

Algorithms in Bioinformatics I, WS2006/7

Assignment sheet # 12

Christian Rausch

January 15, 2007

1 3D structure alignment (1.5 points)

Compare two protein structures of closely related proteins from the last assignment (or alternatively hemoglobin and leghemoglobin as shown in the lecture) using SSAP (Sequential Structure Alignment Program) at <http://www.cathdb.info/cgi-bin/cath/SsapServer.pl>. Visualize the result with Rasmol and try different representations (Menu → Display → e.g. Ribbons).

2 Classification of protein structures (1.5 points)

Please report and explain SCOP classification obtained for the following protein: *p53 tumor suppressor, DNA-binding domain from Human (Homo sapiens)*. Can you find this protein (or its characteristic domain) in CATH and DALI as well?

3 Phylogeny

The second part of this week's task is to write code for reading, writing and drawing phylogenetic trees. To solve the following problems, please download:

<http://www-ab.informatik.uni-tuebingen.de/teaching/ws06/albi1/assign/java/java12.zip>,

and modify and run the file `PhyloTree.java`, located in the package `albi.phylo`.

In the three following problems, the goal is to implement a class `PhyloTree` that can

1. read a phylogenetic tree in Newick format,
2. write a phylogenetic tree in Newick format, and
3. compute and draw an embedding of a phylogenetic tree.

The `PhyloTree` class is based on `jloda.graph.Graph`, which is a general graph implementation. Additionally, it makes use of `jloda.graphview.GraphEditor` that is used to draw the computed embedding. Apply your program to all given files ending on `.top`.

3.1 Reading phylogenetic trees in Newick format (2 points)

Implement the `read` method.

3.2 Writing phylogenetic trees in Newick format (2 points)

Implement the `write` method.

3.3 Drawing a phylogenetic tree (3 points)

Implement the `computeCoordinates` method.

Assignments 1-3.3 are due by 10am, Monday, 22 Jan 2007

The following assignment is counted for next week's assignment sheet:

3.4 Trees with edge lengths (3 points)

Modify your program so that trees containing edge lengths can also be read, written and drawn to scale. Apply the modified program to all given files ending on `.new`.

Assignment 3.4 is due by 10am, Monday, 29 Jan 2007