CS447: Natural Language Processing

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

Lecture 28: Final Exam Review

Julia Hockenmaier

juliahmr@illinois.edu

3324 Siebel Center

Projects and Literature Reviews

Final report due **Friday, Dec 14, 11:59 PM**

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

No 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 12 in class

- Covers only materials after midterm (lectures 15—27)
- Does not cover lecture 20 (when Julia was out of town)
- Almost the same format as midterm:
  closed book, short questions **(now with multiple choice questions)**

Focus will be on conceptual questions (no need to remember complex mathematical formulas)

- Reason: we covered a lot of material in the second half of the semester only superficially, with the goal of giving you breadth rather than depth in a particular topic

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

***Discuss/Argue* whether …**

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

New: Multiple choice questions!

Semantics

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.

CCG with semantics

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

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

Verb semantics

Explain what we mean by thematic roles.

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

Distributional similarities

What is the distributional hypothesis?

Define what we mean by distributional similarities.

Define pointwise mutual information.

Why do we use pointwise mutual information instead of raw frequencies?

How do traditional vector-space semantic representations differ from neural word embeddings?

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?

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.

Discourse, Generation, Dialog

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 the basic architecture of an NLG system

What are the disadvantages of a finite-state dialog manager?

How do frame-based dialog systems differ from finite-state models?

Deep Learning for NLP

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.

RNNs and seq2seq models

What are the advantages of using an RNN for language modeling rather than a feedforward net?

What is the basic architecture of seq2seq models?

Good luck!!

Email/use Piazza for questions

Thank you!