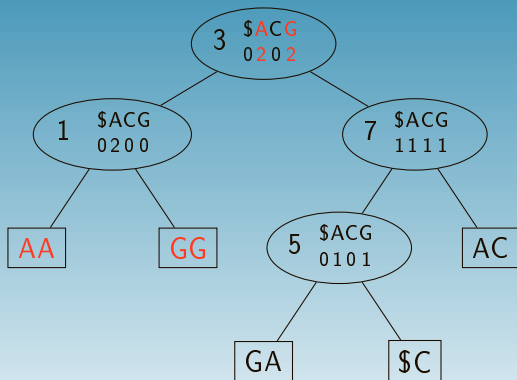
i	0	1	2	3	4	5	6	7	8
TS	8	5	1	3	0	6	2	7	4



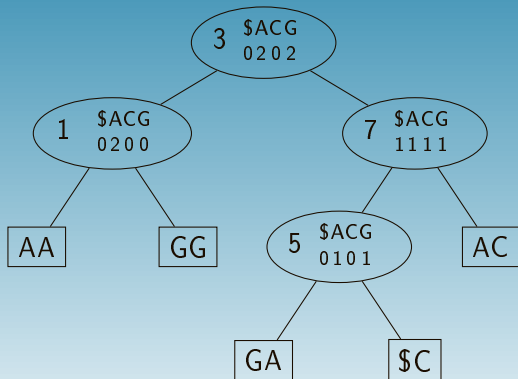
$$L = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & A & A & G & G & G & A & \$ & C & A & C \end{matrix}$$



$L = \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ & A & A & G & G & G & A & \$ & C & A & C \end{matrix}$



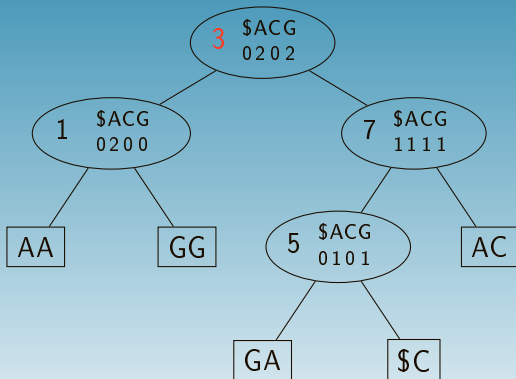
0 1 2 3 4 5 6 7 8 9
 $L = AAGGGA\$CAC$



Comptage

Nombre de A jusqu'à la position 6 :

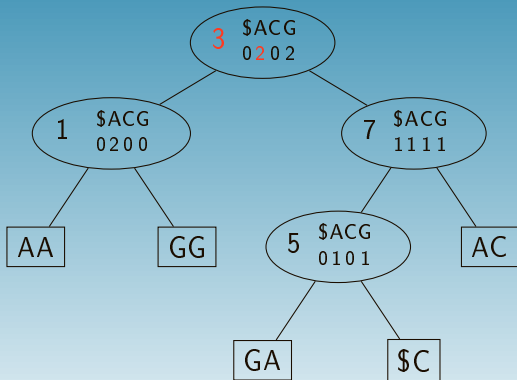
0 1 2 3 4 5 6 7 8 9
 $L = AAGGGA\$CAC$



Comptage

Nombre de A jusqu'à la
 position 6 :
 position 3 < 6

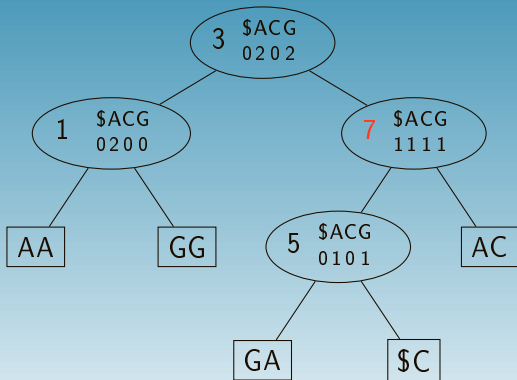
0 1 2 3 4 5 6 7 8 9
 $L = AAGGGA\$CAC$



Comptage

Nombre de A jusqu'à la position 6 :
 position 3 < 6 \Rightarrow 2A

0 1 2 3 4 5 6 7 8 9
 $L = AAGGGA\$CAC$



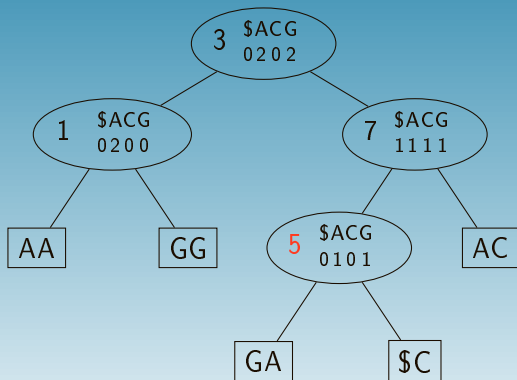
Comptage

Nombre de A jusqu'à la position 6 :

position 3 < 6 ⇒ 2A

position 7 > 6 ⇒ sous-arbre gauche

0 1 2 3 4 5 6 7 8 9
 $L = AAGGGA\$CAC$



Comptage

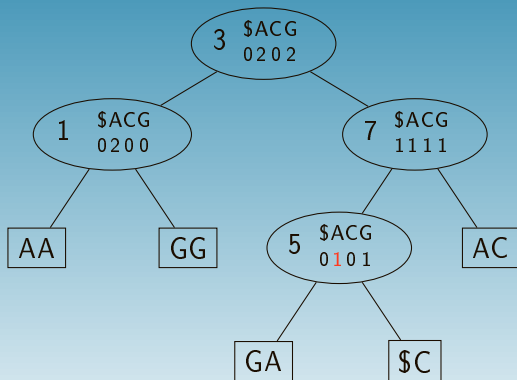
Nombre de A jusqu'à la position 6 :

position 3 < 6 ⇒ 2A

position 7 > 6 ⇒ sous-arbre gauche

position 5 < 6

$L = \overset{0}{A} \overset{1}{A} \overset{2}{G} \overset{3}{G} \overset{4}{G} \overset{5}{G} \overset{6}{\$} \overset{7}{C} \overset{8}{A} \overset{9}{C}$



Comptage

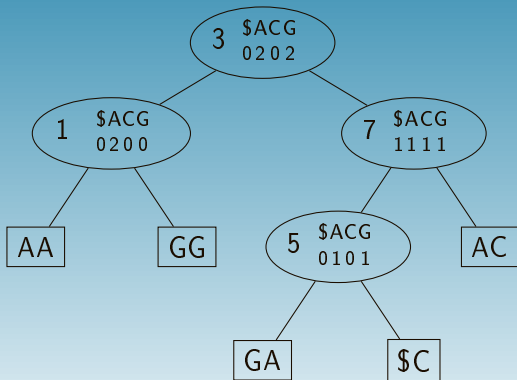
Nombre de A jusqu'à la position 6 :

position 3 < 6 ⇒ 2A

position 7 > 6 ⇒ sous-arbre gauche

position 5 < 6 ⇒ 1A

0 1 2 3 4 5 6 7 8 9
 $L = AAGGGA\$CAC$



Comptage

Nombre de A jusqu'à la position 6 :

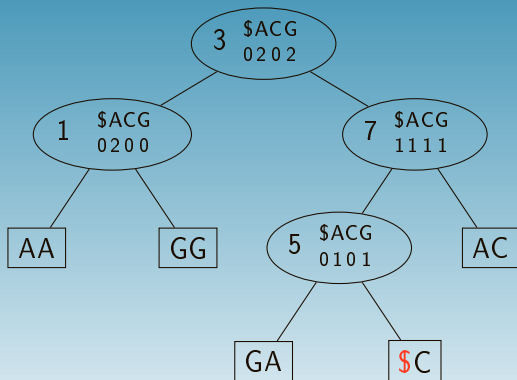
position 3 < 6 ⇒ 2A

position 7 > 6 ⇒ sous-arbre gauche

position 5 < 6 ⇒ 1A

sous-arbre droit ⇒ feuille

0 1 2 3 4 5 6 7 8 9
 $L = AAGGGA\$CAC$



Comptage

Nombre de A jusqu'à la position 6 :

position 3 < 6 ⇒ 2A

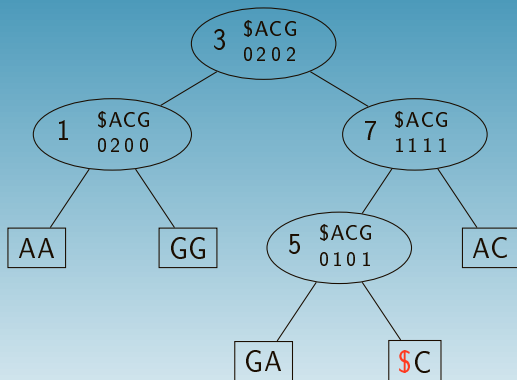
position 7 > 6 ⇒ sous-arbre gauche

position 5 < 6 ⇒ 1A

sous-arbre droit ⇒ feuille

$\$ \neq A$

0 1 2 3 4 5 6 7 8 9
 $L = AAGGGA\$CAC$



Comptage

Nombre de A jusqu'à la position 6 :

position 3 < 6 ⇒ 2A

position 7 > 6 ⇒ sous-arbre gauche

position 5 < 6 ⇒ 1A

sous-arbre droit ⇒ feuille \$ ≠ A

⇒ 3A