

Übungen zur Phylogenetik Vorlesung

Universität Bielefeld, WS 2009/2010
 Dipl.-Inform. Roland Wittler · Dipl.-Inform. Peter Husemann
<http://wiki.techfak.uni-bielefeld.de/gi/GILectures/2009winter/Phylogenetik>

Blatt 8 vom 9.12.2009
Abgabe in einer Woche zu Beginn der Vorlesung.

Aufgabe 1 Rekonstruktion additiver Bäume.

(4 Punkte)

Die folgende Distanzmatrix ist *additiv*:

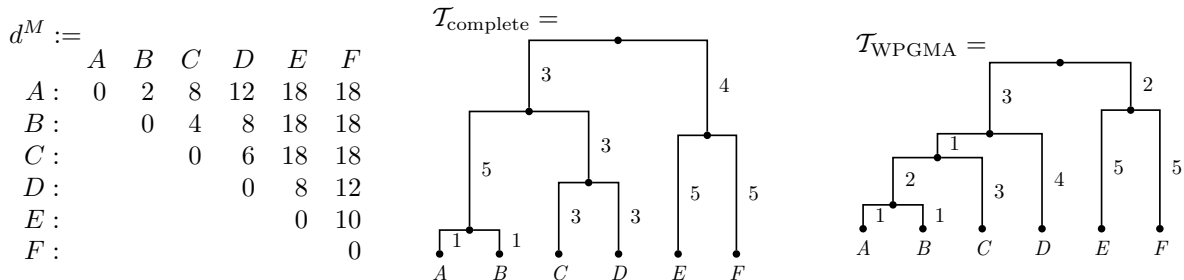
	A	B	C	D	E	F
A:	0	9	7	9	7	9
B:		0	5	5	9	5
C:			0	5	7	5
D:				0	9	3
E:					0	9
F:						0

Rekonstruiere den entsprechenden additiven Baum mit Hilfe des Algorithmus von Waterman (Skript, Abschnitt 6.3.1). Gib alle Zwischenschritte an.

Aufgabe 2 Fitch-Margoliash.

(4 Punkte)

Auf dem Übungsblatt 7 wurden *Single linkage* und *UPGMA* als Clustering-Verfahren zur Baumrekonstruktion benutzt. Werden anstatt dessen *Complete linkage* und *WPGMA* Clustering verwendet, kommen zu der gegebenen Matrix d^M die folgenden Bäume $\mathcal{T}_{\text{complete}}$ und $\mathcal{T}_{\text{WPGMA}}$ heraus:



Berechne den *Least Squares* Fehler $E := \|\vec{d}^T - \vec{d}^M\|^2$ (nach Fitch und Margoliash) für beide Bäume. Dazu gib zuerst d^M in Vektorschreibweise an. Danach führe die folgenden Schritte jeweils für beide Bäume durch:

$$\mathcal{T} \in \{\mathcal{T}_{\text{complete}}, \mathcal{T}_{\text{WPGMA}}\}$$

1. Schreibe M^T und \vec{w} bezüglich \mathcal{T} auf.
2. Berechne daraus zunächst \vec{d}^T .
3. Berechne schließlich $E := \|\vec{d}^T - \vec{d}^M\|^2$.

Welcher der beiden Bäume ist „besser“?