Tuesday Sep 24, 2024

Name:

Discuss and complete the following questions with the person nearest you. You **may** be asked to share your thoughts with the class.

1. Suppose we didn't use the end-symbol like . Train an unsmoothed bigram grammar on the following training corpus without using the end-symbol :

   ```
   <s> a   b
   <s> b   b
   <s> b   a
   <s> a   a
   ```

   a. Calculate the total probability of the four possible 2 word sentences over the alphabet {a, b}

   b. Calculate the total probability of the 8 possible 3 word sentences over the alphabet {a, b}

2. With your midterm replication teammate identify what parts of the NLPScholar toolkit you need to use to replicate the experiment. Create config files for all the parts and submit them on gradescope.