

# Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, SS 2019  
Dr. Daniel Dörr · Michel T. Henrichs

<https://gi.cebitec.uni-bielefeld.de/teaching/2019summer/sa>

**Übungsblatt 6 vom 13.5.2019**  
**Abgabe am 20.5.2019 bis 12:00 Uhr (mittags)**

## Aufgabe 1 (Approximative Textsuche)

(4 Punkte)

1. Finde die Endpositionen aller Vorkommen des Patterns  $x = \text{ATAT}$  im Text  $y = \text{GGTGATATGTAAAC}$  mit maximal  $k = 1$  Fehlern. Verwende die Cutoff-Variante von Sellers' Algorithmus mit Einheitskosten und markiere die last essential indices.
2. Gib zu jeder gefundenen Endposition alle zugehörigen Alignments an.
3. Warum ist es sinnvoll, in der Praxis nicht alle gefundenen Endpositionen auszugeben? (Stichwort: *Runs*.) Welche Endpositionen würde man im obigen Beispiel nicht ausgeben?

## Aufgabe 2 (Forward-Backward Technik)

(6 Punkte)

Gegeben seien die beiden Strings  $x = \text{TGCAT}$  und  $y = \text{TAT}$ . Die Indizierung beginnt bei 1.

1. Berechne die Matrizen  $D, D^{rev} = D^{-1}$  und  $C$  unter dem Einheitskostenmodell.
2. Lies alle optimalen Alignments aus  $C$  ab und gib diese an.
3. Welche zusätzlichen Kosten hat ein Alignment unter der Einschränkung, dass die Präfixe  $x[1..3]$  und  $y[1..2]$  aligniert sind?
4. Gib alle Alignments an, die im Vergleich zum optimalen Alignment zusätzliche Kosten von 1 haben.

## Aufgabe 3 (Backtracking)

(5 Punkte)

1. Schreibe Backtracking-Algorithmus als Pseudocode auf, wie er im Skript auf Seite 22 beschrieben wird.
2. Führe deinen Algorithmus am Beispiel dieser Backtracing-Matrix vor, die auf einem lokalen Alignment zweier Sequenzen basiert. Nehme dabei an, dass Knoten  $v_E$  auf Knoten  $v_{4,6}$  zeigt.

$i \setminus j$	0	1	2	3	4	5	6	7
0	$v_S$	$v_S$	$v_S$	$v_S$	$v_S$	$v_S$	$v_S$	$v_S$
1	$v_S$	$v_{0,0}, v_S$	$v_{0,1}, v_S$	$v_{0,2}, v_S$	$v_{1,3}$	$v_{0,4}, v_{1,4}, v_S$	$v_{0,5}, v_S$	$v_{0,6}, v_S$
2	$v_S$	$v_{1,0}$	$v_{2,1}$	$v_{1,3}$	$v_{1,3}$	$v_{1,4}, v_{2,4}$	$v_{1,5}, v_{2,5}, v_S$	$v_{1,6}, v_S$
3	$v_S$	$v_{2,1}$	$v_{2,1}$	$v_{2,2}, v_{3,2}$	$v_{2,3}, v_{2,4}$	$v_{2,4}$	$v_{3,5}$	$v_{3,6}$
4	$v_S$	$v_{3,0}, v_{3,1}, v_S$	$v_{3,1}$	$v_{3,2}, v_{4,2}$	$v_{3,3}$	$v_{3,5}$	$v_{3,5}$	$v_{3,6}, v_{4,6}$