Practical 6: LCA Mapping

Name Surname Name Surname

24/10/2023, submission deadline 30/10/2023

Solve the following exercise in groups of two students. Write the Python scripts, perform the computations, and make the graphics that are asked for (if any) in the practical below. Write your solution in a LaTeX document and generate a PDF file with your solution. Take care to number your answers exactly as in this exercise. Upload your solution in PDF format to the web page of the course at `raco.fib.upc.edu` no later than the submission deadline.

You can make use of the Python package **networkx** (and other packages) to compute your answers, as you please. The datasets (if any) can be downloaded from the web page of the course at `raco.fib.upc.edu`.

1. (40 points) Given a file **nodes.dmp** for the NCBI taxonomy and a file **mapping.txt** of mappings of sequence reads to NCBI taxonomic identifiers, write a Python script to find the LCA mapping for each sequence read. Give the code of your Python script as your answer to this question, using the LaTeX package **listings**.
2. (5 points) What is the highest taxonomic rank (that is, toward *kingdom*) for these LCA mappings?
3. (5 points) What is the lowest taxonomic rank (that is, toward *species*) for these LCA mappings?
4. (40 points) Given a file **nodes.dmp** for the NCBI taxonomy and a file **mapping.txt** of mappings of sequence reads to NCBI taxonomic identifiers, write a Python script to find the optimal (in terms of the F-measure) taxonomic assignment for each sequence read. Give the code of your Python script as your answer to this question, using the LaTeX package **listings**.
5. (5 points) What is the highest taxonomic rank (that is, toward *kingdom*) for these taxonomic assignments?
6. (5 points) What is the lowest taxonomic rank (that is, toward *species*) for these taxonomic assignments?

```
\documentclass[12pt,a4paper]{article}
\usepackage{listings}
\usepackage{mathptmx}
\usepackage{savetrees}
\title{Practical 6: LCA Mapping}
\author{Name Surname \and Name Surname}
\date{24/10/2023, submission deadline 30/10/2023}
\begin{document}
\maketitle
\begin{enumerate}
\item ...
\item ...
\item ...
\item ...
\item ...
\item ...
\end{enumerate}
\end{document}
```