Natural Language Processing

Lecture 27: Conclusion
Levels of Linguistic Knowledge

spoken
  phonetics
  phonology
written
  orthography
  morphology
  syntax
  semantics
  pragmatics
discourse

“shallower”

“deeper”
“(behaving) as if you are among those whom we could not civilize”
uygarlaştıramadıklarımızdan mıssınızcasına
“(behaving) as if you are among those whom we could not civilize”

uygar  “civilized”
+laş  “become”
+tir  “cause to”
+ama  “not able”
+dık  past participle
+lar  plural
+ımız  first person plural possessive (“our”)  
+dan  second person plural (“y’all”)  
+miş  past
+ sınız  ablative case (“from/among”)  
+casına  finite verb → adverb (“as if”)
Finite-State Automaton

- **Q**: a finite set of states
- **q₀ ∈ Q**: a special start state
- **F ⊆ Q**: a set of final states
- **Σ**: a finite alphabet
- **Transitions**: Encodes a set of strings that can be recognized by following paths from **q₀** to some state in **F**.

- Encodes a **set** of strings that can be recognized by following paths from **q₀** to some state in **F**.
Levels of Linguistic Knowledge

spoken
phonetics
phonology

written
orthography

morphology
syntax
semantics
pragmatics
discourse

“shallower”

“deeper”

ambiguity
Noisy Channel

source \[\rightarrow y \rightarrow x\]

What you want
channel
What you see

\[y^* = \arg \max_y p(y \mid x)\]

\[= \arg \max_y \frac{p(x \mid y) \times p(y)}{p(x)}\]

\[= \arg \max_y p(x \mid y) \times p(y)\]
\[ y^* = \arg \max_y p(y | x) \]
\[ = \arg \max_y \frac{p(x | y) \times p(y)}{p(x)} \]
\[ = \arg \max_y p(x | y) \times p(y) \]
**Noisy Channel**

$y^* = \arg \max_y p(y \mid x)$

$= \arg \max_y \frac{p(x \mid y) \times p(y)}{p(x)}$

$= \arg \max_y p(x \mid y) \times p(y)$

**source**

How are you?

$\rightarrow y \rightarrow \text{channel} \rightarrow X$

你好吗？

decode
Noisy Channel

\[ y^* = \arg \max_y p(y \mid x) \]

\[ = \arg \max_y \frac{p(x \mid y) \times p(y)}{p(x)} \]

\[ = \arg \max_y p(x \mid y) \times p(y) \]
Starting and Stopping

Unigram model:

\[
\prod_{i=1}^{L+1} p(w_i)
\]

Bigram model:

\[
\prod_{i=1}^{L+1} p(w_i \mid w_{i-1})
\]

Trigram model:

\[
\prod_{i=1}^{L+1} p(w_i \mid w_{i-2}, w_{i-1})
\]
Language Modeling Questions

• Why do we use context?
• What does smoothing do, and why is it necessary?
• What do we use to evaluate language models?
Tagging

Bill directed plays about English kings

PropN	Adj	Verb	Prep	Adj	PIN
Verb	Verb	PIN	Adv	Noun	Verb
Noun
Broad POS categories

open classes
- nouns
- verbs
- adjectives
- adverbs

closed classes
- prepositions
- determiners
- pronouns
- conjunctions
- auxiliary verbs
- particles
- numerals
A hearing is scheduled on the issue today.
Parsing

• CKY vs. Earley’s Algorithm
  – Both dynamic programming
  – CNF vs. general forms
<table>
<thead>
<tr>
<th>Noun, Verb</th>
<th>Noun, Verb</th>
<th>VP, S</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>book</td>
<td>Det</td>
<td>NP</td>
<td>NP</td>
</tr>
<tr>
<td>this</td>
<td>Noun</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>flight</td>
<td>Prep</td>
<td>PP</td>
<td></td>
</tr>
<tr>
<td>through</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Houston</td>
</tr>
</tbody>
</table>
CKY Equations

\[
C[i - 1, i, w_i] = \text{TRUE}
\]

\[
C[i - 1, i, V] = \begin{cases} 
\text{TRUE} & \text{if } V \rightarrow w_i \\
\text{FALSE} & \text{otherwise}
\end{cases}
\]

\[
C[i, j, V] = \begin{cases} 
\text{TRUE} & \text{if } \exists j, Y, Z \text{ such that } V \rightarrow YZ \\
& \text{and } C[i, k, Y] \\
& \text{and } C[k, j, Z] \\
& \text{and } i < k < j \\
\text{FALSE} & \text{otherwise}
\end{cases}
\]

\[
\text{goal} = C[0, n, S]
\]
Semantics
Where’s the beef?

1. fertility. Organ meats such as beef and chicken liver, tongue and hear
2. controlling scour. _HOW TO FEED: BEEF AND DAIRY CALVES_ — 0.2 gram Dy
3. ing process discolors the treated beef and liquid accumulates in prepackag
4. say. He did say she could get her beef and vegetables in cans this summer
5. and feed efficiency of fattening beef animals. _HOW TO FEED:_ At the
6. steaks, chops, chicken and prime beef as well as Tom’s favorite dish, stu
7. ross from him was surmounted by a beef barrel with ends knocked out. In t
8. counter of boards laid across two beef barrels. There was, of course, no
9. Because Holstein cattle weren’t a beef breed, they were rarely seen on a
10. 2–5 grams of phenothiazine daily; beef calves—.5 to 1.5 grams daily depe
11. ties of this drug. _HOW TO FEED: BEEF CATTLE (FINISHING RATION)_ — To
12. dairy cows and lesser amounts to beef cattle and poultry. About 90 percen
13. raises enough poultry, pigs, and beef cattle for most of their needs. Lo
14. on of liver abscesses in feed-lot beef cattle. Prevention of bacterial pne
15. pal feed bunk types for dairy and beef cattle: (1) Fence-line bunks— catt
16. es feed efficiency. _HOW TO FEED: BEEF CATTLE_ — 10 milligrams of diet
17. the rations you are feeding your beef, dairy cattle, and sheep are adequa
18. tive business more profitable for beef, dairy, and sheep men. The tar
19. o bear. She was ready to kill the beef, dress it out, and with vegetables
20. . She had raised a calf, grown it beef-fat. She had, with her own work-wea
21. with feeding low-moisture corn in beef-feeding programs. Several firms ar
22. he shelf life (at 35 F) of fresh beef from 5 days to 5 or 6 weeks. Howeve
23. canned pork products. Tests with beef have been largely unsuccessful beca
24. for eggs, pigs to eat garbage, a beef herd and wastes of all kinds. Separ
25. their money’s worth. A good many beef-hungry settlers were accepting the
chicken

The irradiated and refrigerated chicken. Acceptance of radiopasteurization toehouse". Glendora dropped a chicken and a flurry of feathers, and went will specialize in steaks, chops, chicken and prime beef as well as Tom's fa yard as the one concerned with the chicken and the egg. Which came first? Is he millions of buffalo and prairie chicken and the endless seas of grass that "! "Come on, there's some cold chicken and we'll see what else". They wen ves to extend the storage life of chicken at a low cost of about 0.5 cent per CHICKEN CADILLAC# Use one 6-ounce chicken breast for each guest. Salt and pe ion juice, to about half cover the chicken breasts. Bake slowly at least one- d, in butter. Sprinkle over top of chicken breasts. Serve each breast on a th around, they had a hard time". #CHICKEN CADILLAC# Use one 6-ounce chicken successful, and the shelf life of chicken can be extended to a month or more ay from making a cake, building a chicken coop, or producing a book, to found , they decided, but a deck full of chicken coops and piggens was hardly suita im. "Johnny insisted on cooking a chicken dinner in my honor- he's always bee nutes. Kid Ory, the trombonist y Johnson reaching around the wire chicken yes glittering behind dull silver chicken fencing, which half covered the tr wine in the pot roast or that the chicken had been marinated in brandy, and 2 yed this same game and called it "Chicken". He could not go through the f f the Mexicans hiding in a little chicken house had passed through his head, 2 I'll never forget him cleaning the chicken in the tub". A story, no doubt . Organ meats such as beef and chicken liver, tongue and heart are planne p. "Miss Sarah, I can't cut up no chicken. Miss Maude say she won't". Aga 25 pot. "What is it"? he asked. "Chicken", Mose said, and theatrically lick im"? Adam shook his head. "Chicken", Mose said. She was a child too m
Synsets for dog (n)

- S: (n) dog, domestic dog, Canis familiaris (a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds) "the dog barked all night"
- S: (n) frump, dog (a dull unattractive unpleasant girl or woman) "she got a reputation as a frump"; "she's a real dog"
- S: (n) dog (informal term for a man) "you lucky dog"
- S: (n) cad, bounder, blackguard, dog, hound, heel (someone who is morally reprehensible) "you dirty dog"
- S: (n) frank, frankfurter, hotdog, hot dog, dog, wiener, wienerwurst, weenie (a smooth-textured sausage of minced beef or pork usually smoked; often served on a bread roll)
- S: (n) pawl, detent, click, dog (a hinged catch that fits into a notch of a ratchet to move a wheel forward or prevent it from moving backward)
- S: (n) andiron, firedog, dog, dog-iron (metal supports for logs in a fireplace) "the andirons were too hot to touch"
Entity Linking

Mary picked up the ball. She threw it to me.
Semantic Roles

PropBank is a set of verb-sense-specific “frames” with informal descriptions for their arguments.

Consider the word “Agree”

• ARG0: agreeer
• ARG1: proposition
• ARG2: other entity agreeing

[The group]$_{ARG0}$ agreed [it wouldn’t make an offer]$_{ARG1}$.
Usually [John]$_{ARG0}$ agrees [with Mary on everything]$_{ARG2}$.
“Fall (move downward)” in PropBank

- **arg1**: logical subject, patient, thing falling
- **arg2**: extent, amount fallen
- **arg3**: starting point
- **arg4**: ending point
- **argM-loc**: medium

Sales fell to $251.2 million from $278.8 million.
The average junk bond fell by 4.2%.
The meteor fell through the atmosphere, crashing into Cambridge.
MRL #1: First-Order Logic

DressCode(ThePorch)  Function

Serves(UnionGrill, AmericanFood)  Predicates
Restaurant(UnionGrill)

Have(Speaker, FiveDollars) \land \neg Have(Speaker, LotOfTime)

\forall x \ Person(x) \Rightarrow Have(x, FiveDollars)

\exists x, y \ Person(x) \land Restaurant(y) \land \neg HasVisited(x,y)
First Order Logic: Advantages

• Flexible
• Well-understood
• Widely used
EM

• We often have unlabeled or incomplete data
• EM is an for learning without labels, e.g., “classification” without classes

  • Pick random centroids!

  • Iterate the following:
    - **E-step** • Use centroids to label the data!
    - **M-step** • Compute centroids using the labeled data!

  • Keep doing this until labels don’t change
NLP Uses

Answer questions using the Web
Translate documents from one language to another
Do library research; summarize
Manage messages intelligently
Help make informed decisions
Follow directions given by any user
Fix your spelling or grammar
Grade exams
Write poems or novels
Listen and give advice
Estimate public opinion
Read everything and make predictions
Interactively help people learn
Help disabled people
Help refugees/disaster victims
Document or reinvigorate indigenous languages
More NLP ...

- Language Technologies Minor
  - 4 LT courses plus LT project
- 5\textsuperscript{th} year Masters in Language Technologies
More NLP Courses

• 11-492/692 Speech Processing  
  – Fall: Alan W Black  
  – Practical Systems for Speech
• 11-711 Algorithms and NLP  
  – Fall: Yulia Tsvetkov, Robert Frederking  
  – Research oriented
• 11-727 Computational Semantics  
  – Spring: Ed Hovy, Teruko Mitamura
More NLP Courses

- 11-747 Neural Networks for NLP
  - Spring: Graham Neubig
- 11-830 Computational Ethics for NLP
  - Spring: Yulia Tsvetkov, Alan W Black
- 11-777 Advanced Multimodal ML
  - Fall: Louis-Philippe Morency
  - Visual, Gesture, Speech
- Most Neural Net Classing
  - Always involve NLP