Final Review and Wrap Up

CS4705

Natural Language Processing

Announcements

- Final: December 17th 1:10-4, 1024 Mudd
  - Closed book, notes, electronics
- Don't forget courseworks evaluation: only 4% so far have done it.
- Office hours as usual next week

What's Next?

- Natural Language for the Web (Spring 10)
  - TIME CHANGE: Thursdays 6-8pm
- Spoken Language Processing (Spring 10)
- Statistical natural language (Spring 10)
- Machine translation (Fall 10)

Natural Language for the Web

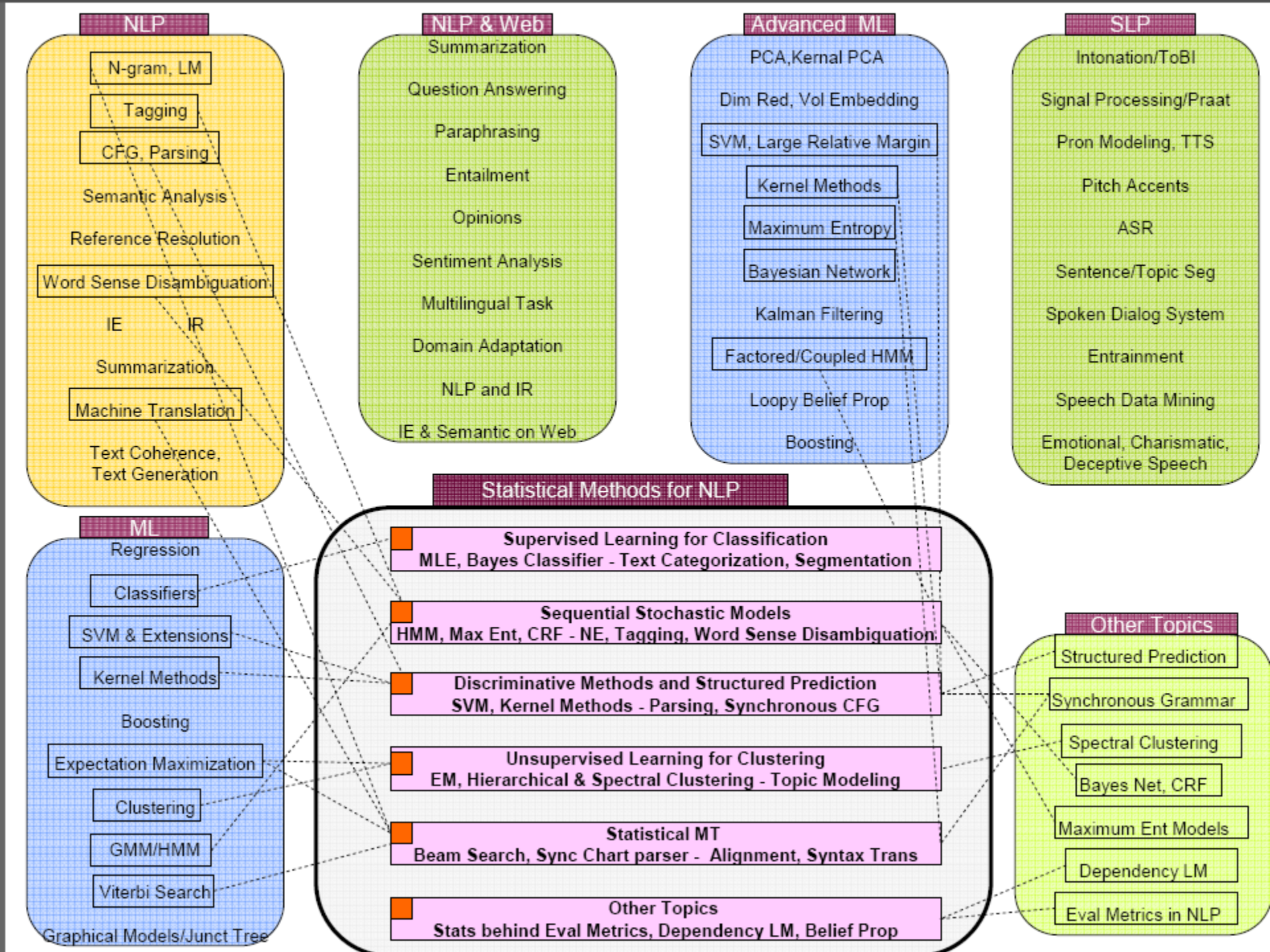
- Seminar style class
- Reading original papers
  - Presentation and discussion
- Semester long project

- The web contains huge amounts of unstructured documents, both written and spoken, in many languages. This class will study applications of natural language processing to the web. We will study search techniques that incorporate language, cross-lingual search, advanced summarization and question answering particularly for new media such as blogs, social networking, sentiment analysis and entailment. For many of these, we will look at multi-lingual approaches.

CS 4706: Spoken Language Processing

- Speech phenomena
  - Acoustics, intonation, disfluencies, laughter
  - Tools for speech annotation and analysis
- Speech technologies
  - Text-to-Speech
  - Automatic Speech Recognition
  - Speaker Identification
  - Dialogue Systems

- Challenges for speech technologies
  - Pronunciation modeling
  - Modeling accent, phrasing and contour
  - Spoken cues to
    - Discourse segmentation
    - Information status
    - Topic detection
    - Speech acts
    - Turn-taking
- Fun stuff: emotional speech, charismatic speech, deceptive speech….


NLP
N-gram, LM
Tagging
CFG, Parsing
Semantic Analysis
Reference Resolution
Word Sense Disambiguation
IE
IR
Summarization
Machine Translation
Text Coherence,
Text Generation
NLP & Web
Summarization
Question Answering
Paraphrasing
Entailment
Opinions
Sentiment Analysis
Multilingual Task
Domain Adaptation
NLP and IR
IE & Semantic on Web
Advanced ML
PCA,Kernal PCA
Dim Red, Vol Embedding
SVM, Large Relative Margin
Kernel Methods
Maximum Entropy
Bayesian Network
Kalman Filtering
Factored/Coupled HMM
Loopy Belief Prop
Boosting
SLP
Intonation/ToBI
Signal Processing/Praat
Pron Modeling, TTS
Pitch Accents
ASR
Sentence/Topic Seg
Spoken Dialog System
Entrainment
Speech Data Mining
Emotional, Charismatic,
Deceptive Speech
Statistical Methods for NLP
Supervised Learning for Classification
MLE, Bayes Classifier - Text Categorization, Segmentation
Sequential Stochastic Models
HMM, Max Ent, CRF - NE, Tagging, Word Sense Disambiguation
Discriminative Methods and Structured Prediction
SVM, Kernel Methods - Parsing, Synchronous CFG
Unsupervised Learning for Clustering
EM, Hierarchical & Spectral Clustering - Topic Modeling
Statistical MT
Beam Search, Sync Chart parser - Alignment, Syntax Trans
Other Topics
Stats behind Eval Metrics, Dependency LM, Belief Prop
ML
Regression
Classifiers
SVM & Extensions
Kernel Methods
Boosting
Expectation Maximization
Clustering
GMM/HMM
Viterbi Search
Graphical Models/Junct Tree
Other Topics
Structured Prediction
Synchronous Grammar
Spectral Clustering
Bayes Net, CRF
Maximum Ent Models
Dependency LM
Eval Metrics in NLP

Stop by and visit

- CS Advising
- Recommendation letters
- Research project
- Advice on applying to graduate school

What is Computational Linguistics? (NLP)

- An experiment done by outgoing ACL President Bonnie Dorr

Form of Final Exam

- Fill-in-the-blank/multiple choice
- Short answer
- Problem solving
- Essay
- Comprehensive (Will cover the full semester)

Semantics

- Meaning Representations
  - Predicate/argument structure and FOPC

$$\exists x,y\{Having(x)\wedge Haver(S,x)\wedge HadThing(y,x)\wedge Car(y)\}$$

- Thematic roles and selectional restrictions

  Agent/ Patient: George hit Bill. Bill was hit by George

  *George assassinated the senator. *The spider assassinated the fly*

- Compositional semantics
  - Rule 2 rule hypothesis
  - E.g. λx λy E(e) (Isa(e,Serving) ^ Server(e,y) ^ Served(e,x))
  - Lambda notation

    λ x P(x): λ + variable(s) + FOPC expression in those variables
- Non-compositional semantics
  - Metaphor: You're the cream in my coffee.
  - Idiom: The old man finally kicked the bucket.
  - Deferred reference: The ham sandwich wants his check.

Sample questions

- Give the FOPC meaning representation for:
  - John showed each girl an apple.
  - All students at Columbia University are tall.

- Given a sentence and a syntactic grammar, give the semantic representation for each word and the semantic annotations for the grammar. Derive the meaning representation for the sentence.

- Representing time:
  - Reichenbach ’47
    - Utterance time (U): when the utterance occurs
    - Reference time (R): the temporal point-of-view of the utterance
    - Event time (E): when events described in the utterance occur

    George is eating a sandwich.

    -- E,R,U →

    George will eat a sandwich?
- Verb aspect
  - Statives, activities, accomplishments, achievements

Word Relations

- Wordnet: pros and cons

- Types of word relations
  - Homonymy: bank/bank
  - Homophones: red/read
  - Homographs: bass/bass
  - Polysemy: Citibank/ The bank on 59th street
  - Synonymy: big/large
  - Hyponym/hypernym: poodle/dog
  - Metonymy: waitress: the man who ordered the ham sandwich wants dessert./the ham sandwich wants dessert.
  - The White House announced the bailout plan.

WordsEye

- What were some problems with WordNet that required creating their own dictionary?
- What are considerations about objects have to be taken into account when generating a picture that depicts an “on” relation?

Implicit Constraint. *The vase is on the nightstand. The lamp is next to the vase.*

Word Sense Disambiguation

Time flies like an arrow.

- Supervised methods
  - Collocational
  - Bag of words
    - What features are used?
    - Evaluation
- Semi-supervised
  - Use bootstrapping: how?
- Baselines
  - Lesk method
  - Most frequent meaning

Robust semantics

- Information Extraction
  - Three types of IE: NER, relation detection, QA
  - Three approaches: statistical sequence labeling, supervised, semi-supervised
  - Learning patterns:
    - Using Wikipedia
    - Using Google
    - Language modeling approach
- Information Retrieval
  - TF/IDF and vector-space model
  - Precision, recall, F-measure

IE Question

- What are the advantages and disadvantages of using exact pattern matching versus using flexible pattern matching for relation detection?
- Given a Wikipedia page for a famous person, show how you would derive the patterns for place of birth.
- If we wanted to use a language modeler to answer definition questions (e.g., “What is a quark?”), how would we do it?

Reference

- Referring expressions, anaphora, coreference, antecedents
- Types of NPs, e.g. pronouns, one-anaphora, definite NPs, ….
- Constraints on anaphoric reference
  - Salience
  - Recency of mention
  - Discourse structure
  - Agreement
  - Grammatical function

- Repeated mention
- Parallel construction
- Verb semantics/thematic roles
- Pragmatics

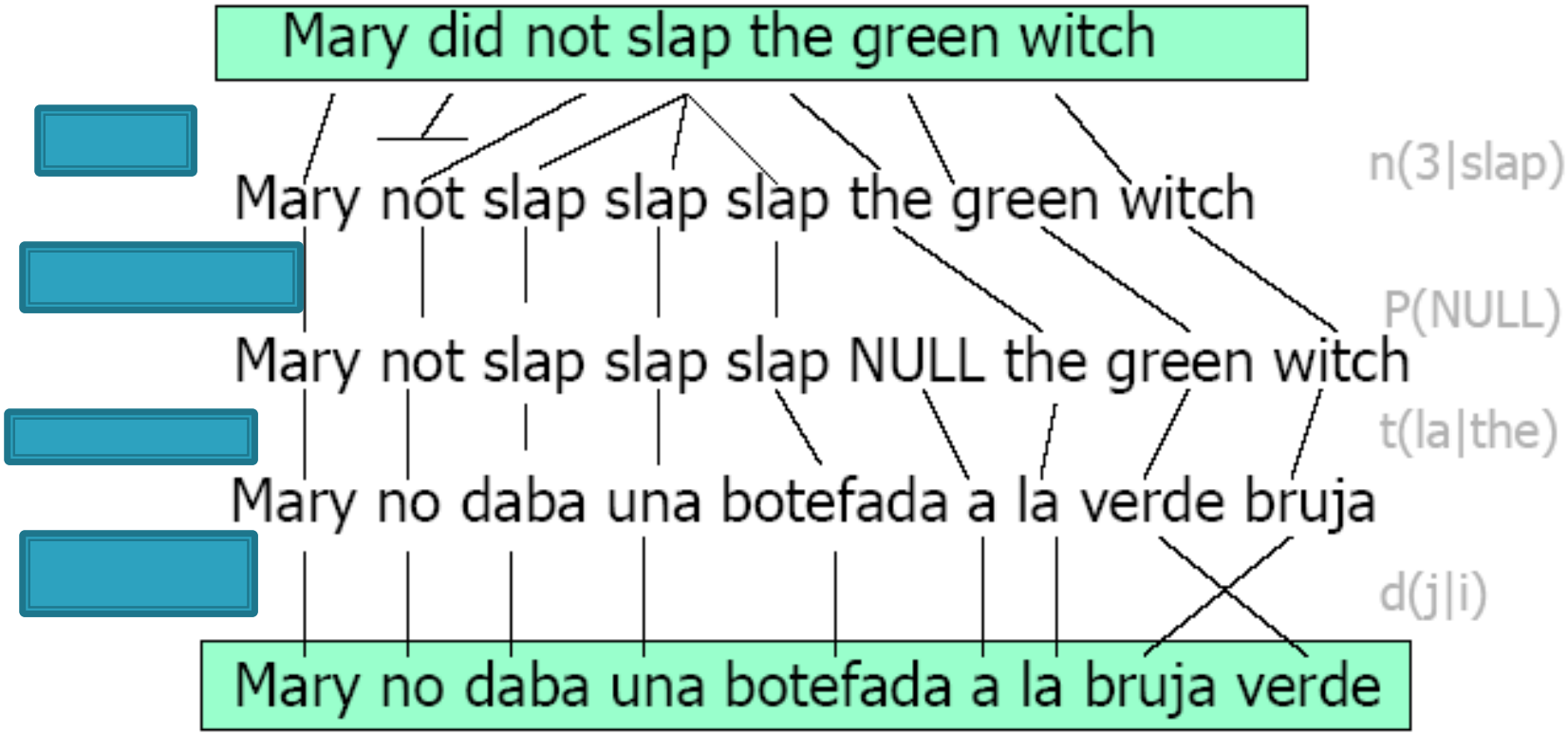
- Algorithms for reference resolution
  - Hobbes – most recent mention
  - Lappin and Leas
  - Centering

MT

- Challenges for MT
  - Orthographical
  - Lexical ambiguity
  - Morphological
  - Translational divergences
- MT Pyramid
  - Surface, transfer, interlingua
  - Statistical?
    - Word alignment
    - Phrase alignment
- Evaluation strategies
  - Bleu
  - Human levels of grading criteria

MT Questions

- How does lexical ambiguity affect MT?
- Compute the Bleu score for the following example, using unigrams and bigrams:
  - Translation: One moment later Alice went down the hole.
  - References: In another moment down went Alice after it,
  - In another minute Alice went into the hole.
  - In one moment Alice went down after it.
  - [never once considering how in the world she was to get out again.]


Mary did not slap the green witch
Mary not slap slap slap the green witch
n(3|slap)
Mary not slap slap slap NULL the green witch
P(NULL)
Mary no daba una botefada a la verde bruja
t(la|the)
d(j|i)
Mary no daba una botefada a la bruja verde

Generation

- Architecture
- Why is generation different from interpretation?
- What are some constraints on syntactic choice? Lexical choice?
- Functional unification grammar
- Statistical language generation
  - Overgenerate and prune
  - Input: abstract meaning representation
  - How are meaning representations linked to English?
  - What kinds of rules generate different forms?

An example grammar

```
((alt GSIMPLE (
   ;; a grammar always has the same form: an alternative
   ;; with one branch for each constituent category.

   ;; First branch of the alternative
   ;; Describe the category clause.
   ((cat clause)
    (agent    ((cat np)))
    (patient ((cat np)))
    (pred   ((cat verb-group)
                (number {agent number})))
    (cset (pred agent patient))
    (pattern  (agent pred patient))

   ;; Second branch: NP
   ((cat np)
    (head ((cat noun) (lex {^ ^ lex})))
    (number  ((alt np-number (singular plural))))
    (alt (  ;; Proper names don't need an article
           ((proper yes)
            (pattern (head)))

           ;; Common names do
           ((proper no)
            (pattern (det head))
            (det ((cat article) (lex "the")))))))

   ;; Third branch: Verb
   ((cat verb-group)
    (pattern (v))
    (aux none)
    (v ((cat verb) (lex {^ ^ lex}))))
))
```

A simple input

- Input to generate: The system advises John.

- I1 = ((cat np)
- (head ((lex “cat")))
- (number plural))

- Show unification with grammar.
- What would be generated?
- Suppose we wanted to change the grammar so that we could generate “a cat” or “cats”?

Discourse

- Structure
  - Topic segmentation
  - Lexical Cues for topic shift
    - Lexical repetition
    - Introduction of new words
    - Lexical chains
    - Possible question: given a discourse, compute the lexical repetition score between each block of 2 sentences
- Coherence
  - Rhetorical Structure
    - Rhetorical relations
    - Nucleus and satellite

Discourse Structure for Generation and Summarization

- Supervised method to learn rhetorical labels
  - Scientific articles
  - Improved summarization

- Content Modeling
  - Learn discourse structures for different topics
  - Domain specific
  - What was the algorithm?
  - How was it used for information ordering and summarization?

- Thank you and good luck on the exam!
- Stop by!

Another take: What is Computational Linguistics?