Lecture 29:  
Final Exam Review

Julia Hockenmaier  
{juliahmr@illinois.edu}  
3324 Siebel Center

Final exam: Sat 12/16, 7:00-8:30pm

Saturday, Dec, 16, 7:00pm-8:30pm  
DCL 1310 and Siebel 0216 (basement)  
* If your UIN ends in an even number (-0,-2,-4,-6,-8), go to DCL 1310  
* If your UIN ends in an odd number (-1,-3,-5,-7,-9), go to Siebel 0216

Covers everything after the midterm review  
(Lecture 13 — today)

Same question format as midterm.  
Closed book. No cheat sheets etc.

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

Final exam topics

Word Senses and Word Sense Disambiguation (Lecture 13)  
Grammars, Parsing (Lectures 14—21):  
- Context-Free Grammars and Probabilistic CFGs  
- CKY parsing algorithm (with and without probabilities)  
- Treebanks and statistical parsing  
- Dependency parsing  
- Feature structure grammars  
- Expressive grammars (TAG, CCG)

Machine Translation (Lectures 22—24):  
- Basics: Linguistic divergences, Vauquois triangle  
- Word-based MT: IBM model 1  
- Phrase-based MT and decoding

Compositional Semantics (Lecture 25)  
Verb Semantics and Discourse (Lecture 26)  
Natural Language Generation (Lecture 27)  
Deep Learning for NLP (Lecture 28)
Word Senses and Word Sense Disambiguation

**WordNet**

What are synsets in WordNet?

Why is the path length in WordNet not a good metric for word similarity?

**WSD**

Describe how you can treat word sense disambiguation as a classification task.

Why does the pseudo-word task provide a good indication of an upper bound on performance for a WSD system?

**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 —> X YP YP  0.75
XP —> XP ZP    0.25

Solution 1:
XP  —>  A  YP  0.75
A   —>  X1 YP  1.00
X1  —>  X     1.00
XP  —>  XP ZP  0.25

Solution 2:
XP  —>  X1 B  0.75
B   —>  YP YP 1.00
X1  —>  X     1.00
XP  —>  XP ZP 0.25

Explain how you can convert a CFG to dependencies.

Answer:
For every rule XP —> L1…Ln X R1…Rn, identify the head child X among the RHS symbols. All other symbols on the RHS are dependents of the head child.
Parsing with CFGs

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.

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?

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 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)

---

**Feature structure grammars, TAGs, CCGs**
Feature structures/unification grammars

Do the following two feature structures unify? If so, give the result. If not, explain why not.
... (example feature structures)

Explain what kind of feature-structure (attribute-value) grammars are equivalent to CFGs.

Explain why we might want to use feature structure grammars.

Tree-Adjoining Grammars

Define the two basic operations of Tree-Adjoining Grammar, substitution and adjunction (you can just draw a picture).

Under what circumstances do these operations make the grammar more expressive than CFGs?

Combinatory Categorial Grammars

Explain why we often refer to CCG (or TAG) as lexicalized grammar formalisms.

Define the basic rules of CCG.

Complete the following CCG derivation:
... (partial CCG derivation, possibly with semantics)

Machine Translation
Machine Translation

Explain what is meant by *lexical* and *syntactic divergences* between languages (give examples).

Explain the purpose of the *language model* for statistical machine translation. If you want to translation from language A to language B, what data would you train this model on?

Explain the purpose of the *translation model* for statistical machine translation. If you want to translation from language A to language B, what data would you train this model on?

Statistical MT

Describe how the IBM models represent word alignment between a sentence in a foreign source language \( F = f_1, \ldots, f_m \) and its target English translation \( E = e_1, \ldots, e_n \).

How do the IBM models define the translation probability for a sentence in a foreign source language \( F = f_1, \ldots, f_m \) and its target English translation \( E = e_1, \ldots, e_n \)?

Describe the algorithm for learning (estimating the parameters) of IBM model 1.

Given a sentence in a foreign source language, explain briefly how to generate a random translation for this sentence. Then explain the purpose of stack-based decoding.

Semantics, Discourse and Generation
**Compositional semantics**

Translate the following sentence to first-order predicate logic:

… example sentence

Explain what natural language phenomena cannot be expressed in first-order predicate logic.

**Verb semantics; Discourse**

Explain what we mean by thematic roles.

Explain what we mean by diathesis alternations. Given an example.

Explain what we mean by rhetorical (discourse) relations. Why are they important for natural language understanding?

**Discourse**

Explain what we mean by coreference resolution, and describe how to build a system that performs coreference resolution.

Explain what a discourse model is, and why we may need it for natural language understanding.

Explain what an anaphoric pronoun is (give an example).

**Generation and Dialog**

Describe the basic architecture of an NLG system.

Describe what we mean by a finite-state dialog manager.
Neural approaches to NLP

Describe the motivation for using neural approaches to NLP.

Explain the advantages of a neural language model over a traditional language model.