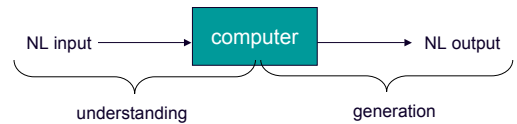


CS674 Natural Language Processing

- Topics for today
 - General introduction to NLP
 - » Why study NLP?
 - » Why is it hard?
 - Handouts
 - » Class description and syllabus
 - » Student info sheet

Natural language and NLP

- “natural” language
 - Languages that people use to communicate with one another
- Ultimate goal
 - To build computer systems that perform as well at using natural language as humans do
- Immediate goal
 - To build computer systems that can process text and speech more intelligently



Dialogue systems

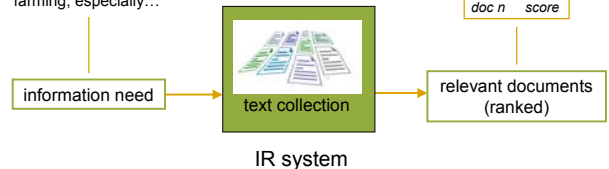
- Require both understanding and generation
 - Dave: Open the pod bay doors, HAL.
 - HAL: I'm sorry Dave, I'm afraid I can't do that.
 - Dave: What's the problem?
 - HAL: I think you know what the problem is just as well as I do.



Why study NLP?

- Useful applications...
 - E.g. information retrieval

Topic: Advantages and disadvantages of using potassium hydroxide in any aspect of organic farming, especially...



Why study NLP?

■ Useful applications...

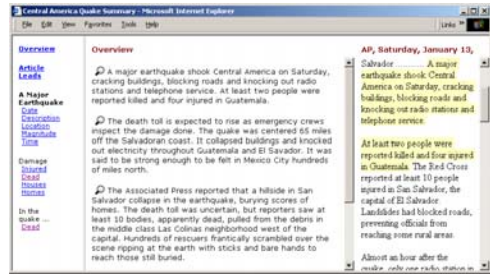
- E.g. question answering systems
 - » How many calories are there in a Big Mac?
 - » Who is the voice of Miss Piggy?
 - » Who was the first American in space?
- Retrieve not just relevant documents, but return the answer



Why study NLP?

■ Useful applications...

- E.g. summarization



[White et al., 2002]

Why study NLP?

■ Useful applications...

- E.g. machine translation
 - » Would clearly facilitate human-human communication
 - » Certainly see a need for it...
 - ◆ The extension of the coverage of the health services to the underserved or not served population of the countries of the region was the central goal of the Ten-Year Plan and probably that of greater scope and transcendence.
 - ◆ Welcome to Chinese Restaurant. Please try your Nice chinese Food With chopsticks, the traditional and typical of Chinese glorious history and cultural. PRODUCT OF CHINA

Bill Gates, 1997 "...now we're betting the company on these natural interface technologies"

Why study NLP?

■ Interdisciplinary...

- Linguistics
 - » models for language
- Psychology and psycholinguistics
 - » models of cognitive processes/language
- Mathematics
 - » studies properties of formal models, methods of inference from these models
- vs. NLP
 - » Computational study of language use
 - » Definite engineering aspect in addition to a scientific one
 - ◆ Engineering: to enable effective human-machine communication
 - ◆ Scientific: to explore the nature of linguistic communication
 - » Emphasis on computational, not cognitive plausibility
 - » Models of language: optional

Why study NLP?

- Challenging...
 - AI-complete
 - » To solve NLP, you'd need to solve all of the problems in AI
 - Turing test
 - » Posits that engaging effectively in linguistic behavior is a sufficient condition for having achieved intelligence.
- ...But little kids can “do” NLP...

Why is NLP such a difficult problem?

Ambiguity!!!! ...at **all** levels of analysis ☹

- Phonetics and phonology
 - Concerns how words are related to the sounds that realize them
 - Important for speech-based systems.
 - » "I scream" vs. "ice cream"
 - » "nominal egg"
 - Moral is:
 - » It's very hard to recognize speech.
 - » It's very hard to wreck a nice beach.
- Morphology
 - Concerns how words are constructed from sub-word units
 - Unionized
 - » un-ionized in chemistry?

Why is NLP such a difficult problem?

Ambiguity!!!! ...at **all** levels of analysis ☹

- Syntax
 - Concerns sentence structure
 - Different syntactic structure implies different interpretation
 - » Squad helps dog bite victim.
 - ◆ [_{np} squad] [_{vp} helps [_{np} dog bite victim]]
 - ◆ [_{np} squad] [_{vp} helps [_{np} dog] [_{inf-clause} bite victim]]
 - » Helicopter powered by human flies.
 - » Visiting relatives can be trying.

Why is NLP such a difficult problem?

Ambiguity!!!! ...at **all** levels of analysis ☹

- Semantics
 - Concerns what words mean and how these meanings combine to form sentence meanings.
 - » Jack invited Mary to the Halloween **ball**.
 - ◆ dance vs. some big sphere with with Halloween decorations?
 - » Visiting relatives can be trying.
 - » Visiting museums can be trying.
 - ◆ Same set of possible syntactic structures for this sentence
 - ◆ But the meaning of **museums** makes only one of them plausible

Why is NLP such a difficult problem?

Ambiguity!!!! ...at **all** levels of analysis ☹

- Discourse
 - Concerns how the immediately preceding sentences affect the interpretation of the next sentence
 - » Merck & Co. formed a joint venture with Ache Group, of Brazil. It will be called Prodome Ltd.
 - » Merck & Co. formed a joint venture with Ache Group, of Brazil. It will own 50% of the new company to be called Prodome Ltd.
 - » Merck & Co. formed a joint venture with Ache Group, of Brazil. It had previously teamed up with Merck in two unsuccessful pharmaceutical ventures.

Why is NLP such a difficult problem?

Ambiguity!!!! ...at **all** levels of analysis ☹

- Pragmatics
 - Concerns how sentences are used in different situations and how use affects the interpretation of the sentence.

“I just came from New York.”

 - » Would you like to go to New York today?
 - » Would you like to go to Boston today?
 - » Why do you seem so out of it?
 - » Boy, you look tired.

Syllabus

- We'll look at the problems that occur at all levels of natural language understanding...
- To the syllabus handout...