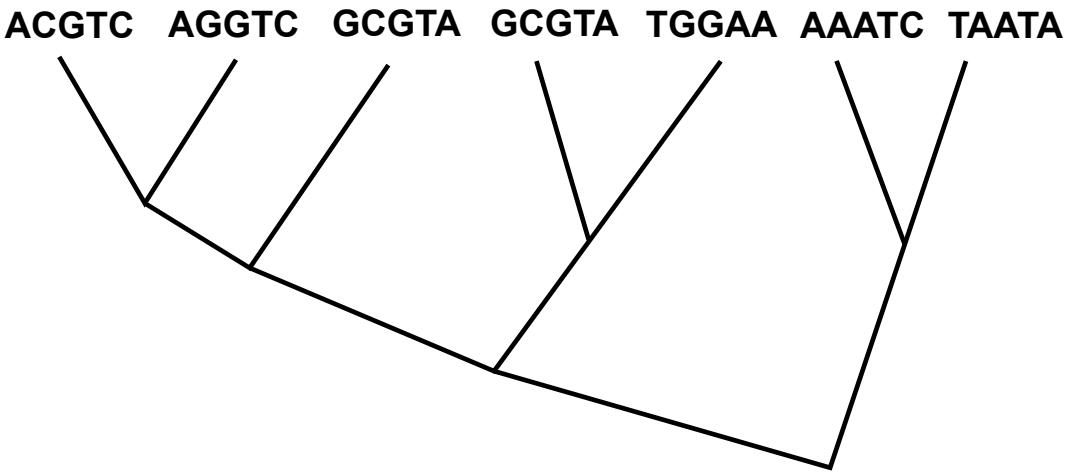
Exercise 1: Apply Fitch's algorithm to find the minimum number of mutations that explain the sequence data with this tree:

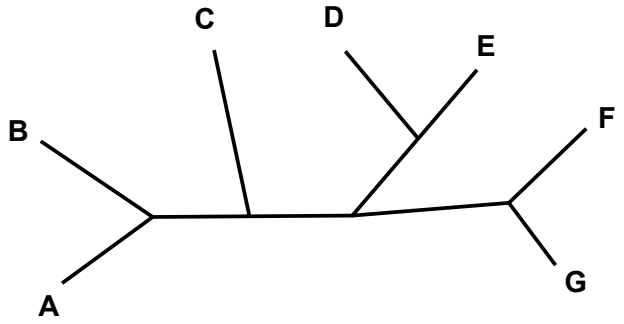


Exercise 2: In the lecture we saw how the Fitch algorithm can be used to find the minimum number of mutations needed to generate given sequences along a given tree. How can you extend this algorithm to also get the corresponding optimal assignment of nucleotides to the inner nodes of the tree?

Exercise 3: Given a tree with tips are labled with homologous sequences and a cost function $S(.,.)$ that assigns a penalty $S(a,b)$ to each possible type of mutations $a \to b$, how can you find the minimal costs needed to generate the sequences by evolution along the tree?

Exercise 4: For taxa labeled with T1, T2, T3, T4, T5, T6, T7, T8 draw the tree that is represented by [2][4][1][0][0] and write this tree in Newick format.

Exercise 5:



(a) Give a tree topology that can be reached from this tree topology with an NNI step.

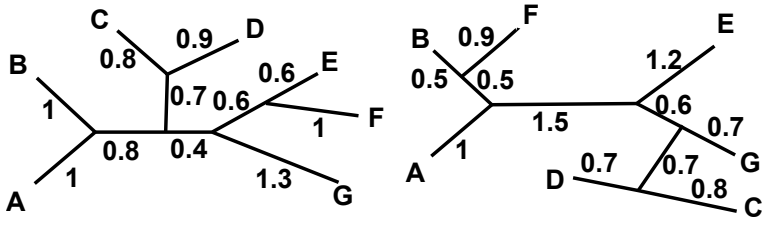
(b) How many different (unrooted) tree topologies are only one NNI step away from this tree topology?

(c) Give a tree topology that can be reached from this tree with an SPR step but not with an NNI step or explain why such a tree cannot exist.

(d) Give a tree topology that can be reached from this tree with an TBR step but not with an SPR step or explain why such a tree cannot exist.

(e) Give a tree topology that can be reached from this tree with an NNI step but not with an TBR step or explain why such a tree cannot exist.

Exercise 6:



Calculate without the help of a phylogenetic sofware the

(a) symmetric distance,

(b) Robinson–Foulds distance,

(c) branch score distance,

(d) quartet distance,

(e) and the NNI distance

beween these two trees.

Exercise 7: Assemble a dataset of published DNA sequences of several primates (including humans, chimpanzees and gorillas) and other mammals to reconstruct the species' phylogenetic tree using NeighborJoining, UPGMA and Maximum Parsimony. Use at least 5, better 10 different loci. Explore how the results depend on the selection of loci. (Several alignment programs are installed on turtle, e.g. mafft, muscle, probcons, and prank).