

CS 440 Midterm 2 Instructions

This is a closed book exam. No notes are allowed. Electronic devices must be put out of sight/reach, ideally in a bag.

The exam will end at 9:50, so you have 45-50 minutes of working time.

If you finish early, turn in your exam at the front and then leave the room.

We will be grading scanned copies of the exams using a tool that automatically separates out answers for each question.

Your answer to each question must be written inside the box provided.

The backs of sheets will not be scanned.

Overlong answers may be continued in the box on the last page. You may use the backs of sheets for scratch work.

You may remove this instruction sheet to use as scratch paper. Other than that do not take the exam apart. (If the sheets come apart accidentally, put your name on any loose sheets and warn the proctors when you turn it in.)

Points may be deducted for solutions which are correct but excessively complicated, hard to understand, poorly explained, or excessively hard to read.

Assume that answers require brief justification/work, unless there is clear indication to the contrary (e.g. it's a multiple choice question, it's asking you to repeat back a standard formula).

Please bring any apparent bugs to the attention of the proctors.

Cheating (e.g. looking at another student's exam) is obviously not allowed. Also, you may not do things that look like cheating, such as talking to your neighbor even if the topic is innocent. If your behavior is suspicious, we may take actions such as issuing warnings, reseating you, or having you take a makeup for the exam.

If you have to do something that could be misinterpreted, e.g. pick up a dropped eraser or silence a ringing cell phone, please do it conspicuously so everyone can easily understand what you're doing.

You may not leave the room and then come back and continue working on your exam. If an emergency (e.g. upset stomach) forces you to leave before you have had time to finish, turn in your partial work and we'll schedule you to take a makeup.

Name: _____ NetID: _____

1. (5 points) What was the Pattern Playback?

2. (5 points) We'd like to create some fake news by splicing a can of coke into historical film footage. Why would it be useful to reconstruct 3D models of the scenes in the film?

Name: _____ NetID: _____

3. (5 points) An HMM part-of-speech tagger uses a matrix of tag transition probabilities. Suppose $P(S \rightarrow T)$ is the probability of a transition from tag S to tag T. In order to be well-formed probabilities, what mathematical constraints must these numbers satisfy?

4. (5 points) How does the choice of k affect the performance of a k nearest neighbor classifier?

Name: _____ NetID: _____

5. (5 points) Mitch is classifying text as German vs. English using a Naive Bayes bag of words classifier. Explain how to model this problem using a Bayes Net.

6. (5 points) What explains the very strong performance of the baseline algorithm for part of speech tagging?

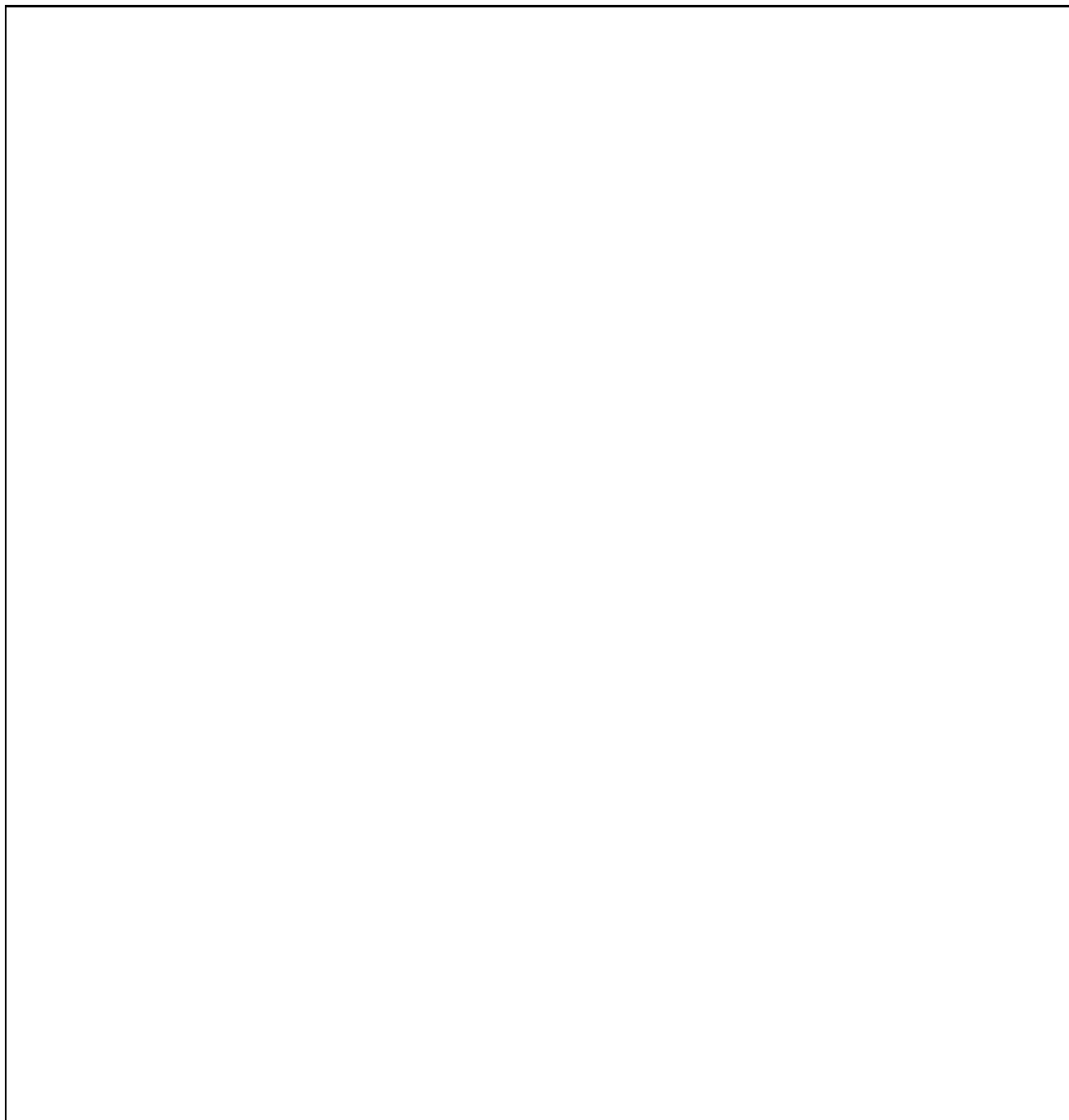
Name: _____ NetID: _____

7. (5 points) What does it mean for a parser to be “lexicalized”? Why would this be a good idea?

8. (5 points) A perceptron classifier can decide among several classes by using one perceptron per class. It outputs the class whose perceptron produces the largest value. Give the rule for updating the weights in this situation, when we see a single training pair (x,y) . (Be sure to define any variables used in your equations.)

Name: _____ NetID: _____

This box is additional answer space. Clearly indicate the problem number.

A large, empty rectangular box with a thin black border, intended for students to provide their answers to the exam questions. The box is currently blank.