

**FM-Index and recursive formulas for pattern matching**

The table below shows the FM-Index for  $S = azcdf rzccbcazctazccazbcd f z r z c \$$ , as well as the F and A columns. First use F to complete the M array. Then use the formulas below to find all the indices of the pattern  $P = azc$  in F. Finally, use A to find the indices of  $P$  in the original string  $S$ .

$i$	F ...	... L	A	occ(\$)	occ(a)	occ(b)	occ(c)	occ(d)	occ(f)	occ(r)	occ(t)	occ(z)
1	\$	c	30	0	0	0	1	0	0	0	0	0
2	a	c	20	0	0	0	2	0	0	0	0	0
3	a	t	16	0	0	0	2	0	0	0	1	0
4	a	\$	1	1	0	0	2	0	0	0	1	0
5	a	c	12	1	0	0	3	0	0	0	1	0
6	b	c	10	1	0	0	4	0	0	0	1	0
7	b	z	22	1	0	0	4	0	0	0	1	1
8	c	z	29	1	0	0	4	0	0	0	1	2
9	c	c	19	1	0	0	5	0	0	0	1	2
10	c	b	11	1	0	1	5	0	0	0	1	2
11	c	c	9	1	0	1	6	0	0	0	1	2
12	c	z	18	1	0	1	6	0	0	0	1	3
13	c	z	8	1	0	1	6	0	0	0	1	4
14	c	z	3	1	0	1	6	0	0	0	1	5
15	c	b	23	1	0	2	6	0	0	0	1	5
16	c	z	14	1	0	2	6	0	0	0	1	6
17	d	c	4	1	0	2	7	0	0	0	1	6
18	d	c	24	1	0	2	8	0	0	0	1	6
19	f	d	5	1	0	2	8	1	0	0	1	6
20	f	d	25	1	0	2	8	2	0	0	1	6
21	r	z	27	1	0	2	8	2	0	0	1	7
22	r	f	6	1	0	2	8	2	1	0	1	7
23	t	c	15	1	0	2	9	2	1	0	1	7
24	z	a	21	1	1	2	9	2	1	0	1	7
25	z	r	28	1	1	2	9	2	1	1	1	7
26	z	a	17	1	2	2	9	2	1	1	1	7
27	z	r	7	1	2	2	9	2	1	2	1	7
28	z	a	2	1	3	2	9	2	1	2	1	7
29	z	a	13	1	4	2	9	2	1	2	1	7
30	z	f	26	1	4	2	9	2	2	2	1	7

Base case: find the start point (sp) and end point (ep) of the *last* character in  $P$  (inclusive, so we subtract 1 from the end point):

$$sp(c) = M[c], \quad ep(c) = M[\text{char alphabetically after } c] - 1$$

Recursion:

$$sp(c\sigma) = M[c] + occ(c, sp(\sigma) - 1)$$

$$ep(c\sigma) = M[c] + occ(c, ep(\sigma)) - 1$$

$c$	$M[c]$
\$	
$a$	
$b$	
$c$	
$d$	
$f$	
$r$	
$t$	
$z$	

1.  $\text{sp}(c) =$

$\text{ep}(c) =$

2.  $\text{sp}(zc) =$

$\text{ep}(zc) =$

3.  $\text{sp}(azc) =$

$\text{ep}(azc) =$

4. Use A to find the indices of  $P = azc$  in the original string.