CS447: Natural Language Processing

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

Lecture 28: Review

Julia Hockenmaier

juliahmr@illinois.edu

3324 Siebel Center

Projects and Literature Reviews

Final report due Sunday**, Dec 15, 11:59 PM**

(PDF written in LaTeX; no length restrictions, but 8–10 pages single spaced, single column, 11pt recommended; submission through Compass)

No further extensions will be given

Projects:

Read and describe a few (2–3) NLP papers on a particular task, implement a system for this task, and describe it in a written report.

Literature surveys:

Read and describe several (5-7) NLP papers on a particular task or topic, and produce a written report that compares and critiques these approaches.

Rubrics for these reports:

https://courses.engr.illinois.edu/CS447/LiteratureReviewRubric.pdf

https://courses.engr.illinois.edu/CS447/FinalProjectRubric.pdf

Final exam

Wednesday, Dec 11 in class
Only materials after midterm
Same format as midterm

Today's lecture: Review session
These questions don't represent an exhaustive sample of the kinds of questions that will be on the exam, but are similar to what will be on the exam.

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.... f_m and its target English translation E = e_1….e_n.

How do the IBM models define the translation probability for a sentence in a foreign source language F = f_1.... f_m and its target English translation E = e_1….e_n?

Representing word alignments

		1	2	3	4	5	6	7	8
		Marie	a	traversé	le	lac	à	la	nage
0	NULL						■	■	
1	Mary	■							
2	swam								■
3	across		■	■					
4	the				■				
5	lake					■			

Position	1	2	3	4	5	6	7	8
Foreign	Marie	a	traversé	le	lac	à	la	nage
Alignment	1	3	3	4	5	0	0	2

Every source word **f[i]** is aligned to **one** target word **e[j]** (incl. NULL). We represent alignments as a vector **a** (of the same length as the source) with **a[i] = j**

Statistical MT

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 with a phrase-based model.
Then explain the purpose of stack-based decoding.

More on RNNs

Define what we mean by a recurrent neural net.

Explain what we mean by a seq2seq architecture, and how that can be used for machine translation.

Explain what we mean by an attention mechanism in a seq2seq model.

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

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 —> 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

CFG basics:

Explain how you can convert a CFG to dependencies.

Answer:
For every rule XP —> $L_1 \ldots Ln$ X $R_1 \ldots R_n$, 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)

Expressive 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)

Semantics

CCG with semantics

$$\dfrac{\dfrac{\dfrac{\textit{John}}{\mathbf{NP} : \textit{John}} \quad \dfrac{\textit{sees}}{(\mathbf{S}\backslash\mathbf{NP})/\mathbf{NP} : \lambda x.\lambda y.sees(x,y)} \quad \dfrac{\textit{Mary}}{\mathbf{NP} : \textit{Mary}}}{\mathbf{S}\backslash\mathbf{NP} : \lambda y.sees(Mary,y)}>}{\mathbf{S} : sees(Mary,John)}<$$

Possible question:
Fill in the blank(s) to complete the derivation

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

Explain what we mean by thematic roles. Give examples.

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

Describe how thematic roles are expressed in Propbank.

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?

Discourse, Dialogue

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).

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

Generation and Dialog

Describe what we mean by a finite-state dialog manager. What are the advantages and disadvantages of this approach?

Describe what we mean by a frame-based dialog manager. What are the advantages and disadvantages of this approach?

What is a dialog act?

What do we mean by common ground in dialogue?

Good luck!!

Email/use Piazza for questions

Thank you!