W b Da a Ma a

C a S a

L 3: S a a ba

W a Pa Ma ?

D :

a T a a a P,

a

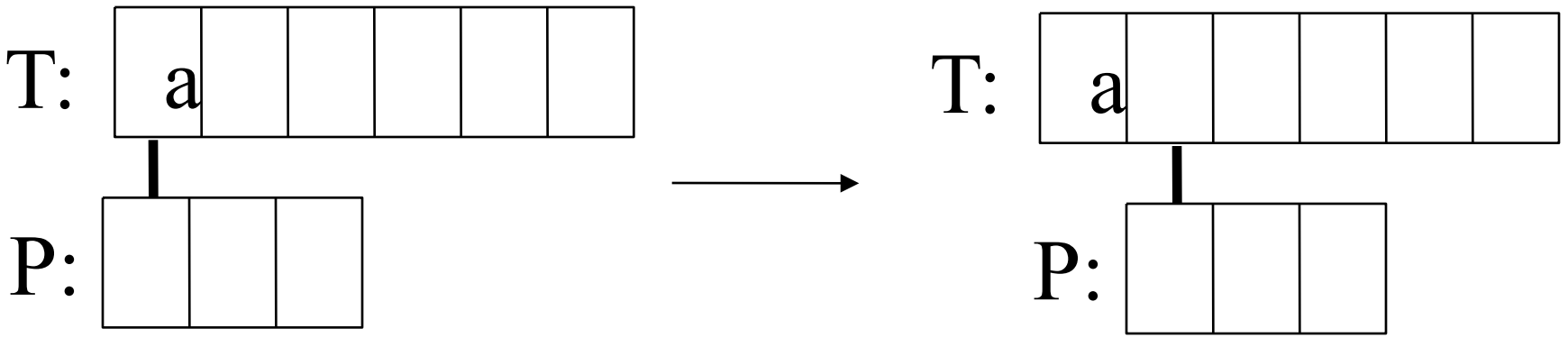
T: a a a a a

P:

T B F A

C a T

a P a a



P 1 a a a T

....

T b

E a a a :

T: "aaaaaaaaaaaaaaaaaaaaaaaaaaaa "

P: "aaa "

E a a a a a :

T: "a a a a a "

P: " "

T KMP A

T K -M -P a (KMP) a
a a

S a

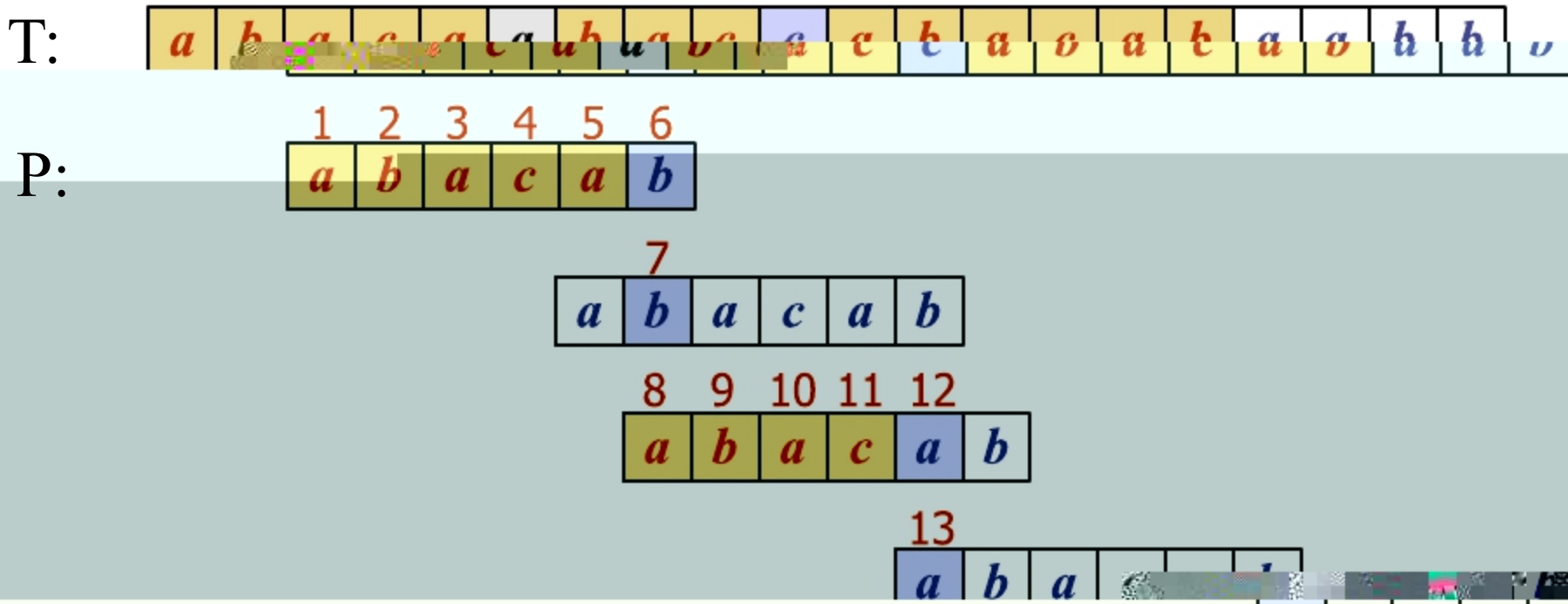
I a a b a
a P a P[ ], a a
a a a
a ?

S a

I a a b a
a P a P[ ], a a
a a a
a ?

: a P[0 .. -1] a
a P[1 .. -1]

E a



	0	1	23	4
***F*( )**	0	0	10	1

KMP A a a

KMP a : O(+)

a

T a

ba a , T

a a

a a a a a

a a

KMP D a a a

KMP a

a ab a

a a a (b

a)

a a a ,

b KMP a a

a

T B -M A

T B -M A

T B -M a a
a ba .

1. T
P T b P,
a a

2. T

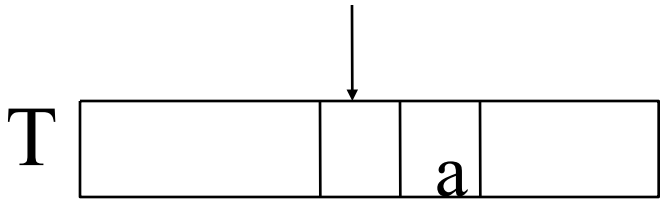
a a a T[ ] ==

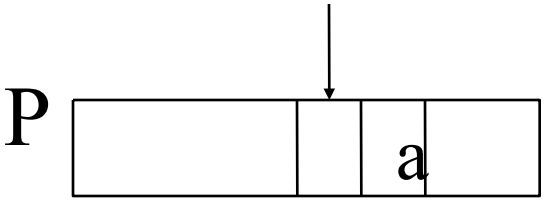
a a a P[ ]

a a T[ ]

T a 3 b

a .



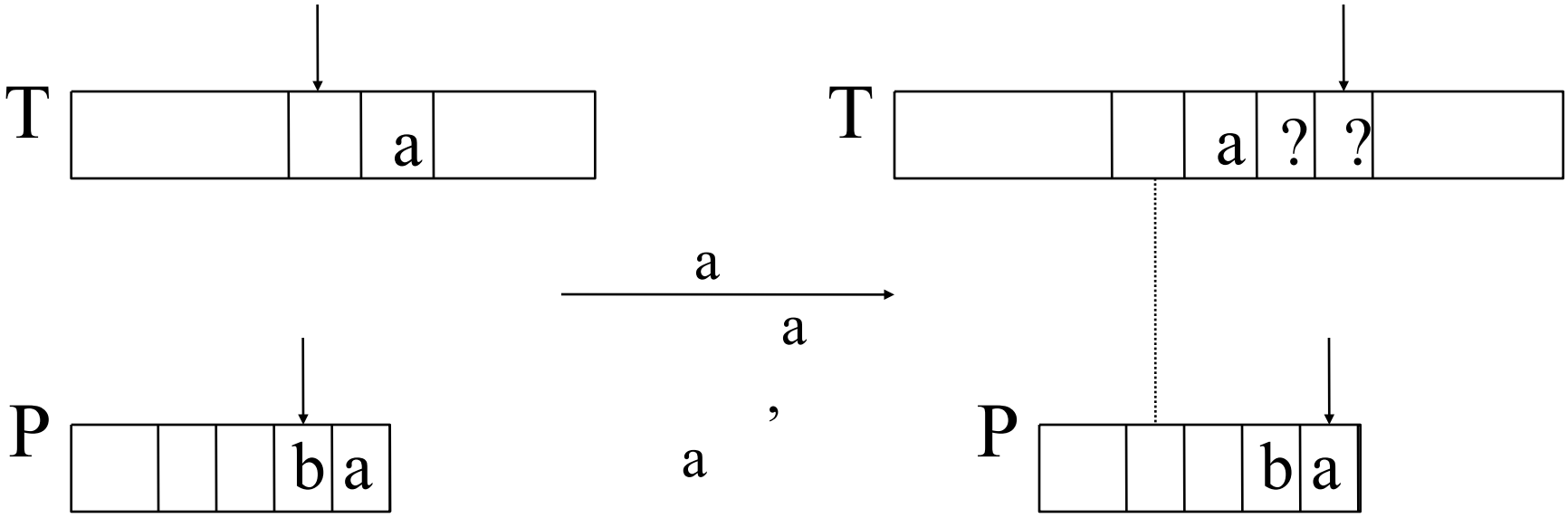


Ca 1

I P a ,

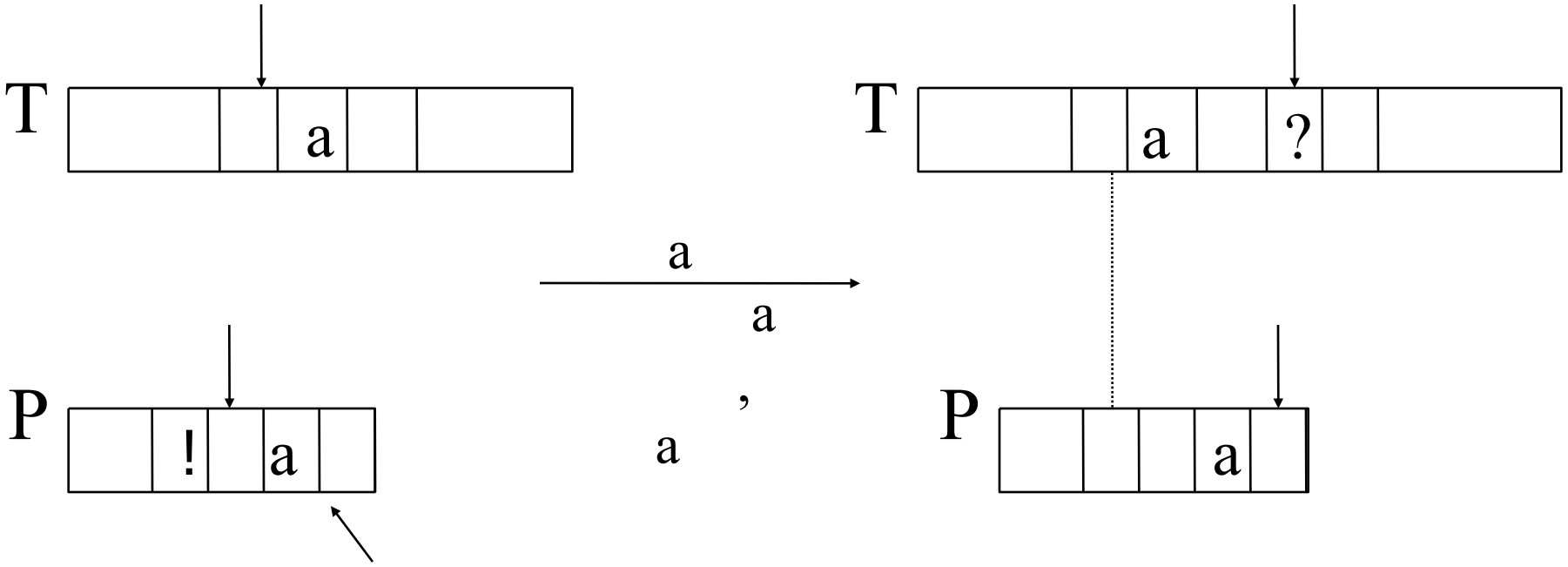
a a

P T[ ].



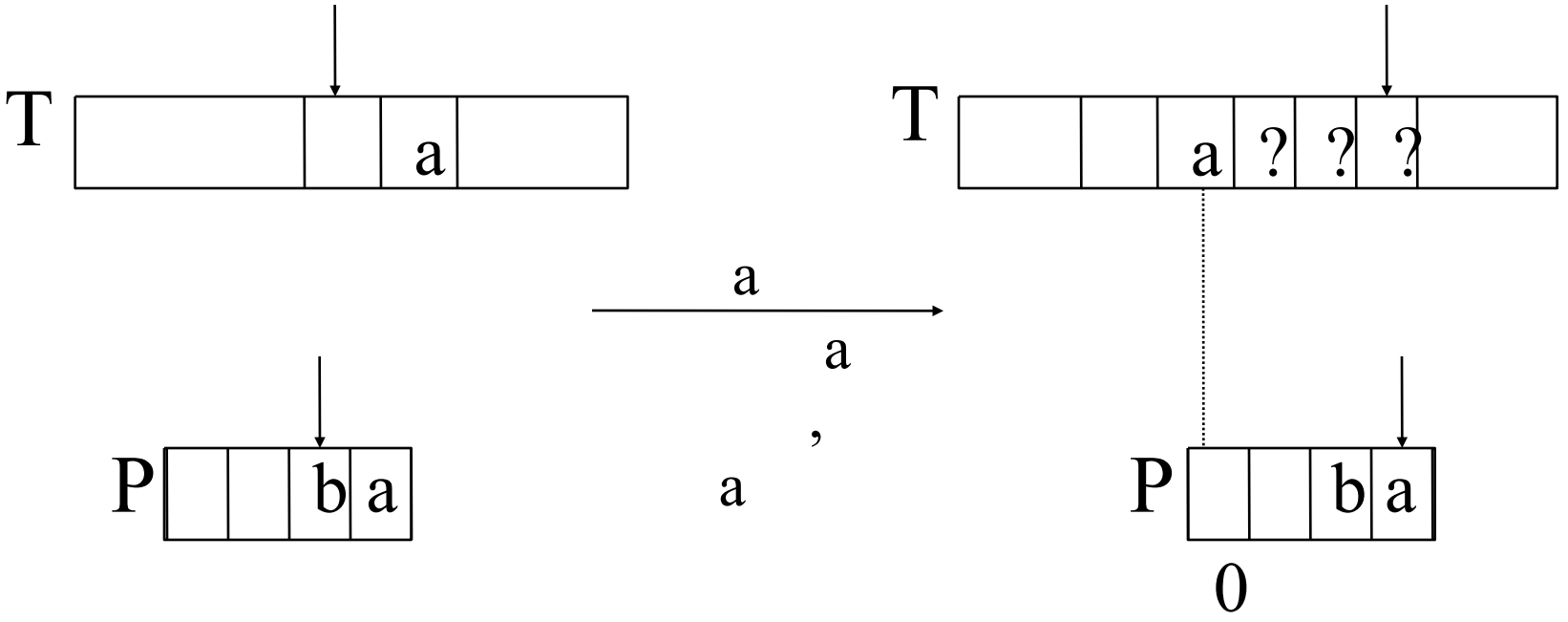
Ca 2

I P a , b a
a b ,
b 1 a a T[+1].

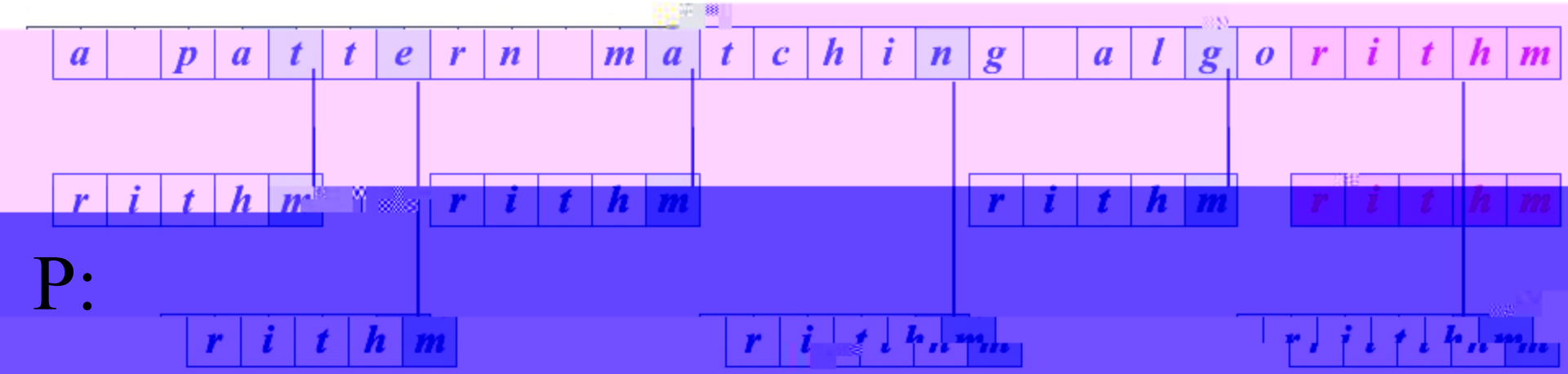


Ca 3

I a 1 a 2 a , P
a P[0] T[+1].



B -M E a (1)


T:
a p a t t e r n m a t c h i n g a l g o r i t h m
P:
r i t h m
r i t h m
r i t h m
r i t h m
r i t h m

La O F

B -M a
a P a a ab A b a a
L()
L() a a A

L( ) a : // a A
a a P[ ] == ,
-1

L() E a

A = a, b, ,

P: "aba ab"

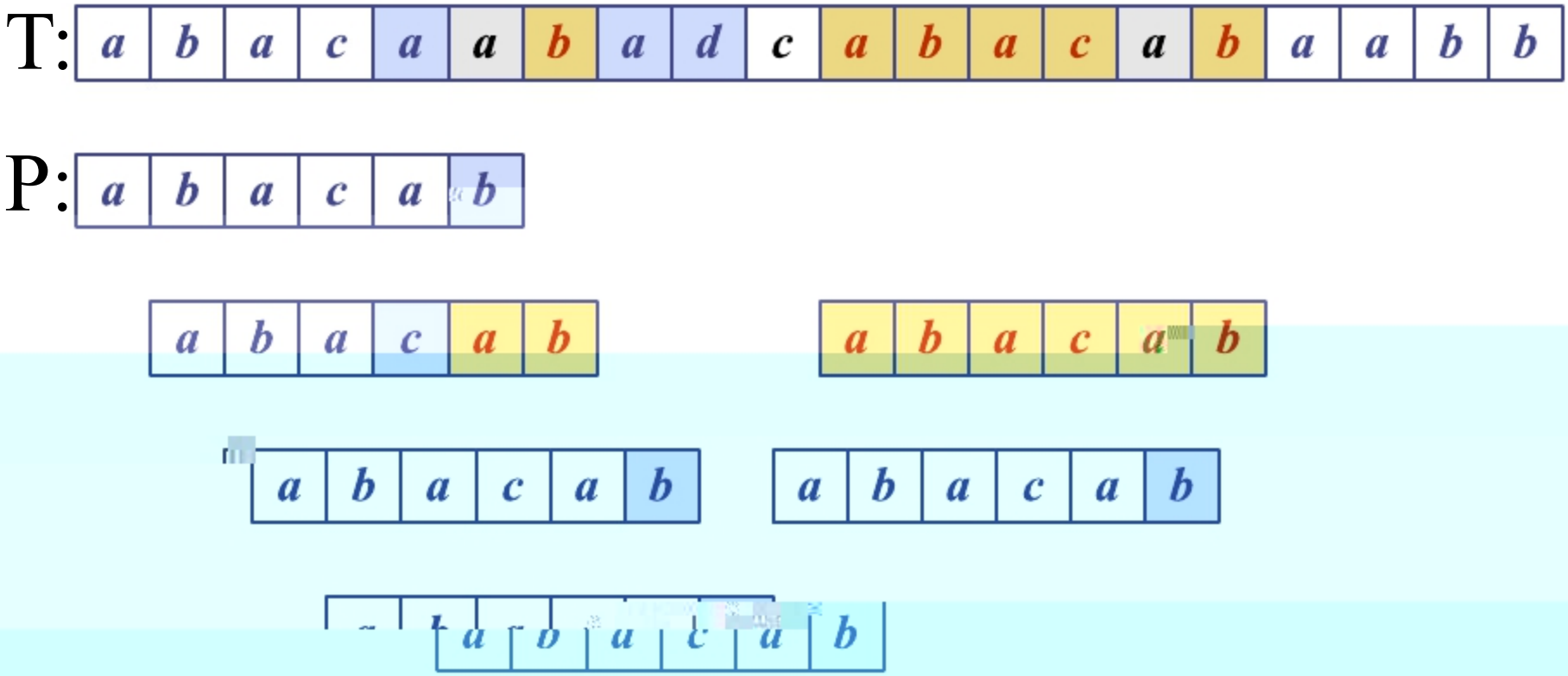
P	a	b	a		a	b
	0	1	2	3	4	5

	a	***b***		***d***
L(r)	4	5	3	-1

L()

P[]

B -M E a (2)



	a	*b*		*d*
L( )	4	5	3	–1

A a

B -M a
O(+ A)

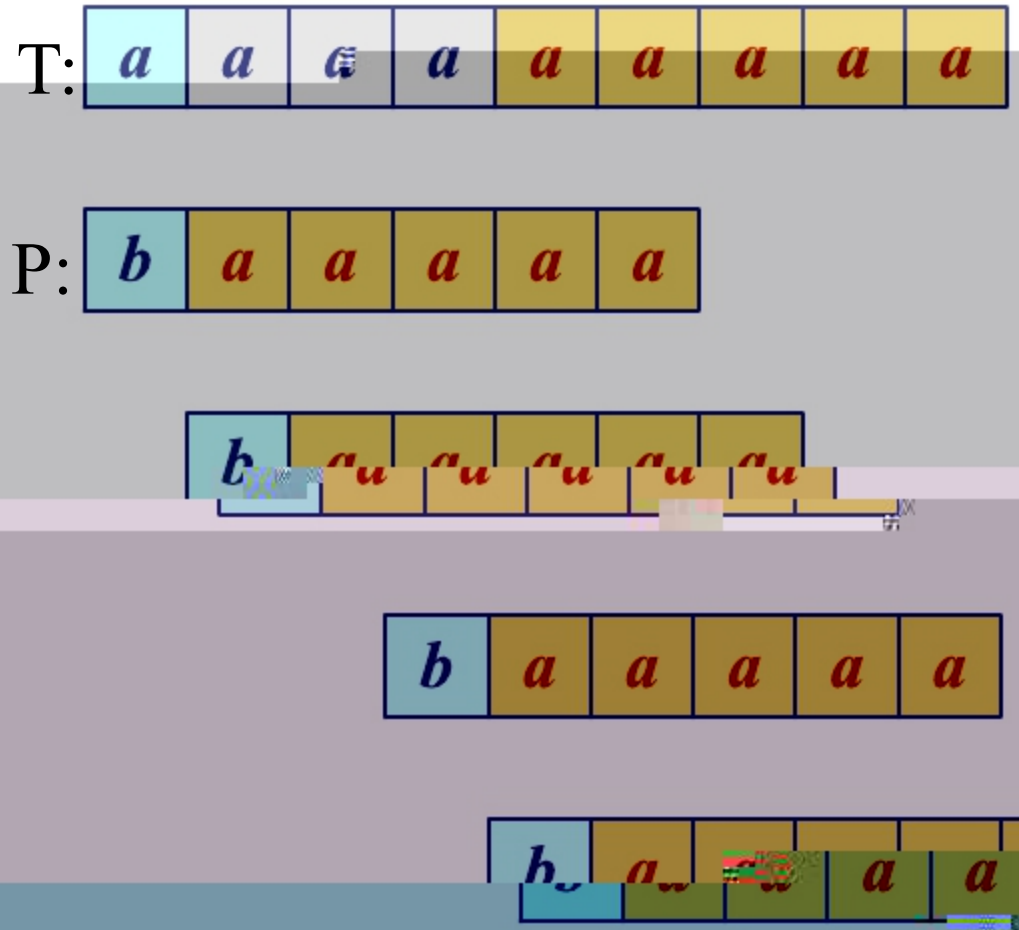
B , B -M a a ab (A)
a , a ab a .
. . E , b a

B -M
a E .

W Ca E a

T: "aaaaa a"

P: "baaaaa"



R a E

N a a a a

D a a

S a a a a .

F a a a a a a a

Ca ab b a a a NFA

T a a a a , RE = NFA = DFA

W a a a b a a

D a R a E

R a a :

a ab ∑, a a a a

2. ε, a a a ε
3. , a a a
4. R_1+R_2, R_1 a R_2 a a , a + ( )
5. R_1R_2, R_1 a R_2 a a a a a
6. R*, R a a a
7. (R), R a a , a a R a a a

T a a , b 1-3 ba

P : Pa a ,

b *, a a , a .

U R a E

R a a a a a a ' .

B Ja a, P , U , P , ...

RE E a

L( ) = 001

L(*) = 0, 1, 10, 100, 1000, 10000,

L(*) = 1, 01, 10, 010, 0010, . . a a a 1

L(∑∑)* = a

L(( ))*) = ε, 00, 01, 0000, 0001, 0100, 0101,

L((**ε** **ε**)) = ε, 0, 1, 01

L(1) = ; a a a .

Rε = R

R+ = R

N a R+ε a a a R (a a ε a a)

N a R a R

E 1

L ∑ b a b

∑ = 10, 11 , ∑* = ?

A

A $: \sum^* =$ ϵ, 10, 11, 1010, 1011, 1110, 1111,

E 2

L1 = 10, 1 , L2 = 011, 11 , L1L2 = ?

A

L1L2 = 10011, 1011, 111

E 3

W RE

A 0 a 1

A 0 a 1 a a 2

0

A 0 a 1 b 1 a

a 0

A

(0 1)*

A 0 a 1

(0 1)*00(0 1)*

A 0 a 1 a a 2

0

(1+10)*

A 0 a 1 b 1

a a 0

M E

1) (0 1)*011
2) 0*1*2*
3) 00*11*22*

M E (A)

1) (0 1)*011

A : a 0 a 1
011

2) 0*1*2*

A : a b 0 b
a b 1 b a b
2

3) 00*11*22*

A : 0*1*2 a a
a b

NFA

D F A a a (DFA)

S a N a .

B a a .

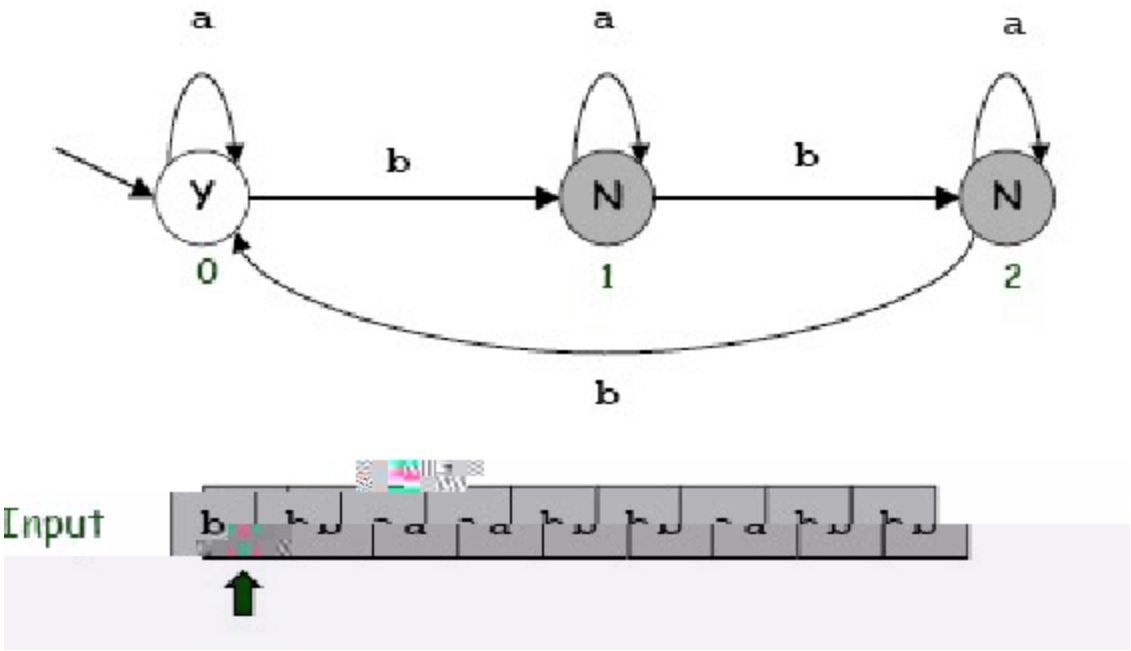
R a b .

M a ,
a a b .

R a a b a .

A ab
a a .

DFA



T DFA a RE

RE. C a b a

.

DFA. Ma a

a .

: a DFA, a

a b a

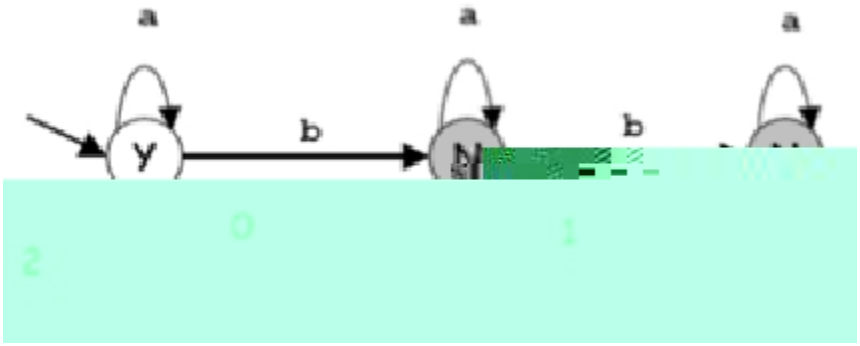
; a a ,

a DFA a

a .

D a E a

DFA 3 b :



RE 3 b :

```
(a*ba*ba*ba*)* a*
```

F a a Q

W a a CANNOT b b
b a RE?

S a b a b
0 a 1 .

S a a a
b .

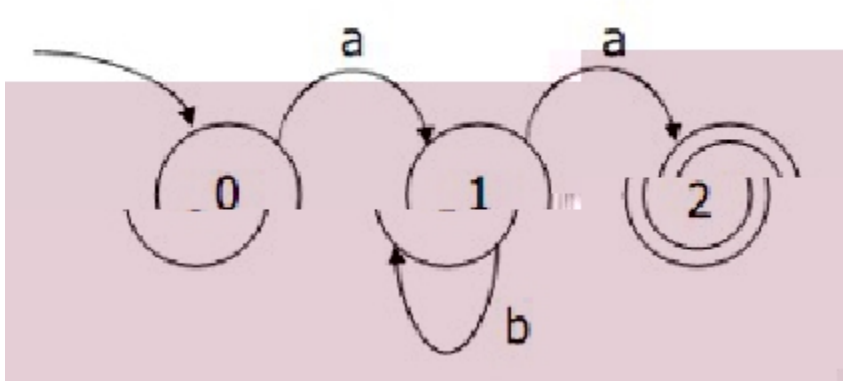
Ma . . . .

P b 1

Ma a DFA a a
a a b a
ab*a

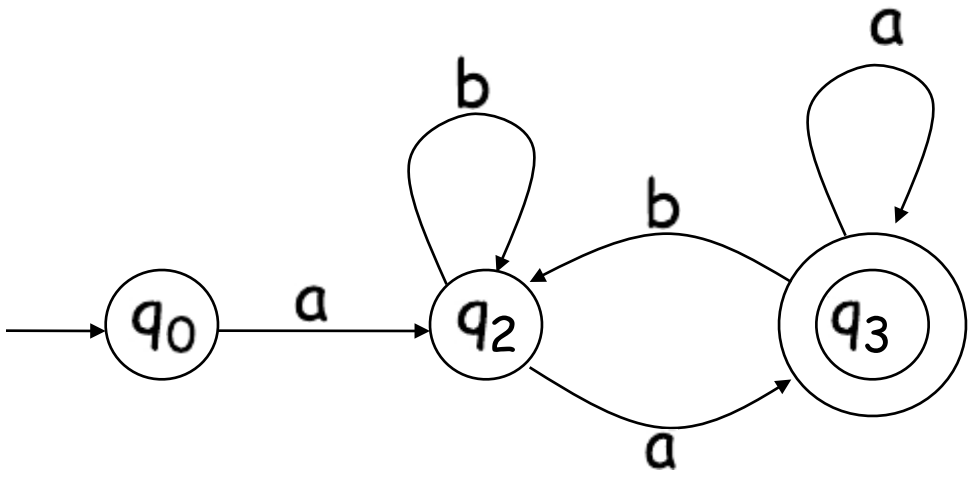
S

ab*a:



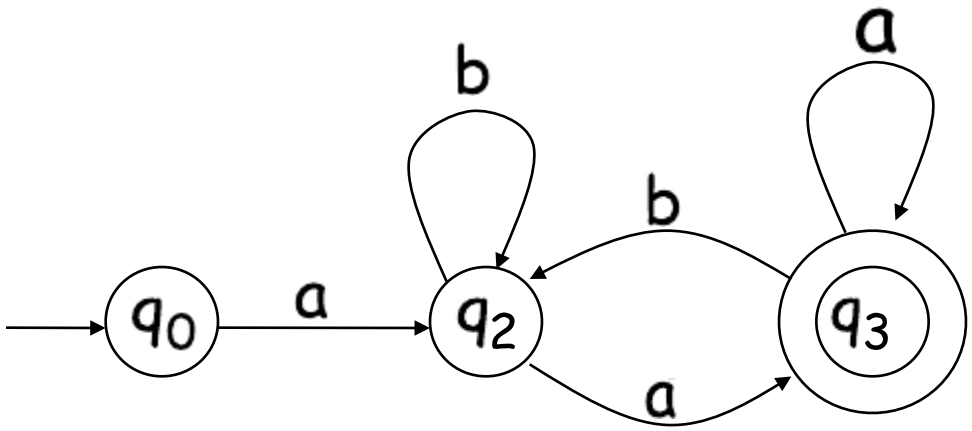
P b 2

W RE a a a:



S

a(a b)*a



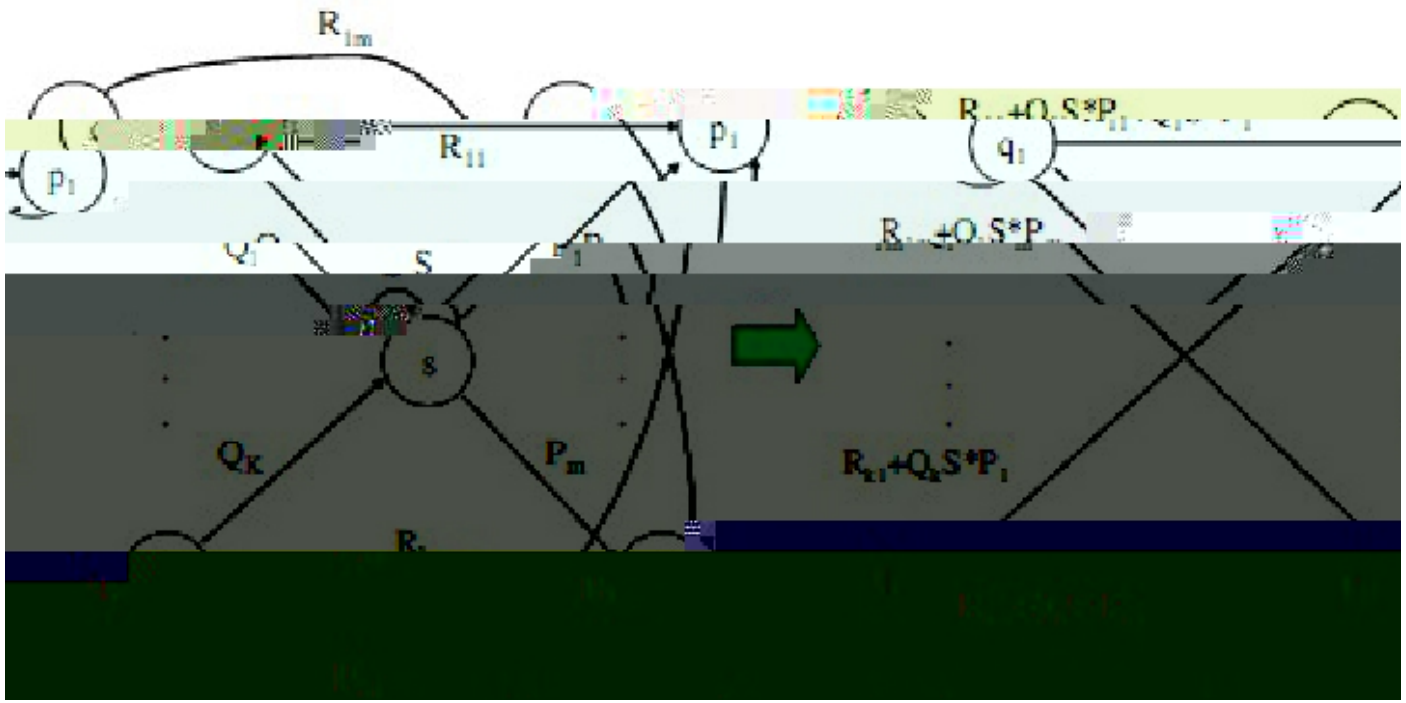
DFA RE: S a E a

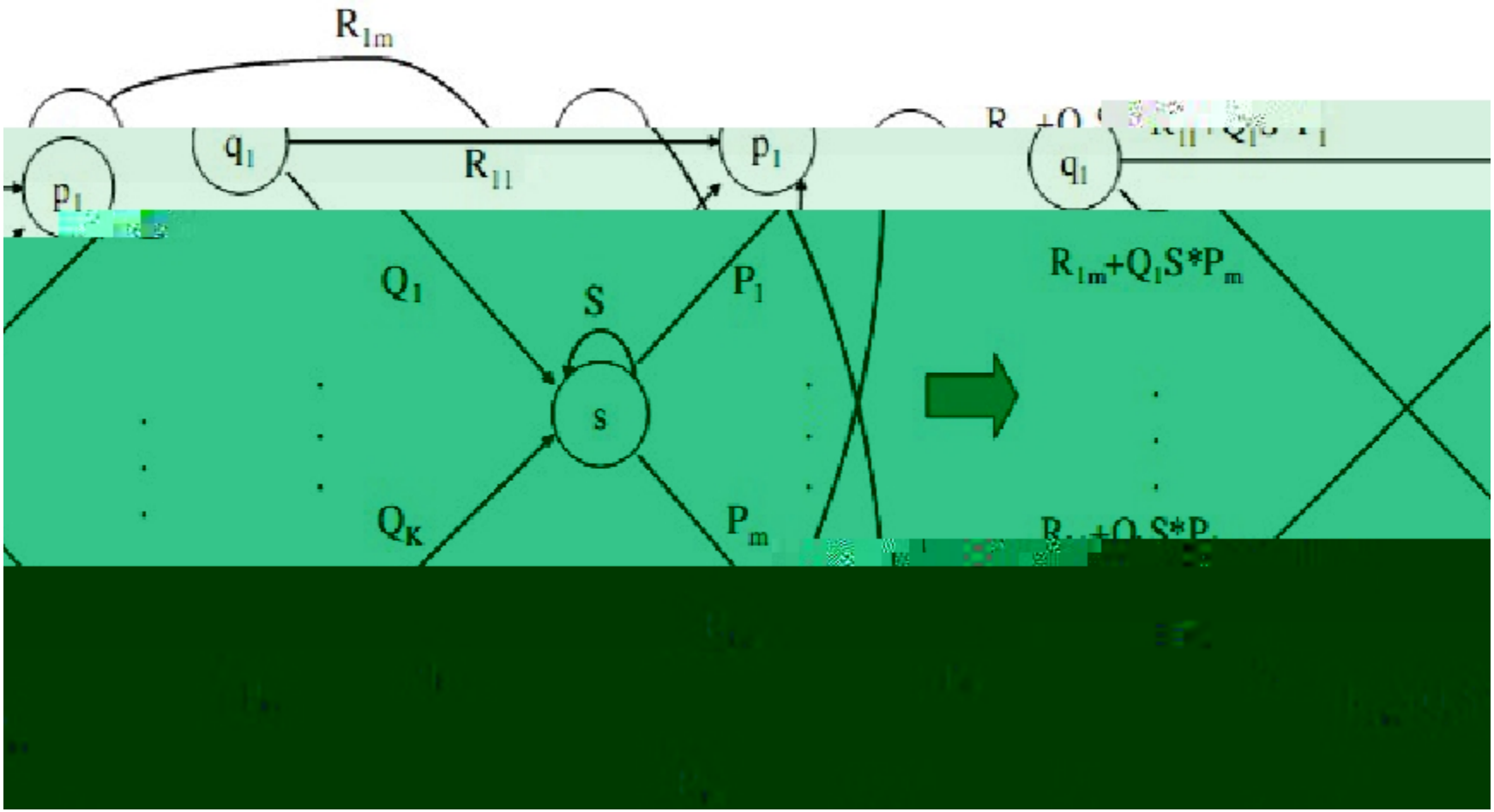
E a a a a a
a a
a b a
a a .

E a a
a a a a , a
a a a RE

S a E a

C b ,



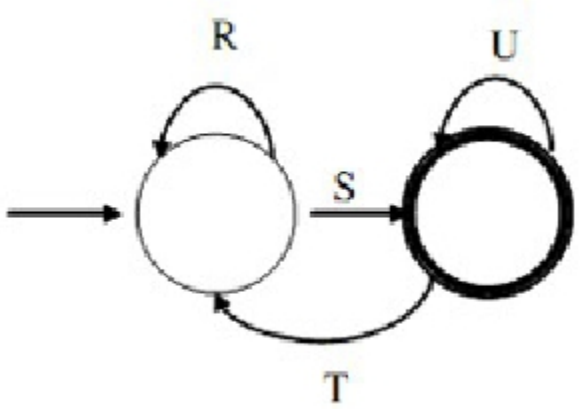

R1m
q1
R11
p1
p1
q1
Q1
S
P1
R1m+Q1S*Pm
s
QK
Pm

DFA RE a S a E a (1)

S a a a a
a a , a a
a a
a a a a a
ab .
T b a a
a a a a a a a
a .

DFA RE S a E a (2)

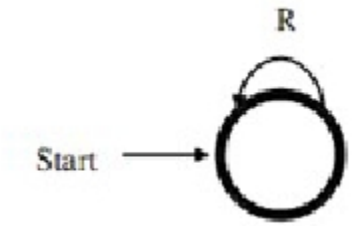
I a a , a a
a a a :



W a b a a a : (R SU*T)*SU*

DFA RE S a E a (3)

I a a a a a a ,



DFA RE S a E a (4)

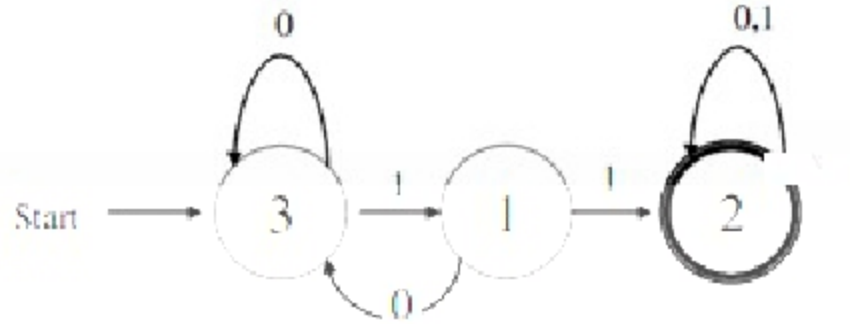
I a a a ,
a ab a a
a a
, R1, R2, R .

F a a a
a a -a .

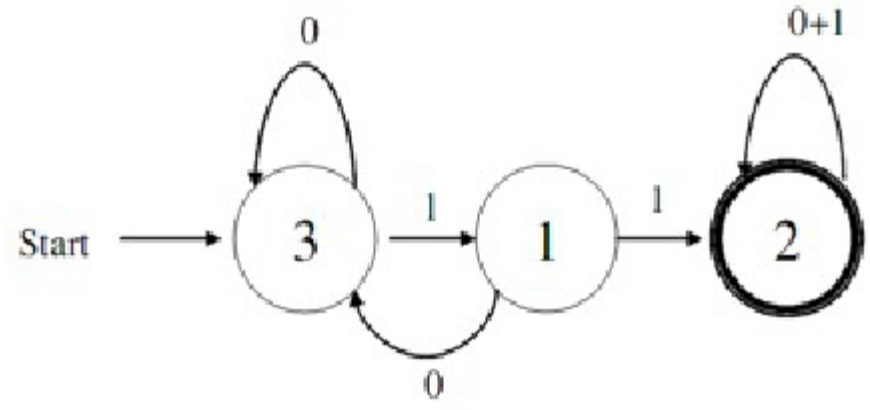
T a
a a a
a : R1 U R2 U RN

DFA->RE E a

C a RE:

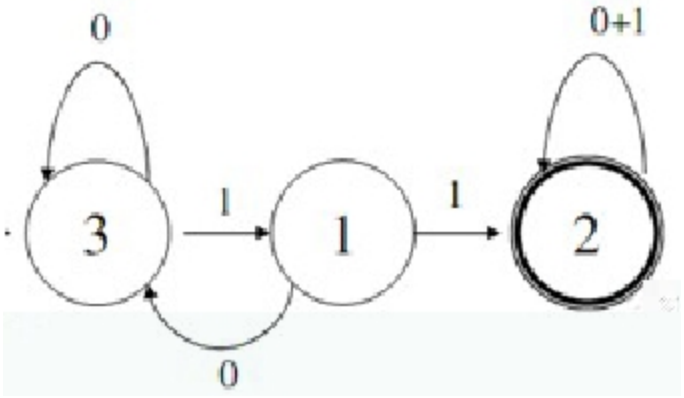


F RE :

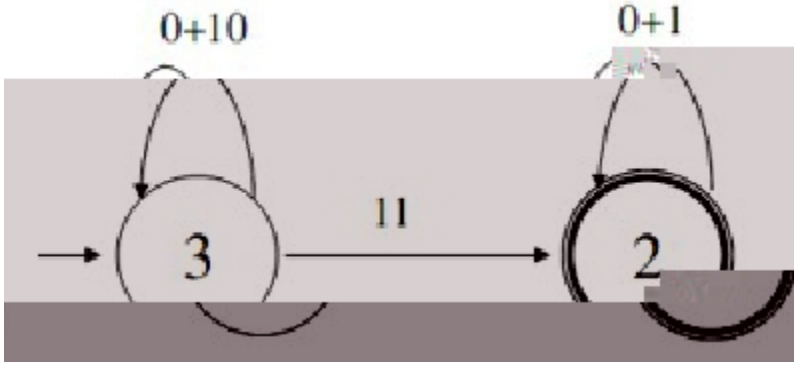


DFA -> RE E a (2)

E a S a 1:



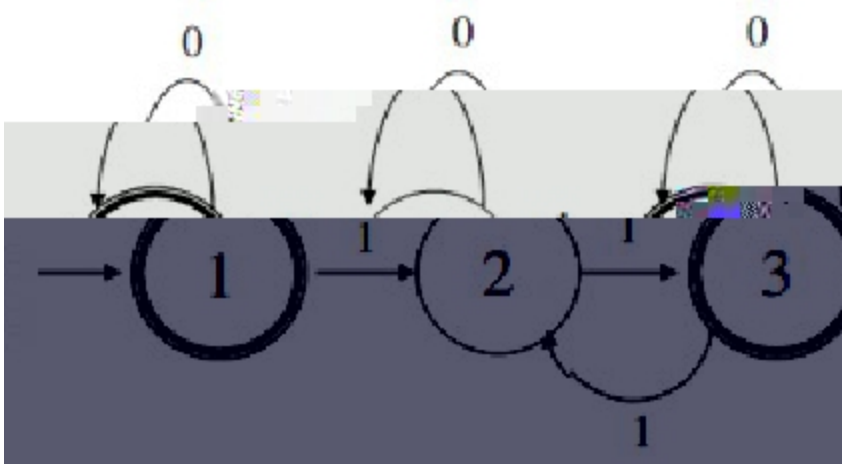
N 3->3



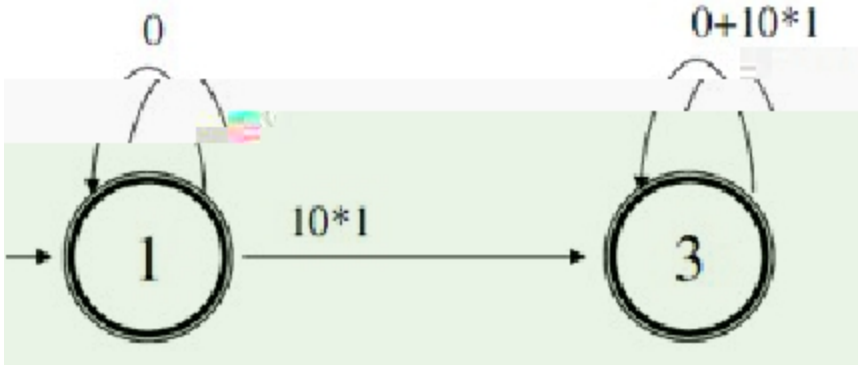
A : (0+10)*11(0+1)*

S E a

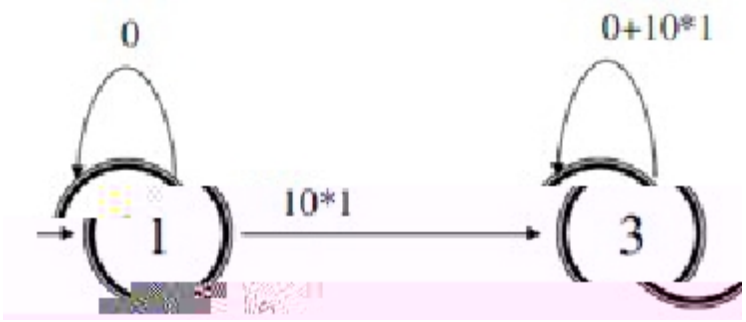
A a a a
a b
1



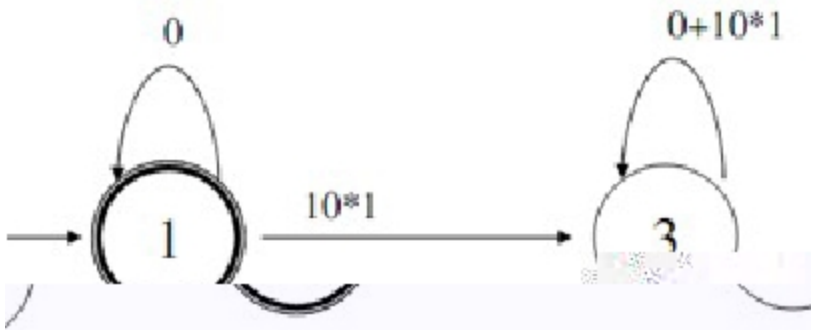
E a a 2:



S E a (2)

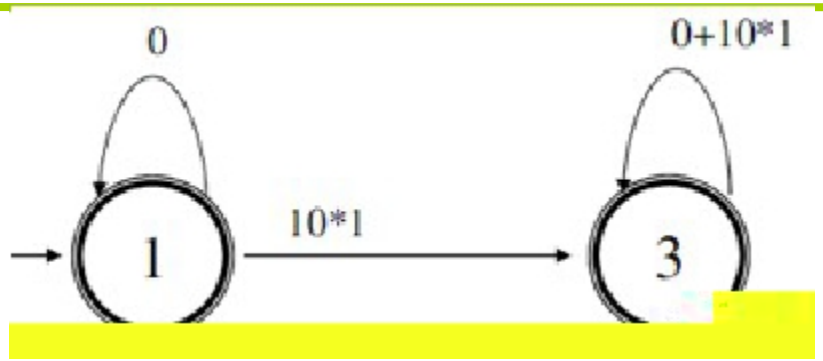


T a a , a 3

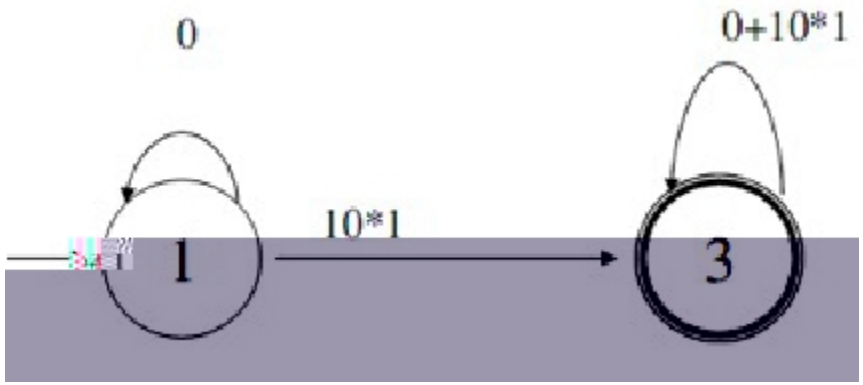


T 0*; a a 3

S E a (3)



T a 1 :



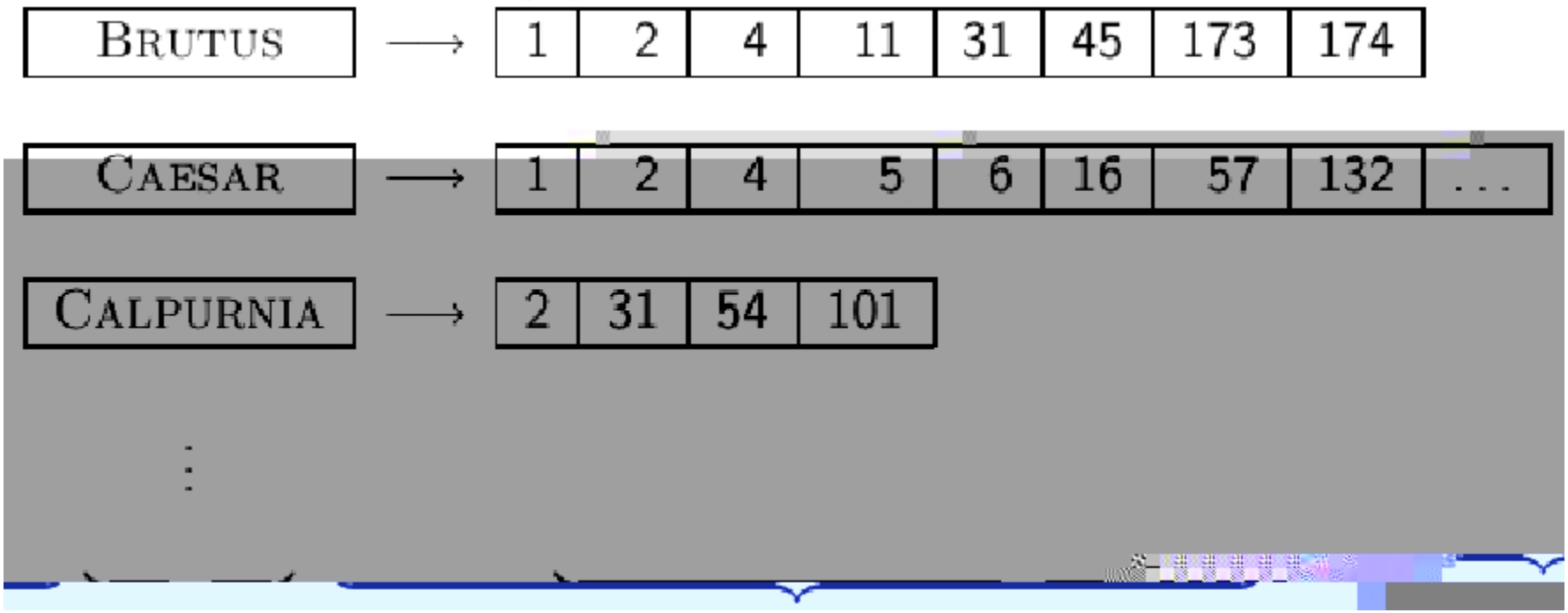
T 0*10*1(0 10*1)*

C b 0* 0*10*1(0 10*1)*

T a

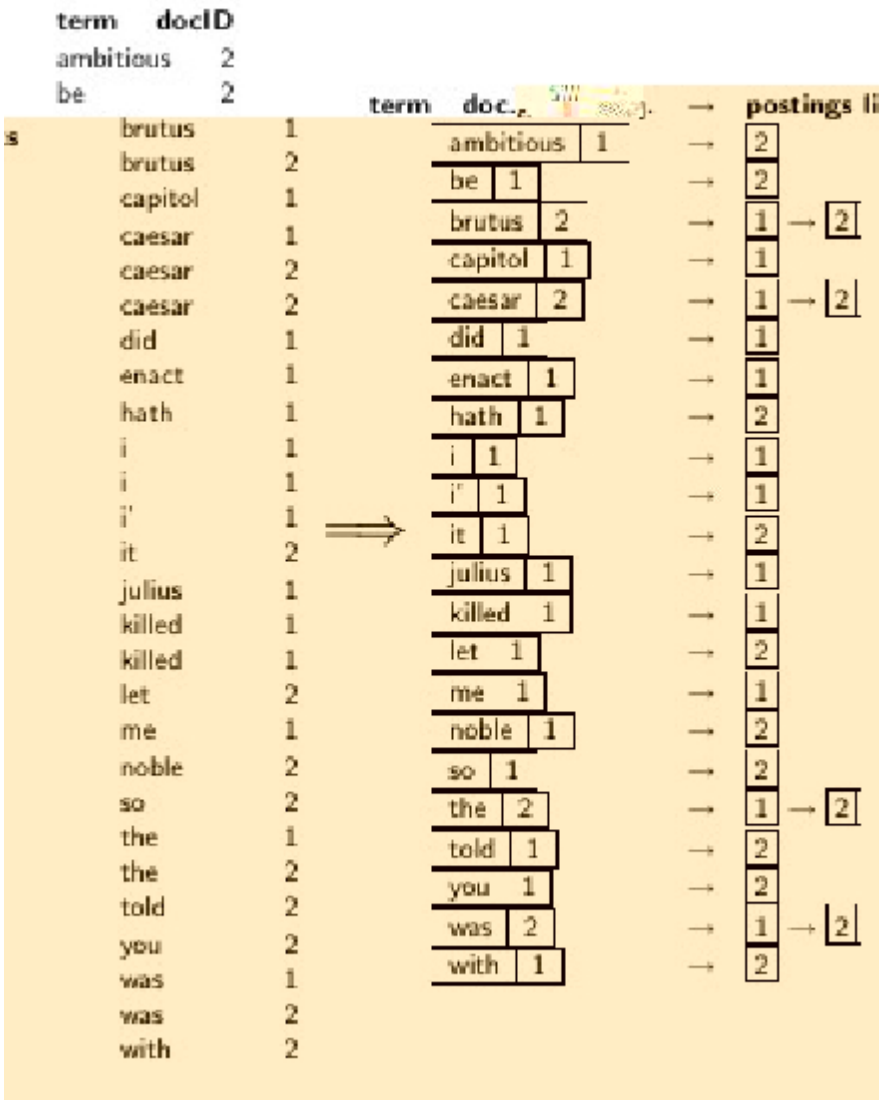
I I

F a , a a a a .


BRUTUS
1 2 4 11 31 45 173 174
CAESAR
1 2 4 5 6 16 57 132 . . .
CALPURNIA
2 31 54 101

a

C a ,


term docID
ambitious 2
be 2
brutus 1
brutus 2
capitol 1
caesar 1
caesar 2
caesar 2
did 1
enact 1
hath 1
i 1
i 1
i' 1
it 2
julius 1
killed 1
killed 1
let 2
me 1
noble 2
so 2
the 1
the 2
told 2
you 2
was 1
was 2
with 2
term doc.
postings li
ambitious 1 → 2
be 1 → 2
brutus 2 → 1 → 2
capitol 1 → 1
caesar 2 → 1 → 2
did 1 → 1
enact 1 → 1
hath 1 → 2
i 1 → 1
i' 1 → 1
it 1 → 2
julius 1 → 1
killed 1 → 1
let 1 → 2
me 1 → 1
noble 1 → 2
so 1 → 2
the 2 → 1 → 2
told 1 → 2
you 1 → 2
was 2 → 1 → 2
with 1 → 2

P a

-
-

P a : E a

1 2

, 2

1 , 1 , , 2, , 2 1

2 1, 1 , , 222, 2

, 1 , 1 0, 2 ,

,

1 , 2 , 1 1 . . .

, 1 2

1 1 , 2

1 , 1 1, 2 1, 0,

1 , 1 , 101 . . .

!

S a

D

W - ba a .

U a .

S ab a .

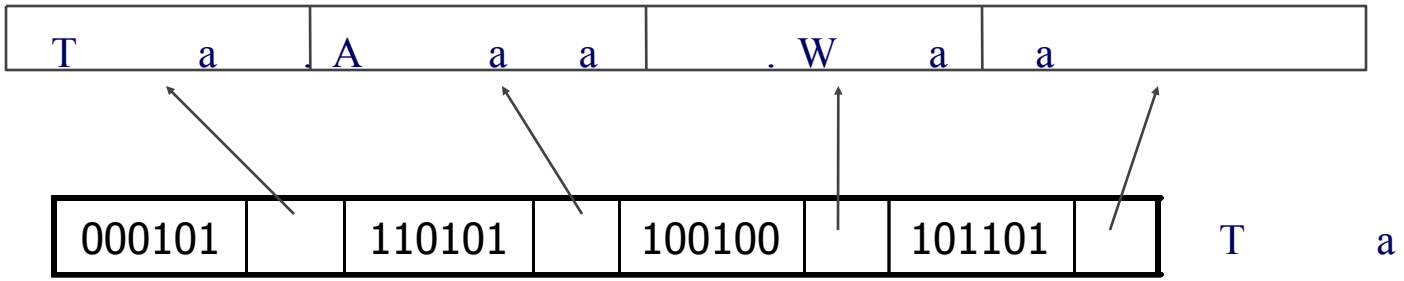
S

Ba a Ha a a b a .

T b .

S a

E a :

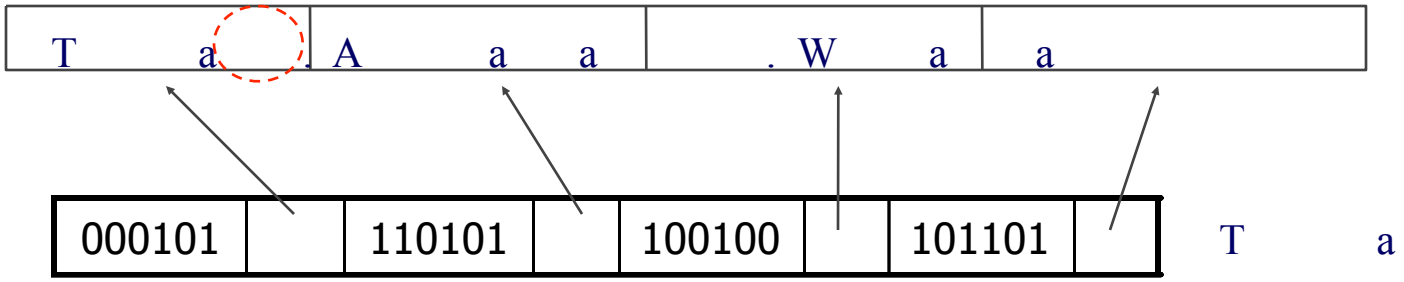


( )	= 000101
(a)	= 110000
( )	= 100100
(a)	= 001100
( )	= 100001

S a

S a

E a :

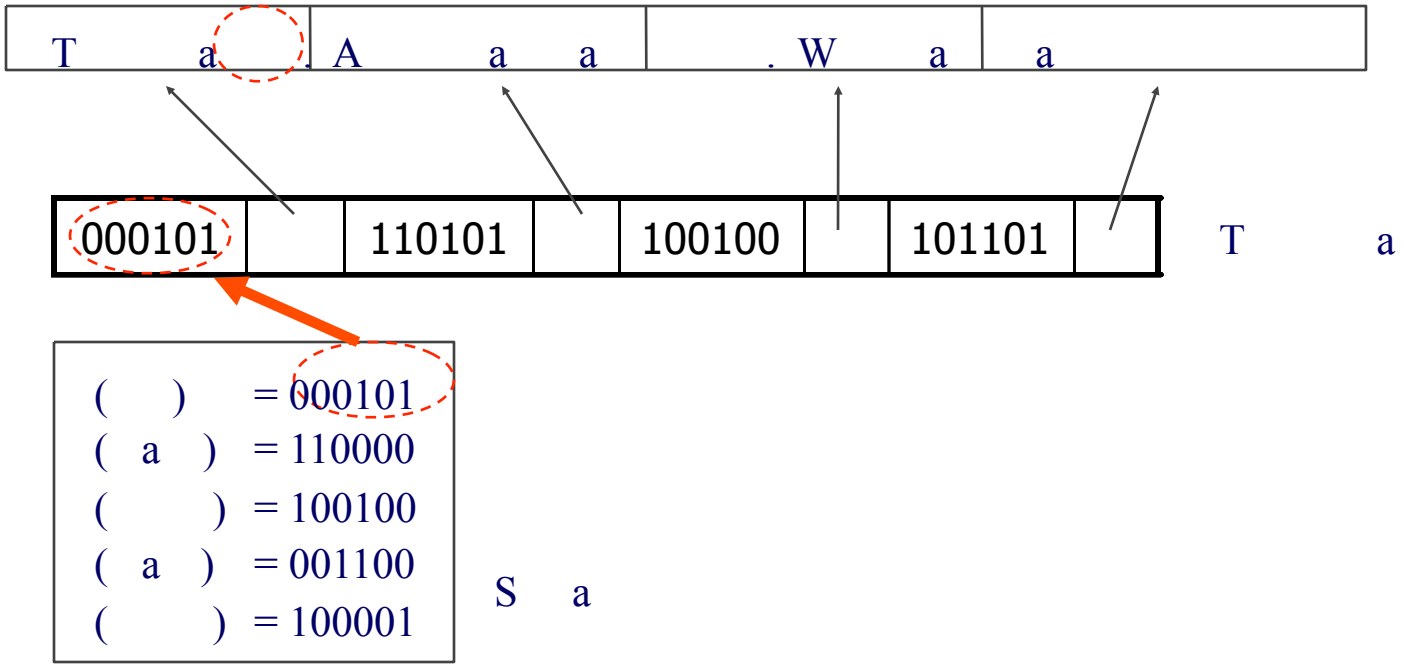


( )	= 000101
(a)	= 110000
( )	= 100100
(a)	= 001100
( )	= 100001

S a

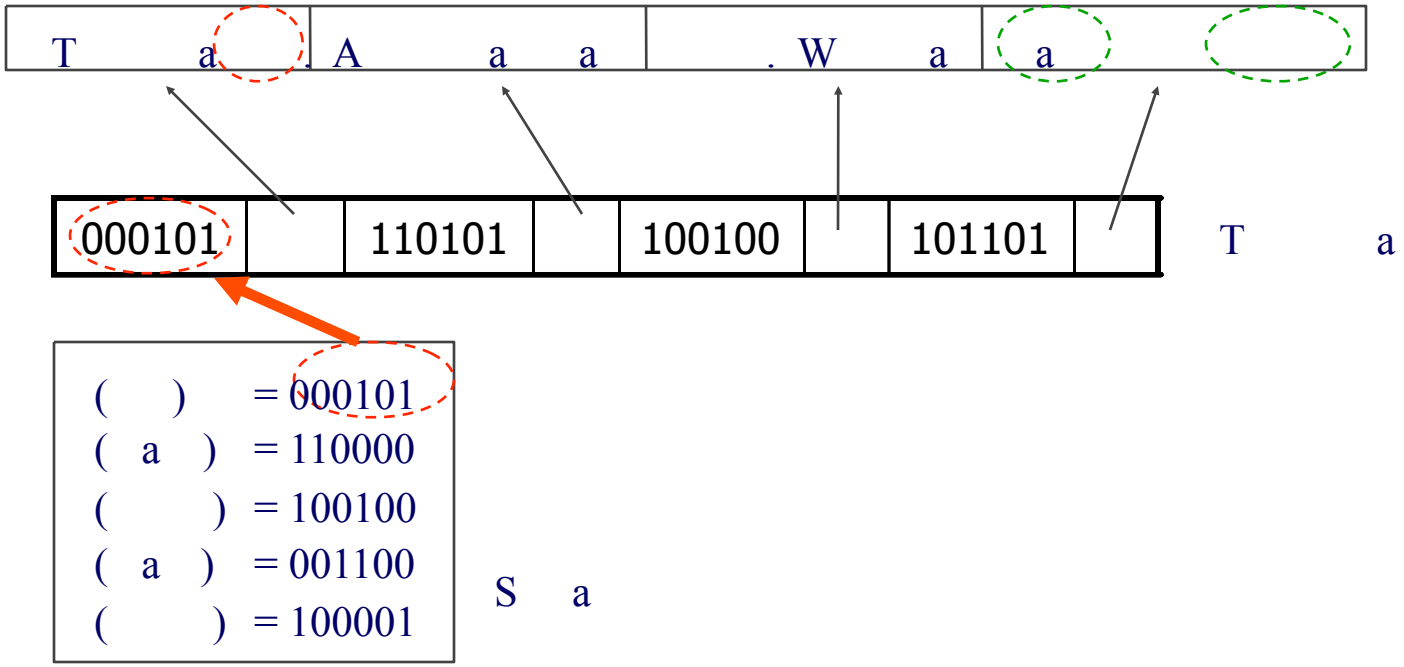
S a

E a :


T a . A a a . W a a
000101
110101
100100
101101
T a
( ) = 000101
(a) = 110000
( ) = 100100
(a) = 001100
( ) = 100001
S a

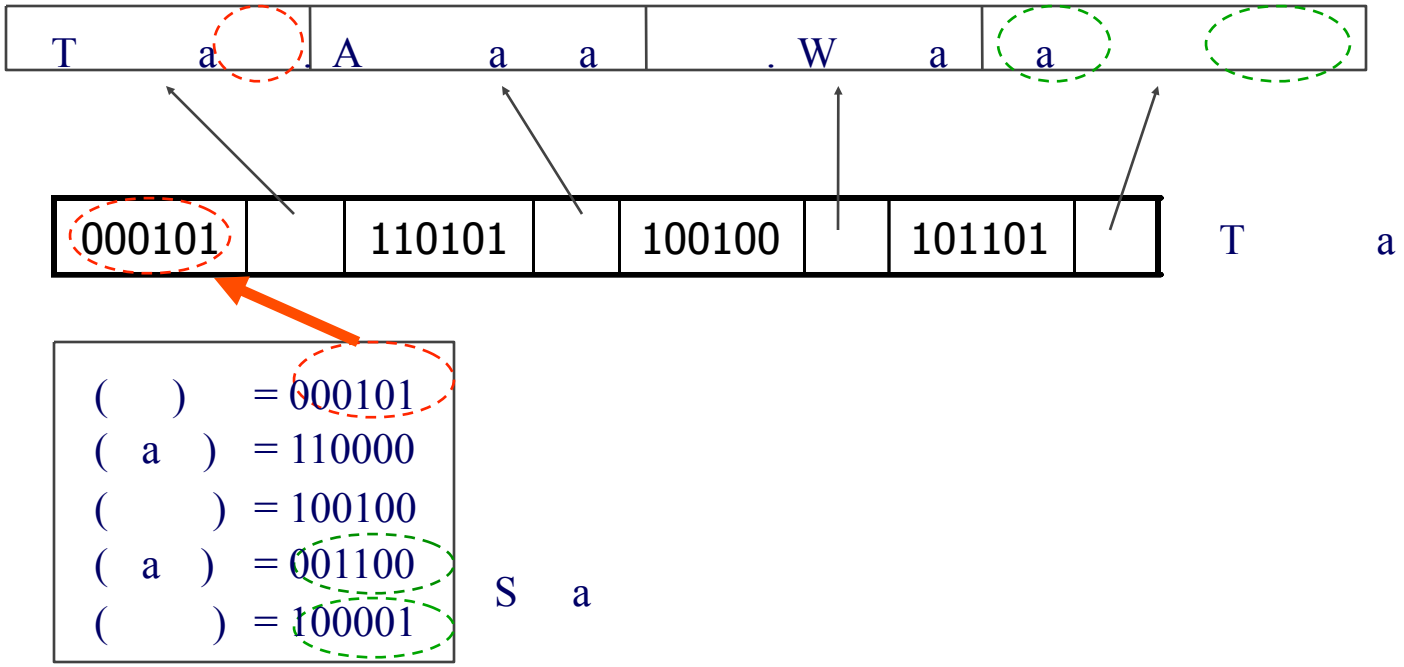
S a

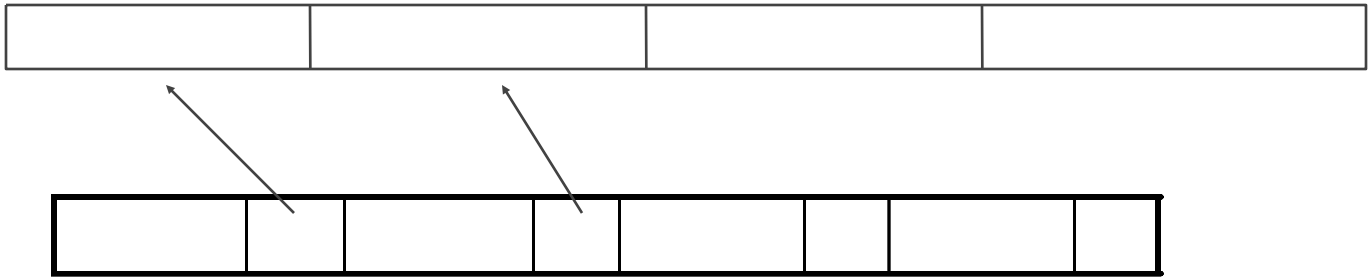
E a :


T a . A a a . W a a
000101
110101
100100
101101
T a
( ) = 000101
(a) = 110000
( ) = 100100
(a) = 001100
( ) = 100001
S a

S a

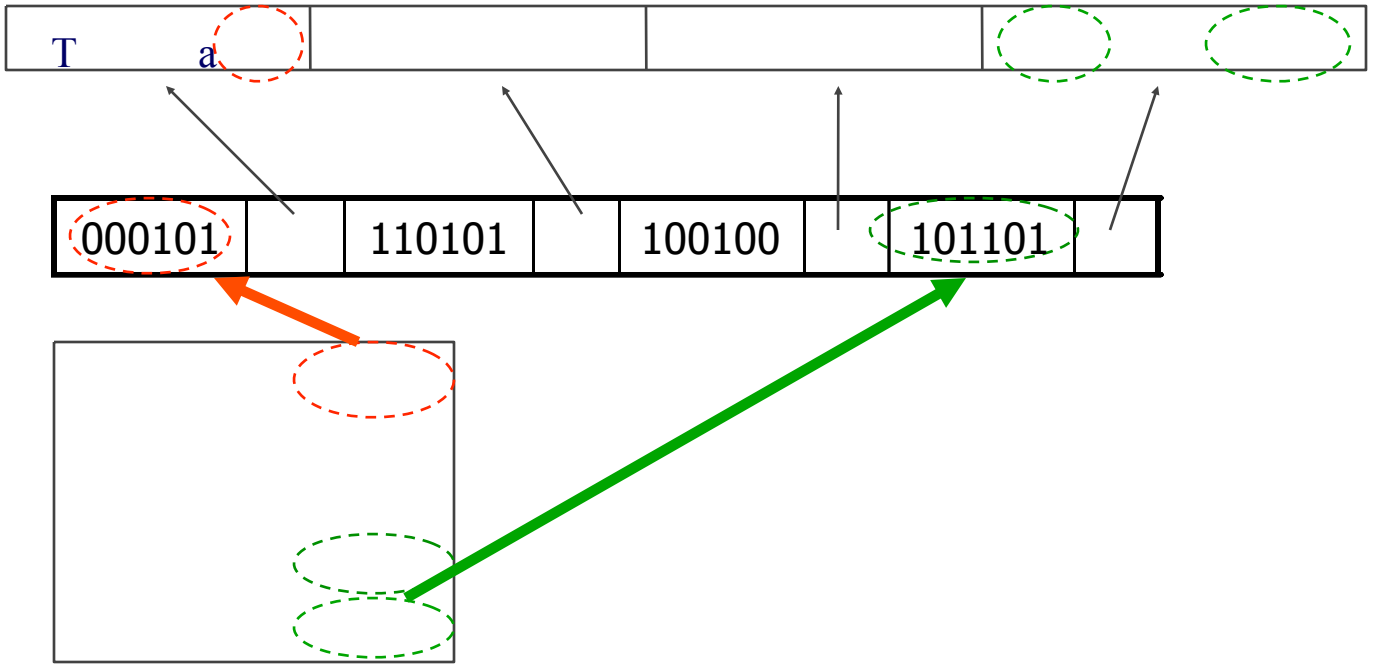
E a :


T a . A a a . W a a
000101
110101
100100
101101
T a
( ) = 000101
(a) = 110000
( ) = 100100
(a) = 001100
( ) = 100001
S a



S a

E a :



S a

Fa P b

T a

!

T a

a .

A S a b a a

b .

bab .

S a

S a

1. F a , Ha a b a W.
2. F a ,
   1) Ha a b a .
   2) B OR a a a b a W.
3. C a W b a B a b .
   I a b W a a B, b a a .
4. F a a a b , a a a b a a a a .

C

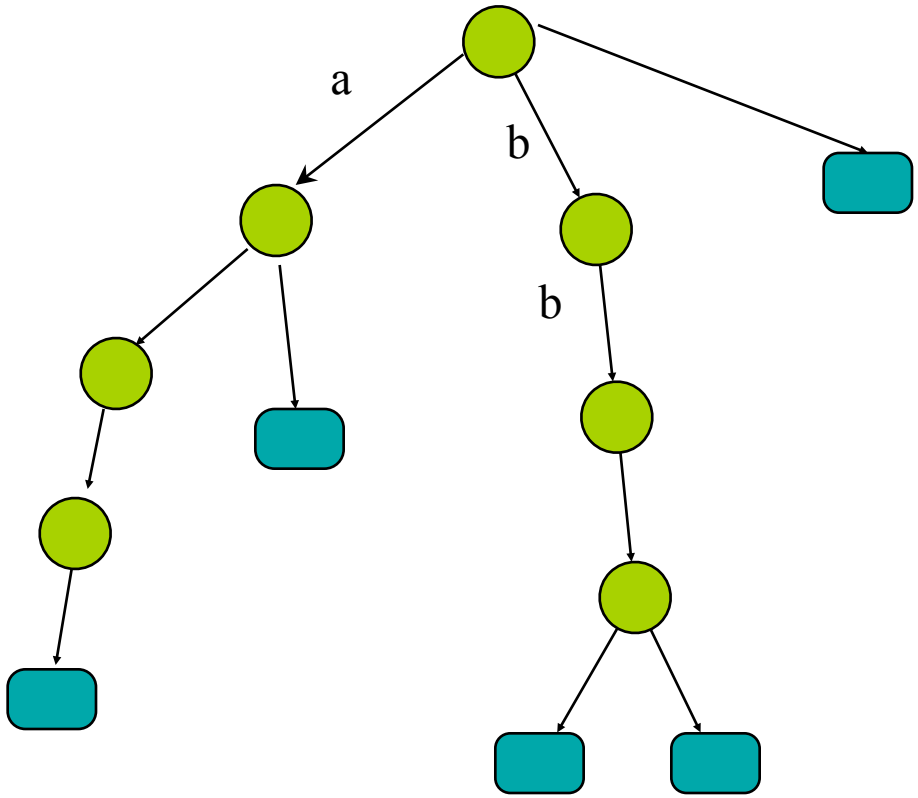
1. C b .
2. G a a a a b .
   T b OR a a b .

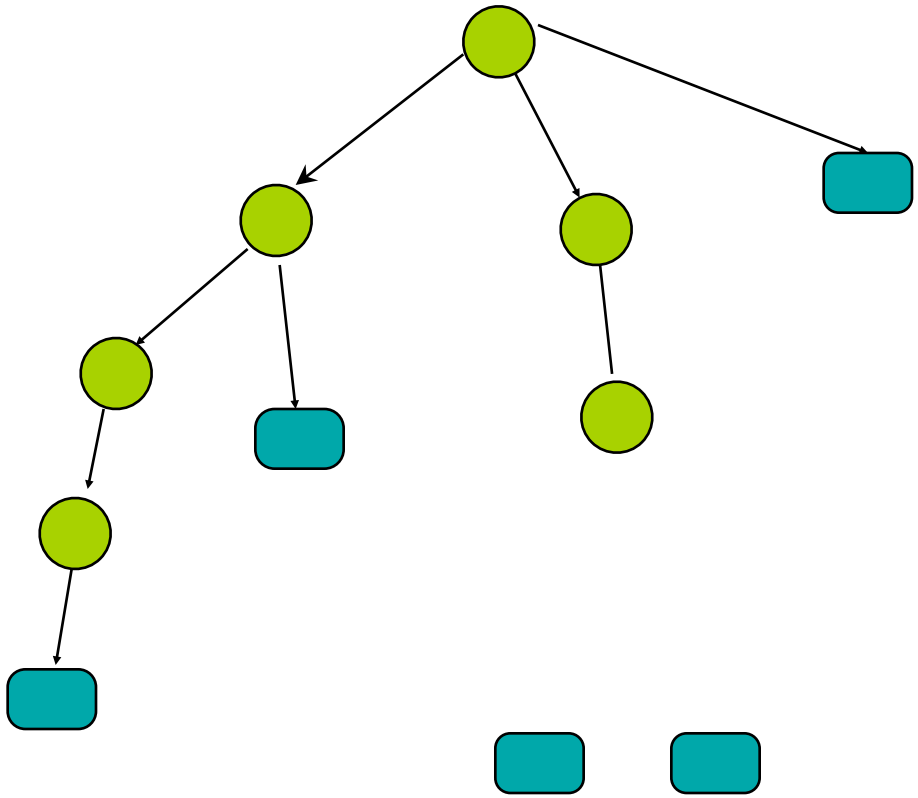
S a a a

T

A a .

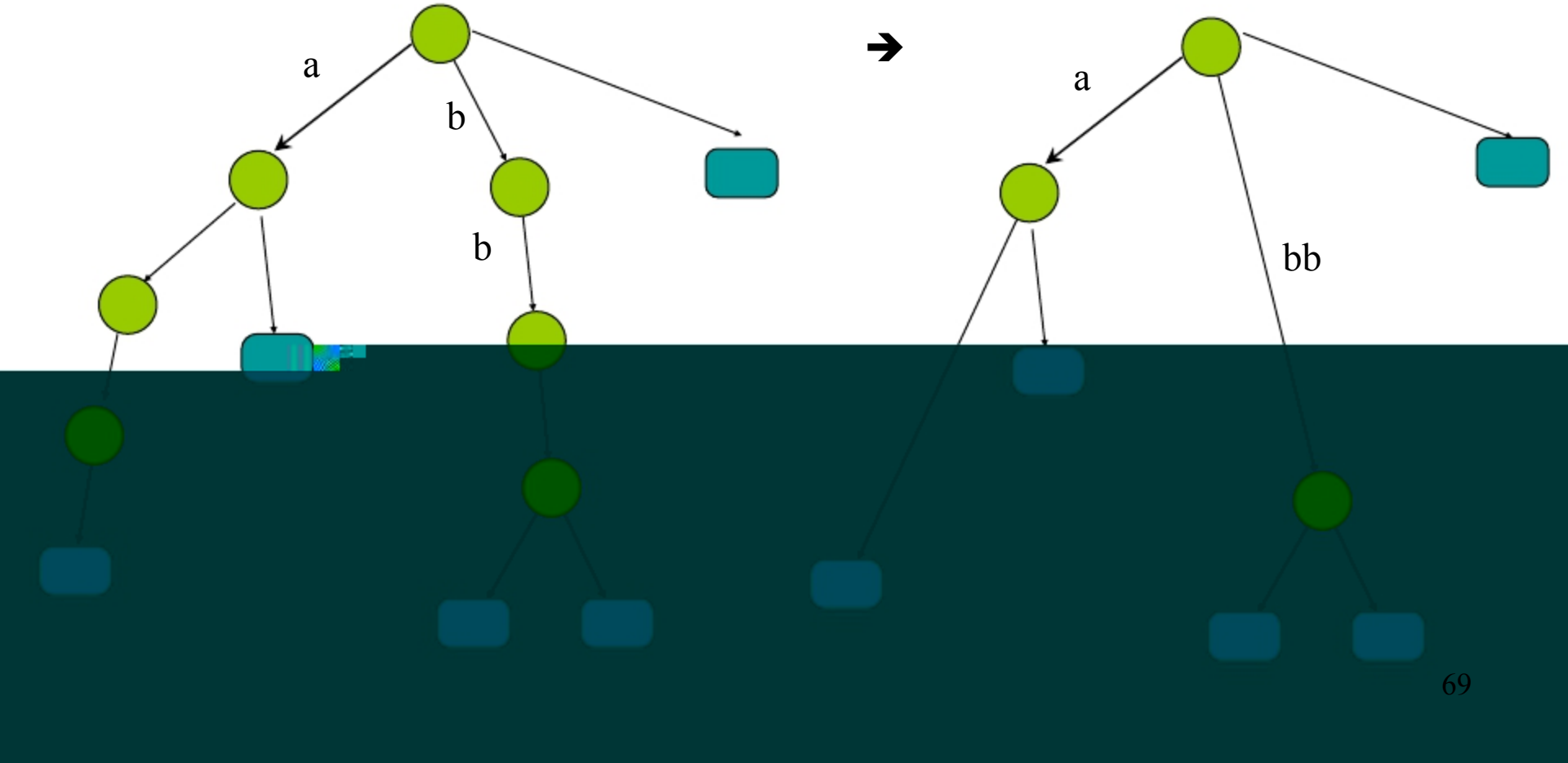
a
a
bb
bb





C T

C a , ab b

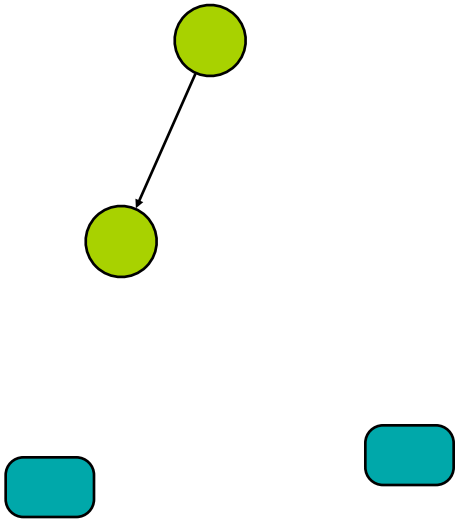


S

G a a a

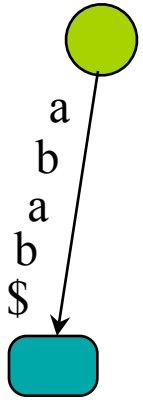
a

T a -

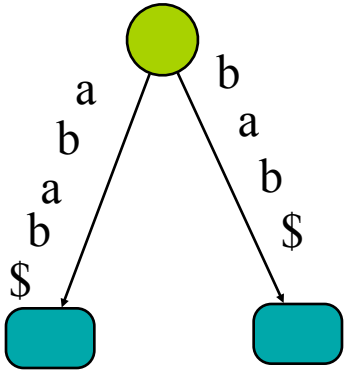


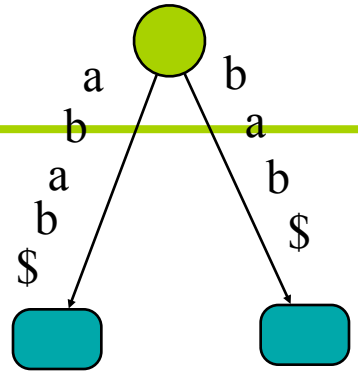
T a a b a S

P a

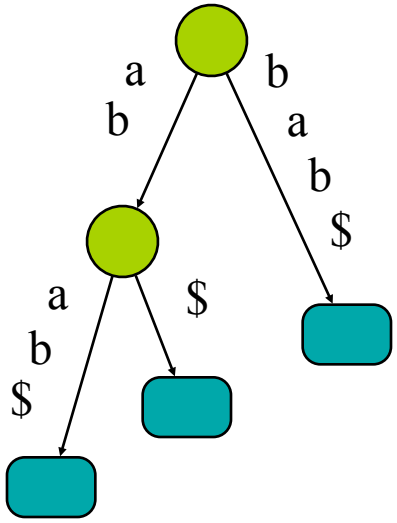


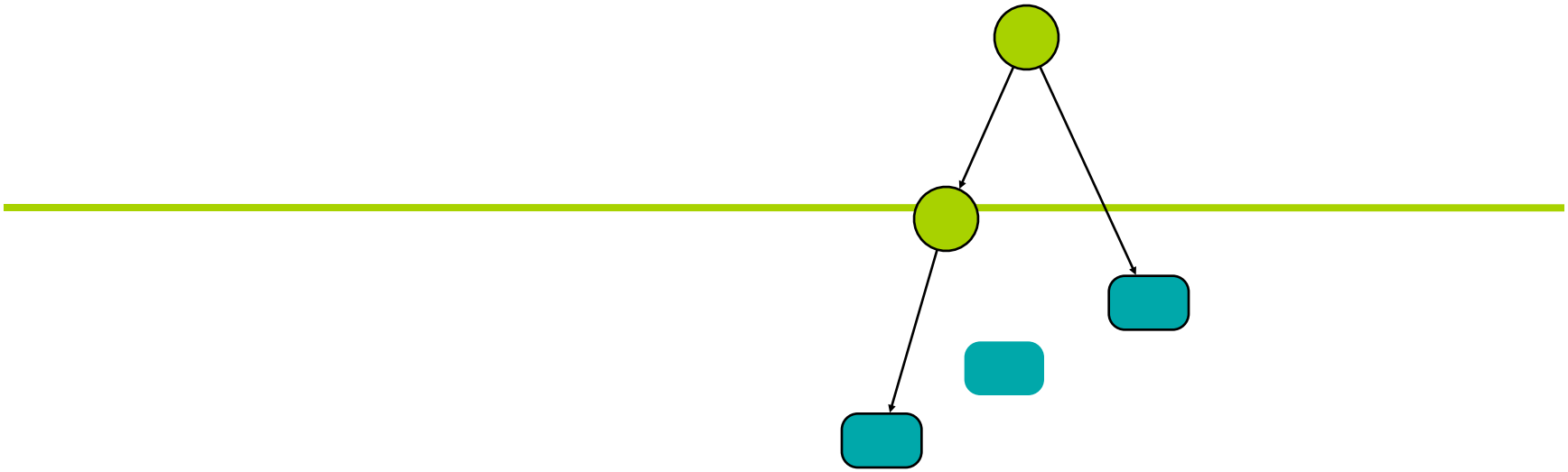
P bab$

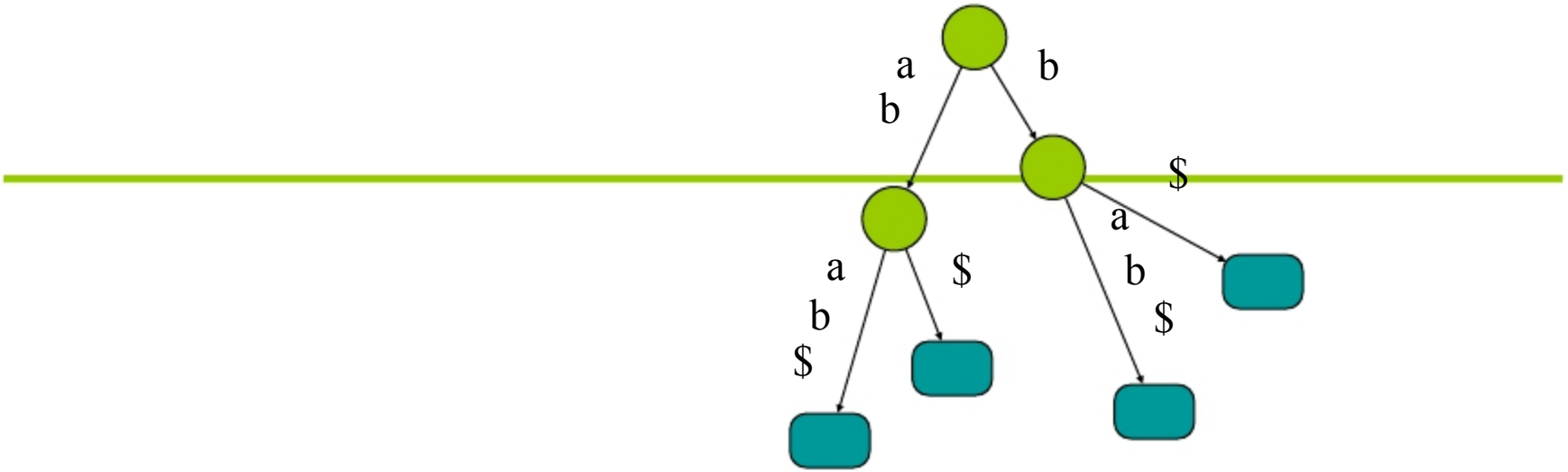




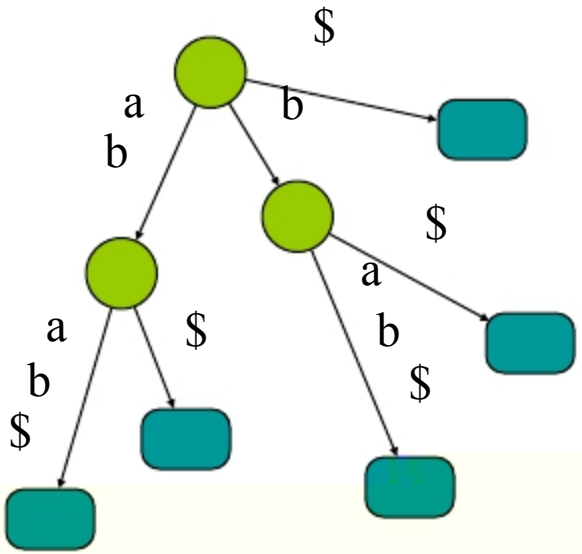
P ab$

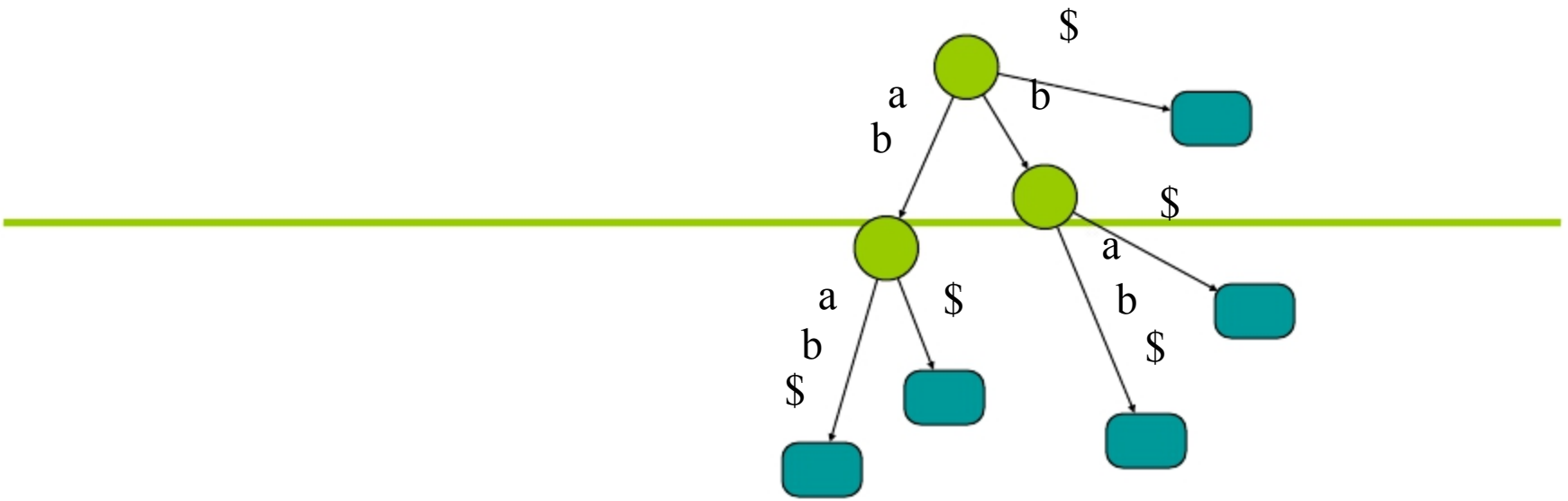




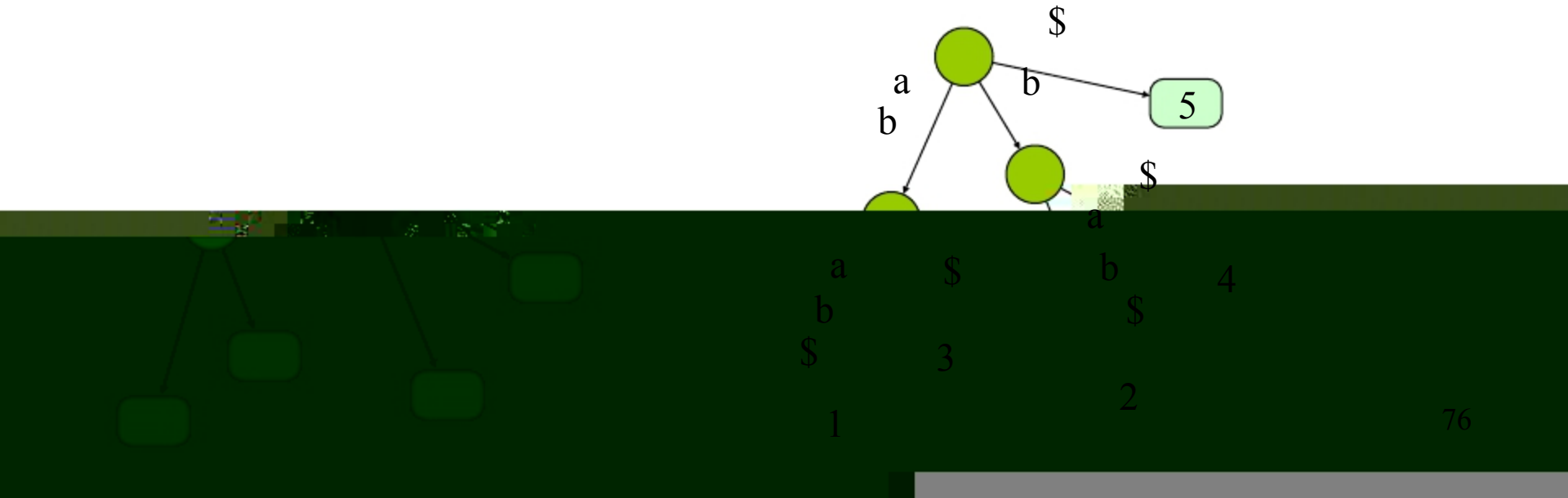


P $





W a ab a a a . .



A a

Ta O(2) b .

W O( )

W a a ?

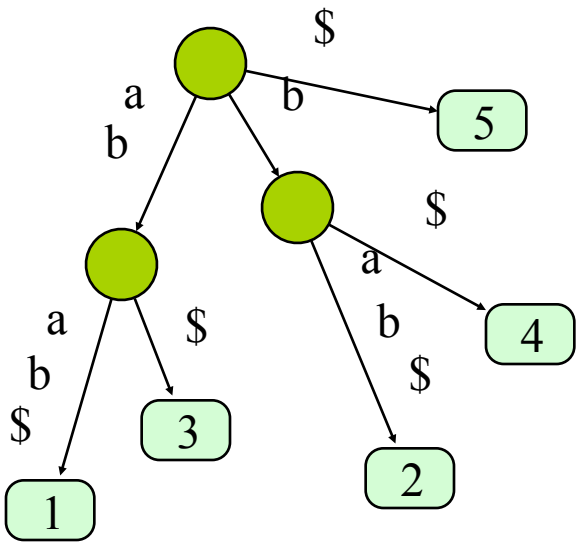
E a a :

Given a Text T, |T| = n, preprocess it such that when a pattern P, |P|=m, arrives you can quickly decide when it occurs in T.

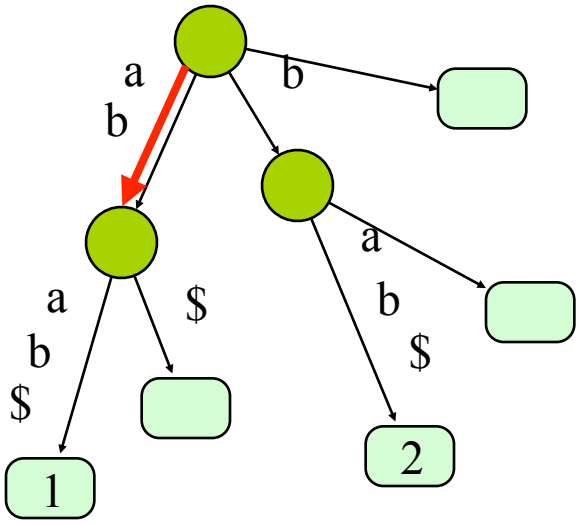
W e may also want to find all occurrences of P in T

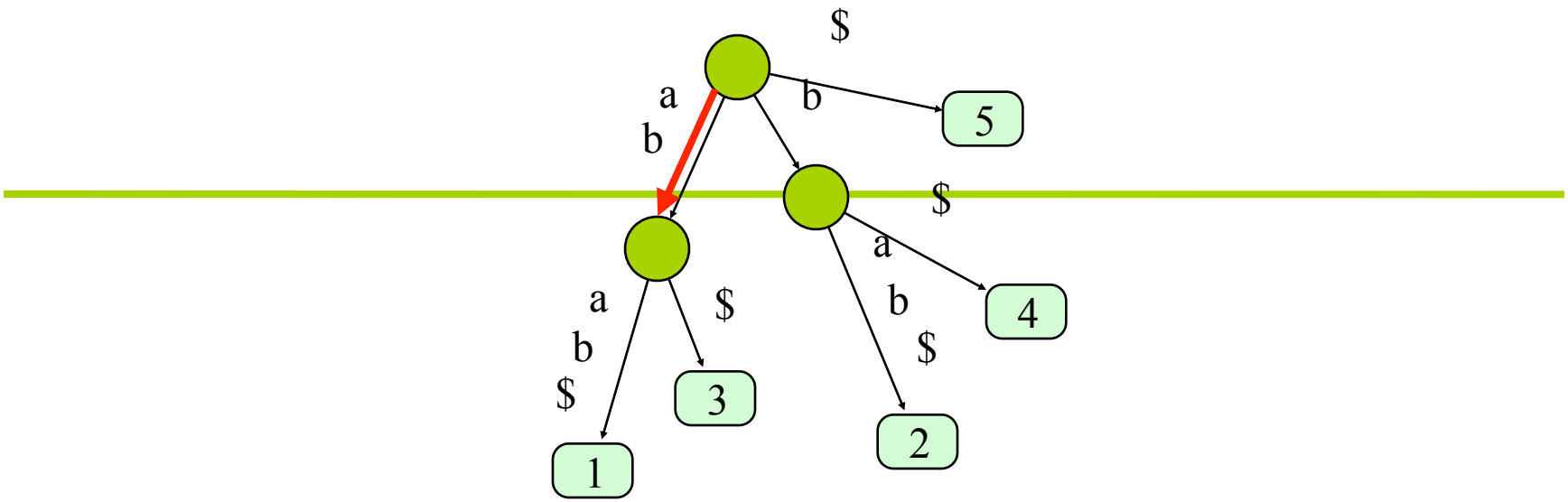
E a a

I b a O( )



G a a P =


a
b
b
a
$
b
$
a
b
$
1
2



I a a a .

Ea a b b a a .

B a b a O(+)

G a

G a S a a
S a a
∈ S

T a - a a
a a , a $, a

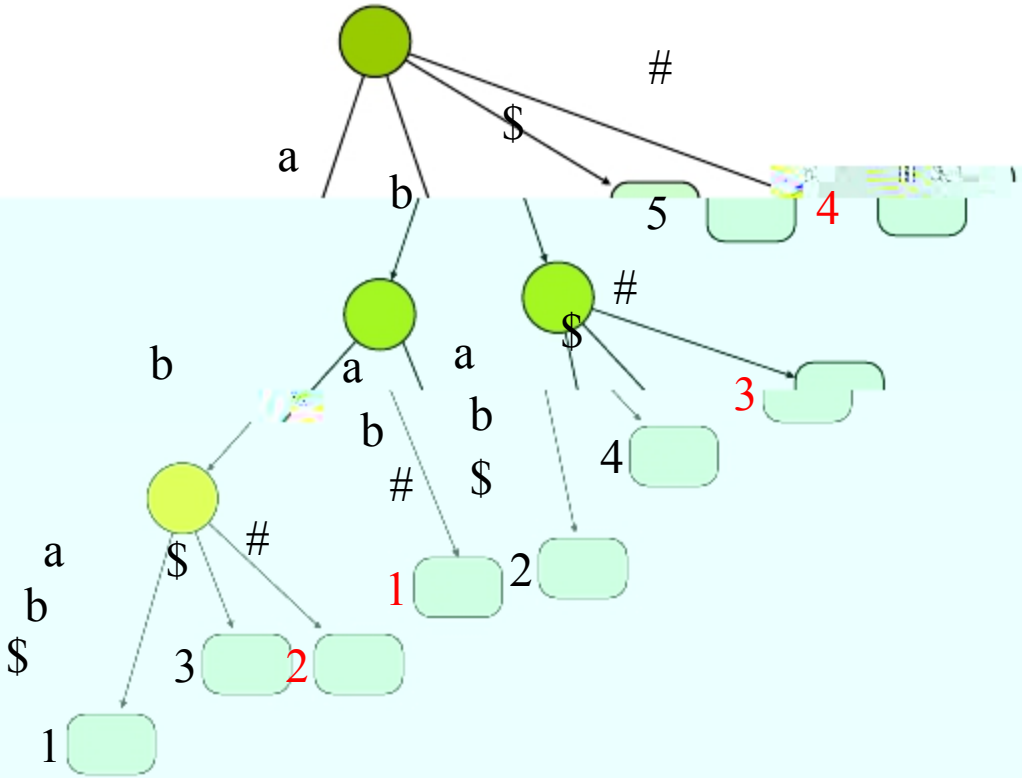
T a a a a
S a a a a a

G a (E a)

L 1=abab a 2=aab a a

1a 2

$	#
b$	b#
ab$	ab#
bab$	aab#
abab$	



S a a ?

Ma a a a a a a aba

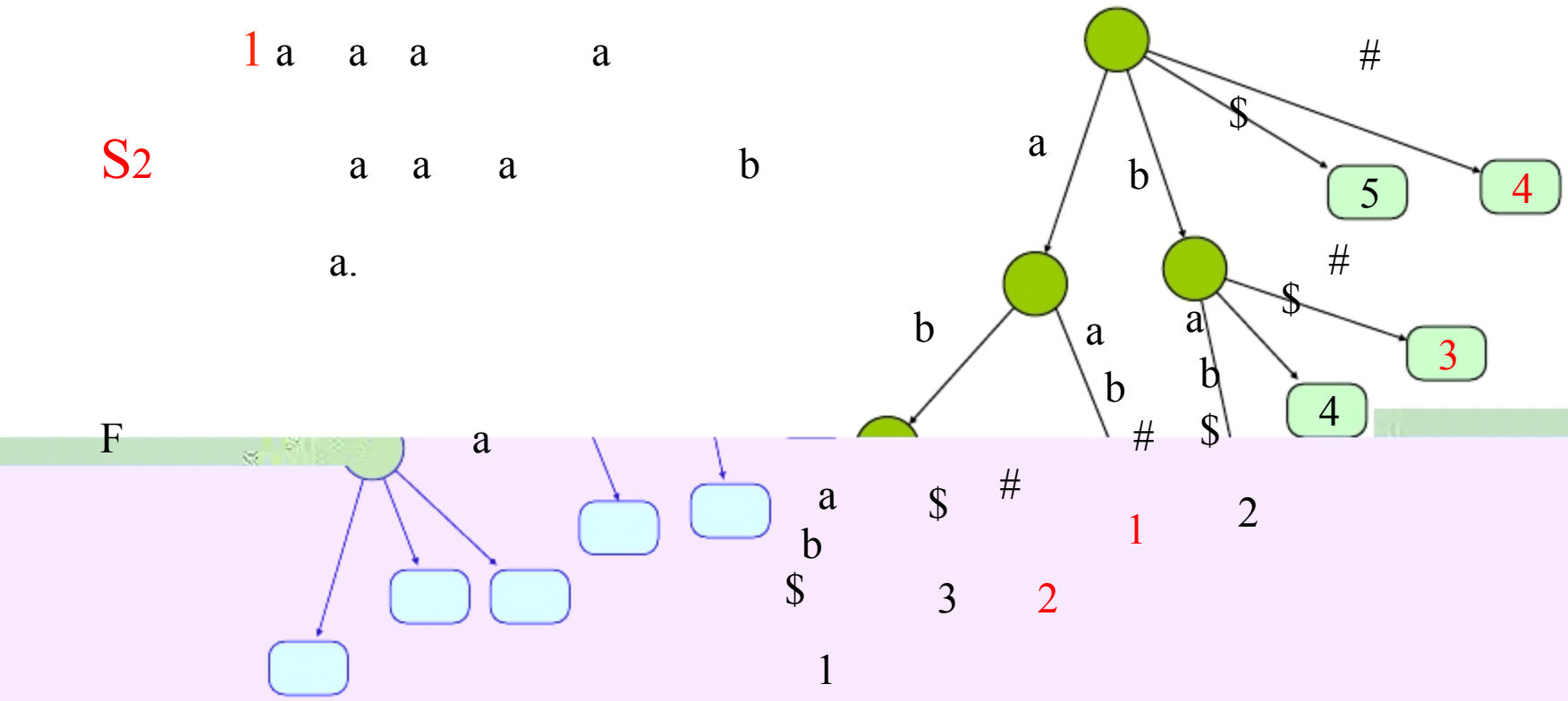
L b ( )

E a a a

1 a a a a

S2 a a a b

a.



L b ( )

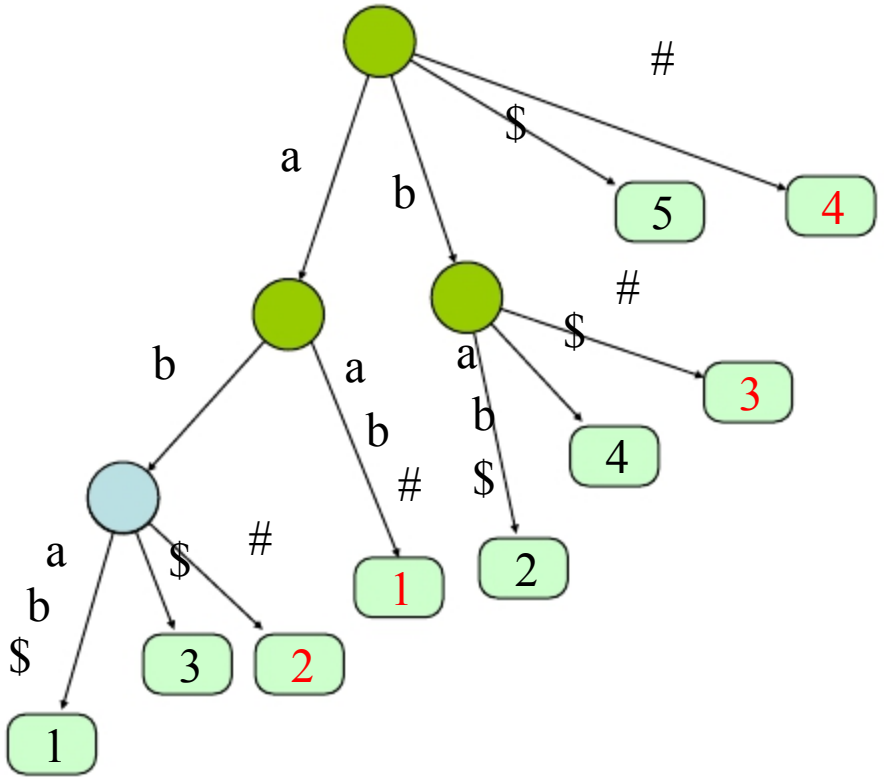
E a a a

1 a a a a

S2 a a a b

a a.

F a

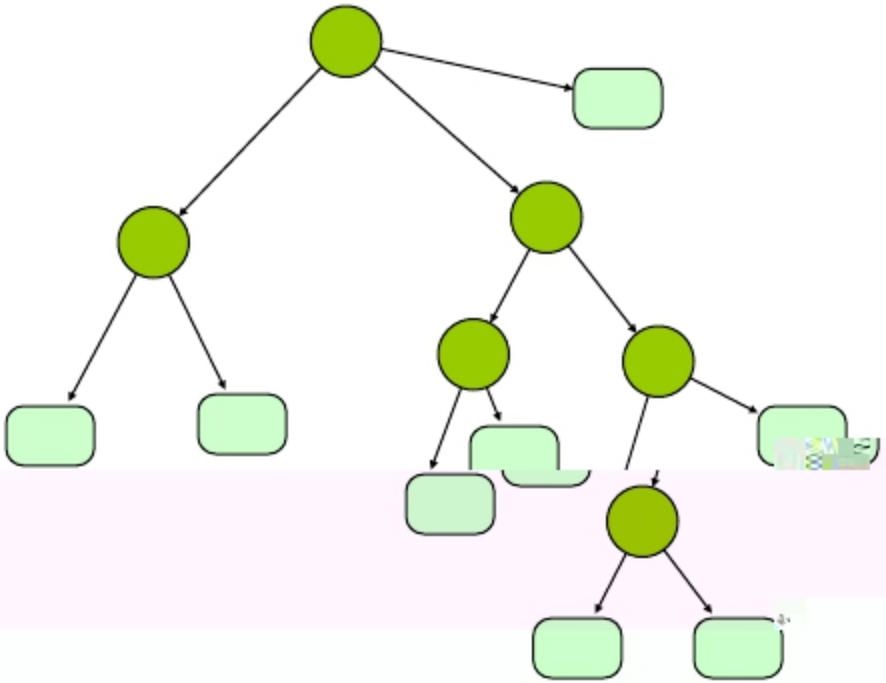


L a

A a b a

a a a

LCA

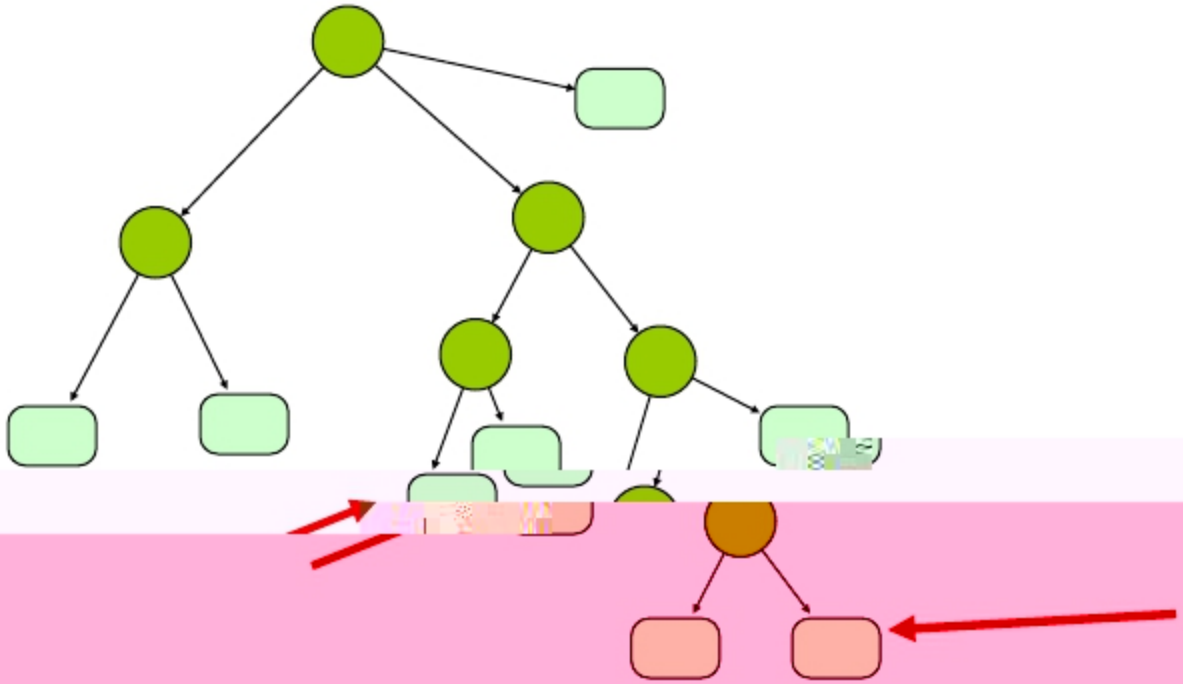


L a

A a b a

a a a

LCA

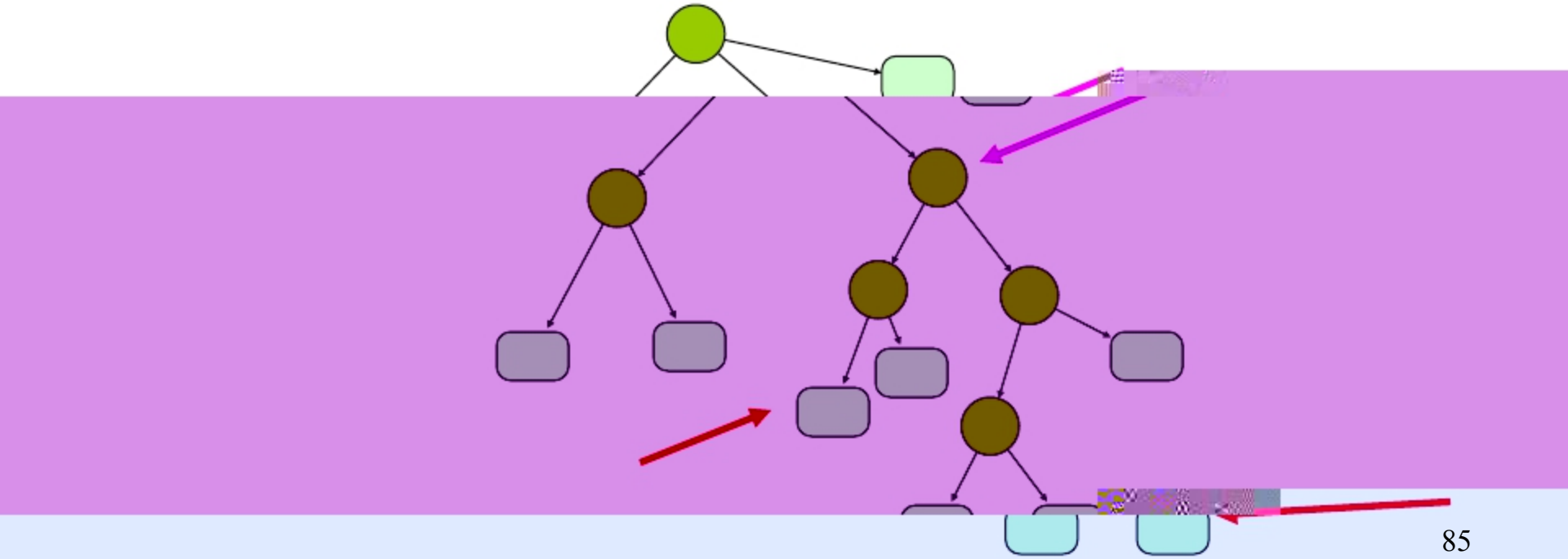


L a

A a b a

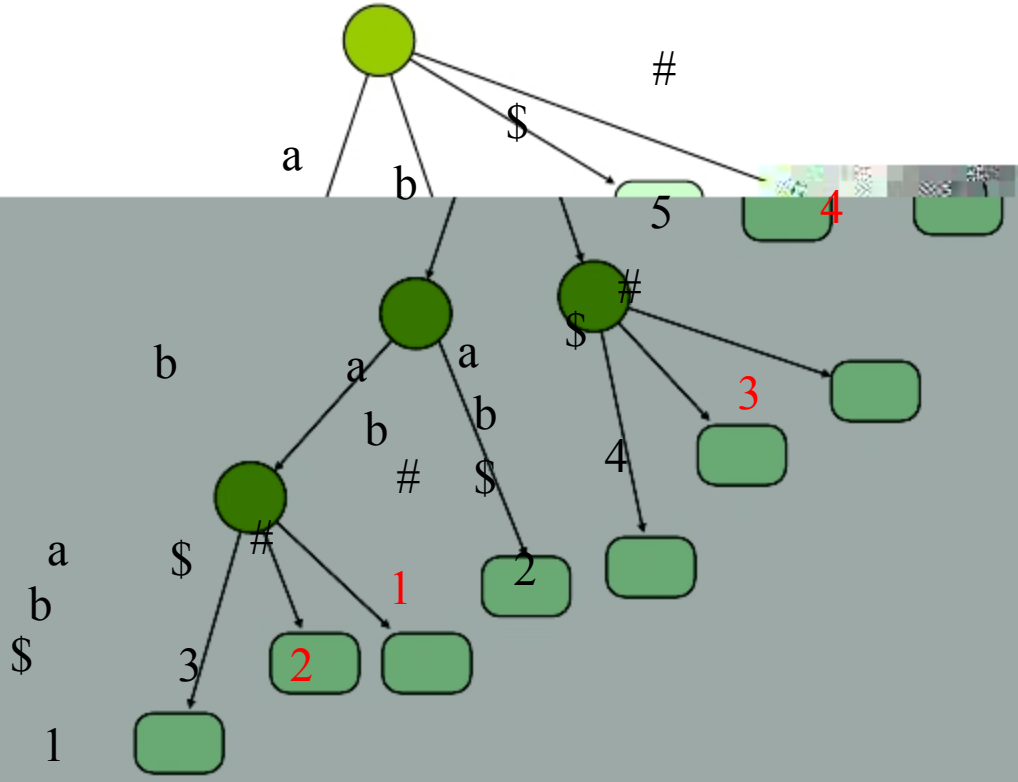
a a a

LCA



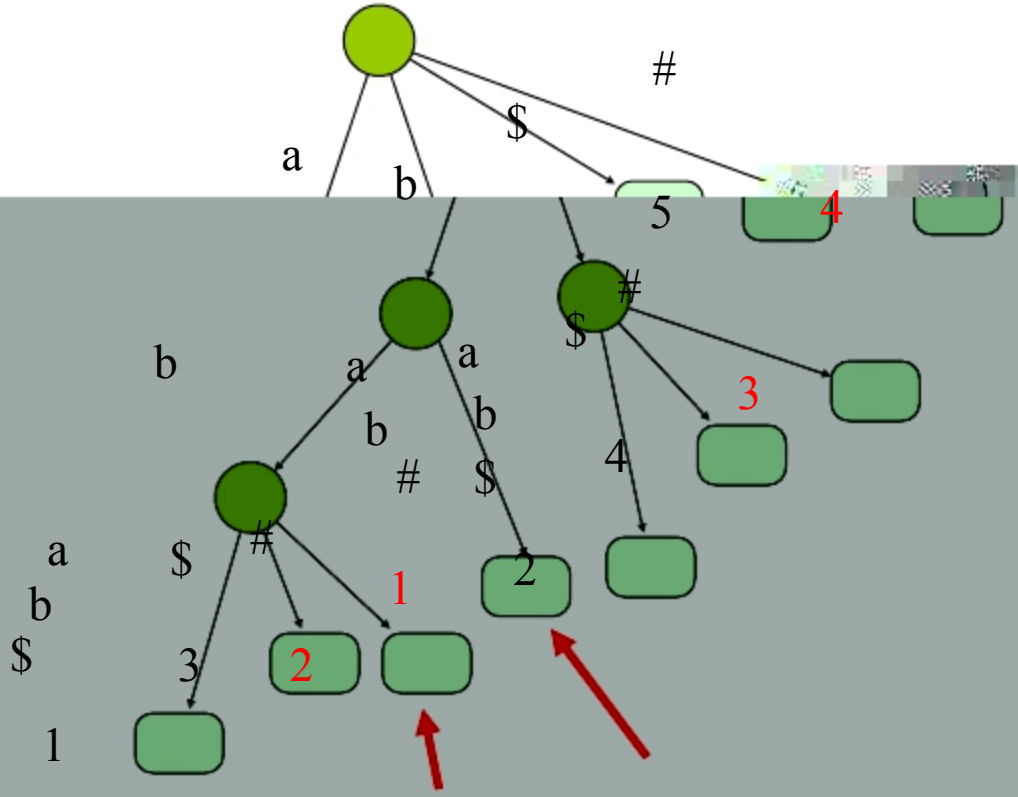
W ?

T LCA a
(LCP) 2



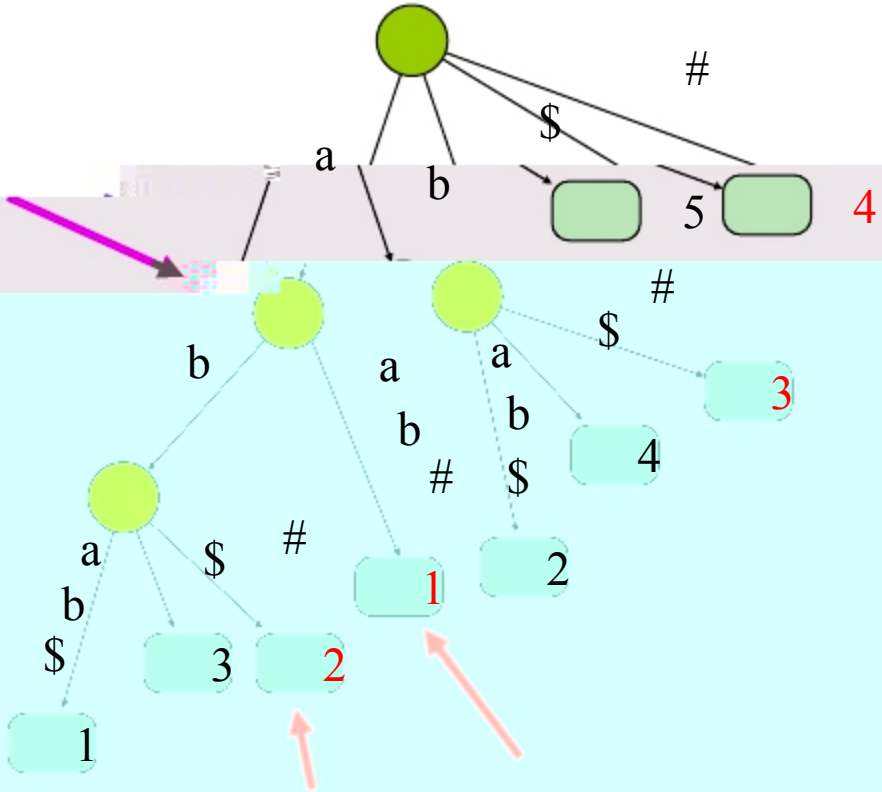
W ?

T LCA a
(LCP) 2



W ?

T LCA a
(LCP) 2



F a a a

A a : aabaa , baab

Wa a a a a a

L = baaba

T a a a b -1 a LCP a

a a - +1

Ma a a a

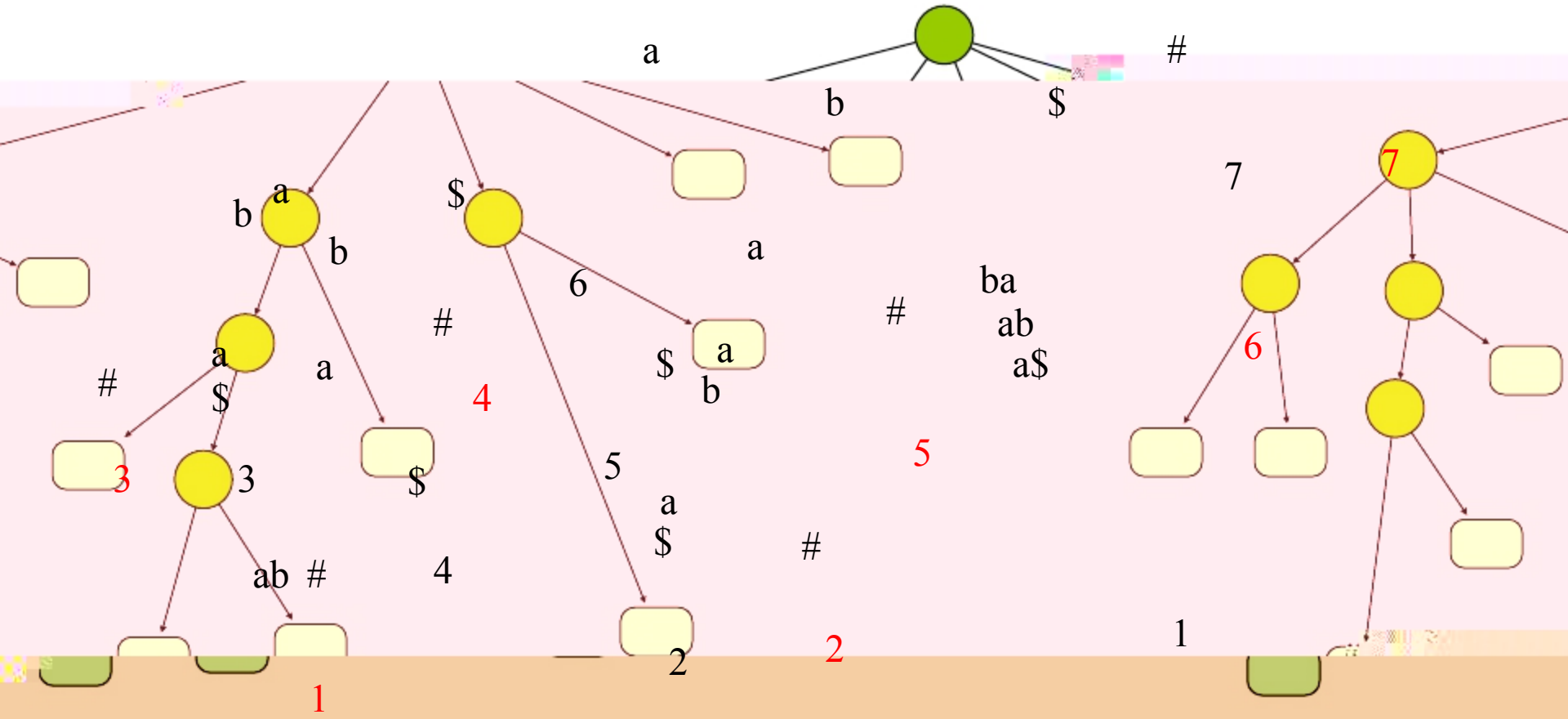
P a a a

= baaba$ a = abaab #

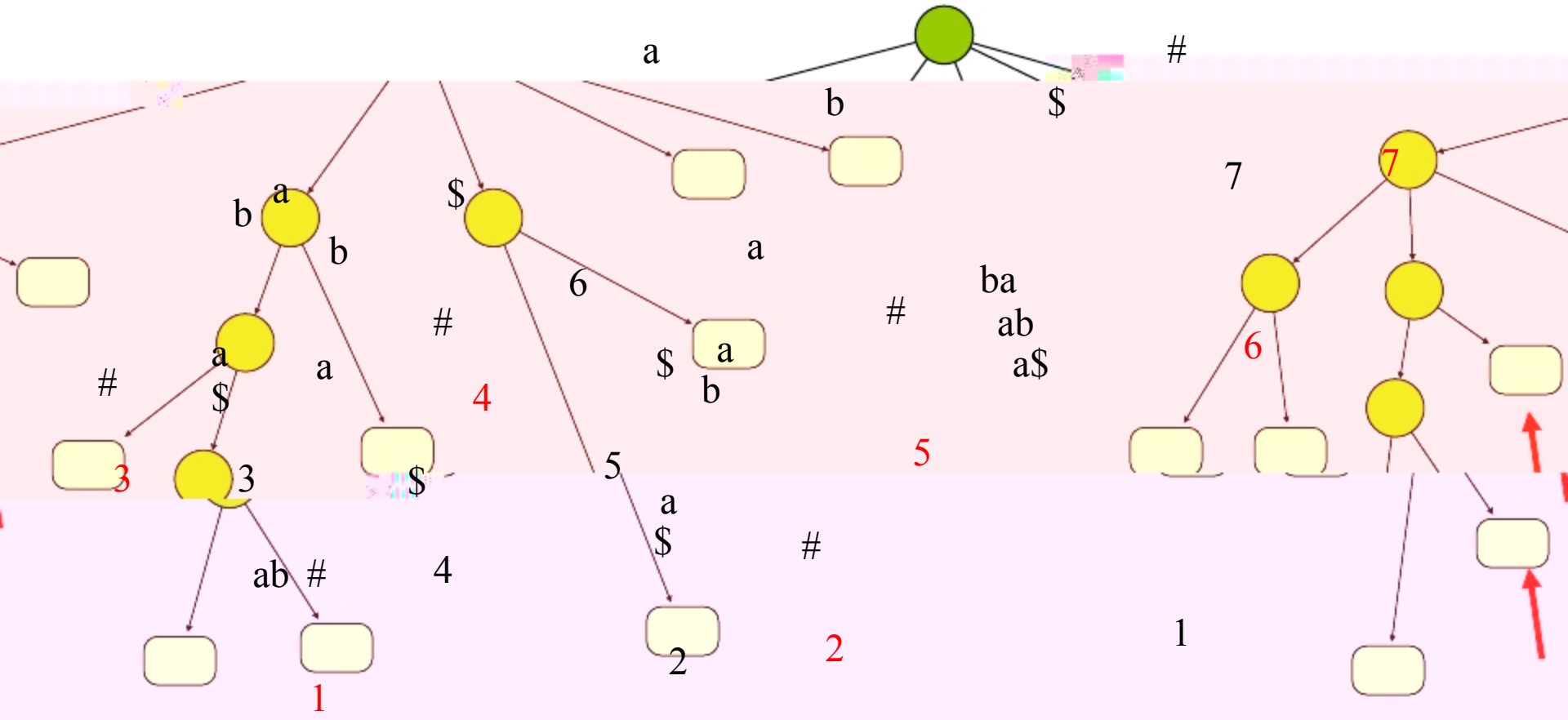
F LCA a

- +1

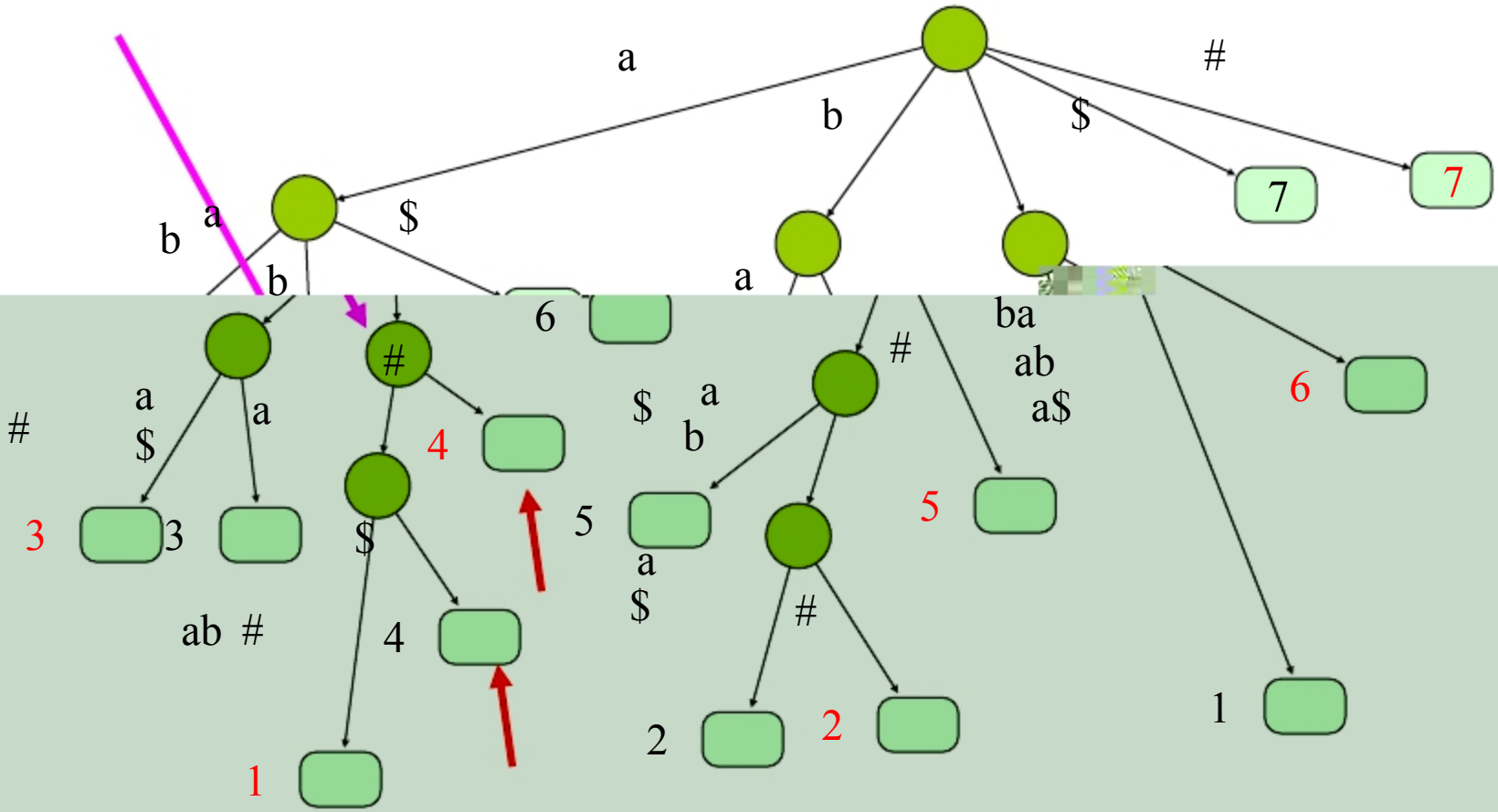
L = baaba$ = abaab



L = baaba$ = abaab #



L = baaba$ = abaab #



A a

O( ) a a

D a ba

S a a

I O( ) b a b

N a a a
a O(1) a
a a . Σ a

S a a

W a b
a a .

L = abab

S a a :

ab, abab, b, bab

T a a

3	1	4	2

H b ?

B a

T a DFS, a a

a

a a a .

O( )

H a a a ?

I P T a a
a a .

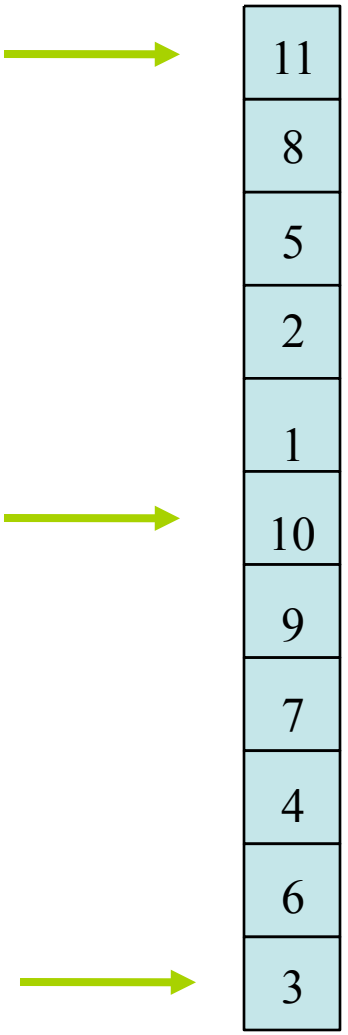
D a b a a a a

Ta O( )

E a

L S =

L P = a



S a

S

a a

.

S a a a

a a .

I a a , b a a a

b a b a a .

S a a a a b

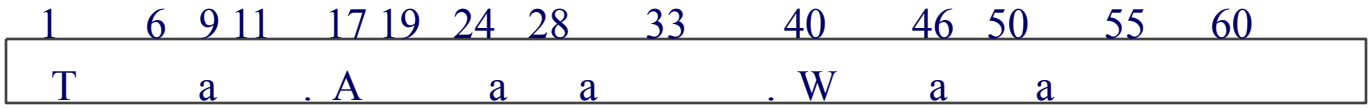
a a .

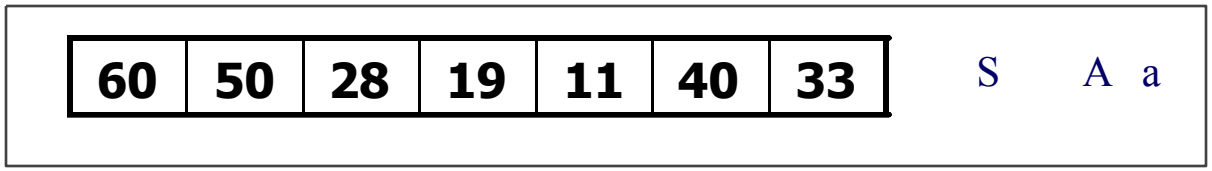
T a ,

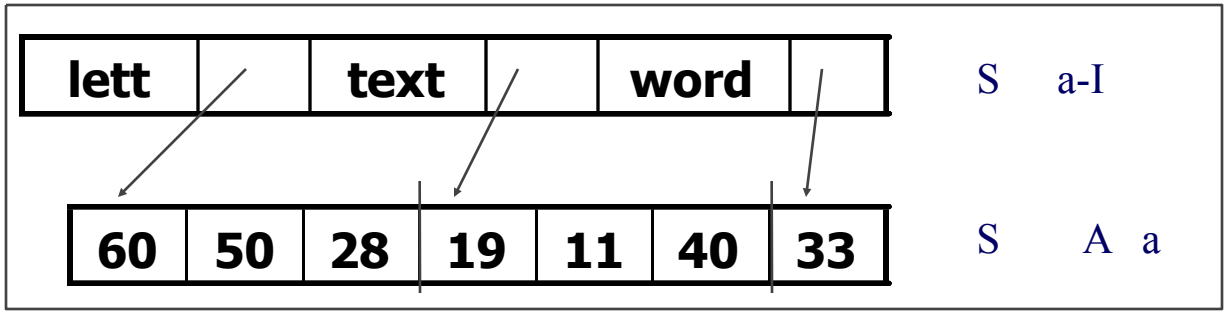
a a a b .

S a

E a

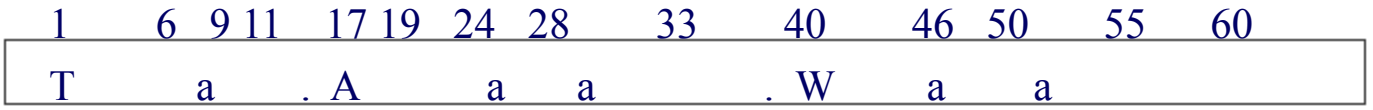

1 6 9 11 17 19 24 28 33 40 46 50 55 60
T a . A a a . W a a

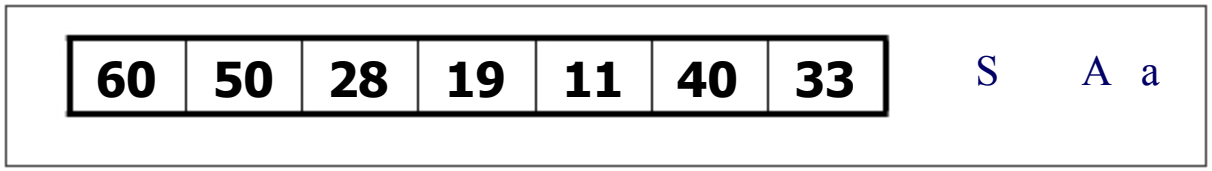

60 50 28 19 11 40 33
S A a

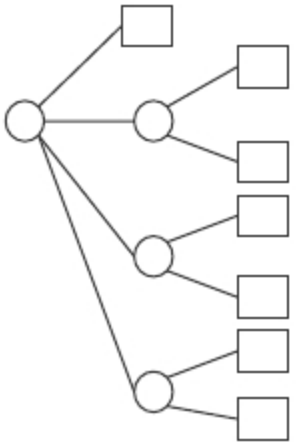

lett text word
S a-I
60 50 28 19 11 40 33
S A a

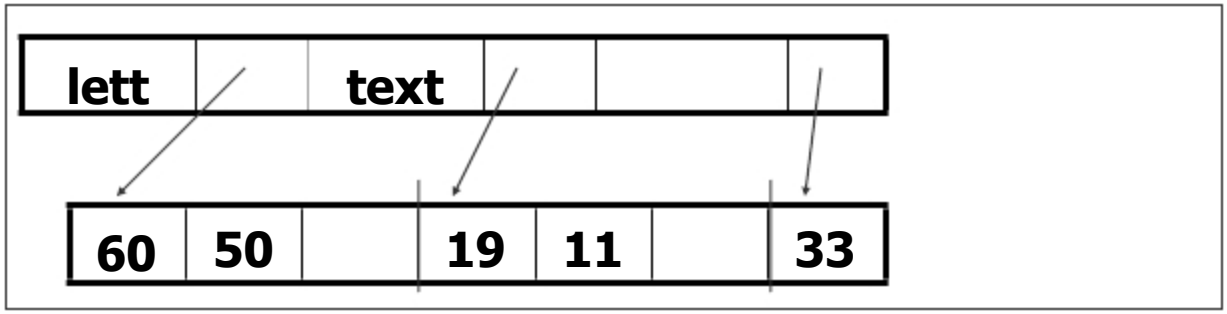
S a

E a


1 6 9 11 17 19 24 28 33 40 46 50 55 60
T a . A a a . W a a


60 50 28 19 11 40 33
S A a




lett
text
60 50 19 11 33