

## Algorithmen und Datenstrukturen in der Bioinformatik

### Drittes Übungsblatt WS 09/10

Abgabe Montag 09.11.2008 12:00

Name:

Übungsgruppe:

A  B  C

Matrikelnummer:

Ich kann Aufgabe \_\_\_\_\_ nicht vorrechnen.

---

#### Exercise 7: Dotplot

- Erstellen sie den Dotplot für die Sequenzen  $S_1 = \text{ACCCT}$  und  $S_2 = \text{GACCT}$
- Erstellen sie einen Dotplot, in dem sie die Sequenz  $S_3 = \text{MADAMIMADAM}$  mit sich selbst vergleichen. Zeigen Sie an Hand des Dotplots, welche Besonderheiten diese Sequenz aufweist.

---

#### Exercise 8: Needleman-Wunsch algorithm for global alignment

Berechnen Sie für die Sequenzen  $S_1$  (horizontal) und  $S_2$  (vertikal) aus Aufgabe 7 ein globales Alignment nach dem Algorithmus von Needleman-Wunsch. Ein Match trägt 3 zum Score des Alignments bei, ein Mismatch  $-5$  und ein Gap  $-2$ .

- Berechnen Sie die DP-Matrix, und
- tragen Sie dabei jeweils *alle* Traceback-Kanten ein (also eventuell mehrere pro Zelle).
- Heben Sie *die beim Traceback benutzten* Kanten besonders hervor, und
- notieren Sie das resultierende *globale Alignment*.


---

### Exercise 9: Searching in Suffix Arrays using lcp values

Suchen Sie in der Sequenz  $S = \text{acbaaabaa}$  (Von letzter Woche) das Pattern  $P = \text{aa}$  mit Hilfe der lcp-Werte. Berechnen Sie vorher nur die 15 lcp-Werte, die bei der Suche auftreten könnten. Geben Sie die Anzahl der Buchstabenvergleiche an. (Nicht die Vergleiche zwischen lcp-Werten berechnen)