Name: _____
CS159: Natural Language Processing
PReview #10
Due in class Apr 20, 2015

Review

The file `http://hmc-cs159-spring2015.github.io/site/handouts/viterbi.py` is partially impelemented for you. In particular, it uses the nltk brown corpus to build the transition and observation probabilities for an HMM. The decode() function is also partially implemented – it initializes a DP table and follows back-pointers through the completed DP table to get the best POS tag sequence.

Your job is to fill in the actual DP part of the code.

Accoring to your algorithm, what is the most likely tag sequence for each of the following?
1. i see the sheep
2. the sheep is in the field
3. who went to the store with you
4. the plane will land on the land
5. the horse raced past the barn fell

1.

2.

3.

4.

5.

■

Preview

Five reviews of the movie "Frozen" are linked from our course website: `http://hmc-cs159-spring2015.github.io/site/handouts/reviews.pdf`.

Based on the reviews, can you predict what rating (on a 1-5 scale) each reviewer gave the movie? Briefly justify your choice for each.

Review 1

Review 2

Review 3

Review 4

Review 5

■

If you were going to automate the process of predicting scores from the text, what features might you use? For example, are there specific words or phrases that are more likely to be associated with positive or negative reviews? Are there any problems you can foresee with using those features?

■