

BIOINFORMATICA

TUMMINELLI CHIARA

ESERCIZI

BIOINFORMATICA

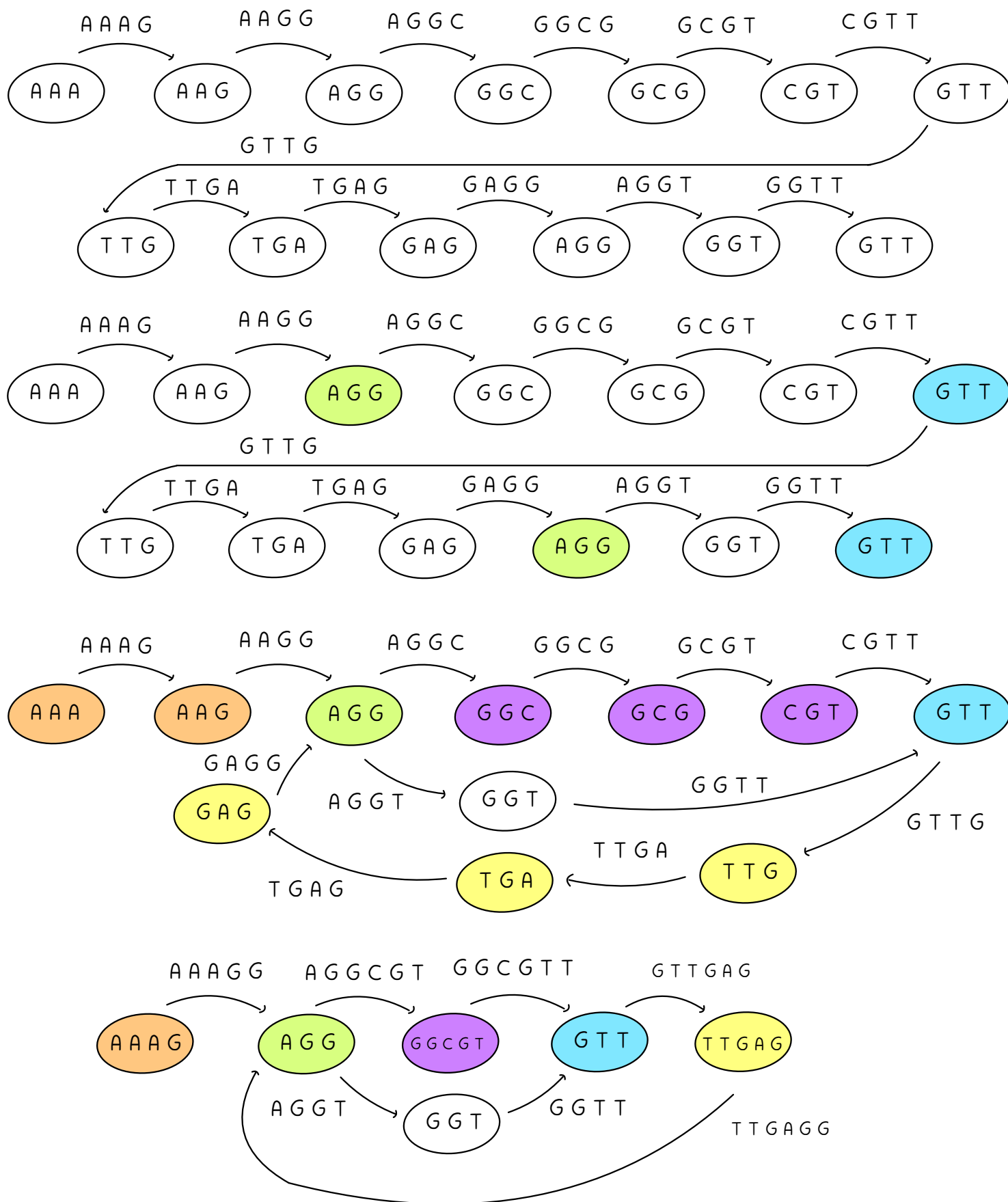
2024-2025



ESERCIZIO 1

1) CREATE DE BRUIJN GRAPH
FROM THE FOLLOWING READ:
AAAGGCGTTGAGGT
WITH $k=4$

2) APPLY A SIMPLIFICATION STEP
ON THE DE BRUIJN GRAPH
OBTAINED



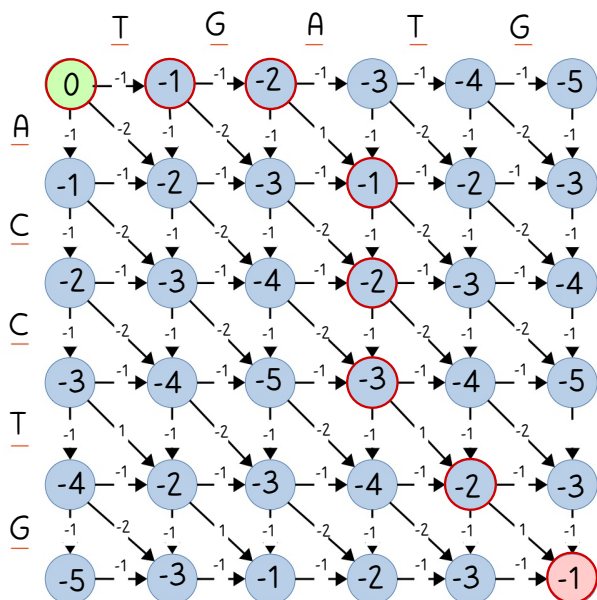
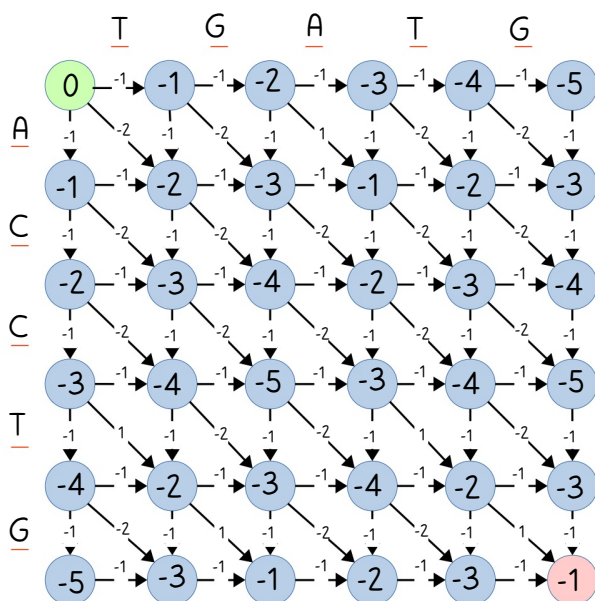
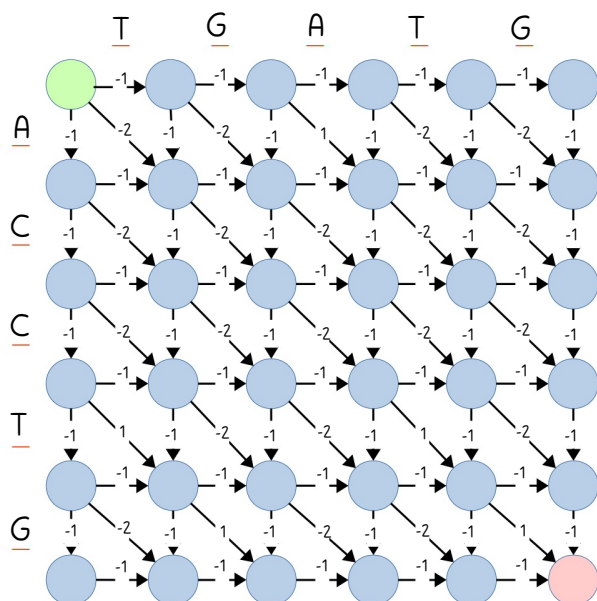
ESERCIZIO 2

Try to align globally the following sequences:

- ACCTG and TGATG;

score matrix

	A	C	G	T	GAP
A	1	-2	-2	-2	-1
C	-2	1	-2	-2	-1
G	-2	-2	1	-2	-1
T	-2	-2	-2	1	-1
GAP	-1	-1	-1	-1	-



RISPOSTA:

- A C C T G
T G A - - T G

PROCEDIMENTO:

PER DETERMINARE I VALORI DEGLI ARCHI, È NECESSARIO ESAMINARE LA MATRICE DI SCORE. GLI ARCHI ORIZZONTALI E VERTICALI SONO TRATTATI COME GAP, MENTRE PER GLI ARCHI DIAGONALI SI VERIFICANO I MATCH O MISMATCH.

PER CALCOLARE I VALORI DEI NODI, SI PARTE DAL PRIMO NODO, CHE HA VALORE 0. SUCCESSIVAMENTE, PER OGNI NODO, SI SOMMANO I VALORI DI OGNI ARCO CHE LO COLLEGA AGLI ALTRI NODI; IL VALORE DEL NODO È IL MASSIMO TRA QUESTI.

IL PERCORSO SI OTTIENE ESEGUENDO IL BACKTRACKING A PARTIRE DALL'ULTIMO NODO, SCEGLIENDO SEMPRE IL NODO CON IL VALORE PIÙ ALTO. PER COSTRUIRE L'ALLINEAMENTO, SE L'ARCO È DIAGONALE, SI SCRIVONO ENTRAMBE LE LETTERE NELLE RISPETTIVE STRINGHE; SE L'ARCO È VERTICALE, SI SCRIVE SOLO NELLA PRIMA STRINGA, MENTRE SE È ORIZZONTALE, SI SCRIVE SOLO NELLA SECONDA.

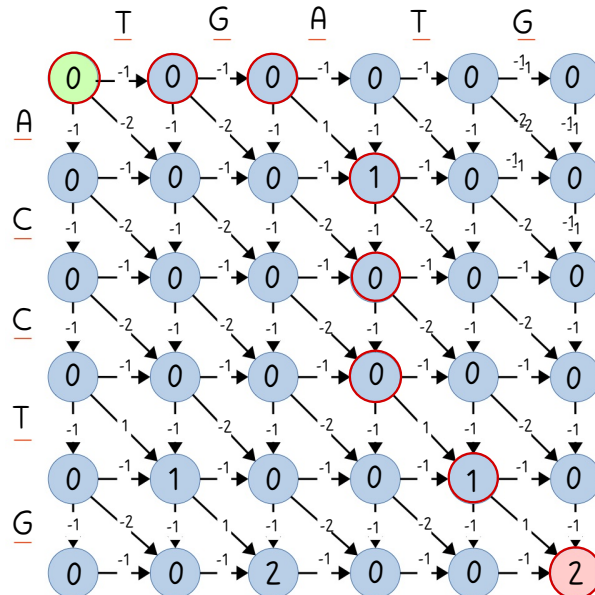
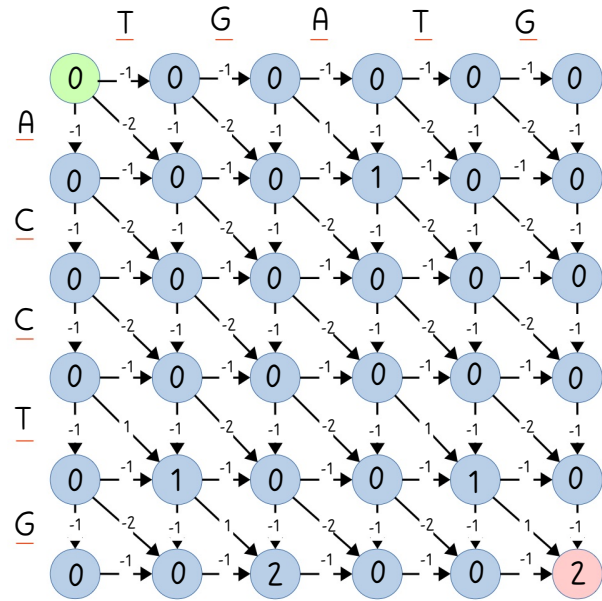
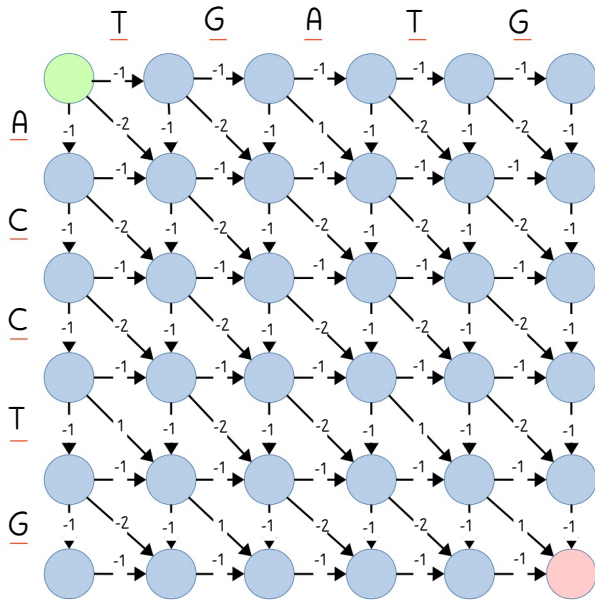
ESERCIZIO 3

Try to align **locally** the following sequences:

- ACCTG and TGATG;

score matrix

	A	C	G	T	GAP
A	1	-2	-2	-2	-1
C	-2	1	-2	-2	-1
G	-2	-2	1	-2	-1
T	-2	-2	-2	1	-1
GAP	-1	-1	-1	-1	-



RISPOSTA:

- - ACCTG
TG A - - TG

PROCEDIMENTO:

IL PROCEDIMENTO È SIMILE A QUELLO DELL'ALLINEAMENTO GLOBALE, MA CON UNA DIFFERENZA IMPORTANTE: SE DURANTE LA VALORIZZAZIONE DEI NODI SI OTTIENE UN VALORE NEGATIVO, IL NODO VIENE IMPOSTATO A 0. QUESTO RIFLETTE IL FATTO CHE, NELL'ALLINEAMENTO LOCALE, I VALORI NEGATIVI NON SONO CONSIDERATI UTILI E INDICANO L'INIZIO DI UNA NUOVA SEQUENZA DI ALLINEAMENTO.

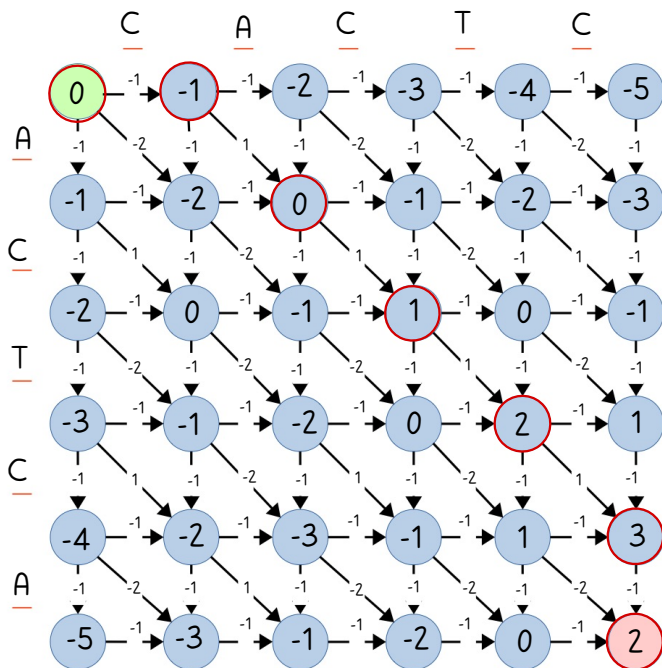
ESERCIZIO 4

Try to align **globally** the following sequences:

- ACTCA and CACTC.

score matrix $\Rightarrow G$

	A	C	G	T	GAP
A	1	-2	-2	-2	-1
C	-2	1	-2	-2	-1
G	-2	-2	1	-2	-1
T	-2	-2	-2	1	-1
GAP	-1	-1	-1	-1	-



- A C T C A

C A C T C -

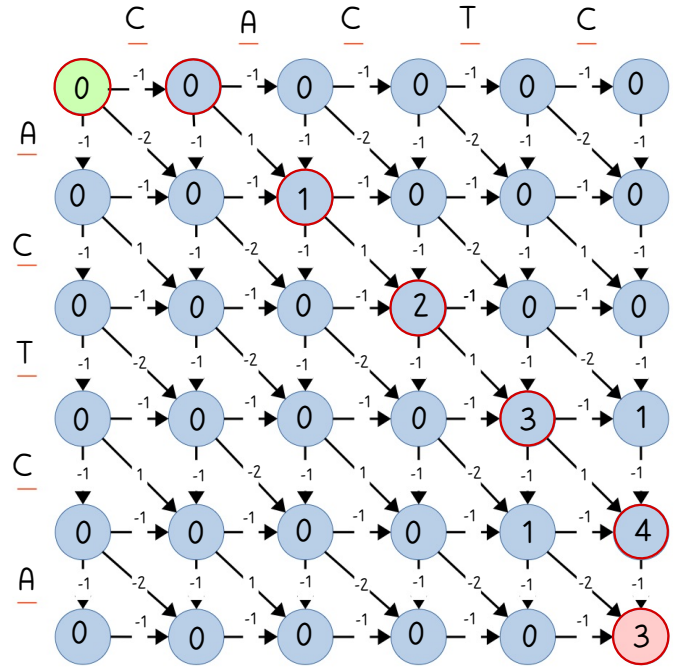
ESERCIZIO 5

Try to align **locally** the following sequences:

- ACTCA and CACTC.

score matrix $\Rightarrow G$

	A	C	G	T	GAP
A	1	-2	-2	-2	-1
C	-2	1	-2	-2	-1
G	-2	-2	1	-2	-1
T	-2	-2	-2	1	-1
GAP	-1	-1	-1	-1	-



- A C T C A

C A C T C -

ESERCIZIO 6

- For reference **bioinformatics** create:

- ▶ BWT;
- ▶ FirstOccurence;
- ▶ CheckPointArray_b with b=5;
- ▶ SuffixArray_a with a=5;
- ▶ Search Pattern: **ioi**

BIOINFORMATICS\$

ROTAZIONI:

BIOINFORMATICS\$
IOINFORMATICS\$B
OINFORMATICS\$BI
NFORMATICS\$BIO
NFORMATICS\$BIOI
FORMATICS\$BIOIN
FORMATICS\$BIOINF
FORMATICS\$BIOINFO
MATIC5\$BIOINFOR
ATIC5\$BIOINFORM
TICS\$BIOINFORMA
ICS\$BIOINFORMAT
CS\$BIOINFORMATI
5\$BIOINFORMATIC
\$BIOINFORMATICS

ROTAZIONI ORDINATE:

\$BIOINFORMATICS
ATICS\$BIOINFORM
BIOINFORMATICS\$
CS\$BIOINFORMATI
FORMATICS\$BIOIN
ICS\$BIOINFORMAT
INFORMATICS\$BIO
IOINFORMATICS\$B
MATICS\$BIOINFOR
NFORMATICS\$BIOI
OINFORMATICS\$BI
ORMATICS\$BIOINF
RMATICS\$BIOINFO
S\$BIOINFORMATIC
TICS\$BIOINFORMA

BWT MATRIX:

\$	BIOINFORMATIC	5
A	TICS\$BIOINFOR	M
B	IOINFORMATICS	\$
C	5\$BIOINFORMAT	I
F	ORMATICS\$BIOI	N
I	C5\$BIOINFORMA	T
I	NFORMATICS\$BI	O
I	OINFORMATICS\$	B
M	ATICS\$BIOINFO	R
N	FORMATICS\$BIO	I
O	INFORMATICS\$B	I
O	RMATICS\$BIOIN	F
R	MATICS\$BIOINF	O
S	\$BIOINFORMATI	C
T	ICS\$BIOINFORM	A

BWT = SM\$INTOBRIIFOCA

BWT MATRIX:

\$ BIOINFORMATICS
A TICS\$BIOINFOR
B IOINFORMATICS
C S\$BIOINFORMAT
F ORMATICS\$BIOI
I CS\$BIOINFORMA
I NFORMATICS\$BI
I OINFORMATICS\$
M ATICS\$BIOINFO
N FORMATICS\$BIO
O INFORMATICS\$B
O RMATICS\$BIOIN
R MATICS\$BIOINF
S \$BIOINFORMATI
T IC\$BIOINFORM

FIRSTOCCURRENCE:

S	←	0
M		1
\$		2
I		3
N		4
T	←	5
O		
B		
R		8
I		9
I	←	10
F		
O		12
C		13
A		14

CHECKPOINT ARRAY

\$ A B C F I M N O R S T
0 0 0 0 0 0 0 0 0 0 0 0

1 0 0 0 0 1 1 1 0 0 1 0

1 0 1 0 0 2 1 1 1 1 1 1

SUFFIX ARRAY

0

5

10

Top = 0
Bottom = 14

Top = 5 + 0 = 5
Bottom = 5 + 3 - 1 = 7

Top = 5
Bottom = 7

Top = 10 + 0 = 10
Bottom = 10 + 1 - 1 = 10

Top = 10
Bottom = 10

Top = 5 + 2 = 7
Bottom = 5 + 3 - 1 = 7

FOUND PREFIX!