

# 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>

## Wiederholungen des bisherigen Stoffes

### Aufgabe 1 (Suffixbaum/Suffixarray)

- Zeichne den Suffixbaum von  $s\$ = \text{bcabcabc}\$$ . Annotiere die Blätter mit den Suffix-Startpositionen und bilde so das Suffixarray  $\text{pos}$ . Beachte dabei die Ordnung  $\$ < b < a < c$ .
- Bestimme das lcp-Array. Wie kann man es aus dem Baum ablesen?
- Wie kann man in Linearzeit das “inverse” Suffixarray  $\text{rank}$  aus  $\text{pos}$  berechnen?

**Aufgabe 2 (RNA-Strukturvorhersage)** Definiere das Problem der RNA-Sekundärstrukturvorhersage. Welcher Algorithmus löst das Problem und wie funktioniert er?

### Aufgabe 3 (Paarweises Alignment mit linearem Speicherbedarf)

- Zeige, dass die Berechnung eines globalen Alignments mit linearem Platzbedarf etwa doppelt so viel Zeit erfordert wie dieselbe Berechnung mit der kompletten Edit-Matrix.
- Welche Probleme treten auf, wenn man die Hirschberg-Technik nicht mit linearen Gapkosten, sondern mit affinen oder gar beliebigen Gapkosten kombinieren möchte? Lassen sich die Probleme lösen?

### Aufgabe 4 (Längennormalisiertes Alignment)

- Wie ist der längennormalisierte Alignmentsscore definiert und welche Vorteile bietet er?
- Konstruiere ein beispielhaftes Sequenzalignment für den Mosaik-Effekt.
- Wie kann man effizient ein normalisiertes Alignment finden?

**Aufgabe 5 (Paarweises Alignment)** Diskutiere folgende Algorithmen (Vor- und Nachteile, praktische Relevanz) und nenne jeweils einen Vertreter:

- Standard-Version des globalen/lokalen Alignments (lineare Gapkosten, quadratischer Platzbedarf, quadratische Laufzeit),
- globales/lokales Alignment mit affinen Gapkosten (quadratischer Platzbedarf, quadratische Laufzeit),
- globales/lokales Alignment mit beliebigen Gapkosten (quadratischer Platzbedarf, kubische Laufzeit),
- globales/lokales Alignment mit linearen Gapkosten und der Hirschberg-Technik (linearer Platzbedarf, quadratische Laufzeit),
- globales/lokales Alignment mit affinen Gapkosten und der Hirschberg-Technik (linearer Platzbedarf, quadratische Laufzeit),
- BLAST (Heuristik, affine Gapkosten, nahezu lineare Laufzeit, kaum zusätzlicher Platzbedarf)