

Übungen zur Vorlesung Sequenzanalyse II

Universität Bielefeld, SoSe 2012

Dr. Alexander Sczyrba · Nina Luhmann · Linda Sundermann

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2012summer/SequenzAnalyse>

Übungsblatt 2 vom 20.04.2012

Abgabe in einer Woche vor Beginn der Vorlesung.

Aufgabe 1 (Simple RNA-Strukturen)

(4 Punkte)

Gegeben sei eine RNA-Struktur S mit den drei Basenpaaren $\{i, j\} \in S$, $\{k, l\} \in S$ und $\{m, n\} \in S$. Zeichne für die folgenden drei Fälle auf, was für eine Struktur sich aus der Anordnung der Basenpaare ergibt. Sind diese Strukturen nach der Definition 8.2 im Skript (S. 78) gültig? Wenn nein, warum würde es Sinn machen diese Strukturen doch zu berechnen und warum macht man es trotzdem nicht?

1. $i < k < m < n < l < j$
2. $i < k < l < j < m < n$
3. $i < k < j < l < m < n$

Aufgabe 2 Darstellung einfacher RNA-Sekundärstrukturen

(4 Punkte)

Folgende RNA-Struktur sei in *dot-bracket*-Notation dargestellt:

```
5' ACUGGAAACAUAAGGGUCGGAUCCUUUUAAGGGUUC 3'
...((...)((...))((...)).((...)).
```

„Übersetze“ diese Struktur in eine *mountain plot*-Schreibweise. Diskutiere die Vor- und Nachteile der *dot-bracket*-Notation und dem *mountain plot* für RNA Sekundärstrukturen.

Aufgabe 3 Kontextfreie Grammatiken

(3 Punkte)

Die folgende kontextfreie Grammatik ist mehrdeutig:

- $S \rightarrow \epsilon$
- $S \rightarrow .S$
- $S \rightarrow S.$
- $S \rightarrow (S)$
- $S \rightarrow SS$

Erkläre, was Mehrdeutigkeit bedeutet und denke dir ein Beispiel aus, das die Mehrdeutigkeit dieser Grammatik verdeutlicht.

Aufgabe 4 Nussinov-Algorithmus

(4 Punkte)

Gegeben sei die RNA-Sequenz $s = \text{GCUAGCUGUUCA}$, $\delta = 2$ und das Scoring-Schema $\text{score}(C, G) := 3$, $\text{score}(A, U) := 2$ und $\text{score}(G, U) := 1$. Wofür steht das δ ? Berechne mit Hilfe des Nussinov-Algorithmus alle Strukturen von s mit maximalem Score und gib diese in *dot-bracket*-Notation an.