

Algorithmische Bioinformatik Übungsblatt 2

Ausgabe: 15. Oktober 2019 · Besprechung: 29. Oktober (2 Wochen!)

Aufgabe 2.1 Gegeben sind die Wahrscheinlichkeiten p, q, r durch

$$\ln p = 0, \quad \ln q = -36, \quad \ln r = -746.$$

Berechne jeweils möglichst exakt die Logarithmen von

1. pqr ,
2. $p + q + r$,
3. $q^{21} + r/746$.

Aufgabe 2.2 Betrachte die DNA-Sequenzen

$$s = \text{GATTACA}, \quad t = \text{ACATTAG}, \quad u = (\text{TACG})^{800},$$

und berechne (bei fester Sequenzlänge) jeweils ihre Wahrscheinlichkeiten unter den Textmodellen

1. M00 (Gleichverteilung),
2. M0 mit Verteilung $p = (p_A, p_C, p_G, p_T) = (1/6, 2/6, 2/6, 1/6)$,
3. M1 mit der Information, dass das erste Nukleotid gleichverteilt ist und im weiteren Verlauf nach einem A oder G jeweils mit W'keit $2/6$ ein C oder T folgt und mit jeweils W'keit $1/6$ ein A oder G, und nach C oder T entsprechend umgekehrt. (Tipp: Schreibe zunächst die Übergangsmatrix auf.)

Aufgabe 2.3 Wie lautet die stationäre Verteilung des M1-Modells aus der vorigen Aufgabe?

Aufgabe 2.4 Eine Medizinstudentin schreibt eine Klausur, die 300 Minuten dauert und aus 300 Aufgaben besteht. Tatsächlich löst sie pro Minute genau eine Aufgabe. Dabei ändert sich ihr Zustand zwischen erschöpft, wach und hochkonzentriert; sie beginnt die Prüfung wach. Abhängig von ihrem Zustand beantwortet sie jede Prüfungsfrage mit einer gewissen Wahrscheinlichkeit richtig oder falsch. Ist sie erschöpft, liegt ihre Fehlerwahrscheinlichkeit bei 50%; ist sie wach, bei 10%; ist sie hochkonzentriert, bei nur 2%. Ihr Zustand kann sich von Minute zu Minute ändern, bleibt aber zu 90% gleich. Ist sie hochkonzentriert, ist sie in der nächsten Minute mit 10% Wahrscheinlichkeit nur noch wach. Ist sie wach, dann ist sie in der nächsten Minuten mit je 5% in einem der anderen Zustände. Ist sie erschöpft, dann ist sie in der nächsten Minute mit 10% Wahrscheinlichkeit wieder wach.

1. Berechne die Wahrscheinlichkeit, dass sie die ersten hundert Aufgaben richtig beantwortet, die zweiten hundert Aufgaben falsch und die dritten hundert Aufgaben wieder richtig.
2. Man besteht mit 50% der möglichen richtigen Antworten. Schätze grob die Wahrscheinlichkeit dafür ab, dass die Studentin besteht. (Tipp: Nicht anfangen zu rechnen!)
3. (schwierig, wird im Lauf der Vorlesung leicht!) Berechne die Wahrscheinlichkeit, dass sie mindestens 200 der 300 Aufgaben richtig beantwortet.