How to use?

Run python master.py to use the grammar checker.

This file will structure theHMMusing two corpora (function in viterbi\_tagger.py), build the

averaged perceptron model (in perceptron\_tagger.py), generate the context-free-grammar

based on the treebank (in cfg.py), and then ask for user input. After the user enters a sentence,

the system gets the tag sequence from either the Viterbi algorithm or AP tagger (depending

whether all input words exist in the corpora). The tag sequence is then passed into the contextfree-

grammar where the nltk ChartParser algorithm tests it and returns True/False for the validity

of the grammar.

If desired, run python test.py to test the viterbi tagger. This also trains and tests the averaged

perceptron tagger. It has a long runtime, but produces the results in Table 1 and Table 2.

If desired, run dataAP/buildTrainTestFiles.py to generate the .txt to retrain the Averaged

Perceptron (AP) tagger. Users can change the training set in dataAP/train.txt.