

Name: \_\_\_\_\_

1. What is the key advantage of a log-linear POS tagger over an HMM-based tagger?

**Answer:** Log-linear models allow for a much richer feature representation

2. Assume we have a trigram log-linear POS tagger:

We have  $w_1 \dots w_3 = \text{the dog barks}$

We would like:

- $p(D \ N \ V | \text{the dog barks}) = 0.5$
- $p(D \ N \ N | \text{the dog barks}) = 0.5$

What are the values for the following parameters?

- $p(D | \text{the dog barks, START})$
- $p(N | \text{the dog barks, START, D})$
- $p(V | \text{the dog barks, D, N})$
- $p(N | \text{the dog barks, D, N})$

**Answer:**

- $p(D | \text{the dog barks, START}) = 1$
- $p(N | \text{the dog barks, START, D}) = 1$
- $p(V | \text{the dog barks, D, N}) = 0.5$
- $p(N | \text{the dog barks, D, N}) = 0.5$

3. Suppose you are training a Naïve Bayes classifier for the Word Sense Disambiguation task for the word “bank” with two classes (“financial institution” and “river bank”).
  - (a) Name 5 features that you will want to specify.
  - (b) What are the assumptions made by the Naïve Bayes model?

**Answer:**

**Sample features:** word preceding “bank”, word following “bank”, two preceding words, two following words, bag-of-words features, POS of the word before, POS of the word after, etc.

**Assumptions:** Every feature is conditionally independent of the rest of the features given the class label.