

CS447: Natural Language Processing

<http://courses.engr.illinois.edu/cs447>

Lecture 13

Midterm Review

Julia Hockenmaier

juliahmr@illinois.edu

3324 Siebel Center

Midterm Exam

When: Friday, October 12, 2018 in class

Where: DCL 1310

What: Closed book exam:

- You are not allowed to use any cheat sheets, computers, calculators, phones etc.
(you shouldn't have to anyway)
- Only the material covered in lectures
- Bring a pen (black/blue) or pencil
- **Short questions — we expect short answers!**
- **Tip:** If you can't answer a question, move on to the next one.
You may not be able to complete the whole exam in the time given — there will be a lot of questions, so first do the ones you know how to answer!

Today's lecture

Quick run through the material we've covered, with some example questions.

(Not an exhaustive list of possible questions!)

Question types

Define X:

Provide a mathematical/formal definition of X

Explain X; Explain what X is/does:

Use plain English to define X and say what X is/does

Compute X:

Return X; Show the steps required to calculate it

Draw X:

Draw a figure of X

Show/Prove that X is true/is the case/...:

This may require a (typically very simple) proof.

Discuss/Argue whether ...

Use your knowledge (of X, Y, Z) to argue your point

Fundamentals

Basics about language

Explain Zipf's law and why it makes NLP difficult.

Explain why we often use statistical models in NLP.

Give two examples of ambiguity and explain how they make natural language understanding difficult.

Basics about language

Explain Zipf's law and why it makes NLP difficult.

Zipf's law says that a few words are very frequent, and most words are very rare. This makes NLP difficult because we will always come across rare/unseen words.

Explain why we often use statistical models in NLP.

To handle ambiguity (and make NLP systems more robust/to deal with the coverage problem).

Give two examples of ambiguity and explain why we have to resolve them.

POS ambiguity: back = noun or verb? Need to resolve this to understand the structure of sentences.

Word sense ambiguity: bank = river bank or institution. Need to resolve this to understand the meaning of sentences.

Morphology and finite-state transducers

Morphology

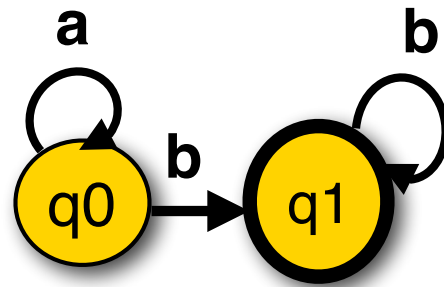
Explain what we mean by derivational morphology, and given an example in a language of your choice.

Draw a finite-state automaton for the language $\{a^n b^m\}$

Explain how we can use finite-state transducers for the morphological analysis of irregular verbs in English.

Morphology

Draw a finite-state automaton for the language $\{a^n b^m\}$



Language modeling

Language modeling

Explain: What is a language model?

Explain and define: What is an n-gram language model?

Discuss the advantages and disadvantages of bigram language models over unigram models

Explain and define how to estimate the parameters of a bigram model

Explain and define how evaluate the quality of a language model

Smoothing

Explain what smoothing is, and why it is necessary.

Define add-one smoothing and explain when it can be used.

Discuss the advantages/disadvantages of add-one smoothing.

Define how smoothing can be done via linear interpolation and explain when this technique can be used.

Hidden Markov Models and POS tagging

POS tagging

Discuss how you would define a POS tag set.

Explain the differences between open and closed word classes.

Explain how to do a quantitative evaluation of a POS tagger.

HMMs

Give the mathematical definition of a bigram HMM.

Explain how to estimate the parameters of a bigram HMM from labeled data.

Explain how the Viterbi algorithm is used for POS tagging with an HMM.

Find the most likely tag sequence for the following sentence (given some HMM).

Sequence labeling

Sequence labeling

Define the BIO encoding for NP chunking.

Define Maximum Entropy Markov Models.

Explain why MEMMs may be more suitable for named entity recognition than HMMs.

Draw the graphical model of MEMMs

Syntax and Context-Free Grammars

Syntax basics

Explain how to determine whether a string is a constituent.

Explain the distinction between arguments and adjuncts.

CFG basics:

Convert the following PCFG rules to Chomsky Normal Form (and preserve the rule probabilities)

(Nonterminals: XP, YP, ZP, Terminals: X, Y, Z)

$$XP \rightarrow X \ YP \ YP \quad 0.75$$
$$XP \rightarrow XP \ ZP \quad 0.25$$

Explain how you can convert a CFG to dependencies.

CFG basics:

Convert the following PCFG rules to Chomsky Normal Form (and preserve the rule probabilities)

(Nonterminals: XP, YP, ZP, Terminals: X, Y, Z)

$XP \rightarrow X \ YP \ YP \quad 0.75$

$XP \rightarrow XP \ ZP \quad 0.25$

Solution 1:

$XP \rightarrow A \ YP \quad 0.75$

$A \rightarrow X1 \ YP \quad 1.00$

$X1 \rightarrow X \quad 1.00$

$XP \rightarrow XP \ ZP \quad 0.25$

Solution 2:

$XP \rightarrow X1 \ B \quad 0.75$

$B \rightarrow YP \ YP \quad 1.00$

$X1 \rightarrow X \quad 1.00$

$XP \rightarrow XP \ ZP \quad 0.25$

CFG basics:

Explain how you can convert a CFG to dependencies.

Answer:

For every rule $XP \rightarrow L_1 \dots L_n X R_1 \dots R_n$, identify the head child X among the RHS symbols. All other symbols on the RHS are dependents of the head child.

Parsing with CFGs

CKY Questions

Given the following grammar and the following input sentence, fill in the CKY parse chart:

... (input sentence)

... (CFG)

How many parse trees does the input sentence have?

What is the most likely parse tree for this sentence?

More on PCFGs

Define how to compute the probability of a parse tree under a PCFG.

Define how to compute the probability of a string under a PCFG.

Statistical Parsing/Penn Treebank

Define the Parseval metrics for evaluating statistical (PCFG) parsers.

Explain why basic PCFGs do not perform well, and describe one way to improve their performance.
(NB: we've covered several such methods in class).

Dependency Grammars

Dependency Grammar Basics

Explain the difference between projective and nonprojective dependencies.

Explain how dependency grammar represents syntactic structures.

Draw the correct dependency tree for the following sentence:

.... (example sentence)

Transition-based parsing

Define what we mean by a parser configuration in the context of transition-based parsing.

Describe the actions that a transition-based parser can perform.

Show the sequence of actions that a transition-based parser has to perform to return the correct dependency tree for the following sentence:
... (short input sentence)

Good luck!