

Prof. Dr. Knut Reinert
Sascha Meiers
Thimo Wellner

Institut für Informatik
AG Algorithmische Bioinformatik

Algorithmen und Datenstrukturen in der Bioinformatik

Achtes Übungsblatt WS 13/14

Abgabe Montag, 13.01.2014, 15:00 Uhr

Name:

Matrikelnummer:

Niveau I

Aufgabe1: q -gram-Lemma

Beweisen Sie das q -gram-Lemma: Zwei Strings der Länge w , die sich in höchstens k Stellen unterscheiden, haben mindestens $w + 1 - (k + 1)q$ gemeinsame q -gramme.

Aufgabe 2: Quasar

Bauen Sie das Suffix-Array und die q -gram-Tabelle für $k = 2$ und das Suchpattern $P = \text{ACGACACA}$ auf. Erklären Sie, wie der q -gram-Index bei der Suche im Quasar-Algorithmus eingesetzt wird.

Aufgabe 3: WSOP

Gegeben sei folgendes Scoring scheme:

$$\begin{aligned} &+4 \text{ for } x = y; \\ &-2 \text{ for } x \neq y, (x, y) \in \text{Alphabet} \times \text{Alphabet}; \\ &0 \text{ for } x = y = '-'; \\ &-1 \text{ otherwise;} \end{aligned}$$

Bewerten Sie folgendes Alignment mit WSOP:

$A1 = \text{AC} - \text{CTCGTA} - \text{ATG}$, $A2 = \text{TCCCAGGT} - -\text{C} - \text{G}$, $A3 = \text{ACCTTTGA} - -\text{GGT}$

Niveau II

Aufgabe 4:Progressives Alignment

Erklären Sie in 50-100 Wörtern, wie das MSA-Problem heute heuristisch mit progressivem Alignment gelöst wird. Nutzen Sie dazu das Material 'MultiMobitex.pdf' S. 31-60.