

# Sequence Analysis SS 2015

Freie Universität Berlin, Institut für Informatik  
Knut Reinert, Enrico Siragusa  
Sommersemester 2015

8. Exercise sheet, 5. June 2015  
Discussion: 9. June 2015

*Exercise 1 (PEX Algorithm).*

The script (Chapter 13.4) gives the pseudo-codes of the preprocessing and search phase of the PEX algorithm.

Apply them to:

1. construct a search tree for the pattern BRAUN within edit distance  $k = 2$ .
2. use the above search tree to match the pattern in the text BLAUKRAUT.

*Exercise 2.*

Prove Lemma 1 from the script (Chapter 13.1).

*Exercise 3.*

Consider the alphabet  $\Sigma = \{A, C, G, T\}$ , a text of length  $n = 3 \cdot 10^9$  and a pattern of length  $m = 100$ , both uniformly generated over  $\Sigma$ .

Given hamming distance  $k = 5$ , compute the expected number of verifications for:

1. filtration with exact seeds ( $s = 6$ );
2. filtration with 1-approximate seeds ( $s = 3$ ).