

Übungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, SS 2019

Dr. Daniel Dörr · Michel T. Henrichs

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

Übungsblatt 12 vom 24.6.2019

Abgabe am 1.7.2019 bis 12:00 Uhr (mittags)

Aufgabe 1 (Carrillo-Lipman-Heuristik)

(3 Punkte)

1. Wie viele Carrillo-Lipman-Schranken $U_{x,y}$ müssen berechnet werden, um k Sequenzen zu alignieren? Wie ist die asymptotische Laufzeit für die Berechnung der Schranken?
2. Für welche Sequenzen funktioniert die Carrillo-Lipman-Heuristik gut bzw. schlecht?

Aufgabe 2 (Center-Star-Approximation)

(5 Punkte)

Gegeben sind die Sequenzen $s_1 = \text{ATACT}$, $s_2 = \text{ATCT}$ und $s_3 = \text{GTGT}$. 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 mithilfe deiner Lösung über die Kosten des optimalen Alignments sagen?
4. Beschreibe in eigenen Worten die Laufzeit- und Speicherplatzkomplexität der Center-Star-Approximation. Unterscheide dabei zuerst die einzelnen Phasen und erkläre dann das Gesamtergebnis.

Aufgabe 3 (Längste gemeinsame Subsequenz)

(6* Punkte)

Schreibe ein Programm in einer Programmiersprache deiner Wahl, das zwei Sequenzen x und y entgegennimmt und mittels dynamischer Programmierung folgende Angaben berechnet:

1. Die Länge der längsten gemeinsamen Subsequenz von x und y .
2. Eine längste gemeinsame Subsequenz von x und y .
3. Alle längsten gemeinsamen Subsequenzen von x und y .