

Präsenzübungen zur Vorlesung Sequenzanalyse

Universität Bielefeld, SS 2023
Prof. Dr. Jens Stoye · Tizian Schulz

<https://gi.cebitec.uni-bielefeld.de/teaching/2023summer/sa>

Präsenzübungsblatt 10, vom 12./13.6.2023

Aufgabe 1 (Algorithmen-Tabelle)

Mache dir klar, welche Laufzeiten die folgenden Algorithmen haben:

Berechnung der q -gram-Distanz; WOTD-Algorithmus; Berechnung der Maximal-Matches-Distanz; Needleman-Wunsch-Algorithmus; Smith-Waterman-Algorithmus; Algorithmus für Alignment mit free end gaps; Gotoh-Algorithmus; Waterman-Eggert-Algorithmus; Sellers' Algorithmus; Hirschberg-Algorithmus; Algorithmus, um einen String exakt in einem Suffixbaum zu finden; Algorithmus, um mit Hilfe eines Suffixbaums den kürzesten eindeutigen String zu finden; Algorithmus, um maximale Repeats mit einem Suffixbaum zu finden; Konstruktion des Suffixarrays pos , wenn der Suffixbaum gegeben ist; Berechnung der Burrows-Wheeler-Transformation (BWT) einer Sequenz; Rücktransformation einer BWT-Sequenz; Fülle dazu eine Tabelle in der unten stehenden Form aus. Nenne den Namen des Algorithmus und beschreibe kurz, was er macht, wenn man es nicht schon am Namen erkennen kann. Schreibe die asymptotische Laufzeit auf und erkläre die Bedeutung der Parameter. Verstehe, wie die Laufzeit zu Stande kommt, aber schreibe es nicht auf.

Algorithmus	Zweck	Laufzeit	Parametererklärung
Berechnung der q -gram-Distanz	–	$\mathcal{O}(\sigma^q + m + n)$	σ : Alphabetgröße, q : Länge des q -grams, m und n : Länge der Sequenzen
...			
Needleman - Wunsch - Algorithmus	berechnet ein globales Alignment zwischen zwei Sequenzen	$\mathcal{O}(n \cdot m)$	m und n : Länge der Sequenzen
...			