

Algorithms in Comparative Genomics

Bielefeld University, WS 2019/20

Dr. Roland Wittler

<https://gi.cebitec.uni-bielefeld.de/teaching/2019winter/cg>

Exercise sheet 7, 28.11.2019

Exercise 1 (SPP under SCJ continued)

Recall the setting from the previous exercise sheet (Exercise 2 on Sheet 6): Given the phylogenetic tree $((A, B), (C, D))$; with the following gene orders at its leaves, perform the bottom-up phase of the Fitch Algorithm for each adjacency (presence/absence).

$A = 1\ 2\ 3\ 4$

$B = 1\ 3\ 2\ 4$

$C = 1\ 2\ 3\ 4$

$D = 2\ 1\ 3\ 4$

1. Let X be the set of all adjacencies with candidate set $\{1\}$ or $\{0,1\}$ at the root node. Formulate a corresponding instance of the MAX-ROW-component-Ci1P, (manually) find a maximum weight consistent subset $X' \subseteq X$, and provide the corresponding set of adjacencies.
2. Compare your result to the result from the previous version of this exercise (Exercise 2.2 on Sheet 6).
3. Perform the top-down phase of Fitch's algorithm starting with the root labeling obtained in exercise 2.1.

Exercise 2 (Small Parsimony under SCJ — from sparse to dense)

In its original formulation (P. Feijao and J. Meidanis. SCJ: a breakpoint-like distance that simplifies several rearrangement problems. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 8(5):1318–1329, 2011.), the authors choose the absence of an adjacency at the root node in case of ambiguity, and prove in Lemma 6.2 that then no conflicting adjacencies would be assigned to any node in the tree during the top-down phase.

Slightly modify this proof such that it also holds when choosing a maximum consistent subset of adjacencies at the root node.