Übungen Algorithmen der Bioinformatik II
Wintersemester 2005/06
Prof. Dr. Stefan Posch, Dr. Birgit Möller

Institut für Informatik
Universität Halle

Blatt 10

Aufgabe 10.1 Wir modellieren Sequenzen der Länge L über einem festen Alphabet der Größe D mit einem inhomogenen Markov Modell 1-ter Ordnung. Als Daten stehen uns N Sequenzen der Länge L zur Verfügung, die – wie üblich – als paarweise unabhängige Realisierung diese iMM(1) angenommen werden. Als a priori Verteilung der Parameter des iMM(1) nehmen wir statistisch unabhängige Dirichlet-Verteilungen an.

Geben Sie die explizit die ML-, MAP- und MP-Schätzwerte der Parameter des iMM(1) an.

Aufgabe 10.2 Wir betrachten N Sequenzen über einem festen Alphabet, die nun aber unterschiedliche Längen L_n haben.

Geben Sie für folgende Modelle für Sequenzen der Länge $L, L > 5$, jeweils die ML-Schätzer an:

(a) ein homogenes Markov Modell 5-ter Ordnung

(b) ein "periodisches" inhomogenes Markov Modell 5-ter Ordnung

(also ein iMM(5), für das gilt:
$P_l(X_l|X_{l-M}, \cdots, X_{l-1}) = P_k(X_k|X_{k-M}, \cdots, X_{k-1})$, falls $k = l\%5$)

Wir wünschen Euch ein frohes Weihnachtsfest und einen guten Rutsch!

Abgabe: 12.01.06