

Prof. Dr. Knut Reinert
Kathrin Trappe
René Rahn
Kathleen Gallo
Thomas Krannich

Institut für Informatik
AG Algorithmische Bioinformatik

Algorithmen und Datenstrukturen in der Bioinformatik

Dreizehntes Übungsblatt WS 12/13

Abgabe Montag, 01.02.2013, 15:00 Uhr

Name:

Übungsgruppe:

A B C

Matrikelnummer:

Niveau I

Aufgabe 1: Viterbi und Posterior Decoding

Betrachten Sie das HMM mit den Transitionswahrscheinlichkeiten A und den Emissionswahrscheinlichkeiten e der Abb.2.

- Berechnen Sie für die Sequenz xyx den Viterbi- und den posterior decoding Pfad.
 - Erklären Sie, warum sich die beiden unterscheiden. Warum kann das Ergebnis des "falschen" Pfades trotzdem von Interesse sein?
-

Aufgabe 2: Training von HMM's

Gegeben die Sequenz $cdcccd$ und die dazugehörige Zustandsfolge $AABAAB$

- Bestimmen sie mit der *maximum likelihood*-Methode die Parameter für einen HMM, der diese Sequenz erzeugt
 - Welche Probleme können auftreten, wenn man einen HMM mit zu wenigen Sequenzen trainiert? Wie kann man diese beheben?
-

Aufgabe 3: Modellierung von Markov-Chains in der Biologie

Überlegen Sie sich, wie sie die Aminosäure codierenden Basentriplets der DNA mit einer Markovkette modellieren würden. Gehen Sie dabei kurz auf die Eigenschaft(en) Ihrer Markovkette ein und erklären Sie Ihre Überlegungen an einem kleinen graphischen Beispiel.

Niveau II

Aufgabe 4: Das wilde Waschmaschinenrennen

Im Badezimmer von WG Y wird jede Woche ein Rennen zwischen Waschmaschine und Trockner abgehalten: Die Waschmaschine bewegt sich jede Minute *unabhängig* von der Bewegung der Vorminute entweder 1cm nach vorne (v) oder 1cm rückwärts (r), mit jeweils gleichen Wahrscheinlichkeiten.

Der Trockner bewegt sich auch entweder 1cm nach vorne (v) oder 1cm zurück (r), er behält jedoch mit 75% Wahrscheinlichkeit die Richtung bei.

Beide Geräte starten jeweils mit 50% Wahrscheinlichkeit mit einer Vorwärtsbewegung. Die Ziellinie ist in 2 cm Entfernung, beide Geräte haben beliebig viel Platz in beide Richtungen und beliebig viel Zeit.

- a) Modellieren sie beide Geräte als *Markov-Ketten* (Keine HMMs!).
- b) Mit welcher Wahrscheinlichkeit wird die Kette *rvvv* von beiden Geräten erzeugt?
- c) Mit welcher Wahrscheinlichkeit kommt die Waschmaschine ans Ziel? Wie lange braucht sie dafür durchschnittlich?
- d) (Intuition) Auf welches der beiden Geräte würden Sie ihr Geld setzen?

Zur Erinnerung: Am 07.02. findet das 2.Review zur Vorlesung statt. Es können alle Themen ab dem 1.Review (ab FASTA) abgefragt werden.