Exercises – Phylogenetics

Universität Bielefeld, SS 2018
Dr. Roland Wittler
https://gi.cebitec.uni-bielefeld.de/Teaching/2018summer/Phylogenetik

Exercise Sheet 5 — 17.05.2018

Due: 24.05.2018

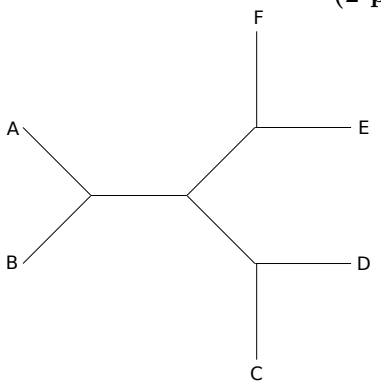
Task 1 Greedy Sequencial Addition. **(2 points)**
Consider the *Greedy Sequencial Addition* (GSA) heuristic for the *maximum parsimony problem*. For given taxa $t_1, t_2, t_3, \ldots, t_n$, the algorithm works as described on page 35 in the lecture notes. See also Figure 6.3.

(a) How many trees are considered before choosing one, if you add the i-th taxon?

(b) How many trees are considered in all steps $i = 4, \ldots, n$? Show that this number increases quadratically in n. (Hint: Use the formula of the "young Gauß".)

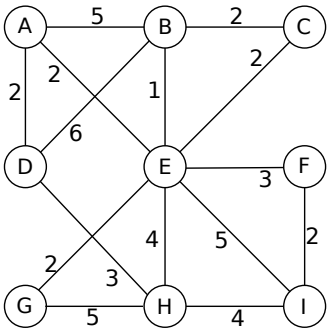
Task 2 Nearest Neighbor Interchange. **(2 points)**

Consider the given tree. List all trees that are *neighbors* according to the *Nearest Neighbor Interchange* approach.



Task 3 Spanning Trees. **(2 points)**

Find all minimum spanning trees in the given graph. Write down the order of the edges you choose. If there are several edges that can be chosen in one step, list them all but choose one (arbitrary) edge to continue. Use *either* Kruskal's algorithm, *or* Prim's algorithm starting with node A.



Task 4 Ackermann function. **(2 points)**
Implement the Ackermann function:

$$\begin{aligned} a(0, m) &= m + 1 \\ a(n + 1, 0) &= a(n, 1) \\ a(n + 1, m + 1) &= a\,(n, a(n + 1, m)) \end{aligned}$$

Try to compute as many values $a(i, i)$ for $i = 1, 2, 3, \ldots$ as possible.

Task 5 DNA Grid Graph. **(2 points)**
What is the number of edges $|E|$ for a DNA grid graph $G = (V, E)$ for sequence length m? Derive a formula for $|E|$ that only depends on m and explain it.