Project: Maximum Agreement Subtree

Due to different data sources, different evolution models, and different computation methods, we may obtain different phylogenetic trees for the same set of taxa. In this assignment, you need to implement a method to compute the maximum agreement subtree of two phylogenetic trees.

Detail of the assignment

You are required to do the following:

1. **Describe an algorithm to compute the size of the maximum agreement subtree (MAST) of two rooted binary trees. What is the time complexity? Please also implement the algorithm which outputs the size to STDOUT.**
   ```
   java mast_score tree1.dnd tree2.dnd
   ```
2. **Describe an algorithm to report the MAST of two rooted binary trees. What is the time complexity? Please also implement the algorithm which outputs the maximum agreement subtree to STDOUT.**
   ```
   java mast tree1.dnd tree2.dnd
   ```
3. **Describe an algorithm to compute the size of the MAST of two unrooted degree-3 trees. What is the time complexity? Please also implement the algorithm which outputs the size to STDOUT.**
   ```
   java umast_score tree1.dnd tree2.dnd
   ```
4. **Describe an algorithm to report the MAST of two unrooted degree-3 trees. What is the time complexity? Please also implement the algorithm which outputs the maximum agreement subtree to STDOUT.**
   ```
   java umast tree1.dnd tree2.dnd
   ```

In the questions, tree1.dnd and tree2.dnd are the two input trees in .dnd format (i.e. Newick tree format. See http://evolution.gs.washington.edu/phylip/newicktree.html).

For questions 1 and 3, please output an integer representing the maximum size of the MAST of the two trees. For questions 2 and 4, please output your tree in .dnd format.

Note that trees in .dnd format can be visualized using the software TreeView (see http://taxonomy.zoology.gla.ac.uk/rod/treeview.html).

Example

For example, consider the following two trees:
tree1.dnd
((A, B), ((C, D), (E, F)));

tree2.dnd
((A, (B, (C, D))), (E, F));

For question 1, the answer is 4.
For question 2, a possible answer is ((A, B), (E, F)).
For question 3, the answer is 5.
For question 4, a possible answer is ((A, (C, D)), (E, F)).

Submission

Please submit a pdf file describing your algorithms for questions 1, 2, 3, and 4 and the time analysis. Also, please submit your four java programs for the four questions.