

7. Dezember 2005

**Abgabe: 14. Dezember 2005**

**Aufgabe 1:** Diskutieren Sie an Beispielen die Unterschiede zwischen dem Sum-of-Pairs-Score und dem Entropie-Score, sowie deren Vor- und Nachteile.

**Aufgabe 2:** Schreiben Sie ein auf dynamischer Programmierung beruhendes Programm für multiples Alignment mit linearer Gap-Penalty und Sum-of-Pairs-Score. Das Programm soll optional auch die MSA-Methode von Carillo und Lipman verwenden können. Vergleichen Sie für verschiedene Sequenz-Anzahlen und -Längen wie viel Laufzeiterparnis die Idee von Carillo und Lipman in der Praxis bringt.

**Aufgabe 3:**

(a) UPGMA kann jeden Baum, der die Molekulare-Uhr-Eigenschaft erfüllt, aus den exakten Distanzen der Blätter rekonstruieren. Beweisen oder widerlegen Sie, dass das auch für hierarchische Cluster-Verfahren gilt, die Paaren von Clustern nicht die mittlere, sondern die maximale (oder minimale) Distanz von Elementen der entsprechenden Cluster zuordnen.

(b) Vervollständigen Sie den Beweis des Lemmas über die Äquivalenz von Molekulare-Uhr- und Ultrametrik-Eigenschaft: Zeigen Sie, dass im Fall  $d(s_{n+1}, x) = d(s_{n+1}, y)$  für alle  $i \leq n$  die Gleichung  $d(s_{n+1}, x) = d(s_{n+1}, s_i)$  aus der Ultrametrik-Eigenschaft folgt.