

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, SS 2017

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

<https://gi.cebitec.uni-bielefeld.de/teaching/2017summer/sa>

Übungsblatt 9 vom 11.07.2017

Abgabe in einer Woche vor Beginn der Vorlesung.

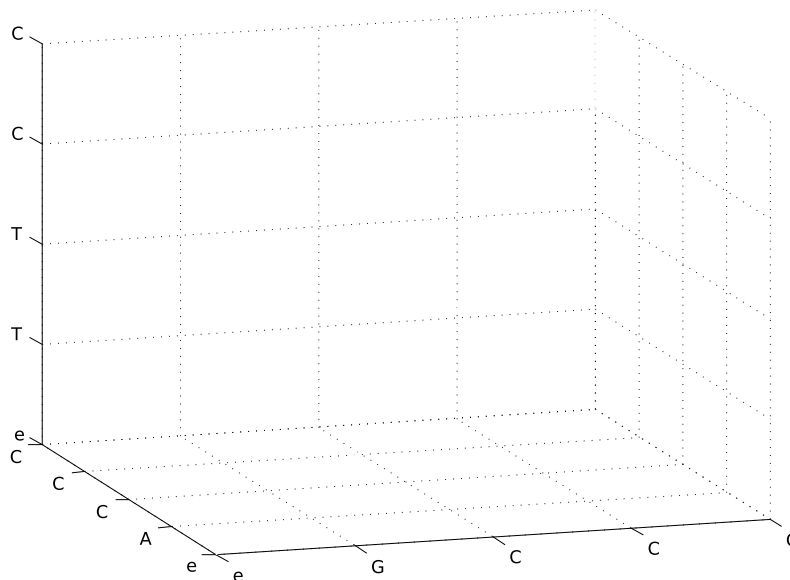
Aufgabe 1 (Geometrische Projektion des Alignments)

(2 Punkte)

Gegeben sei das multiple Alignment der Sequenzen $s_1 = GCCC$, $s_2 = ACCC$, $s_3 = TTCC$:

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

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 $s_1 = ATACT$, $s_2 = ATCT$ und $s_3 = GTGT$. Benutze für deine folgenden Berechnungen Einheitskosten.

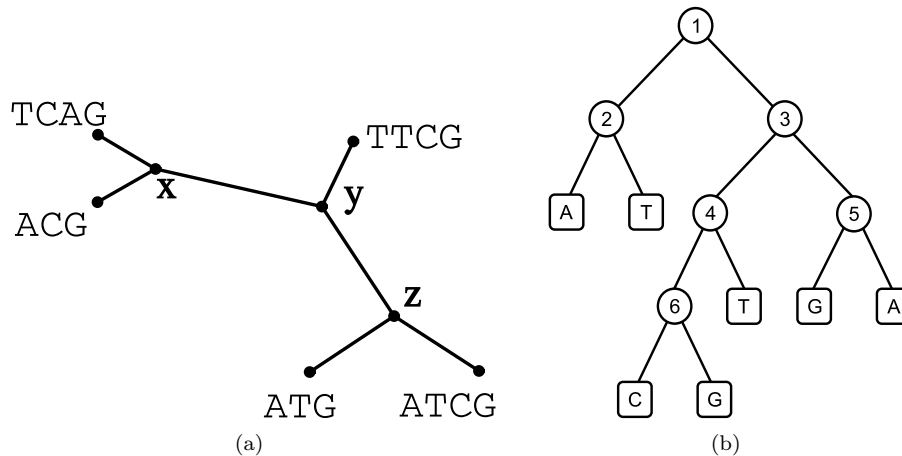
1. Berechne die *Center-Sequenz* s_c .
2. Erstelle das multiple Alignment A_c und gib seine Sum-of-Pairs-Kosten an.
3. Was kannst du über die Optimalität des gefundenen Alignments sagen?
4. Beschreibe in eigenen Worten die Laufzeit- und Speicherkomplexität der Center-Star-Approximation. Unterscheide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

Bitte wenden!

Aufgabe 3 (Baualignment)

(6 Punkte)

Gegeben seien die folgenden Bäume:



1. Um die Kosten für ein Baualignment exakt zu berechnen, kann der Sankoff-Algorithmus verwendet werden. Erkläre in 3–5 Sätzen, wie dieser Algorithmus funktioniert. Erläutere auch, welche Rolle der Fitch-Algorithmus dabei spielt.
2. Beschrifte die inneren Knoten x , y und z des Baumes in Abbildung 1(a) so, dass die Kosten des Baualignments möglichst gering sind. (Hier soll kein Algorithmus angewendet werden, sondern “gut geraten” werden.) Berechne die Kosten des entsprechenden Alignments.
3. Berechne für den phylogenetischen Baum in Abbildung 1(b) die sparsamste Beschriftung der inneren Knoten mit Hilfe des Fitch-Algorithmus. Gib dabei für die inneren Knoten jeweils die mit ihnen assoziierten Informationen der Bottom-Up- und der Top-Down-Phase an.