CSEP 573: Artificial Intelligence

Applications

Hanna Hajishirzi

Many slides over the course adapted from Pieter Abbeel, Luke Zettlemoyer and Dan Klein.
Applications of AI

- Web
- NLP
- Vision
- Robotics
- Games
- Predictions
- Diagnosis
- ...
Web Search

- Information retrieval:
  - Given information needs, produce information
  - Includes, e.g. web search, question answering, and classic IR

- Web search: not exactly classification, but rather ranking

\[ x = \text{“Apple Computers”} \]
Feature-Based Ranking

\[ x = \text{“Apple Computer”} \]

\[ f(x, \text{Apple}) = [0.3 \ 5 \ 0 \ 0 \ \ldots] \]

\[ f(x, \text{Apple Inc.}) = [0.8 \ 4 \ 2 \ 1 \ \ldots] \]
Perceptron For Ranking

- **Inputs** $x$
- **Candidates** $y$
- **Many feature vectors:** $f(x, y)$
- **One weight vector:** $w$
  - Prediction:
    $$y = \arg \max_y w \cdot f(x, y)$$
  - Update (if wrong):
    $$w = w + f(x, y^*) - f(x, y)$$
NLP

- Headlines:
  - Enraged Cow Injures Farmer With Ax
  - Hospitals Are Sued by 7 Foot Doctors
  - Ban on Nude Dancing on Governor’s Desk
  - Iraqi Head Seeks Arms
  - Local HS Dropouts Cut in Half
  - Juvenile Court to Try Shooting Defendant
  - Stolen Painting Found by Tree
  - Kids Make Nutritious Snacks

- Why are these funny?
  - Fundamental goal: analyze and process human language, broadly, robustly, accurately...
  - End systems that we want to build:
    - Ambitious: speech recognition, machine translation, information extraction, dialog interfaces, question answering...
    - Modest: spelling correction, text categorization...
Parsing

Hershey bars protest
Grammar

- Natural language grammars are very ambiguous!
- PCFGs are a formal probabilistic model of trees
  - Each “rule” has a conditional probability (like an HMM)
  - Tree’s probability is the product of all rules used
- Parsing: Given a sentence, find the best tree – search!

```
ROOT
  |- S
    |- NP
      |- PRP
        |- He
        
    |- VP
      |- VBD
        |- was
      
    |- ADJP
      |- JJ
        |- right

ROOT -> S 375/420
S -> NP VP . 320/392
NP -> PRP 127/539
VP -> VBD ADJP 32/401
......
```
Dialogue Systems

• Watson
  
  ▪ A question-answering system (IBM, 2011)
  ▪ Designed for the game of Jeopardy
  ▪ How does it work:
    ▪ Sophisticated NLP: deep analysis of questions, noisy matching of questions to potential answers
    ▪ Lots of data: onboard storage contains a huge collection of documents (e.g. Wikipedia, etc.), exploits redundancy
    ▪ Lots of computation: 90+ servers
  
  ▪ Can beat all of the people all of the time?
Machine Translation

- Translate text from one language to another
- Recombines fragments of example translations
- Challenges:
  - What fragments? [learning to translate]
  - How to make efficient? [fast translation search]
Problem with Dictionary Lookups

頂部  /top/roof/
顶端  /summit/peak/top/apex/
顶头  /coming directly towards one/top/end/
盖    /lid/top/cover/canopy/build/Gai/
盖帽  /surpass/top/
极    /extremely/pole/utmost/top/collect/receive/
尖峰  /peak/top/
面    /fade/side/surface/aspect/top/face/flour/
摘心  /top/topping/

Example from Douglas Hofstadter
Data-Driven Approach

Target language corpus:
- I will get to it soon
- See you later
- He will do it

Sentence-aligned parallel corpus:
- Yo lo haré mañana
  - I will do it tomorrow
- Hasta pronto
  - See you soon
- Hasta pronto
  - See you around

Machine translation system:
- Yo lo haré pronto
  - NOVEL SENTENCE
- Model of translation
- I will do it soon
# Learning to Translate

## Classic Soups

<table>
<thead>
<tr>
<th>No.</th>
<th>Item Description</th>
<th>Sm.</th>
<th>Lg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>57.</td>
<td>House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)</td>
<td>1.50</td>
<td>2.75</td>
</tr>
<tr>
<td>58.</td>
<td>Chicken Rice Soup</td>
<td>1.85</td>
<td>3.25</td>
</tr>
<tr>
<td>59.</td>
<td>Chicken Noodle Soup</td>
<td>1.85</td>
<td>3.25</td>
</tr>
<tr>
<td>60.</td>
<td>Cantonese Wonton Soup</td>
<td>1.50</td>
<td>2.75</td>
</tr>
<tr>
<td>61.</td>
<td>Tomato Clear Egg Drop Soup</td>
<td>1.65</td>
<td>2.95</td>
</tr>
<tr>
<td>62.</td>
<td>Regular Wonton Soup</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>63.</td>
<td>Hot &amp; Sour Soup</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>64.</td>
<td>Egg Drop Soup</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>65.</td>
<td>Egg Drop Wonton Mix</td>
<td>1.10</td>
<td>2.10</td>
</tr>
<tr>
<td>66.</td>
<td>Tofu Vegetable Soup</td>
<td>NA</td>
<td>3.50</td>
</tr>
<tr>
<td>67.</td>
<td>Chicken Corn Cream Soup</td>
<td>NA</td>
<td>3.50</td>
</tr>
<tr>
<td>68.</td>
<td>Crab Meat Corn Cream Soup</td>
<td>NA</td>
<td>3.50</td>
</tr>
<tr>
<td>69.</td>
<td>Seafood Soup</td>
<td>NA</td>
<td>3.50</td>
</tr>
</tbody>
</table>
An HMM Translation Model

E: Thank you, I shall do so gladly.

A: 1 → 3 → 7 → 6 → 8 → 8 → 8 → 8 → 9 → O

F: Gracias, lo haré de muy buen grado.

Model Parameters

Emissions: \( P(F_1 = \text{Gracias} \mid E_A = \text{Thank}) \)

Transitions: \( P(A_2 = 3 \mid A_1 = 1) \)
Levels of Transfer

\[ P(\text{English} | \text{lo haré}) = 0.8 \]

| English (E)       | \( P(\text{E} | \text{lo haré}) \) |
|-------------------|-----------------------------------|
| will do it        | 0.8                               |
| will do so        | 0.2                               |

| English (E)       | \( P(\text{E} | \text{mañana}) \) |
|-------------------|---------------------------------|
| tomorrow          | 0.7                             |
| morning           | 0.3                             |

\( \text{Yo lo haré mañana.} \)
\( \text{I will do it tomorrow.} \)
Coreference Resolution

[Michael Eisner] and [Donald Tsang] announced the grand opening of [Hong Kong Disneyland] yesterday. [Eisner] thanked [Mr. Tsang] and welcomed [fans] to [the park].

• Coreference resolution:
  – Determine when two mentions refer to same individual

• Require semantic knowledge to better coreference
Michael Eisner and Donald Tsang announced the grand opening of [[Hong Kong Disneyland]] yesterday. Eisner thanked Mr. Tsang and welcomed fans to the park.

- Will Eisner
- Kurt Eisner
- Michael Eisner

Michael Eisner: • Person • Businessman • Organization leader

Hong Kong Disneyland: • Location • Tourist attraction • Amusement park • park

• Match mentions to entities in an external knowledge base (Freebase, Wikipedia)
  – Use entity attributes as semantic knowledge
• NEL is challenging
Solving Arithmetic Word Problems

Liz had 9 black kittens. She gave some of her kittens to Joan. Joan has now 11 kittens. Liz has 5 kitten left and 3 has spots. How many kittens did Joan get?

Equation: \[ 9 - x = 5 \]
Solution: \[ x = 5 \] kittens

Liz gave some of her kittens to Joan.
Vision

• Search
• Detection
• Surveillance
• Recognition
Mobile visual search: **Google Goggles**

Google Goggles in Action

Click the icons below to see the different ways Google Goggles can be used.

- Landmark
- Book
- Contact Info
- Artwork
- Places
- Wine
- Logo
Face detection

- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, ...
Smile detection

The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.
Face recognition: Apple iPhoto, Facebook, Google, etc
Object recognition (in supermarkets)

LaneHawk by EvolutionRobotics
“A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk, you are assured to get paid for it... “
Computer alert for drowning girl

A 10-year-old girl has been saved from drowning by a computer system designed to raise the alarm when swimmers get into difficulties.

The girl, from Rochdale, was at the deep end of the pool in Bangor, north Wales, when she sank to the bottom.

The £65,000 system, called Poseidon, detected her on the pool floor and sounded the alarm. A lifeguard pulled her out and she recovered in hospital.
December 4, 2009 (CHICAGO) (WLS) -- Chicago police have closed the case in the death of Chicago School Board President Michael Scott.

Police Supt. Jody Weis says investigators used police cameras in the city to trace Scott's last steps in the hours before his body was found in November.

Scott's death has been ruled a suicide. The medical examiner's office concluded --not long after Scott's body was found -- that he had committed suicide. Police did not dispute the finding but wanted to pursue all the investigative leads they could. They say they have done that and have now reached the same conclusion.
Automotive safety

- **Mobileye**: Vision systems in high-end BMW, GM, Volvo models
  - Pedestrian collision warning
  - Forward collision warning
  - Lane departure warning
  - Headway monitoring and warning

Source: A. Shashua, S. Seitz
Intelligent Suspension system
SIGGRAPH Talks 2011

KinectFusion:
Real-Time Dynamic 3D Surface Reconstruction and Interaction

Shahram Izadi 1, Richard Newcombe 2, David Kim 1,3, Otmar Hilliges 1, David Molyneaux 1,4, Pushmeet Kohli 1, Jamie Shotton 1, Steve Hodges 1, Dustin Freeman 5, Andrew Davison 2, Andrew Fitzgibbon 1

1 Microsoft Research Cambridge 2 Imperial College London
3 Newcastle University 4 Lancaster University
5 University of Toronto
Object Detection
Person model

root filters
coarse resolution

part filters
finer resolution

deformation models
Person detections

high scoring true positives

high scoring false positives
(not enough overlap)
Car

root filters
coarse resolution

part filters
finer resolution

deformation
models
Car detections

high scoring true positives

high scoring false positives
Cat detections

high scoring true positives

high scoring false positives
(not enough overlap)
Autonomous Driving
Failures
An Autonomous Car

GPS

GPS compass

5 Lasers

E-stop

Camera

Radar

Control Screen

Steering motor

6 Computers

IMU

Red Bull

Android
Actions: Steering Control

Reference Trajectory

Error

Steering Angle (with respect to trajectory)

Velocity
Obstacle Detection

Trigger if $|Z^i - Z^j| > 15$cm for nearby $z^i, z^j$

Raw Measurements: 12.6% false positives
Probabilistic Error Model

\[ x_t \rightarrow x_{t+1} \rightarrow x_{t+2} \]

\[ Z_t \rightarrow Z_{t+1} \rightarrow Z_{t+2} \]

GPS IMU

GPS IMU

GPS IMU
HMMs for Obstacle Detection

Measurements: 12.6% false positives

HMM Inference: 0.02% false positives
Road Detection
Now on the Streets