

Algorithmen und Datenstrukturen in der Bioinformatik

Sechstes Übungsblatt WS 10/11

Abgabe Montag, 28.11., 15:00 Uhr

Name: _____ Übungsgruppe: A B C

Matrikelnummer: _____

Niveau I

Aufgabe 1: Banded Alignment

Bilden Sie ein *k*-band-Alignment der beiden Sequenzen $S_1 = \text{GAGTC}$ und $S_2 = \text{GGACT}$. Benutzen Sie als Parameter \mathbf{k} zuersts 0, danach 1. Scores : Match 5, Mismatch -4 und Gap -3 . Wie können Sie prüfen ob ihr berechneter Score optimal ist?

Niveau II

Aufgabe 2: Noch einmal Gapkosten

Für einige Anwendungen kann es sinnvoll sein, Gaps nie schlechter als einen bestimmten Wert zu bewerten. Man spricht hierbei von **capped gap costs** (gedeckelten Gapkosten). Hierzu wird einerseits ein bekanntes Scoringschema verwendet (lineare oder affine Gapkosten), übersteigen die Kosten für eine Gap jedoch den festgelegten Maximalwert g_{max} , so werden die Gaps nur mit g_{max} gewertet.

- Überlegen Sie sich, wie Sie mithilfe dynamischer Programmierung das optimale Alignment von zwei Sequenzen mit **capped gap costs** bei ansonsten **linearen Gapkosten** finden. Sie können sich am Algorithmus vom Gotoh orientieren. Schreiben Sie die **Rekursion** auf.
- Erweitern Sie diesen Algorithmus, sodass er **affine** Gapkosten berücksichtigt und notieren sie die Rekursion.