

---

## Algorithmische Bioinformatik I

---

### Aufgabe 1

Sei  $w : \bar{\Sigma} \times \bar{\Sigma} \rightarrow \mathbb{N}_0$  eine metrische Kostenfunktion für ein Distanzmaß, d.h. die Kostenfunktion ist ganzzahlig und  $w$  ist eine Metrik. Konstruiere einen Algorithmus, der für zwei Sequenzen  $s$  und  $t$  und eine gegebene Zahl  $k$  in Zeit  $O(k(|s| + |t|))$  ein optimales globales Sequence Alignment mit Distanz höchstens  $k$  konstruiert, sofern ein solches existiert.

### Aufgabe 2

- Wie muss der Hirschberg-Algorithmus bei Verwendung von Ähnlichkeitsmaße anstatt von Distanzmaßen modifiziert werden?
- Lässt sich der Hirschberg-Algorithmus für semi-globale Sequence Alignments modifizieren?
- Lässt sich der Hirschberg-Algorithmus für lokale Sequence Alignments modifizieren?

### Aufgabe 3

Betrachte das folgende so genannte *All-Against-All-Problem*:

**Geg.:**  $s, t \in \Sigma^*$  mit  $|s| = n$  und  $|t| = m$ ,  $\vartheta \in \mathbb{R}$  und ein Distanzmaß  $d$ .

**Ges.:** Alle Paare  $(s', t')$  mit  $d(s', t') \leq \vartheta$ , wobei  $s'$  bzw.  $t'$  ein Teilwort von  $s$  bzw.  $t$  ist.

Konstruiere einen Algorithmus zur Lösung des All-Against-All-Problems, der auf einem hybriden Verfahren aus Dynamischer Programmierung und Suffix-Bäumen basiert, und analysiere diesen. Die Laufzeit wird dabei auch von der Anzahl der „Treffer“ abhängen:

$$h = |\{(s', t') \mid d(s', t') \leq \vartheta \wedge s' \sqsubseteq s \wedge t' \sqsubseteq t\}|.$$

Hierbei gilt  $s' \sqsubseteq s$ , wenn  $s'$  ein Teilwort von  $s$  ist.