

Problem Set 5 Report

- Testing the algorithm on the hard-coded graphs from the problem set description showed no errors, indicating correctness of our code.
- Testing the algorithm on the “example” files produced no errors.
- Testing the algorithm on the “simple” files produced 5 errors, as indicated in the problem set description.
- When running the algorithm on the Brown corpus, we observed 1557 errors (corresponding to an accuracy rate of 95.7%). As we increased the unseen-word penalty, we observed more errors. Decreasing it further than -100 produced no differences in the number of errors.
 - In the sentence: *“the bill , which daniel said he drafted personally , would force banks , insurance firms , pipeline companies and other corporations to report such property to the state treasurer .”*, our algorithm got an error in detecting the 7th word (“drafted”). Our model analyzed it as “V” whereas the correct tag is “VD”. This error likely occurred as these are simply two different occurrences of verb forms, which may be used interchangeably.
 - In the sentence: *“attorneys for the mayor said that an amicable property settlement has been agreed upon .”*, our model produced the expected tags, showing that our training model was accurate.