# COMP9319 Web Data Compression and Search

An Occ Implementation,
RLFM (Compressed FM Index) Revisit

1

1

An example Occ implementation

2

2

## FM Index (L(x) = c)

| | F | L | c | C | |
|---|---|---|---|---|---|
| 0 | # | i | 0 | # | 0 |
| 1 | i | p | 0 | i | 1 |
| 2 | i | s | 0 | m | 5 |
| 3 | i | s | 1 | p | 6 |
| 4 | i | m | 0 | s | 8 |
| 5 | m | # | 0 | | |
| 6 | p | p | 1 | | |
| 7 | p | i | 1 | | |
| 8 | s | s | 2 | | |
| 9 | s | s | 3 | | |
| 10 | s | i | 2 | | |
| 11 | s | i | 3 | | |

3

3

## FM Index (when reversing from L[5])

| | F | L | c | | C | |
|---|---|---|---|---|---|---|
| 0 | # | i | 0 | | # | 0 |
| 1 | i | p | 0 | | i | 1 |
| 2 | i | s | 0 | | m | 5 |
| 3 | i | s | 1 | | p | 6 |
| 4 | i | m | 0 | LF[5] = 0+0 = 0, i | s | 8 |
| 5 | m | # | 0 | | | |
| 6 | p | p | 1 | | | |
| 7 | p | i | 1 | | | |
| 8 | s | s | 2 | | | |
| 9 | s | s | 3 | | | |
| 10 | s | i | 2 | | | |
| 11 | s | i | 3 | | | |

4

4

## FM Index (when reversing from L[5])

| | F | L | c | | C | |
|---|---|---|---|---|---|---|
| 0 | # | i | 0 | | # | 0 |
| 1 | i | p | 0 | | i | 1 |
| 2 | i | s | 0 | | m | 5 |
| 3 | i | s | 1 | | p | 6 |
| 4 | i | m | 0 | LF[5] = 0+0 = 0, i | s | 8 |
| 5 | m | # | 0 | LF[0] = 1+0 = 1, p | | |
| 6 | p | p | 1 | LF[1] = 6+0 = 6, p | | |
| 7 | p | i | 1 | LF[6] = 6+1 = 7, i | | |
| 8 | s | s | 2 | LF[7] = 1+1 = 2, s | | |
| 9 | s | s | 3 | LF[2] = 8+0 = 8, s | | |
| 10 | s | i | 2 | LF[8] = 8+2 = 10, i | | |
| 11 | s | i | 3 | LF[10] = 1+2 = 3, s | | |
| | | | | LF[3] = 8+1 = 9, s | | |
| | | | | LF[9] = 8+3 = 11, i | | |
| | | | | LF[11] = 1+3 = 4, m | | |

5

5

## FM Index (L(x) ≠ c)

| | F | L | i | m | p | s | C | |
|---|---|---|---|---|---|---|---|---|
| 0 | # | i | 1 | 0 | 0 | 0 | # | 0 |
| 1 | i | p | 1 | 0 | 1 | 0 | i | 1 |
| 2 | i | s | 1 | 0 | 1 | 1 | m | 5 |
| 3 | i | s | 1 | 0 | 1 | 2 | p | 6 |
| 4 | i | m | 1 | 1 | 1 | 2 | s | 8 |
| 5 | m | # | 1 | 1 | 1 | 2 | | |
| 6 | p | p | 1 | 1 | 2 | 2 | | |
| 7 | p | i | 2 | 1 | 2 | 2 | | |
| 8 | s | s | 2 | 1 | 2 | 3 | | |
| 9 | s | s | 2 | 1 | 2 | 4 | | |
| 10 | s | i | 3 | 1 | 2 | 4 | | |
| 11 | s | i | 4 | 1 | 2 | 4 | | |

6

6

## FM Index (con't)

pssi

| | F | L | i | m | p | s | C |
|---|---|---|---|---|---|---|---|
| 0 | # | i | 1 | 0 | 0 | 0 | # 0 |
| 1 | i | p | 1 | 0 | 1 | 0 | i 1 |
| 2 | i | s | 1 | 0 | 1 | 1 | m 5 |
| 3 | i | s | 1 | 0 | 1 | 2 | p 6 |
| 4 | i | m | 1 | 1 | 1 | 2 | s 8 |
| 5 | m | # | 1 | 1 | 1 | 2 | |
| 6 | p | p | 1 | 1 | 2 | 2 | |
| 7 | p | i | 2 | 1 | 2 | 2 | |
| 8 | s | s | 2 | 1 | 2 | 3 | |
| 9 | s | s | 2 | 1 | 2 | 4 | |
| 10 | s | i | 3 | 1 | 2 | 4 | |
| 11 | s | i | 4 | 1 | 2 | 4 | |

Fst=1
Lst=4

7

7

## FM Index (con't)

pssi

| | F | L | i | m | p | s | C |
|---|---|---|---|---|---|---|---|
| 0 | # | i | 1 | 0 | 0 | 0 | # 0 |
| 1 | i | p | 1 | 0 | 1 | 0 | i 1 |
| 2 | i | s | 1 | 0 | 1 | 1 | m 5 |
| 3 | i | s | 1 | 0 | 1 | 2 | p 6 |
| 4 | i | m | 1 | 1 | 1 | 2 | s 8 |
| 5 | m | # | 1 | 1 | 1 | 2 | |
| 6 | p | p | 1 | 1 | 2 | 2 | |
| 7 | p | i | 2 | 1 | 2 | 2 | |
| 8 | s | s | 2 | 1 | 2 | 3 | |
| 9 | s | s | 2 | 1 | 2 | 4 | |
| 10 | s | i | 3 | 1 | 2 | 4 | |
| 11 | s | i | 4 | 1 | 2 | 4 | |

Fst=8+0
Lst=(8+2)-1

8

8

## FM Index (con't)

pssi

| | F | L | i | m | p | s | C |
|---|---|---|---|---|---|---|---|
| 0 | # | i | 1 | 0 | 0 | 0 | # 0 |
| 1 | i | p | 1 | 0 | 1 | 0 | i 1 |
| 2 | i | s | 1 | 0 | 1 | 1 | m 5 |
| 3 | i | s | 1 | 0 | 1 | 2 | p 6 |
| 4 | i | m | 1 | 1 | 1 | 2 | s 8 |
| 5 | m | # | 1 | 1 | 1 | 2 | |
| 6 | p | p | 1 | 1 | 2 | 2 | |
| 7 | p | i | 2 | 1 | 2 | 2 | |
| 8 | s | s | 2 | 1 | 2 | 3 | |
| 9 | s | s | 2 | 1 | 2 | 4 | |
| 10 | s | i | 3 | 1 | 2 | 4 | |
| 11 | s | i | 4 | 1 | 2 | 4 | |

Fst=8+0
Lst=(8+2)-1

9

9

## FM Index (con't)

pssi

| | F | L | i | m | p | s | C |
|---|---|---|---|---|---|---|---|
| 0 | # | i | 1 | 0 | 0 | 0 | # 0 |
| 1 | i | p | 1 | 0 | 1 | 0 | i 1 |
| 2 | i | s | 1 | 0 | 1 | 1 | m 5 |
| 3 | i | s | 1 | 0 | 1 | 2 | p 6 |
| 4 | i | m | 1 | 1 | 1 | 2 | s 8 |
| 5 | m | # | 1 | 1 | 1 | 2 | |
| 6 | p | p | 1 | 1 | 2 | 2 | |
| 7 | p | i | 2 | 1 | 2 | 2 | |
| 8 | s | s | 2 | 1 | 2 | 3 | |
| 9 | s | s | 2 | 1 | 2 | 4 | |
| 10 | s | i | 3 | 1 | 2 | 4 | |
| 11 | s | i | 4 | 1 | 2 | 4 | |

Fst=8+2
Lst=(8+4)-1

10

10

## FM Index (con't)

pssi

| | F | L | i | m | p | s | C |
|---|---|---|---|---|---|---|---|
| 0 | # | i | 1 | 0 | 0 | 0 | # 0 |
| 1 | i | p | 1 | 0 | 1 | 0 | i 1 |
| 2 | i | s | 1 | 0 | 1 | 1 | m 5 |
| 3 | i | s | 1 | 0 | 1 | 2 | p 6 |
| 4 | i | m | 1 | 1 | 1 | 2 | s 8 |
| 5 | m | # | 1 | 1 | 1 | 2 | |
| 6 | p | p | 1 | 1 | 2 | 2 | |
| 7 | p | i | 2 | 1 | 2 | 2 | |
| 8 | s | s | 2 | 1 | 2 | 3 | |
| 9 | s | s | 2 | 1 | 2 | 4 | |
| 10 | s | i | 3 | 1 | 2 | 4 | |
| 11 | s | i | 4 | 1 | 2 | 4 | |

Fst=8+2
Lst=(8+4)-1

11

11

## FM Index (con't)

pssi

| | F | L | i | m | p | s | C |
|---|---|---|---|---|---|---|---|
| 0 | # | i | 1 | 0 | 0 | 0 | # 0 |
| 1 | i | p | 1 | 0 | 1 | 0 | i 1 |
| 2 | i | s | 1 | 0 | 1 | 1 | m 5 |
| 3 | i | s | 1 | 0 | 1 | 2 | p 6 |
| 4 | i | m | 1 | 1 | 1 | 2 | s 8 |
| 5 | m | # | 1 | 1 | 1 | 2 | |
| 6 | p | p | 1 | 1 | 2 | 2 | |
| 7 | p | i | 2 | 1 | 2 | 2 | |
| 8 | s | s | 2 | 1 | 2 | 3 | |
| 9 | s | s | 2 | 1 | 2 | 4 | |
| 10 | s | i | 3 | 1 | 2 | 4 | |
| 11 | s | i | 4 | 1 | 2 | 4 | |

Fst=6+2
Lst=(6+2)-1

12

12

## FM Index (con't)

| | | F | L | i | m | p | s | C |
|---|---|---|---|---|---|---|---|---|
| pssi | 0 | # | i | 1 | 0 | 0 | 0 | # 0 |
| | 1 | i | p | 1 | 0 | 1 | 0 | i 1 |
| | 2 | i | s | 1 | 0 | 1 | 1 | m 5 |
| | 3 | i | s | 1 | 0 | 1 | 2 | p 6 |
| | 4 | i | m | 1 | 1 | 1 | 2 | s 8 |
| | 5 | m | # | 1 | 1 | 1 | 2 | |
| | 6 | p | p | 1 | 1 | 2 | 2 | |
| | 7 | p | i | 2 | 1 | 2 | 2 | |
| | 8 | s | s | 2 | 1 | 2 | 3 | |
| | 9 | s | s | 2 | 1 | 2 | 4 | |
| ➡ | 10 | s | i | 3 | 1 | 2 | 4 | |
| ➡ | 11 | s | i | 4 | 1 | 2 | 4 | |

Fst=6+2
Lst=(6+2)-1

Fst > Lst => No match

13

## FM Index (con't)

| | | F | L | i | m | p | s | C |
|---|---|---|---|---|---|---|---|---|
| pssi | 0 | # | i | | | | | # 0 |
| | 1 | i | p | | | | | i 1 |
| | 2 | i | s | | | | | m 5 |
| | 3 | i | s | 1 | 0 | 1 | 2 | p 6 |
| | 4 | i | m | | | | | s 8 |
| | 5 | m | # | | | | | |
| | 6 | p | p | | | | | |
| | 7 | p | i | 2 | 1 | 2 | 2 | |
| | 8 | s | s | | | | | |
| | 9 | s | s | | | | | |
| | 10 | s | i | | | | | |
| | 11 | s | i | 4 | 1 | 2 | 4 | |

To reduce space

14

14

## FM Index (con't)

| | | F | L | i | m | p | s | C |
|---|---|---|---|---|---|---|---|---|
| pssi | 0 | # | i | | | | | # 0 |
| | 1 | i | p | | | | | i 1 |
| | 2 | i | s | | | | | m 5 |
| | 3 | i | s | 1 | 0 | 1 | 2 | p 6 |
| | 4 | i | m | | | | | s 8 |
| | 5 | m | # | | | | | |
| | 6 | p | p | | | | | |
| | 7 | p | i | 2 | 1 | 2 | 2 | |
| | 8 | s | s | | | | | |
| | 9 | s | s | | | | | |
| | 10 | s | i | | | | | |
| | 11 | s | i | 4 | 1 | 2 | 4 | |

To reduce space

15

15

## Similar when L(x) = c

| | F | L | i | m | p | s | C |
|---|---|---|---|---|---|---|---|
| 0 | # | i | | | | | # 0 |
| 1 | i | p | | | | | i 1 |
| 2 | i | s | | | | | m 5 |
| 3 | i | s | 1 | 0 | 1 | 2 | p 6 |
| 4 | i | m | | | | | s 8 |
| 5 | m | # | | | | | |
| 6 | p | p | | | | | |
| 7 | p | i | 2 | 1 | 2 | 2 | |
| 8 | s | s | | | | | |
| 9 | s | s | | | | | |
| 10 | s | i | | | | | |
| 11 | s | i | 4 | 1 | 2 | 4 | |

To reduce space

16

16

RLFM Index (Revisit)

17

17

## RLFM Index (Derive B' from LF)

| | F | L | C |
|---|---|---|---|
| 1 | a | c | a 0 |
| 2 | a | c | c 3 |
| 3 | a | c | g 6 |
| 4 | c | a | t 8 |
| 5 | c | a | |
| 6 | c | g | |
| 7 | g | g | |
| 8 | g | a | |
| 9 | t | t | |
| 10 | t | t | |

| | B | S | Fs | Cs |
|---|---|---|---|---|
| 1 | 1 | c | a | a 0 |
| 2 | 0 | a | a | c 2 |
| 3 | 0 | g | c | g 3 |
| 4 | 1 | a | g | t 4 |
| 5 | 0 | t | t | |
| 6 | 1 | | | |
| 7 | 0 | | | |
| 8 | 1 | | | |
| 9 | 1 | | | |
| 10 | 0 | | | |

18

18

## RLFM Index (con't from the prev lecture)

| | F | L | C |
|---|---|---|---|
| 1 | a | c | a 0 |
| 2 | a | c | c 3 |
| 3 | a | c | g 6 |
| 4 | c | a | t 8 |
| 5 | c | a | |
| 6 | c | g | |
| 7 | g | g | |
| 8 | g | a | |
| 9 | t | t | |
| 10 | t | t | |

| | B | S | Fs | Cs | B′ |
|---|---|---|---|---|---|
| 1 | 1 | c | a | a 0 | |
| 2 | 0 | a | a | c 2 | |
| 3 | 0 | g | c | g 3 | |
| 4 | 1 | a | g | t 4 | 1 |
| 5 | 0 | t | t | | 0 |
| 6 | 1 | | | | 0 |
| 7 | 0 | | | | |
| 8 | 1 | | | | |
| 9 | 1 | | | | |
| 10 | 0 | | | | |

19

19

## RLFM Index (con't from the prev lecture)

| | F | L | C |
|---|---|---|---|
| 1 | a | c | a 0 |
| 2 | a | c | c 3 |
| 3 | a | c | g 6 |
| 4 | c | a | t 8 |
| 5 | c | a | |
| 6 | c | g | |
| 7 | g | g | |
| 8 | g | a | |
| 9 | t | t | |
| 10 | t | t | |

| | B | S | Fs | Cs | B′ |
|---|---|---|---|---|---|
| 1 | 1 | c | a | a 0 | 1 |
| 2 | 0 | a | a | c 2 | 0 |
| 3 | 0 | g | c | g 3 | |
| 4 | 1 | a | g | t 4 | 1 |
| 5 | 0 | t | t | | 0 |
| 6 | 1 | | | | 0 |
| 7 | 0 | | | | |
| 8 | 1 | | | | |
| 9 | 1 | | | | |
| 10 | 0 | | | | |

20

20

## RLFM Index (con't from the prev lecture)

| | F | L | C |
|---|---|---|---|
| 1 | a | c | a 0 |
| 2 | a | c | c 3 |
| 3 | a | c | g 6 |
| 4 | c | a | t 8 |
| 5 | c | a | |
| 6 | c | g | |
| 7 | g | g | |
| 8 | g | a | |
| 9 | t | t | |
| 10 | t | t | |

| | B | S | Fs | Cs | B′ |
|---|---|---|---|---|---|
| 1 | 1 | c | a | a 0 | 1 |
| 2 | 0 | a | a | c 2 | 0 |
| 3 | 0 | g | c | g 3 | |
| 4 | 1 | a | g | t 4 | 1 |
| 5 | 0 | t | t | | 0 |
| 6 | 1 | | | | 0 |
| 7 | 0 | | | | 1 |
| 8 | 1 | | | | 0 |
| 9 | 1 | | | | |
| 10 | 0 | | | | |

21

21

## RLFM Index (con't from the prev lecture)

| | F | L | C |
|---|---|---|---|
| 1 | a | c | a 0 |
| 2 | a | c | c 3 |
| 3 | a | c | g 6 |
| 4 | c | a | t 8 |
| 5 | c | a | |
| 6 | c | g | |
| 7 | g | g | |
| 8 | g | a | |
| 9 | t | t | |
| 10 | t | t | |

| | B | S | Fs | Cs | B′ |
|---|---|---|---|---|---|
| 1 | 1 | c | a | a 0 | 1 |
| 2 | 0 | a | a | c 2 | 0 |
| 3 | 0 | g | c | g 3 | 1 |
| 4 | 1 | a | g | t 4 | 1 |
| 5 | 0 | t | t | | 0 |
| 6 | 1 | | | | 0 |
| 7 | 0 | | | | 1 |
| 8 | 1 | | | | 0 |
| 9 | 1 | | | | |
| 10 | 0 | | | | |

22

22

## RLFM Index (no L & F, nor LF)

| | B | S | B′ |
|---|---|---|---|
| 1 | 1 | c | |
| 2 | 0 | a | |
| 3 | 0 | g | |
| 4 | 1 | a | |
| 5 | 0 | t | |
| 6 | 1 | | |
| 7 | 0 | | |
| 8 | 1 | | |
| 9 | 1 | | |
| 10 | 0 | | |

If only B and S are stored and given… then how ???

23

23

## RLFM Index (no L & F, nor LF)

| | B | S | Fs | B′ |
|---|---|---|---|---|
| 1 | 1 | c | a | |
| 2 | 0 | a | a | |
| 3 | 0 | g | c | |
| 4 | 1 | a | g | |
| 5 | 0 | t | t | |
| 6 | 1 | | | |
| 7 | 0 | | | |
| 8 | 1 | | | |
| 9 | 1 | | | |
| 10 | 0 | | | |

If only B and S are stored and given… then how ???

24

24

## RLFM Index (no L & F, nor LF)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | |
| 2 | 0 | a | a | c | 2 | |
| 3 | 0 | g | c | g | 3 | |
| 4 | 1 | a | g | t | 4 | |
| 5 | 0 | t | t | | | |
| 6 | 1 | | | | | |
| 7 | 0 | | | | | |
| 8 | 1 | | | | | |
| 9 | 1 | | | | | |
| 10 | 0 | | | | | |

same

25

25

## RLFM Index (No LF mapping)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | |
| 2 | 0 | a | a | c | 2 | |
| 3 | 0 | g | c | g | 3 | |
| 4 | 1 | a | g | t | 4 | |
| 5 | 0 | t | t | | | |
| 6 | 1 | | | | | |
| 7 | 0 | | | | | |
| 8 | 1 | | | | | |
| 9 | 1 | | | | | |
| 10 | 0 | | | | | |

26

26

## RLFM Index (No LF mapping)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | |
| 2 | 0 | a | a | c | 2 | |
| 3 | 0 | g | c | g | 3 | |
| 4 | 1 | a | g | t | 4 | |
| 5 | 0 | t | t | | | |
| 6 | 1 | | | | | |
| 7 | 0 | | | | | |
| 8 | 1 | | | | | |
| 9 | 1 | | | | | |
| 10 | 0 | | | | | |

27

27

## RLFM Index (No LF mapping)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | |
| 2 | 0 | a | a | c | 2 | |
| 3 | 0 | g | c | g | 3 | |
| 4 | 1 | a | g | t | 4 | |
| 5 | 0 | t | t | | | |
| 6 | 1 | | | | | |
| 7 | 0 | | | | | |
| 8 | 1 | | | | | |
| 9 | 1 | | | | | |
| 10 | 0 | | | | | |

28

28

## RLFM Index (No LF mapping)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | |
| 2 | 0 | a | a | c | 2 | |
| 3 | 0 | g | c | g | 3 | |
| 4 | 1 | a | g | t | 4 | |
| 5 | 0 | t | t | | | |
| 6 | 1 | | | | | |
| 7 | 0 | | | | | |
| 8 | 1 | | | | | |
| 9 | 1 | | | | | |
| 10 | 0 | | | | | |

29

29

## RLFM Index (No LF mapping)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | |
| 4 | 1 | a | g | t | 4 | |
| 5 | 0 | t | t | | | |
| 6 | 1 | | | | | |
| 7 | 0 | | | | | |
| 8 | 1 | | | | | |
| 9 | 1 | | | | | |
| 10 | 0 | | | | | |

30

30

## RLFM Index (No LF mapping)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | |
| 4 | 1 | a | g | t | 4 | |
| 5 | 0 | t | t | | | |
| 6 | 1 | | | | | |
| 7 | 0 | | | | | |
| 8 | 1 | | | | | |
| 9 | 1 | | | | | |
| 10 | 0 | | | | | |

31

31

## RLFM Index (No LF mapping)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | |
| 5 | 0 | t | t | | | |
| 6 | 1 | | | | | |
| 7 | 0 | | | | | |
| 8 | 1 | | | | | |
| 9 | 1 | | | | | |
| 10 | 0 | | | | | |

32

32

## RLFM Index (No LF mapping)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | |
| 5 | 0 | t | t | | | |
| 6 | 1 | | | | | |
| 7 | 0 | | | | | |
| 8 | 1 | | | | | |
| 9 | 1 | | | | | |
| 10 | 0 | | | | | |

33

33

## RLFM Index (No LF mapping)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | |
| 5 | 0 | t | t | | | |
| 6 | 1 | | | | | |
| 7 | 0 | | | | | |
| 8 | 1 | | | | | |
| 9 | 1 | | | | | |
| 10 | 0 | | | | | |

34

34

## RLFM Index (No LF mapping)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | 1 |
| 5 | 0 | t | t | | | 0 |
| 6 | 1 | | | | | 0 |
| 7 | 0 | | | | | |
| 8 | 1 | | | | | |
| 9 | 1 | | | | | |
| 10 | 0 | | | | | |

35

35

## RLFM Index (No LF mapping)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | 1 |
| 5 | 0 | t | t | | | 0 |
| 6 | 1 | | | | | 0 |
| 7 | 0 | | | | | |
| 8 | 1 | | | | | |
| 9 | 1 | | | | | |
| 10 | 0 | | | | | |

36

36

## RLFM Index (No LF mapping)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | 1 |
| 5 | 0 | t | t | | | 0 |
| 6 | 1 | | | | | 0 |
| 7 | 0 | | | | | 1 |
| 8 | 1 | | | | | 0 |
| 9 | 1 | | | | | |
| 10 | 0 | | | | | |

37

37

## RLFM Index (No LF mapping)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | 1 |
| 5 | 0 | t | t | | | 0 |
| 6 | 1 | | | | | 0 |
| 7 | 0 | | | | | 1 |
| 8 | 1 | | | | | 0 |
| 9 | 1 | | | | | 1 |
| 10 | 0 | | | | | 0 |

38

38

Now we have B, S, **B′**

Let's **reverse (decode)** using LF mapping

39

39

## CHANGES TO FORMULAS

- Recall that we need to compute $C_T[c]+rank_c(L,i)$ in the backward search.
- **Theorem:** $C[c]+rank_c(L,i)$ is equivalent to $select_1(B',C_S[c]+1+rank_c(S,rank_1(B,i)))-1$, when $L[i] \neq c$ (e.g., when backward search) , and otherwise (e.g., when reverse, sometimes backward search too) to $select_1(B',C_S[c]+rank_c(S,rank_1(B,i)))+ i-select_1(B,rank_1(B,i))$.

You can apply these formulas to do reversing & backward search.

40

40

## CHANGES TO FORMULAS

- Recall that we need to compute $C_T[c]+rank_c(L,i)$ in the backward search.
- **Theorem:** $C[c]+rank_c(L,i)$ is equivalent to $select_1(B',C_S[c]+1+rank_c(S,rank_1(B,i)))-1$, when $L[i] \neq c$ (e.g., when backward search) , and otherwise (e.g., when reverse, sometimes backward search too) to $select_1(B',C_S[c]+rank_c(S,rank_1(B,i)))+ i-select_1(B,rank_1(B,i))$.

But I promised that I would explain why/how these formulas actually work

41

41

## RLFM Index (con't from the prev lecture)

| | F | L | C | | | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | a | c | a | 0 | | 1 | c | a | a | 0 | 1 |
| 2 | a | c | c | 3 | | 0 | a | a | c | 2 | 0 |
| 3 | a | c | g | 6 | | 0 | g | c | g | 3 | 1 |
| 4 | c | a | t | 8 | | 1 | a | g | t | 4 | 1 |
| 5 | c | a | | | | 0 | t | t | | | 0 |
| 6 | c | g | | | | 1 | | | | | 0 |
| 7 | g | g | | | | 0 | | | | | 1 |
| 8 | g | a | | | | 1 | | | | | 0 |
| 9 | t | t | | | | 1 | | | | | 1 |
| 10 | t | t | | | | 0 | | | | | 0 |

Suppose reverse from L[8]

42

42

## RLFM Index (con't from the prev lecture)

| | B | S | Fs | Cs | B′ |
|---|---|---|---|---|---|
| 1 | 1 | c | a | a 0 | 1 |
| 2 | 0 | a | a | c 2 | 0 |
| 3 | 0 | g | c | g 3 | 1 |
| 4 | 1 | a | g | t 4 | 1 |
| 5 | 0 | t | t | | 0 |
| 6 | 1 | | | | 0 |
| 7 | 0 | | | | 1 |
| 8 | 1 | | | | 0 |
| 9 | 1 | | | | 1 |
| 10 | 0 | | | | 0 |

rank$_a$(S, rank$_1$(B, 8)) = 2

43

43

## RLFM Index (con't from the prev lecture)

| | B | S | Fs | Cs | B′ |
|---|---|---|---|---|---|
| 1 | 1 | c | a | a 0 | 1 |
| 2 | 0 | a | a | c 2 | 0 |
| 3 | 0 | g | c | g 3 | 1 |
| 4 | 1 | a | g | t 4 | 1 |
| 5 | 0 | t | t | | 0 |
| 6 | 1 | | | | 0 |
| 7 | 0 | | | | 1 |
| 8 | 1 | | | | 0 |
| 9 | 1 | | | | 1 |
| 10 | 0 | | | | 0 |

rank$_a$(S, rank$_1$(B, 8)) = 2

44

44

## RLFM Index (con't from the prev lecture)

| | B | S | Fs | Cs | B′ |
|---|---|---|---|---|---|
| 1 | 1 | c | a | a 0 | 1 |
| 2 | 0 | a | a | c 2 | 0 |
| 3 | 0 | g | c | g 3 | 1 |
| 4 | 1 | a | g | t 4 | 1 |
| 5 | 0 | t | t | | 0 |
| 6 | 1 | | | | 0 |
| 7 | 0 | | | | 1 |
| 8 | 1 | | | | 0 |
| 9 | 1 | | | | 1 |
| 10 | 0 | | | | 0 |

select$_1$(B′, C$_s$[a]+rank$_a$(S, rank$_1$(B, 8)))

45

45

## RLFM Index (con't from the prev lecture)

| | F | L | C |
|---|---|---|---|
| 1 | a | c | a 0 |
| 2 | a | c | c 3 |
| 3 | a | c | g 6 |
| 4 | c | a | t 8 |
| 5 | c | a | |
| 6 | c | g | |
| 7 | g | g | |
| 8 | g | a | |
| 9 | t | t | |
| 10 | t | t | |

| | B | S | Fs | Cs | B′ |
|---|---|---|---|---|---|
| 1 | 1 | c | a | a 0 | 1 |
| 2 | 0 | a | a | c 2 | 0 |
| 3 | 0 | g | c | g 3 | 1 |
| 4 | 1 | a | g | t 4 | 1 |
| 5 | 0 | t | t | | 0 |
| 6 | 1 | | | | 0 |
| 7 | 0 | | | | 1 |
| 8 | 1 | | | | 0 |
| 9 | 1 | | | | 1 |
| 10 | 0 | | | | 0 |

select$_1$(B′, C$_s$[a]+rank$_a$(S, rank$_1$(B, 8))) = 3

46

46

Good, but not good enough

47

47

## RLFM Index (con't from the prev lecture)

| | B | S | Fs | Cs | B′ |
|---|---|---|---|---|---|
| 1 | 1 | c | a | a 0 | 1 |
| 2 | 0 | a | a | c 2 | 0 |
| 3 | 0 | g | c | g 3 | 1 |
| 4 | 1 | a | g | t 4 | 1 |
| 5 | 0 | t | t | | 0 |
| 6 | 1 | | | | 0 |
| 7 | 0 | | | | 1 |
| 8 | 1 | | | | 0 |
| 9 | 1 | | | | 1 |
| 10 | 0 | | | | 0 |

select$_1$(B′, C$_s$[c]+rank$_c$(S, rank$_1$(B, 3)))
=select$_1$(B′, 2 + 1 )=4

48

48

## RLFM Index (con't from the prev lecture)

| | F | L | C | |
|---|---|---|---|---|
| 1 | a | c | a | 0 |
| 2 | a | c | c | 3 |
| 3 | a | c | g | 6 |
| 4 | c | a | t | 8 |
| 5 | c | a | | |
| 6 | c | g | | |
| 7 | g | g | | |
| 8 | g | a | | |
| 9 | t | t | | |
| 10 | t | t | | |

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | 1 |
| 5 | 0 | t | t | | | 0 |
| 6 | 1 | | | | | 0 |
| 7 | 0 | | | | | 1 |
| 8 | 1 | | | | | 0 |
| 9 | 1 | | | | | 1 |
| 10 | 0 | | | | | 0 |

$select_1(B', C_s[c]+rank_c(S, rank_1(B, 3)))$
$=select_1(B', 2 + 1)=4$ ?

49

49

## RLFM Index (con't from the prev lecture)

| | F | L | C | |
|---|---|---|---|---|
| 1 | a | c | a | 0 |
| 2 | a | c | c | 3 |
| 3 | a | c | g | 6 |
| 4 | c | a | t | 8 |
| 5 | c | a | | |
| 6 | c | g | | |
| 7 | g | g | | |
| 8 | g | a | | |
| 9 | t | t | | |
| 10 | t | t | | |

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | 1 |
| 5 | 0 | t | t | | | 0 |
| 6 | 1 | | | | | 0 |
| 7 | 0 | | | | | 1 |
| 8 | 1 | | | | | 0 |
| 9 | 1 | | | | | 1 |
| 10 | 0 | | | | | 0 |

$select_1(B', C_s[c]+rank_c(S, rank_1(B, 3)))$
$=select_1(B', 2 + 1)=4 + 2$

50

50

## RLFM Index (con't from the prev lecture)

| | F | L | C | |
|---|---|---|---|---|
| 1 | a | c | a | 0 |
| 2 | a | c | c | 3 |
| 3 | a | c | g | 6 |
| 4 | c | a | t | 8 |
| 5 | c | a | | |
| 6 | c | g | | |
| 7 | g | g | | |
| 8 | g | a | | |
| 9 | t | t | | |
| 10 | t | t | | |

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | 1 |
| 5 | 0 | t | t | | | 0 |
| 6 | 1 | | | | | 0 |
| 7 | 0 | | | | | 1 |
| 8 | 1 | | | | | 0 |
| 9 | 1 | | | | | 1 |
| 10 | 0 | | | | | 0 |

$select_1(B', C_s[c]+rank_c(S, rank_1(B, 3)))$
$=select_1(B', 2 + 1)=4 + (i - rank_1(B, i))$

51

51

Another example, LF[5] = ?

52

52

## RLFM Index (con't from the prev lecture)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | 1 |
| 5 | 0 | t | t | | | 0 |
| 6 | 1 | | | | | 0 |
| 7 | 0 | | | | | 1 |
| 8 | 1 | | | | | 0 |
| 9 | 1 | | | | | 1 |
| 10 | 0 | | | | | 0 |

$select_1(B', C_s[a]+rank_a(S, rank_1(B, 5)))$
$=select_1(B', 0 + 1)=1 + (i - rank_1(B, i))$

53

53

## RLFM Index (con't from the prev lecture)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | 1 |
| 5 | 0 | t | t | | | 0 |
| 6 | 1 | | | | | 0 |
| 7 | 0 | | | | | 1 |
| 8 | 1 | | | | | 0 |
| 9 | 1 | | | | | 1 |
| 10 | 0 | | | | | 0 |

$select_1(B', C_s[a]+rank_a(S, rank_1(B, 5)))$
$=select_1(B', 0 + 1)=1 + (i - rank_1(B, i))$

54

54

## RLFM Index (con't from the prev lecture)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | 1 |
| 5 | 0 | t | t | | | 0 |
| 6 | 1 | | | | | 0 |
| 7 | 0 | | | | | 1 |
| 8 | 1 | | | | | 0 |
| 9 | 1 | | | | | 1 |
| 10 | 0 | | | | | 0 |

$select_1(B', C_S[a]+rank_a(S, rank_1(B, 5)))$
$=select_1(B', 0 + 1 )=1 + (i - rank_1(B, i))$

55

55

## RLFM Index (con't from the prev lecture)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | 1 |
| 5 | 0 | t | t | | | 0 |
| 6 | 1 | | | | | 0 |
| 7 | 0 | | | | | 1 |
| 8 | 1 | | | | | 0 |
| 9 | 1 | | | | | 1 |
| 10 | 0 | | | | | 0 |

$select_1(B', C_S[a]+rank_a(S, rank_1(B, 5)))$
$=select_1(B', 0 + 1 )=1 + (i - select_1(B, rank_1(B, i)))$

56

56

## RLFM Index (con't from the prev lecture)

| | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | c | a | a | 0 | 1 |
| 2 | 0 | a | a | c | 2 | 0 |
| 3 | 0 | g | c | g | 3 | 1 |
| 4 | 1 | a | g | t | 4 | 1 |
| 5 | 0 | t | t | | | 0 |
| 6 | 1 | | | | | 0 |
| 7 | 0 | | | | | 1 |
| 8 | 1 | | | | | 0 |
| 9 | 1 | | | | | 1 |
| 10 | 0 | | | | | 0 |

$select_1(B', C_S[a]+rank_a(S, rank_1(B, 5)))$
$=select_1(B', 0 + 1 )=1 + (i - select_1(B, rank_1(B, i)))$

57

57

## RLFM Index (con't from the prev lecture)

| | F | L | C | | B | S | Fs | Cs | | B′ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | a | c | a | 0 | 1 | c | a | a | 0 | 1 |
| 2 | a | c | c | 3 | 0 | a | a | c | 2 | 0 |
| 3 | a | c | g | 6 | 0 | g | c | g | 3 | 1 |
| 4 | c | a | t | 8 | 1 | a | g | t | 4 | 1 |
| 5 | c | a | | | 0 | t | t | | | 0 |
| 6 | c | g | | | 1 | | | | | 0 |
| 7 | g | g | | | 0 | | | | | 1 |
| 8 | g | a | | | 1 | | | | | 0 |
| 9 | t | t | | | 1 | | | | | 1 |
| 10 | t | t | | | 0 | | | | | 0 |

$select_1(B', C_S[a]+rank_a(S, rank_1(B, 5)))$
$=select_1(B', 0 + 1 )=1 + (i - select_1(B, rank_1(B, i)))$

58

58

## RLFM Index (con't from the prev lecture)

### CHANGES TO FORMULAS

- Recall that we need to compute $C_T[c]+rank_c(L,i)$ in the backward search.
- **Theorem:** $C[c]+rank_c(L,i)$ is equivalent to $select_1(B',C_S[c]+1+rank_c(S,rank_1(B,i)))-1$, when $L[i] \neq c$ (e.g., when backward search) , and otherwise (e.g., when reverse, sometimes backward search too) to $select_1(B',C_S[c]+rank_c(S,rank_1(B,i)))+ i-select_1(B,rank_1(B,i))$.

$select_1(B', C_S[a]+rank_a(S, rank_1(B, 5)))$
$=select_1(B', 0 + 1 )=1 + (i - select_1(B, rank_1(B, i)))$

59

59

## RLFM Index (con't from the prev lecture)

Backward Search

60

60

## Backward search for "si"

| | B | S | $F_S$ | C | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | i | # | # | 0 | 1 |
| 2 | 1 | p | i | i | 1 | 1 |
| 3 | 1 | s | i | m | 4 | 1 |
| 4 | 0 | m | i | p | 5 | 1 |
| 5 | 1 | # | m | s | 7 | 0 |
| 6 | 1 | p | p | | | 1 |
| 7 | 1 | i | p | | | 1 |
| 8 | 1 | s | s | | | 1 |
| 9 | 1 | i | s | | | 1 |
| 10 | 0 | | | | | 0 |
| 11 | 1 | | | | | 1 |
| 12 | 0 | | | | | 0 |

61

61

## Backward search for "si"

| | B | S | $F_S$ | C | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | i | # | # | 0 | 1 |
| 2 | 1 | p | i | i | 1 | 1 |
| 3 | 1 | s | i | m | 4 | 1 |
| 4 | 0 | m | i | p | 5 | 1 |
| 5 | 1 | # | m | s | 7 | 0 |
| 6 | 1 | p | p | | | 1 |
| 7 | 1 | i | p | | | 1 |
| 8 | 1 | s | s | | | 1 |
| 9 | 1 | i | s | | | 1 |
| 10 | 0 | | | | | 0 |
| 11 | 1 | | | | | 1 |
| 12 | 0 | | | | | 0 |

62

62

## Backward search for "si"

| | B | S | $F_S$ | C | | B′ |
|---|---|---|---|---|---|---|
| 1 | 1 | i | # | # | 0 | 1 |
| 2 | 1 | p | i | i | 1 | 1 |
| 3 | 1 | s | i | m | 4 | 1 |
| 4 | 0 | m | i | p | 5 | 1 |
| 5 | 1 | # | m | s | 7 | 0 |
| 6 | 1 | p | p | | | 1 |
| 7 | 1 | i | p | | | 1 |
| 8 | 1 | s | s | | | 1 |
| 9 | 1 | i | s | | | 1 |
| 10 | 0 | | | | | 0 |
| 11 | 1 | | | | | 1 |
| 12 | 0 | | | | | 0 |

c = i
Fst = 2
Lst = 5

63

63

## Backward search for "si"

| | F | L | C | | B | S | $F_S$ | C | | B′ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | # | i | # | 0 | 1 | i | # | # | 0 | 1 |
| 2 | i | p | i | 1 | 1 | p | i | i | 1 | 1 |
| 3 | i | s | m | 5 | 1 | s | i | m | 4 | 1 |
| 4 | i | s | p | 6 | 0 | m | i | p | 5 | 1 |
| 5 | i | m | s | 8 | 1 | # | m | s | 7 | 0 |
| 6 | m | # | | | 1 | p | p | | | 1 |
| 7 | p | p | | | 1 | i | p | | | 1 |
| 8 | p | i | | | 1 | s | s | | | 1 |
| 9 | s | s | | | 1 | i | s | | | 1 |
| 10 | s | s | | | 0 | | | | | 0 |
| 11 | s | i | | | 1 | | | | | 1 |
| 12 | s | i | | | 0 | | | | | 0 |

c = i
Fst = 2
Lst = 5

64

64

## Backward search for "si"

| | F | L | C | | B | S | $F_S$ | C | | B′ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | # | i | # | 0 | 1 | i | # | # | 0 | 1 |
| 2 | i | p | i | 1 | 1 | p | i | i | 1 | 1 |
| 3 | i | s | m | 5 | 1 | s | i | m | 4 | 1 |
| 4 | i | s | p | 6 | 0 | m | i | p | 5 | 1 |
| 5 | i | m | s | 8 | 1 | # | m | s | 7 | 0 |
| 6 | m | # | | | 1 | p | p | | | 1 |
| 7 | p | p | | | 1 | i | p | | | 1 |
| 8 | p | i | | | 1 | s | s | | | 1 |
| 9 | s | s | | | 1 | i | s | | | 1 |
| 10 | s | s | | | 0 | | | | | 0 |
| 11 | s | i | | | 1 | | | | | 1 |
| 12 | s | i | | | 0 | | | | | 0 |

c = s
Fst =
C[c] + Occ(c, Fst – 1) + 1
= ?

65

65

## Backward search for "si"

| | F | L | C | | B | S | $F_S$ | C | | B′ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | # | i | # | 0 | 1 | i | # | # | 0 | 1 |
| 2 | i | p | i | 1 | 1 | p | i | i | 1 | 1 |
| 3 | i | s | m | 5 | 1 | s | i | m | 4 | 1 |
| 4 | i | s | p | 6 | 0 | m | i | p | 5 | 1 |
| 5 | i | m | s | 8 | 1 | # | m | s | 7 | 0 |
| 6 | m | # | | | 1 | p | p | | | 1 |
| 7 | p | p | | | 1 | i | p | | | 1 |
| 8 | p | i | | | 1 | s | s | | | 1 |
| 9 | s | s | | | 1 | i | s | | | 1 |
| 10 | s | s | | | 0 | | | | | 0 |
| 11 | s | i | | | 1 | | | | | 1 |
| 12 | s | i | | | 0 | | | | | 0 |

$T_{x}$
$L_{x}$

c = i
Fst = 2
Lst = 5

66

66

## Backward search for "si"

| | F | L | C | | B | S | $F_S$ | C | | B′ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | # | i | # | 0 | 1 | i | # | # | 0 | 1 |
| ➡ 2 | i | p | i | 1 | 1 | p | i | i | 1 | 1 |
| 3 | i | s | m | 5 | 1 | s | i | m | 4 | 1 |
| 4 | i | s | p | 6 | 0 | m | i | p | 5 | 1 |
| ➡ 5 | i | m | s | 8 | 1 | # | m | s | 7 | 0 |
| 6 | m | # | | | 1 | p | p | | | 1 |
| 7 | p | p | | | 1 | i | p | | | 1 |
| 8 | p | i | | | 1 | s | s | | | 1 |
| 9 | s | s | | | 1 | i | s | | | 1 |
| 10 | s | s | | | 0 | | | | | 0 |
| 11 | s | i | | | 1 | | | | | 1 |
| 12 | s | i | | | 0 | | | | | 0 |

c = s
Fst = ??

67

67

## Backward search for "si"

| | F | L | C | | B | S | $F_S$ | C | | B′ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | # | i | # | 0 | 1 | i | # | # | 0 | 1 |
| ➡ 2 | i | p | i | 1 | 1 | p | i | i | 1 | 1 |
| 3 | i | s | m | 5 | 1 | s | i | m | 4 | 1 |
| 4 | i | s | p | 6 | 0 | m | i | p | 5 | 1 |
| ➡ 5 | i | m | s | 8 | 1 | # | m | s | 7 | 0 |
| 6 | m | # | | | 1 | p | p | | | 1 |
| 7 | p | p | | | 1 | i | p | | | 1 |
| 8 | p | i | | | 1 | s | s | | | 1 |
| 9 | s | s | | | 1 | i | s | | | 1 |
| 10 | s | s | | | 0 | | | | | 0 |
| 11 | s | i | | | 1 | | | | | 1 |
| 12 | s | i | | | 0 | | | | | 0 |

c = s
Fst
Occ of s: $rank_s(S, rank_1(B,2-1)) = 0$

68

68

## Backward search for "si"

| | F | L | C | | B | S | $F_S$ | C | | B′ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | # | i | # | 0 | 1 | i | # | # | 0 | 1 |
| ➡ 2 | i | p | i | 1 | 1 | p | i | i | 1 | 1 |
| 3 | i | s | m | 5 | 1 | s | i | m | 4 | 1 |
| 4 | i | s | p | 6 | 0 | m | i | p | 5 | 1 |
| ➡ 5 | i | m | s | 8 | 1 | # | m | s | 7 | 0 |
| 6 | m | # | | | 1 | p | p | | | 1 |
| 7 | p | p | | | 1 | i | p | | | 1 |
| 8 | p | i | | | 1 | s | s | | | 1 |
| 9 | s | s | | | 1 | i | s | | | 1 |
| 10 | s | s | | | 0 | | | | | 0 |
| 11 | s | i | | | 1 | | | | | 1 |
| 12 | s | i | | | 0 | | | | | 0 |

c = s
Fst
Occ of s: $rank_s(S, rank_1(B,2-1)) = 0$
$select_1(B',7+1+0)$
So Fst = 9

69

69

## Backward search for "si"

| | F | L | C | | B | S | $F_S$ | C | | B′ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | # | i | # | 0 | 1 | i | # | # | 0 | 1 |
| ➡ 2 | i | p | i | 1 | 1 | p | i | i | 1 | 1 |
| 3 | i | s | m | 5 | 1 | s | i | m | 4 | 1 |
| 4 | i | s | p | 6 | 0 | m | i | p | 5 | 1 |
| ➡ 5 | i | m | s | 8 | 1 | # | m | s | 7 | 0 |
| 6 | m | # | | | 1 | p | p | | | 1 |
| 7 | p | p | | | 1 | i | p | | | 1 |
| 8 | p | i | | | 1 | s | s | | | 1 |
| 9 | s | s | | | 1 | i | s | | | 1 |
| 10 | s | s | | | 0 | | | | | 0 |
| 11 | s | i | | | 1 | | | | | 1 |
| 12 | s | i | | | 0 | | | | | 0 |

c = s
Fst
Occ of s: $rank_s(S, rank_1(B,2-1)) = 0$
$select_1(B',7+1+0)$
So Fst = 9

70

70

## Backward search for "si"

| | F | L | C | | B | S | $F_S$ | C | | B′ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | # | i | # | 0 | 1 | i | # | # | 0 | 1 |
| 2 | i | p | i | 1 | 1 | p | i | i | 1 | 1 |
| 3 | i | s | m | 5 | 1 | s | i | m | 4 | 1 |
| 4 | i | s | p | 6 | 0 | m | i | p | 5 | 1 |
| 5 | i | m | s | 8 | 1 | # | m | s | 7 | 0 |
| 6 | m | # | | | 1 | p | p | | | 1 |
| 7 | p | p | | | 1 | i | p | | | 1 |
| 8 | p | i | | | 1 | s | s | | | 1 |
| 9 | s | s | | | 1 | i | s | | | 1 |
| 10 | s | s | | | 0 | | | | | 0 |
| 11 | s | i | | | 1 | | | | | 1 |
| 12 | s | i | | | 0 | | | | | 0 |

c = s
Lst
Occ of s: $rank_s(S, rank_1(B,5)) = 1$

71

71

## Backward search for "si"

| | F | L | C | | B | S | $F_S$ | C | | B′ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | # | i | # | 0 | 1 | i | # | # | 0 | 1 |
| 2 | i | p | i | 1 | 1 | p | i | i | 1 | 1 |
| 3 | i | s | m | 5 | 1 | s | i | m | 4 | 1 |
| 4 | i | s | p | 6 | 0 | m | i | p | 5 | 1 |
| 5 | i | m | s | 8 | 1 | # | m | s | 7 | 0 |
| 6 | m | # | | | 1 | p | p | | | 1 |
| 7 | p | p | | | 1 | i | p | | | 1 |
| 8 | p | i | | | 1 | s | s | | | 1 |
| 9 | s | s | | | 1 | i | s | | | 1 |
| 10 | s | s | | | 0 | | | | | 0 |
| 11 | s | i | | | 1 | | | | | 1 |
| 12 | s | i | | | 0 | | | | | 0 |

c = s
Lst
Occ of s: $rank_s(S, rank_1(B,5)) = 1$
$select_1(B',7+1+1) = 11$
$11 - 1 = 10$
So Lst = 10

-1: since inclusively, e.g., Lst-Fst+1 = #matches

72

72

## Backward search for "ssi"

| | F | L | C |
|---|---|---|---|
| 1 | # | i | # 0 |
| 2 | i | p | i 1 |
| 3 | i | s | m 5 |
| 4 | i | s | p 6 |
| 5 | i | m | s 8 |
| 6 | m | # | |
| 7 | p | p | |
| 8 | p | i | |
| 9 | s | s | |
| 10 | s | s | |
| 11 | s | i | |
| 12 | s | i | |

| | B | S | $F_S$ | C | B′ |
|---|---|---|---|---|---|
| 1 | 1 | i | # | # 0 | 1 |
| 2 | 1 | p | i | i 1 | 1 |
| 3 | 1 | s | i | m 4 | 1 |
| 4 | 0 | m | i | p 5 | 1 |
| 5 | 1 | # | m | s 7 | 0 |
| 6 | 1 | p | p | | 1 |
| 7 | 1 | i | p | | 1 |
| 8 | 1 | s | s | | 1 |
| 9 | 1 | i | s | | 1 |
| 10 | 0 | | | | 0 |
| 11 | 1 | | | | 1 |
| 12 | 0 | | | | 0 |

c = s
Fst
Occ of s: $rank_S$(S, $rank_1$(B,9-1)) = 1
$select_1$(B',7+1+1)
So Fst = 11



73

## Backward search for "ssi"

| | F | L | C |
|---|---|---|---|
| 1 | # | i | # 0 |
| 2 | i | p | i 1 |
| 3 | i | s | m 5 |
| 4 | i | s | p 6 |
| 5 | i | m | s 8 |
| 6 | m | # | |
| 7 | p | p | |
| 8 | p | i | |
| 9 | s | s | |
| 10 | s | s | |
| 11 | s | i | |
| 12 | s | i | |

| | B | S | $F_S$ | C | B′ |
|---|---|---|---|---|---|
| 1 | 1 | i | # | # 0 | 1 |
| 2 | 1 | p | i | i 1 | 1 |
| 3 | 1 | s | i | m 4 | 1 |
| 4 | 0 | m | i | p 5 | 1 |
| 5 | 1 | # | m | s 7 | 0 |
| 6 | 1 | p | p | | 1 |
| 7 | 1 | i | p | | 1 |
| 8 | 1 | s | s | | 1 |
| 9 | 1 | i | s | | 1 |
| 10 | 0 | | | | 0 |
| 11 | 1 | | | | 1 |
| 12 | 0 | | | | 0 |

c = s
Lst
Occ of s: $rank_S$(S, $rank_1$(B,10))
= 1
Since L[i]=c, $select_1$(B',Cs[c]+$rank_c$(S,$rank_1$(B,i)))+ i-$select_1$(B,$rank_1$(B,i)).
$select_1$(B',7+2)
= 11
11 + 1 = 12
So Lst = 12



74

## Backward search for "issi"

| | F | L | C |
|---|---|---|---|
| 1 | # | i | # 0 |
| 2 | i | p | i 1 |
| 3 | i | s | m 5 |
| 4 | i | s | p 6 |
| 5 | i | m | s 8 |
| 6 | m | # | |
| 7 | p | p | |
| 8 | p | i | |
| 9 | s | s | |
| 10 | s | s | |
| 11 | s | i | |
| 12 | s | i | |

| | B | S | $F_S$ | C | B′ |
|---|---|---|---|---|---|
| 1 | 1 | i | # | # 0 | 1 |
| 2 | 1 | p | i | i 1 | 1 |
| 3 | 1 | s | i | m 4 | 1 |
| 4 | 0 | m | i | p 5 | 1 |
| 5 | 1 | # | m | s 7 | 0 |
| 6 | 1 | p | p | | 1 |
| 7 | 1 | i | p | | 1 |
| 8 | 1 | s | s | | 1 |
| 9 | 1 | i | s | | 1 |
| 10 | 0 | | | | 0 |
| 11 | 1 | | | | 1 |
| 12 | 0 | | | | 0 |

c = i
Fst
Occ of i: $rank_S$(S, $rank_1$(B,11-1))
= 2
$select_1$(B',1+1+2)
So Fst = 4



75

## Backward search for "issi"

| | F | L | C |
|---|---|---|---|
| 1 | # | i | # 0 |
| 2 | i | p | i 1 |
| 3 | i | s | m 5 |
| 4 | i | s | p 6 |
| 5 | i | m | s 8 |
| 6 | m | # | |
| 7 | p | p | |
| 8 | p | i | |
| 9 | s | s | |
| 10 | s | s | |
| 11 | s | i | |
| 12 | s | i | |

| | B | S | $F_S$ | C | B′ |
|---|---|---|---|---|---|
| 1 | 1 | i | # | # 0 | 1 |
| 2 | 1 | p | i | i 1 | 1 |
| 3 | 1 | s | i | m 4 | 1 |
| 4 | 0 | m | i | p 5 | 1 |
| 5 | 1 | # | m | s 7 | 0 |
| 6 | 1 | p | p | | 1 |
| 7 | 1 | i | p | | 1 |
| 8 | 1 | s | s | | 1 |
| 9 | 1 | i | s | | 1 |
| 10 | 0 | | | | 0 |
| 11 | 1 | | | | 1 |
| 12 | 0 | | | | 0 |

c = i
Lst
Occ of i: $rank_S$(S, $rank_1$(B,12))
= 3
Since L[i]=c, $select_1$(B',Cs[c]+$rank_c$(S,$rank_1$(B,i)))+ i-$select_1$(B,$rank_1$(B,i)).
$select_1$(B',1+3)
= 4
4 + 1 = 5
So Lst = 5



76

## Therefore …

### CHANGES TO FORMULAS

- Recall that we need to compute $C_T[c]+rank_c(L,i)$ in the backward search.
- **Theorem:** $C[c]+rank_c(L,i)$ is equivalent to $select_1(B',C_S[c]+1+rank_c(S,rank_1(B,i)))-1$, when $L[i] \neq c$ (e.g., when backward search) , and otherwise (e.g., when reverse, sometimes backward search too) to $select_1(B',C_S[c]+rank_c(S,rank_1(B,i)))+$ $i\text{-}select_1(B,rank_1(B,i))$.

77