

Algorithmen und Datenstrukturen in der Bioinformatik

Achtes Übungsblatt WS 14/15

Abgabe Donnerstag 12:00

Niveau I

Aufgabe 1: Quasar: shape Gegeben seien zwei Worte mit Länge 9, die sich an Positionen 3 und 6 unterscheiden (==X==X==).

- Wie viele Treffer gibt es mit der Shape $\#.###$?
- Finden sie eine Shape mit mehr Treffern, die $size = 4$ und $span = 6$ hat.
- Wie viele mögliche Shapes gibt es mit $size = 4$ und $span = 7$?

Aufgabe 2: WSOP

Berechnen Sie den Score des folgenden MSA mittels des WSOP-Algorithmus aus der Vorlesung. Gegeben seien folgende Sequenzen (Basen unabhängig, gleichverteilt): $a1=AA-CG-GCCCTT$, $a2=ACGTA-AGGTGA$, $a3=GATGGTC-CC-C$

Nutzen Sie dazu folgendes Scoringschema:

align	score
$x = y$	5
$x! = y$	-3
$x = y = ' -'$	0
else	-1

Niveau II

Aufgabe 3: Multiple string matching

Geben Sie einen weiteren Algorithmus an (nicht Aho-Corasick), der viele Pattern gleichzeitig in einem Text sucht (Beschreibung reicht, kein Pseudocode verlangt). Sie können dafür einen bekannten Algorithmus abwandeln (z.B. *Shift-Or* oder *Horspool*) oder sich einen neuen ausdenken.

Welche Laufzeit hat ihr Algorithmus?
