

Übungen zur Vorlesung Sequenzanalyse II

Universität Bielefeld, SS 2010
Prof. Dr. Jens Stoye · Dr. Inke Herms

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2010summer/SequenzAnalyse>

Blatt 7 vom 28.05.2010

Abgabe in einer Woche vor Beginn der Vorlesung.
Bitte gib auch den Namen deines Tutors an.

Aufgabe 1 (Parametrisches Alignment)

3 Punkte

Wir betrachten Alignments mit affinen Gapkosten ($g(\ell) = -d - e(\ell - 1)$) und den Alignment-Score $S_A(d, e)$ als lineare Funktion in den Parametern d und e . Ausgehend von einem Punkt p im (d, e) -Parameterraum mit einem optimalen globalen Alignment A findet Newton's Ray-Search-Algorithmus auf einem von p ausgehenden Strahl h den Punkt r^* , an dem A gerade noch zu den optimalen Alignments gehört. Erkläre in eigenen Worten, wie r^* gefunden wird.

Aufgabe 2 (Ray-Search)

4 Punkte

Gegeben sind $s = \text{AGA}$ und $t = \text{GGCCA}$ mit festen Scores $+3$ für Matches und -3 für Mismatches. Sei $p := (6, 2)$ ein Punkt im (d, e) -Parameterraum und $h(d) = d - 4$ ein Strahl durch p .

- Berechne ein optimales globales Alignment A von s und t in p , sowie den am weitesten von p entfernten Punkt p^* auf h (in Richtung $(-1, -1)$), an dem A noch optimal ist. (Kontrolle: $p^* = (4, 0)$)
- Ausgehend von p^* , berechne (gemäß Algorithmus 11.7 im Skript) den Punkt r^* und die Gerade, die den Parameterraum teilt.

Aufgabe 3 (XPARAL-Software)

3 Punkte

Sieh dir die XPARAL-Software auf <http://www.cs.ucdavis.edu/~gusfield/xparall/index.html> an.

- Welche Möglichkeiten bietet das Programm?
- Betrachte die Zerlegung der (d, e) -Ebene für das globale Alignment der Sequenzen $s = \text{GCCGAATTGATC}$ und $t = \text{AGCGCACTAATTC}$ mit affinen Gapkosten, wobei die Scores $+3$ für Matches und -3 für Mismatches fest sind. In wieviele Polygone wird der Parameterraum unterteilt, und wieviele co-optimale Alignments gibt es jeweils?