

## 6. Übung zur Vorlesung “Algorithmische Phylogenetik”

Sebastian Böcker, Markus Fleischauer

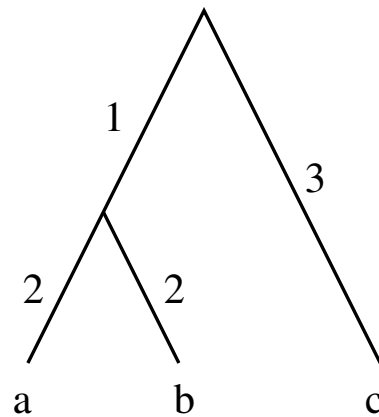
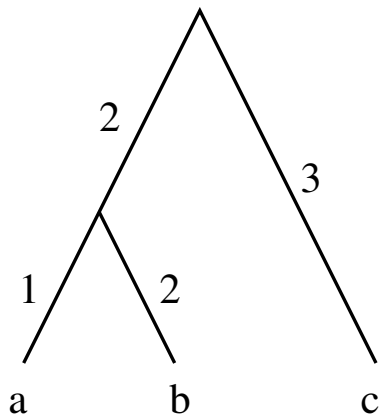
Abgabe: 16.01.2017

### Aufgabe 1 (Likelihood) (10 Punkte)

Gegeben sei das Alphabet  $\Sigma = \{a, b, c, d\}$ , die Anfangsverteilung  $\pi^{(0)} = (0.3 \ 0.2 \ 0.3 \ 0.2)$  und die folgende Übergangsmatrix

$$P = \begin{pmatrix} 0.7 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.7 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.7 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.7 \end{pmatrix}.$$

Berechnen Sie die Likelihood der folgenden Bäume. Beachten Sie, dass die Wahrscheinlichkeiten von  $d$  durchaus eine Rolle spielen, auch wenn  $d$  nicht als Label vorkommt.



**Aufgabe 2 (Jukes-Cantor)** (5 Punkte)

Gegeben sei das multiple Alignment

$$\begin{pmatrix} A & C & G & T & A & T & T & C \\ A & G & G & T & A & A & T & C \\ A & C & C & A & T & A & A & C \\ T & A & A & A & T & G & C & C \end{pmatrix}$$

Berechnen Sie die Jukes-Cantor korrigierten Distanzen der paarweisen Alignments und diskutieren Sie die Ergebnisse.

**Aufgabe 3 (Generalised time reversible, GTR)** (2 Punkte)

Warum ist das GTR-Modell das "beste" Modell der Sequenzevolution? Und warum ist es oftmals sinnvoll, simplere Modelle zu benutzen?

**Aufgabe 4 (Bootstrapping)** (3 Punkte)

Gegeben sei das Alignment

$$A = \begin{pmatrix} A & G & T & G \\ C & G & T & T \\ A & G & C & G \end{pmatrix}$$

Welche der folgenden Alignments wurden durch Bootstrapping aus  $A$  erstellt? Warum?

$$B = \begin{pmatrix} A & T & G \\ C & T & T \\ A & C & G \end{pmatrix} \quad C = \begin{pmatrix} A & T & G & A \\ C & T & G & C \\ A & G & G & A \end{pmatrix} \quad D = \begin{pmatrix} G & G & G & G \\ G & G & G & G \\ G & G & G & G \end{pmatrix}$$

**Aufgabe 5 (MC MC)** (Bonusaufgabe)

Warum lässt man bei Monte Carlo Markov Chain (MC MC) am Anfang den Prozess einbrennen?