

Boyer-Moore: Putting it together

Use bad character or good suffix rule, *whichever skips more*

- Step 1: $T: \text{GTTATAGCTGATCGCGGCGTAGCGGCGAA}$
 $P: \text{GTAGCGGGCG}$ bc: 6, gs: 0 bad character
- Step 2: $T: \text{GTTATAGCTGATCGCGGCGTAGCGGCGAA}$
 $P: \text{GTAGCGGGCG}$ bc: 0, gs: 2 good suffix
- Step 3: $T: \text{GTTATAGCTGATCGCGGCGTAGCGGCGAA}$
 $P: \text{GTAGCGGGCG}$ bc: 2, gs: 7 good suffix
- Step 4: $T: \text{GTTATAGCTGATCGCGGCGTAGCGGCGAA}$
 $P: \text{GTAGCGGGCG}$

11 characters of T we ignored



Step 1: $T:$ GTTATAGCT**T**GATCGCGGCGTAGCGGCGAA
 $P:$ **G**TAGCGGG**C**G

Step 2: $T:$ GTTATAGCTGAT**CGCG**GC**G**TAGCGGCGAA
 $P:$ **G**TAG**C**GG**C**G

Step 3: $T:$ GTTATAGCTGAT**CGCGGCG**TAGCGGCGAA
 $P:$ **G**T**A**CGGG**C**G

Step 4: $T:$ GTTATAGCTGATCGCGG**C**G**T**AGCGGCGAA
 $P:$ **G**T**A**CGGG**C**G
A horizontal sequence of 16 small squares. The first 7 squares are blue and aligned under the 7 characters of T that are being aligned with the first character of P. The next 2 squares are red and aligned under the 2 characters of P that are being aligned with the second character of T. The final 7 squares are blue and aligned under the 7 characters of T that are being aligned with the last character of P.

Skipped 15 alignments

Boyer-Moore: Preprocessing

Pre-calculate skips. For bad character rule, $P = \text{TCGC}$:

P

	T	C	G	C
A	0	1	2	3
C	0	-	0	-
G	0	1	-	0
T	-	0	1	2

Σ

$T: \text{AATC} \color{red}{A} \text{ATAGC}$
 $P: \color{orange}{T} \color{black}{C} \color{red}{G} \color{green}{C}$

