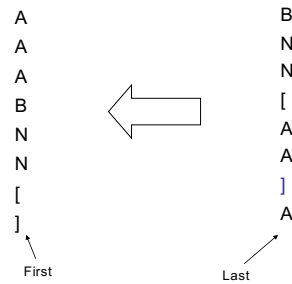


# COMP9319 Web Data Compression and Search

BWT revisit  
Backward Search overview

1

Recall: Last column = BWT

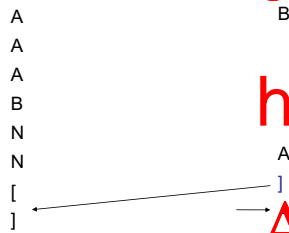


2

Assignment Project Exam Help

<https://eduassistpro.github.io/>

Add WeChat edu\_assist\_pro

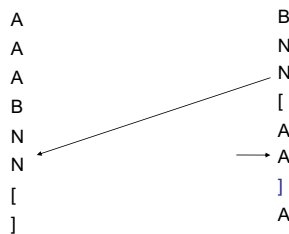


3



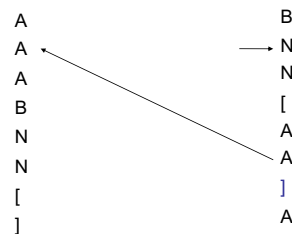
4

ANA]

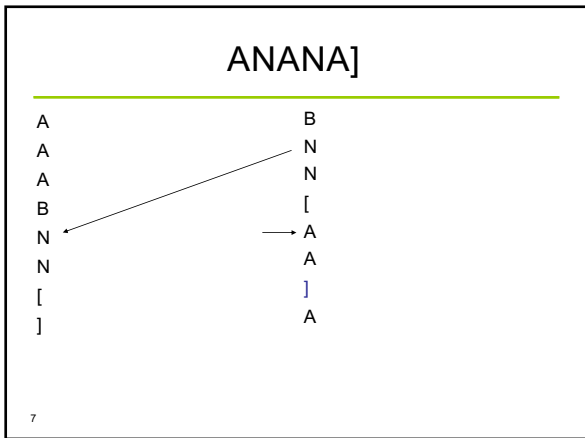


5

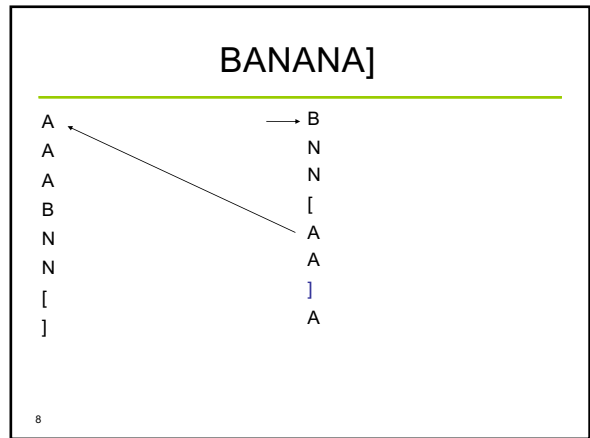
NANA]



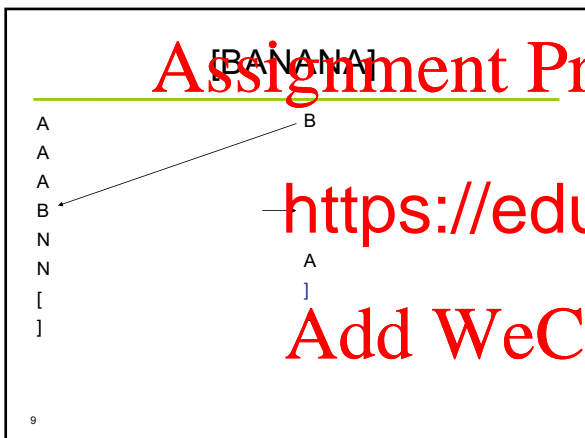
6



7



8



9

Example using C[] & Occ[]

Position	Symbol	# Matching
0		
1		
2		
3		
4		
5		
6		
7		

Symbol	# LessThan
A	0
B	3
N	4
[	6
]	7

10

10

**???????]**

Position	Symbol	# Matching
0	B	0
1	N	0
2	N	1
3	[	0
4	A	0
5	A	1
6	]	0
7	A	2

Symbol	# LessThan
A	0
B	3
N	4
[	6
]	7

11

11

**???????A]**

Position	Symbol	# Matching
0	B	0
1	N	0
2	N	1
3	[	0
4	A	0
5	A	1
6	]	0
7	A	2

Symbol	# LessThan
A	0
B	3
N	4
[	6
]	7

12

12

?????NA]

Position	Symbol	# Matching
0	B	0
1	N	0
2	N	1
3	[	0
4	A	0
5	A	1
6	]	0
7	A	2

Symbol	# LessThan
A	0
B	3
N	4
[	6
]	7

13

13

????ANA]

Position	Symbol	# Matching
0	B	0
1	N	0
2	N	1
3	[	0
4	A	0
5	A	1
6	]	0
7	A	2

Symbol	# LessThan
A	0
B	3
N	4
[	6
]	7

14

14

???NANA]

Position	Symbol	# Matching
0	B	0
1	N	0
2	N	1
3	[	0
4	A	0
5	A	1
6	]	0
7	A	2

Symbol	# LessThan
A	0
B	3
N	4
[	6
]	7

15

15

??NANA]

Position	Symbol	# Matching
0	B	0
1	N	0
2	N	1
3	[	0
4	A	0
5	A	1
6	]	0
7	A	2

Symbol	# LessThan
A	0
B	3
N	4
[	6
]	7

16

16

?BANANA]

Position	Symbol	# Matching
0	B	0
1	N	0
2	N	1
3	[	0
4	A	0
5	A	1
6	]	0
7	A	2

Symbol	# LessThan
A	0
B	3
N	4
[	6
]	7

17

17

[BANANA]

Position	Symbol	# Matching
0	B	0
1	N	0
2	N	1
3	[	0
4	A	0
5	A	1
6	]	0
7	A	2

Symbol	# LessThan
A	0
B	3
N	4
[	6
]	7

18

18

Assignment Project Exam Help

<https://eduassistpro.github.io/>

Add WeChat edu\_assist\_pro

### [BANANA]

Position	Symbol	# Matching
0	B	0
1	N	0
2	N	1
3	[	0
4	A	0
5	A	1
6	]	0
7	A	2

Occ / Rank

Symbol	# LessThan
A	0
B	3
N	4
[	6
]	7

C [ ]

19

### C[ ] & Occ()

Position	Symbol	# Matching
0	B	0
1	N	0
2	N	1
3	[	0
4	A	0
5	A	1
6	]	0
7	A	2

Occ(Symbol, Pos) => # Matching

Symbol	# LessThan
A	0
B	3
N	4
[	6
]	7

C[Symbol] => (startPos, endPos)

20

### C[ ] & Occ()

C[Symbol] => (startPos, endPos)

Occ(Symbol, Pos)

Can these two functions such that they can return the result in constant time ?  
Yes, have a precomputed table.  
Can they be precomputed efficiently ?  
Yes, a single pass.

21

### Backward Search for ANA

A	B
	N
	N
	[
	A
	A
N	

22

### Backward Search for NAN

A	B
A	N
A	N
B	[
N	A
N	A
[	]
]	A

23

### Backward Search for ANA

Position	Symbol	# Matching
0	B	0
1	N	0
2	N	1
3	[	0
4	A	0
5	A	1
6	]	0
7	A	2

Occ(Symbol, Pos) => # Matching

Symbol	# LessThan
A	0
B	3
N	4
[	6
]	7

C[Symbol] => (startPos, endPos)

24

### Backward Search for NAN

Position	Symbol	# Matching
0	B	0
1	N	0
2	N	1
3	[	0
4	A	0
5	A	1
6	]	0
7	A	2

Occ(Symbol, Pos) => # Matching

Symbol	# LessThan
A	0
B	3
N	4
[	6
]	7

C[Symbol] => (startPos, endPos)

25

25

### Why not Forward Search: ANA

A	B
A	N
A	N
B	[
N	A
N	A
[	]
]	A

26

26

# Assignment Project Exam Help

<https://eduassistpro.github.io/>

Add WeChat edu\_assist\_pro