25. Oktober 2006 **Abgabe: 1. November 2006**

Aufgabe 1: Wir interessieren uns nun für lokale Alignments, die die Eigenschaft haben, dass es kein Alignment mit höherem Score gibt, das mit dem selben Basenpaar beginnt oder mit dem selben Basenpaar endet wie das gegebene Alignment. Schreiben Sie ein möglichst effizientes Programm, welches für ein gegebenes Sequenzenpaar den Score jedes solchen Alignments ausgibt, falls dieser eine eingegebene Schranke u übersteigt.

Aufgabe 2: Seien X und Y Zufallsvariablen mit endlichem Wertebereich $U \subset \mathbb{R}$, seien $a, b \in \mathbb{R}$, seien V und W zufällige Vektoren mit endlichem Wertebereich $S \subset \mathbb{R}^n$ und sei M eine $n \times n$-Matrix. Zeigen Sie:

- $\mathbb{E}(a \cdot X) = a \cdot \mathbb{E}X$
- $\mathbb{E}(a \cdot X + b \cdot Y) = a \cdot \mathbb{E}X + b \cdot \mathbb{E}Y$
- $\mathbb{E}(a \cdot V + b \cdot W) = a \cdot \mathbb{E}V + b \cdot \mathbb{E}W$
- $\mathbb{E}(M \cdot V) = M \cdot \mathbb{E}V$

Aufgabe 3: In einer rein zufällig aus $\{\mathtt{A}, \mathtt{C}, \mathtt{G}, \mathtt{T}\}$ zusammengesetzten Sequenz der Länge n werden mit dem Knuth-Morris-Pratt-Algorithmus (einzeln) die Muster `AAA` und `ACA` gesucht. Beantworten Sie folgende Fragen durch wahrscheinlichkeitstheoretische Argumentation oder approximativ mittels Computersimulation:

(a) Geben Sie die jeweilige Funktion π an.

(b) Wie oft kommen die beiden Muster in der Sequenz vor (in Erwartung)?

(c) Berechnen Sie für jedes der beiden Muster die erwartete Anzahl an Positionen bis zu der Stelle, wo das Muster zum ersten Mal beginnt. (Oder begründen Sie zumindest ob die Erwartungswerte für beide Muster gleich sind.)

(d) Wieviele Vergleiche werden im Schnitt (Erwartungswert!) durchgeführt, bis das Muster zum ersten Mal beginnt?

(e) Zwei Spieler wetten, welches der beiden Muster als erstes in der Sequenz vorkommt. Ist das ein faires Spiel? Begründen Sie!