

Computational Linguistics

CSCE 915

Homework 1

Spring 2024

1. (40 points) Implement the UPGMA algorithm. Apply it to the Hw1_words data file to build a hypothetical language family tree.
2. (40 points) Implement the CMM algorithm. Apply it to the Hw1_words data file to build a hypothetical language family tree.
3. (10 points) Which language family tree gives a more detailed explanation of the development of the language family tree. Why that explanation may be important for a historical linguist?
4. (10 points) Which of the two language family trees is more accurate, that is, more likely to reflect the actual history of the development of the languages within the given language family? Why?