More definitions

- Let r_i and l_i be defined as follows:

$$r_i = \max_{\substack{1 < j \leq i \\ \text{s.t. } Z_j > 0}} (j + Z_j - 1)$$

$$l_i = \operatorname{argmax}_{\substack{1 < j \leq i \\ \text{s.t. } Z_j > 0}} (j + Z_j - 1)$$

Position:	1	2	3	4	5	6
String:	x	t	p	x	t	d
Z_i:		0	0	2	0	0
r_i:		0	0	5	5	5
l_i:		0	0	4	4	4

Linear-time calculation of Z_i

Initialize $Z_1 = r = l = 0$ and start at $k = 1$. From left-to-right over S, hence at position k we have already calculated Z_{k-1} and have current values for l and r

- If $k > r$, begin comparing with beginning of S. Length of match is Z_k. If $Z_k > 0$, then $r = k + Z_k - 1$ and $l = k$.
- If $k \leq r$, then $S(k) = S(k')$ where $k' = k - l + 1$

  Further, $S[k, r] = S[k', Z_l]$

  Thus, $Z_k \geq \min(Z_{k'}, |S[k, r]|)$
- If $Z_{k'} < |S[k, r]|$, then $Z_k = Z_{k'}$ and r, l unchanged
- If $Z_{k'} > |S[k, r]|$, then $Z_k = |S[k, r]|$ and r, l unchanged
- Otherwise, begin comparing position $r + 1$ with $|S[k, r]| + 1$

  If mismatch at position q, then $Z_k = q - k, l = k, r = q - 1$

Simple linear-time exact match algorithm

- **Why do we care about Z_i?**
- **Let \$ be a reserved symbol unobserved in P or T**
- **Let $S = P\$T$. Since \$ not in T, $|P| = \max(Z_i)$**
- **When $Z_i = |P|$, there is a match beginning at $S(i)$**
- **Worst case complexity $O(|P|+|T|)$ instead of $O(|P||T|)$**

Simple linear-time algorithm: worst case

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4

↓

a a a a a a $ a a a a a a a a a a a a a a a a a a a a a a a a a a a

Z:

↑

$r = 0 \qquad k > r$

$l = 0$

Comparisons: 1

Calculations: 0

Simple linear-time algorithm: worst case

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
		↓	1																															
	a	a	a	a	a	a	$	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
Z:																																		
			↑																															

$r = 0 \qquad k > r$

$l = 0$

Comparisons: 2

Calculations: 0

Simple linear-time algorithm: worst case

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
		↓	2																														
a	a	a	a	a	a	$	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
z:																																	
			↑																														

$r = 0 \qquad k > r$

$l = 0$

Comparisons: 3

Calculations: 0

Simple linear-time algorithm: worst case

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
			↓	3																													
a	a	a	a	a	a	$	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a

Z:

↑

$r = 0$ $k > r$

$l = 0$

Comparisons: 4

Calculations: 0

Simple linear-time algorithm: worst case

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
				↓	4																												
a	a	a	a	a	a	$	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a

Z:

↑

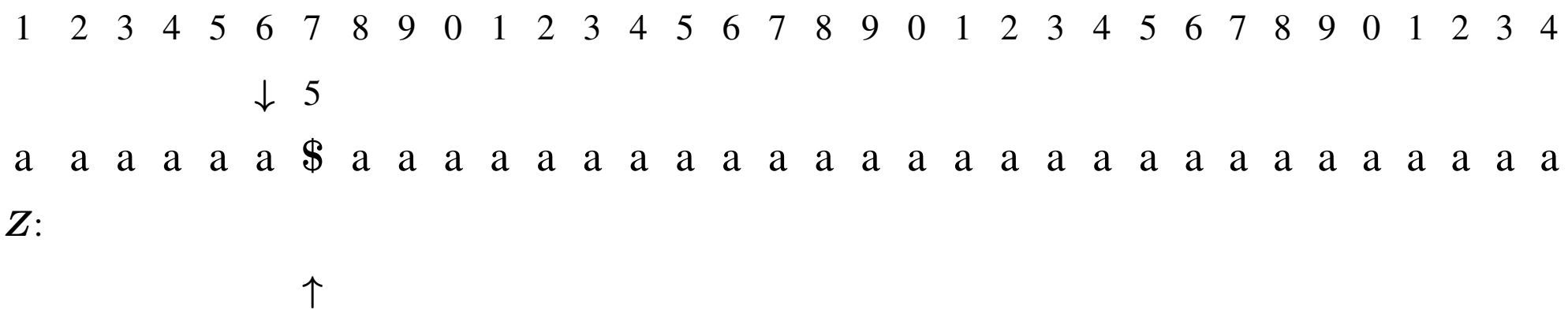
$r = 0$ $k > r$

$l = 0$

Comparisons: 5

Calculations: 0

Simple linear-time algorithm: worst case



$r = 0$ $k > r$

$l = 0$

Comparisons: 6

Calculations: 0

Simple linear-time algorithm: worst case

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
a	a	a	a	a	a	$	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
Z:	5																																
		↑																															

$r = 6 \qquad k < r,\ k' = 2,\ Z_{k'} = 5,\ |S[k, r]| = 4$

$l = 2$

Comparisons: 6

Calculations: 1

Simple linear-time algorithm: worst case

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
a	a	a	a	a	a	$	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
Z:	5	4																															
			↑																														

$r = 6 \qquad k < r,\ k' = 3,\ Z_{k'} = 4,\ |S[k, r]| = 3$

$l = 2$

Comparisons: 6

Calculations: 2

Simple linear-time algorithm: worst case

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
a	a	a	a	a	a	$	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
Z:	5	4	3																														
				↑																													

$r = 6 \qquad k < r,\ k' = 4,\ Z_{k'} = 3,\ |S[k, r]| = 2$

$l = 2$

Comparisons: 6

Calculations: 3

Simple linear-time algorithm: worst case

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
a	a	a	a	a	a	$	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
Z:	5	4	3	2																													
					↑																													

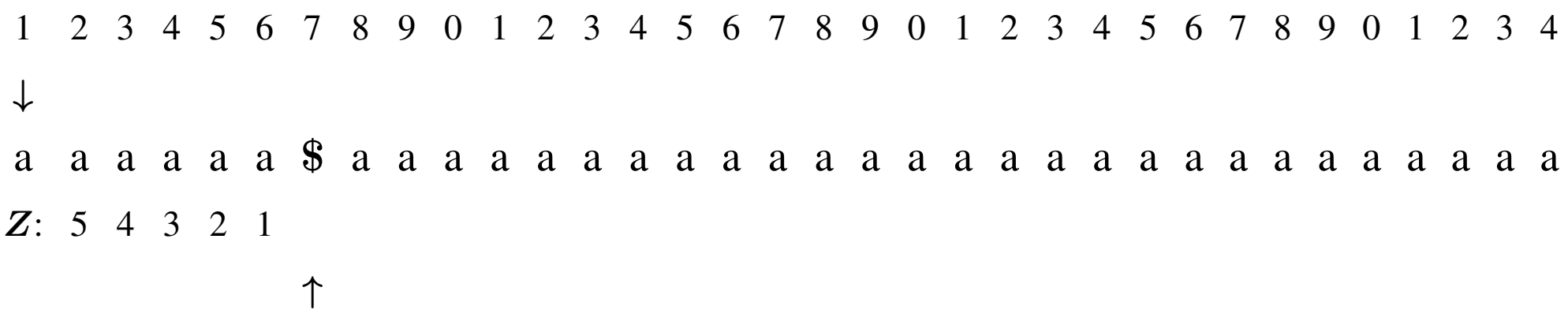
$r = 6 \qquad k = r,\ k' = 5,\ Z_{k'} = 2,\ |S[k, r]| = 1$

$l = 2$

Comparisons: 6

Calculations: 4

Simple linear-time algorithm: worst case



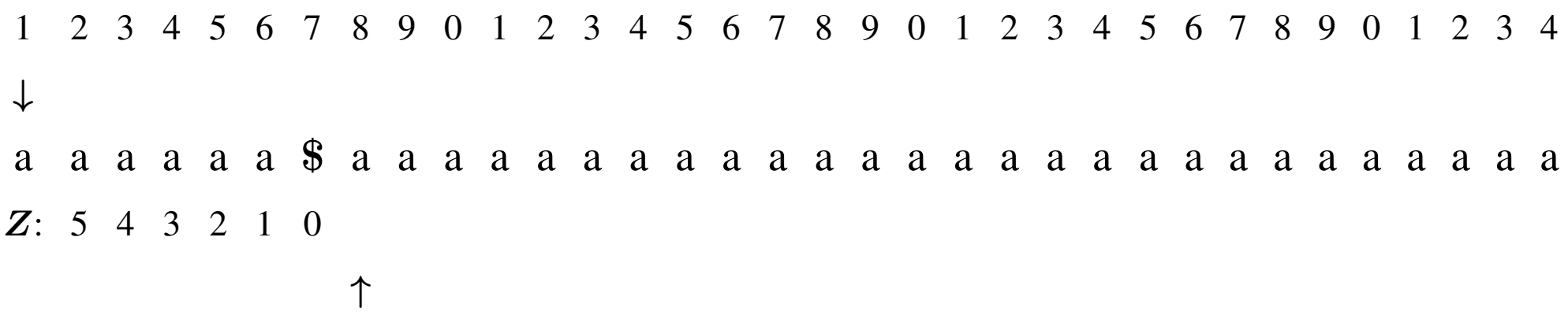
$r = 6 \qquad k > r$

$l = 2$

Comparisons: 7

Calculations: 4

Simple linear-time algorithm: worst case



$r = 6 \qquad k > r$

$l = 2$

Comparisons: 8

Calculations: 4

Simple linear-time algorithm: worst case

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
	↓																																	
	a	a	a	a	a	a	$	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
Z:	5	4	3	2	1	0	6																											
								↑																											

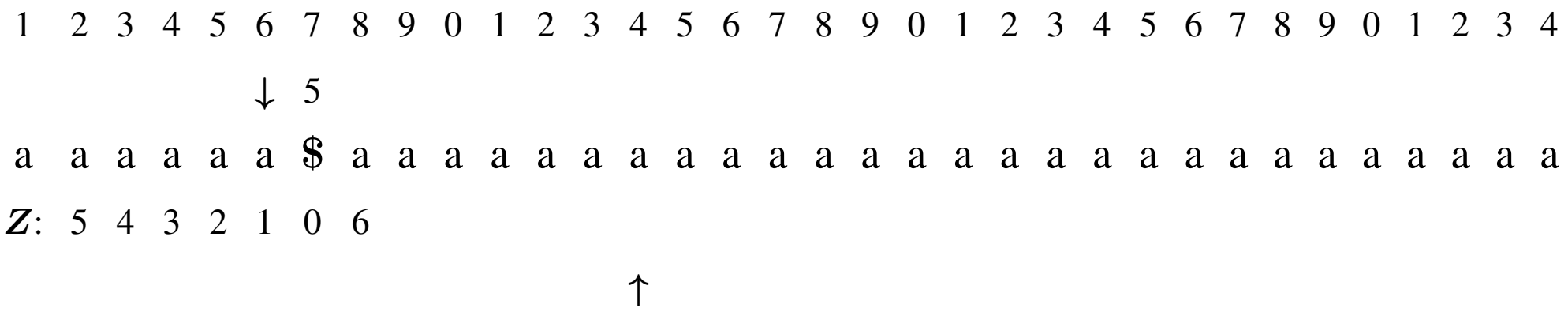
$r = 13 \qquad k < r,\ k' = 2,\ Z_{k'} = 5,\ |S[k, r]| = 5$

$l = 8$

Comparisons: 15

Calculations: 5

Simple linear-time algorithm: worst case



$r = 13 \qquad k < r,\ k' = 2,\ Z_{k'} = 5,\ |S[k, r]| = 5$

$l = 8$

Comparisons: 16

Calculations: 5

Simple linear-time algorithm: worst case

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
a	a	a	a	a	a	$	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
Z:	5	4	3	2	1	0	6	6																									
									↑																									

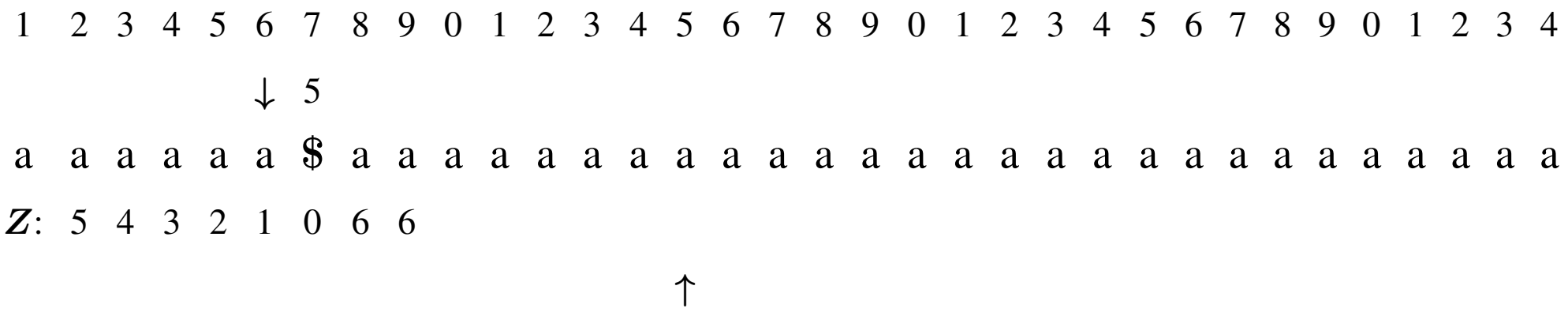
$r = 14 \qquad k < r,\ k' = 2,\ Z_{k'} = 5,\ |S[k, r]| = 5$

$l = 9$

Comparisons: 15

Calculations: 6

Simple linear-time algorithm: worst case



$r = 14 \qquad k < r,\ k' = 2,\ Z_{k'} = 5,\ |S[k, r]| = 5$

$l = 9$

Comparisons: 16

Calculations: 6

Typical case complexity

- Any position with $Z_i = 6$ (length of pattern) starts an exact match
- Clear that this algorithm, in the worst case, will be linear in the length of pattern+text
- Not clear if the improvement over naive will be that great typically
  - Less frequent matching of even short prefixes
  - Less gain from just calculation

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y

Z:

↑

$r = 0 \qquad k > r$

$l = 0$

Comparisons: 1

Calculations: 0

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0																																
	↑																																

$r = 0 \qquad k > r$

$l = 0$

Comparisons: 2

Calculations: 0

Simple linear-time algorithm

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
	↓																																	
	x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0																																
				↑																														

$r = 0 \qquad k > r$

$l = 0$

Comparisons: 3

Calculations: 0

Simple linear-time algorithm

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
		↓	1																															
	x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0																																
					↑																													

$r = 0 \qquad k > r$

$l = 0$

Comparisons: 4

Calculations: 0

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
		↓	2																														
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0																															
					↑																												

$r = 0 \qquad k > r$

$l = 0$

Comparisons: 5

Calculations: 0

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2																														
				↑																													

$r = 5$ $\quad k = r,\ k' = 2,\ Z_{k'} = 0,\ |S[k, r]| = 1$

$l = 4$

Comparisons: 5

Calculations: 1

Simple linear-time algorithm

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
	↓																																	
	x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0																														
						↑																												

$r = 5 \qquad k > r$

$l = 4$

Comparisons: 6

Calculations: 1

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0																												
						↑																												

$\boldsymbol{r = 5}$ $\boldsymbol{k > r}$

$\boldsymbol{l = 4}$

Comparisons: 7

Calculations: 1

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0																											
							↑																											

$\boldsymbol{r = 5}$ $\boldsymbol{k > r}$

$\boldsymbol{l = 4}$

Comparisons: 8

Calculations: 1

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
	↓	1																															
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0																											
								↑																										

$\boldsymbol{r = 5}$ $\boldsymbol{k > r}$

$\boldsymbol{l = 4}$

Comparisons: 9

Calculations: 1

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1																										
								↑																										

$r = 8 \qquad k > r$

$l = 8$

Comparisons: 10

Calculations: 1

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0																									
									↑																									

$r = 8 \qquad k > r$

$l = 8$

Comparisons: 11

Calculations: 1

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0																								
										↑																							

$r = 8 \qquad k > r$

$l = 8$

Comparisons: 12

Calculations: 1

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
	↓	1																															
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0																								
											↑																							

$r = 8 \qquad k > r$

$l = 8$

Comparisons: 13

Calculations: 1

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
		↓	2																														
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0																								
												↑																						

$r = 8 \qquad k > r$

$l = 8$

Comparisons: 14

Calculations: 1

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
			↓	3																													
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0																								
													↑																					

$r = 8 \qquad k > r$

$l = 8$

Comparisons: 15

Calculations: 1

Simple linear-time algorithm

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
					↓	4																													
	x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0						↑																				

$r = 8 \qquad k > r$

$l = 8$

Comparisons: 16

Calculations: 1

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
					↓	5																											
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0																								
															↑																		

$r = 8 \qquad k > r$

$l = 8$

Comparisons: 17

Calculations: 1

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
						↓	6																										
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0																								
																↑																	

$r = 8 \qquad k > r$

$l = 8$

Comparisons: 18

Calculations: 1

Simple linear-time algorithm

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
	x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6																								

↑

$r = 16 \qquad k < r,\ k' = 2,\ Z_{k'} = 0,\ |S[k, r]| = 5$

$l = 11$

Comparisons: 18

Calculations: 2

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0																						

↑

$r = 16 \qquad k < r,\ k' = 3,\ Z_{k'} = 0,\ |S[k, r]| = 4$

$l = 11$

Comparisons: 18

Calculations: 3

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0	0																					

↑

$r = 16 \qquad k < r,\ k' = 4,\ Z_{k'} = 2,\ |S[k, r]| = 3$

$l = 11$

Comparisons: 18

Calculations: 4

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0	0	2																				

↑

$r = 16 \qquad k < r,\ k' = 5,\ Z_{k'} = 0,\ |S[k, r]| = 2$

$l = 11$

Comparisons: 18

Calculations: 5

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0																				

↑

$r = 16 \qquad k = r,\ k' = 6,\ Z_{k'} = 0,\ |S[k, r]| = 1$

$l = 11$

Comparisons: 18

Calculations: 6

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0																		
																↑																	

$r = 16 \qquad k > r$

$l = 11$

Comparisons: 19

Calculations: 6

Simple linear-time algorithm

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
	↓																																	
	x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0																		
																		↑																	

$r = 16 \qquad k > r$

$l = 11$

Comparisons: 20

Calculations: 6

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0																
																		↑															

$r = 16 \qquad k > r$

$l = 11$

Comparisons: 21

Calculations: 6

Simple linear-time algorithm

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
	↓																																	
	x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0																
																				↑														

$r = 16 \qquad k > r$

$l = 11$

Comparisons: 22

Calculations: 6

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0														
																				↑													

$r = 16 \qquad k > r$

$l = 11$

Comparisons: 23

Calculations: 6

Simple linear-time algorithm

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
		↓	1																															
	x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0															
																						↑												

$r = 16 \qquad k > r$

$l = 11$

Comparisons: 24

Calculations: 6

Simple linear-time algorithm

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
			↓	2																														
	x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0															
																							↑											

$r = 16 \qquad k > r$

$l = 11$

Comparisons: 25

Calculations: 6

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
			↓	3																													
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0														
																							↑										

$r = 16 \qquad k > r$

$l = 11$

Comparisons: 26

Calculations: 6

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
				↓	4																												
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0														
																								↑									

$r = 16 \qquad k > r$

$l = 11$

Comparisons: 27

Calculations: 6

Simple linear-time algorithm

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
						↓	5																												
	x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0							↑								

$r = 16 \qquad k > r$

$l = 11$

Comparisons: 28

Calculations: 6

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0	5														

↑

$r = 25 \qquad k < r,\ k' = 2,\ Z_{k'} = 0,\ |S[k, r]| = 4$

$l = 21$

Comparisons: 28

Calculations: 7

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0	5	0												
																						↑											

$r = 25 \qquad k < r,\ k' = 3,\ Z_{k'} = 0,\ |S[k, r]| = 3$

$l = 21$

Comparisons: 28

Calculations: 8

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0	5	0	0											
													↑																					

$r = 25 \qquad k < r,\ k' = 4,\ Z_{k'} = 2,\ |S[k, r]| = 2$

$l = 21$

Comparisons: 28

Calculations: 9

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
		↓	2																														
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0	5	0	0											
															↑																		

$r = 25 \qquad k < r,\ k' = 4,\ Z_{k'} = 2,\ |S[k, r]| = 2$

$l = 21$

Comparisons: 29

Calculations: 9

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0	5	0	0	2											
														↑																				

$r = 25 \qquad k < r,\ k' = 5,\ Z_{k'} = 0,\ |S[k, r]| = 1$

$l = 21$

Comparisons: 29

Calculations: 10

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0	5	0	0	2	0									
															↑																		

$r = 25 \qquad k > r$

$l = 21$

Comparisons: 30

Calculations: 10

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
↓																																	
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0	5	0	0	2	0	0									
																										↑								

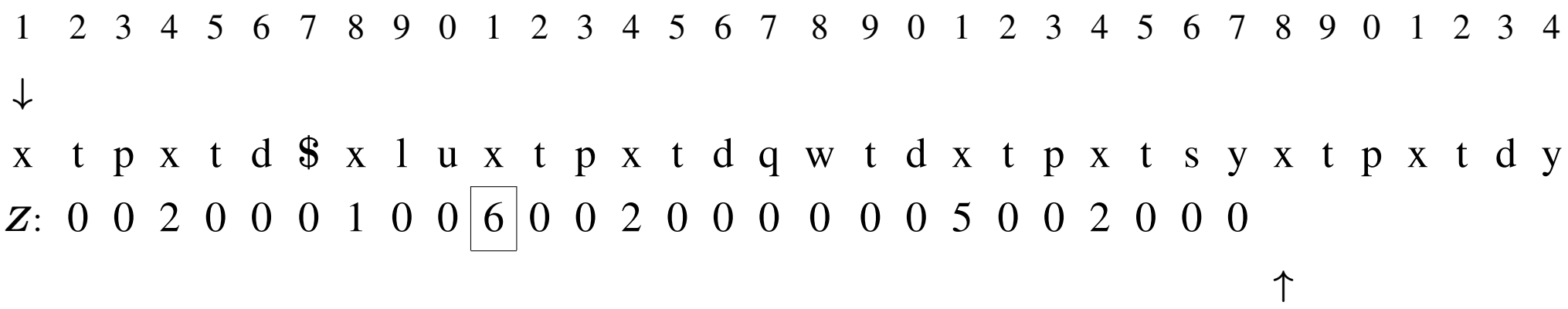
$r = 25 \qquad k > r$

$l = 21$

Comparisons: 31

Calculations: 10

Simple linear-time algorithm



$r = 25 \qquad k > r$

$l = 21$

Comparisons: 32

Calculations: 10

Simple linear-time algorithm

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
		↓	1																															
	x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0	5	0	0	2	0	0	0								
																			↑															

$r = 25 \qquad k > r$

$l = 21$

Comparisons: 33

Calculations: 10

Simple linear-time algorithm

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
			↓	2																														
	x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0	5	0	0	2	0	0	0								
																				↑														

$r = 25 \qquad k > r$

$l = 21$

Comparisons: 34

Calculations: 10

Simple linear-time algorithm

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
				↓	3																													
	x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0	5	0	0	2	0	0	0					↑			

$r = 25 \qquad k > r$

$l = 21$

Comparisons: 35

Calculations: 10

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
				↓	4																												
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0	5	0	0	2	0	0	0							
																															↑		

$r = 25 \qquad k > r$

$l = 21$

Comparisons: 36

Calculations: 10

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
					↓	5																											
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0	5	0	0	2	0	0	0							
																																↑	

$r = 25 \qquad k > r$

$l = 21$

Comparisons: 37

Calculations: 10

Simple linear-time algorithm

1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
						↓	6																										
x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
Z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0	5	0	0	2	0	0	0							
																																↑	

$r = 25 \qquad k > r$

$l = 21$

Comparisons: 38

Calculations: 10

Simple linear-time algorithm

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4
	x	t	p	x	t	d	$	x	l	u	x	t	p	x	t	d	q	w	t	d	x	t	p	x	t	s	y	x	t	p	x	t	d	y
z:	0	0	2	0	0	0	1	0	0	6	0	0	2	0	0	0	0	0	0	0	5	0	0	2	0	0	0	6						
																												↑						

$r = 25 \qquad k > r$

$l = 21$

Comparisons: 38

Calculations: 10

Notes

- **Better worst-case complexity, not necessarily better performance for any given P and T**
  - **For our P and T, naive algorithm 42 comparisons**
  - **Fewer comparisons, but more operations with linear-time algorithm (longer string)**
- **With language, naive algorithm often competitive**
- **Smaller alphabet of bio-sequences makes linear-time improvements more important**
- **More complex linear-time algorithms will yield other improvements**