Exercises – Phylogenetics

Universität Bielefeld, WS 2016/2017,
M. Sc. Tizian Schulz
https://gi.cebitec.uni-bielefeld.de/Teaching/2016winter/Phylogenetik

Exercise Sheet 9 — 10.01.2017

Due: 17.01.2017

Task 1 Neighbor Joining. **(5 points)**
Use *Neighbor Joining* to reconstruct a phylogenetic tree from the following matrix. Write down all steps.

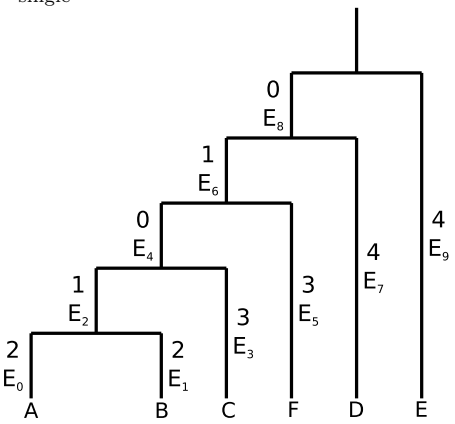
	A	B	C	D	E
A :	0	8	4	8	4
B :		0	8	6	6
C :			0	8	4
D :				0	6
E :					0

Task 2 Fitch-Margoliash. **(3 points)**
On exercise sheet 8, different agglomerative clustering methods have been used for a tree reconstruction. For the given matrix d^M, the methods *single linkage* and *WPGMA* can result in the trees $\mathcal{T}_{\text{single}}$ and $\mathcal{T}_{\text{WPGMA}}$, respectively.

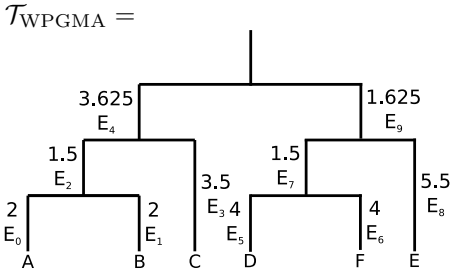
$d^M :=$

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

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



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



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

Which tree is the "better" one?

Please turn over! Bitte wenden!

Task 3 Minimum Evolution. **(5 points)**

Write a linear program (LP) that calculates the *minimum evolution tree* (ME tree) for the distance matrix from Task 2 and the tree topology resulting for the single linkage method.

Use an online solver to get a solution for it, for instance:

- `http://www.zweigmedia.com/RealWorld/simplex.html` or
- `http://www.phpsimplex.com/simplex/simplex.htm?l=en`.

Send your LP to your T.A.[1]

What are the minimum total edge length and the individual edge lengths?

In comparison, the resulting topology for WPGMA has a total edge length of 37. Compare these results and the results from Task 2.

[1] jeschulz[at]techfak... | ssimon[at]techfak...