Übungen zur Phylogenetik Vorlesung

Universität Bielefeld, WS 2011/2012, Dr. Roland Wittler

http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2011winter/Phylogenetik

Blatt 8 vom 1.12.2011

Abgabe in einer Woche zu Beginn der Vorlesung oder vorab im Briefkasten bei U10-151.

Aufgabe 1 Zusammenhänge. **(3 Punkte)**

Bilde mindestens sechs sinnvolle Paarungen aus den folgenden Begriffen/Graphiken und erläutere jeweils in einem Satz den Zusammenhang.

(a) additiv

(b) agglomeratives Clustering

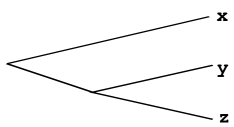
(c) Waterman

(d) UPGMA

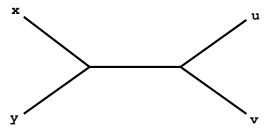
(e) ultrametrisch

(f) Vier-Punkt-Bedingung

(g)



(h)



(i) molecular clock

Aufgabe 2 Rekonstruktion additiver Bäume. **(4 Punkte)**

Die nebenstehende Distanzmatrix ist *additiv*.
Rekonstruiere den entsprechenden Baum mit Hilfe des Algorithmus von Waterman (Skript, Abschnitt 7.3.1). Gib alle Zwischenschritte an.

	A	B	C	D	E
A :	0	9	7	9	10
B :		0	6	6	7
C :			0	6	7
D :				0	3
E :					0

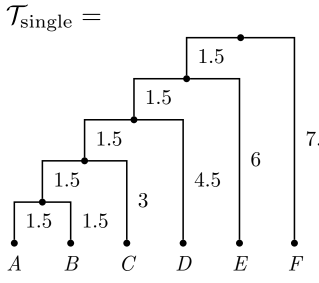
Aufgabe 3 Fitch-Margoliash. **(4 Punkte)**

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

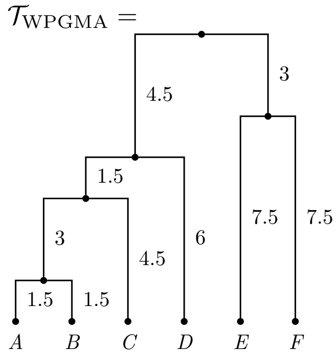
$d^M :=$

	A	B	C	D	E	F
A :	0	3	12	18	27	27
B :		0	6	12	27	27
C :			0	9	27	27
D :				0	12	18
E :					0	15
F :						0

$\mathcal{T}_{\text{single}} =$



$\mathcal{T}_{\text{WPGMA}} =$



Berechne den *Least Squares* Fehler $E := ||\vec{d}^{\mathcal{T}} - \vec{d}^M||^2$ (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:

1. Schreibe $M^{\mathcal{T}}$ und $\vec{w}$ bezüglich $\mathcal{T}$ auf. (Diesen Schritt darfst du für $\mathcal{T}_{\text{WPGMA}}$ auslassen.)
   Tipp: Fasse die zwei Kanten, die inzident mit der Wurzel sind, zu einer Kante zusammen.

2. Berechne daraus $\vec{d}^{\mathcal{T}}$. (Bzw. lies die Werte für $\mathcal{T}_{\text{WPGMA}}$ aus dem Baum ab.)

3. Berechne schließlich $E := ||\vec{d}^{\mathcal{T}} - \vec{d}^M||^2$.

Welcher der beiden Bäume ist „besser“?