

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, WS 2016/2017

Prof. Dr. Jens Stoye · M.Sc. Tizian Schulz

<https://gi.cebitec.uni-bielefeld.de/teaching/2016winter/sa>

Übungsblatt 8 vom 20.12.2016

Abgabe am Dienstag, den 10.01.2017

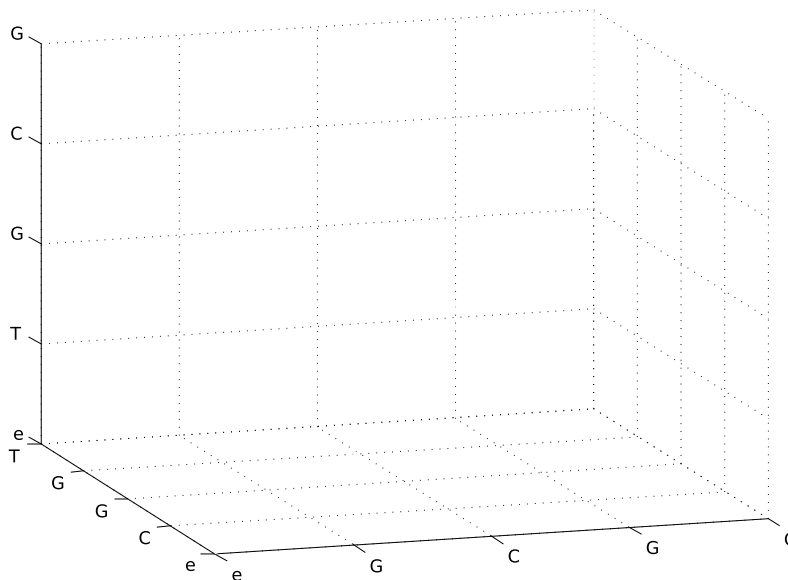
Aufgabe 1 (Geometrische Projektion des Alignments)

(2 Punkte)

Gegeben sei das multiple Alignment der Sequenzen $s_1 = \text{GCGC}$, $s_2 = \text{TGCG}$, $s_3 = \text{CGGT}$:

```
- G C G C
T G C G -
C G - G T
```

Zeichne das Alignment im 3-Dimensionalen Raum und die jeweiligen Projektionen auf den 2-Dimensionalen Unterräumen ein:



Aufgabe 2 (Center-Star-Approximation)

(5 Punkte)

Gegeben sind die Sequenzen $s_1 = \text{TATGA}$, $s_2 = \text{TAGA}$ und $s_3 = \text{CACA}$. Benutze für deine folgenden Berechnungen Einheitskosten.

1. Berechne die *Center-Sequenz* s_c .
2. Erstelle das multiple Alignment A_c und gib seine Sum-of-Pairs-Kosten an.
3. Was kannst du über die Optimalität des gefundenen Alignments sagen?
4. Beschreibe in eigenen Worten die Laufzeit- und Speicherkomplexität der Center-Star-Approximation. Unterscheide dabei zuerst die einzelnen Phasen und erkläre dann das Gesamtergebnis.

Aufgabe 3 (Forward-Backward Technik)

(6 Punkte)

Gegeben seien die beiden Strings $x = \text{ACGTA}$ und $y = \text{ATA}$. 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 minimalen zusätzlichen Kosten hat ein Alignment unter der Einschränkung, dass $x[2]$ und $y[1]$ aligniert sind?
4. Gib alle Alignments an, die im Vergleich zum optimalen Alignment zusätzliche Kosten von 1 haben.