CS 294-115 Algorithmic Human-Robot Interaction **Fall 2015**

Lecture 9: Natural Language with Dan Klein

Scribes: David McPherson, Christopher Lin

9.1 Language Technologies

Goal: Deep understanding

- requires context, linguistic structure, meaning and etc.

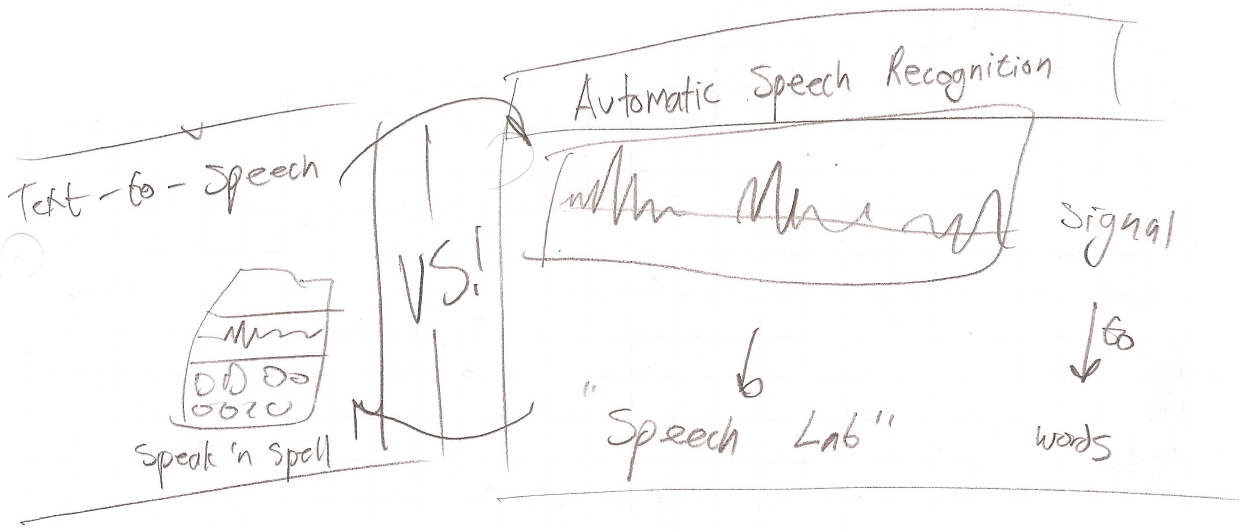
Reality: Shallow Matching (Ex: Grep)

- requires robustness and scale

- amazing success but fundamentally limited

9.2 State of the Art

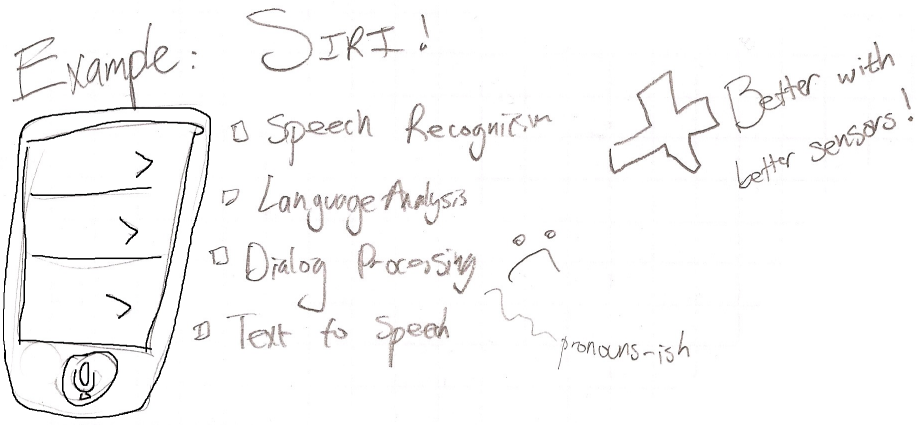
9.2.1 Text to Speech and Speech to Text



Automatic Speech Recognition (ASR)

- 0.3 % failure for digit strings
- 5 % failure for dictation
- 50% + failure for TV

Text To Speech is totally intelligible but oftentimes unnatural



9.2.2 Example: Siri

- contains speech recognition, language analysis, dialogue processing and text to speech
- voice command system, can't converse well
- microphone quality greatly affects accuracy. Apple has an advantage in this regard because hardware is standardized and they know the acoustic profile of the microphone

9.2.3 Example: Watson

- incredible engineering feat (of systems integration)
- effectively combined techniques present in the NLP community (the only novel part was satisfying constraints present in jeopardy questions such as rhymes with)
- would have lost if it was two Watsons vs one human because the two Watsons would split points whereas the human contestant wouldn't have to.

9.3 Iceberg Analogy

Text data is superficial (iceberg above water) but langauge is complex (iceberg below surface).
Lots of hidden structure in text

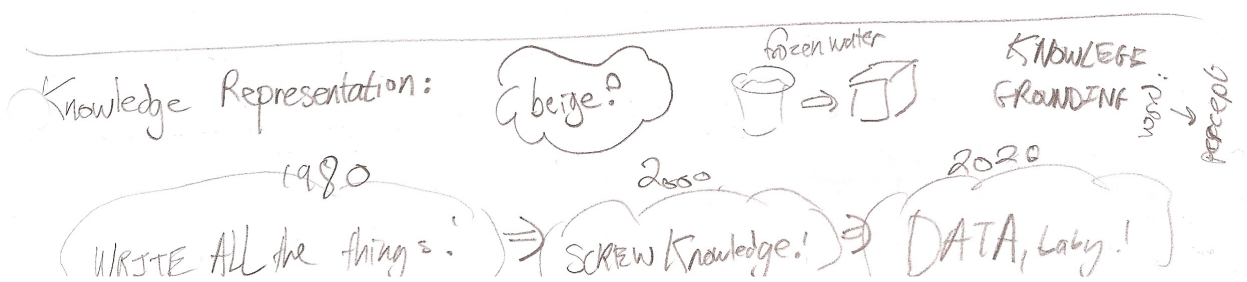
9.4 Machine Translation

- matching CHUNKS of words or phrases between languages
- difficulty in translation between languages vary due to amount of data and degree of alignment in language structure

- having lots of data really helps
- translation between English and French better than other languages due to similarities in sentence structure and lots of data from the Canadian government writing down proceedings in French and English

9.5 Data and Knowledge

- Classic knowledge representation worry: How will a machine ever know that...
  - Ice if frozen water?
  - What Beige looks like?
  - Chairs are solid?



- Answers:
  - 1980: write it all down
  - 2000: get by without it
  - 2020: learn if from data

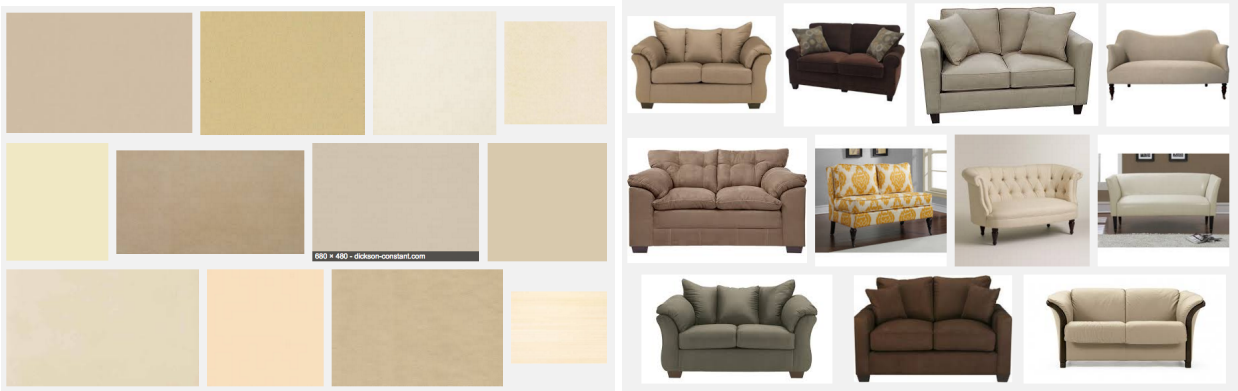
9.6 Deeper Understanding: Reference

Q: Who signed the Serve America Act?
A: Barack Obama
Text: Presdient Barack Obama received the Serve America Act after congress' vote. He signed the bill last Thursday. The president said it would greatly increase service opportunities for the American people.

Information is spread out over sentences. Systems need to understand that 'the bill' refers to the Serve America Act and that 'He' and 'The presdient' refers to Barack Obama.

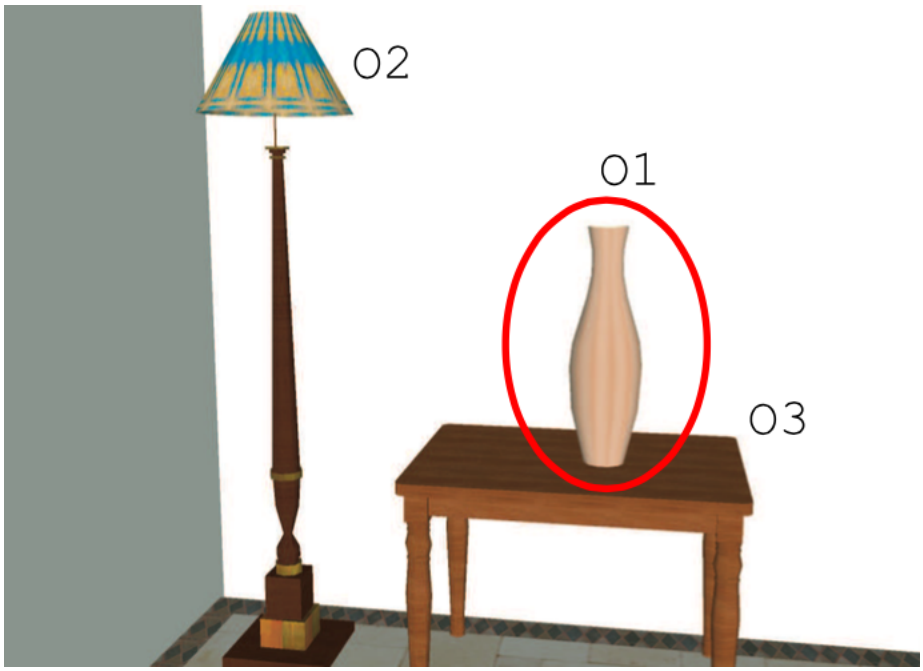
9.7 Grounded Language

9.7.1 Grounding with Natuaral Data

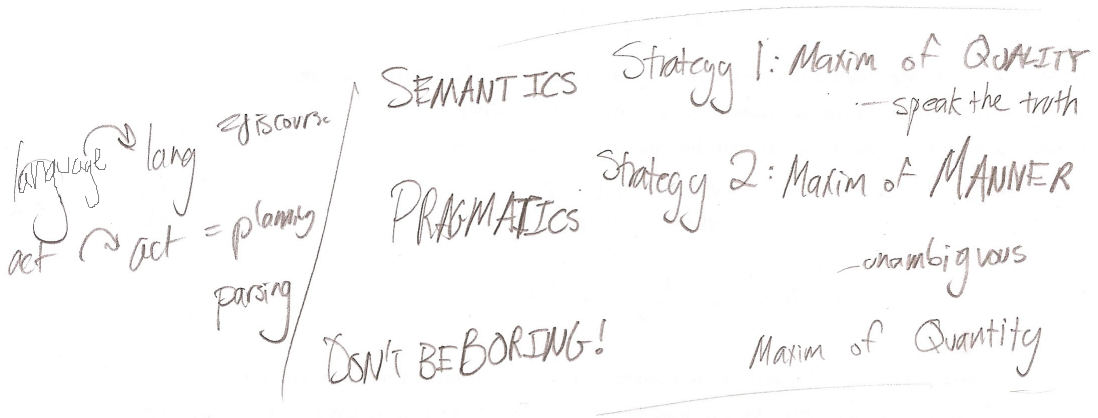


Pictures of beige and loveseat, but how does a system interpret what a Beige loveseat means or is?

9.7.2 Language is about communication



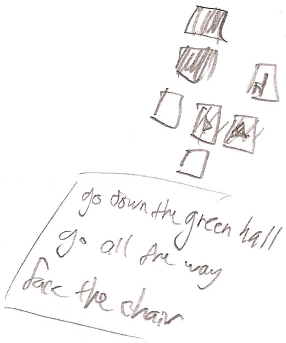
Goal: refer to O1



One approach is to speak the truth (Maxim of Quality). For example, right of O2
However, one problem is ambiguity (right of O2 could refer to two objects)
If we design the system to be unambiguous it may give needlessly long descriptions such as 'under the ceiling, to the right of the gray wall, to the right of O2 and etc...'
Hence, what we want is something unambiguous but at the same time not too long.

9.8 Alignment-based compositional semantics

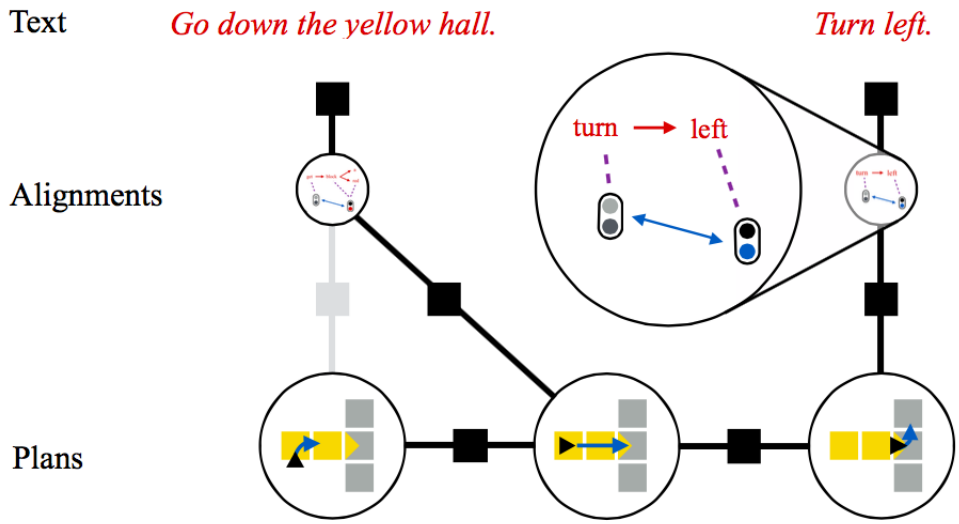
Goal: Follow instructions



Model:

- infers structure to structure alignment between the syntax of a sentence and a graph-based representation of the action

- graph valued conditional random field with alignment potentials to relate instructions and actions and transition potentials to encode environment model



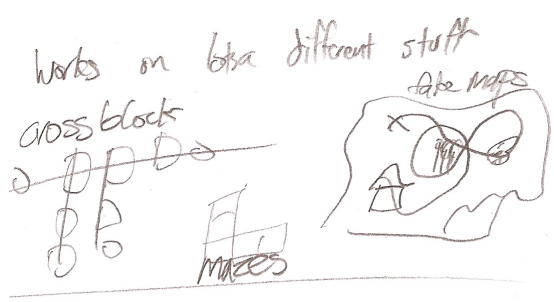
Solving Joint Model
Learning: coordinate ascent

$$\max_{\theta, \text{alignments}} P(actions, alignments | text; \theta)$$

Solve using EM

- solve for alignment using dynamic programming with θ fixed
- solve for θ using L-BFGS

When system fails, it ignores important portions of the instructions.



Can be used on a variety of games:

See 'Alignment-Based Compositional Semantics for Instruction Following' by Jacob Andreas and Dan Klein for details.