

Exercises – Phylogenetics

Universität Bielefeld, WS 2017/2018,
 Dr. Roland Wittler, Prof. Dr. Jens Stoye, M. Sc. Tizian Schulz
<https://gi.cebitec.uni-bielefeld.de/Teaching/2017winter/Phylogenetik>

Exercise Sheet 8 — 30.11.2017

Due: 7.12.2017

Task 1 Reconstruction of Additive Trees.

(4 points)

The distance matrix on the right is *additive*.

Use the algorithm of Waterman (lecture notes, section 7.3.1) to reconstruct the corresponding tree.

Proceed lexicographically, i.e., start with the edge $\{A, B\}$ and add the taxa C, D and E in that order. Try pair $\{A, B\}$ to add a new taxon first. Sometimes you might have to choose another pair.

	A	B	C	D	E
A :	0	11	7	9	10
B :		0	8	8	9
C :			0	6	7
D :				0	3
E :					0

Write down all steps.

Task 2 Fitch-Margoliash.

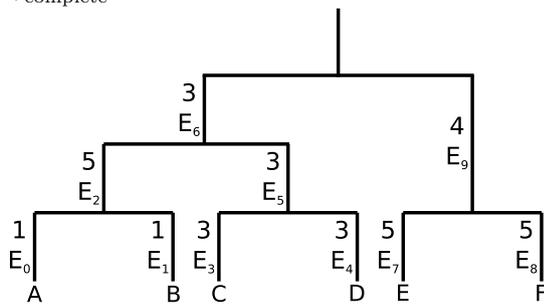
(3 points)

On exercise sheet 7, different agglomerative clustering methods have been used for a tree reconstruction. For the given matrix d^M , the methods *complete linkage* and *UPGMA* can result in the trees $\mathcal{T}_{\text{complete}}$ and $\mathcal{T}_{\text{UPGMA}}$, respectively.

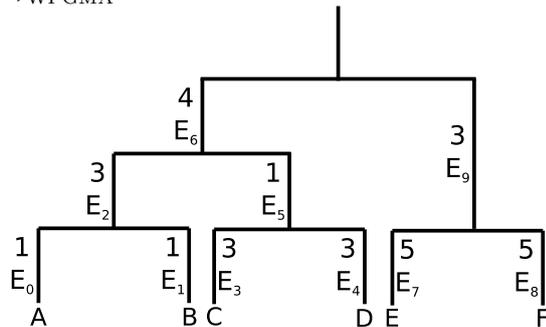
$d^M :=$

	A	B	C	D	E	F
A :	0	2	8	12	18	18
B :		0	4	8	18	18
C :			0	6	18	18
D :				0	8	12
E :					0	10
F :						0

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



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



Calculate the *least squares* error $E := \|\vec{d}^T - \vec{d}^M\|^2$ (according to Fitch and Margoliash) for both trees. You do not need to write down matrix M^T and vector \vec{w} explicitly.

Which tree is the “better” one?