Introduction to Natural Language Processing

Jurafsky & Martin, Chapter 1

Outline

Administration (email), Review and Discussion

Natural Language Processing

Knowledge of Language

Applications

The Ambiguity Problem

Models and Algorithms

Can Machines Think?

History

Review and Discussion

◇ Last time: Environment Specification and Types; Agent Types

◇ Agent Discussion Points

- Both the performance measure and the utility function measure how well an agent is doing. Explain the difference.
- For a dialogue system with speech recognition errors, develop a PEAS description, characterize the environment, and select an agent design.
- Discuss possible agent programs for the following stochastic versions of the vacuum environment:
  - Murphy's law: 25% of the time, the sucking action fails to clean the floor if dirty, and deposits dirt if clean. How is your agent program affected if the dirt sensor gives the wrong answer 10% of the time?
  - Small children: At each time step, each clean square has a 10% chance of becoming dirty.

◇ NLP Demo Debriefing

Natural Language Processing

- The field of *Natural Language Processing (NLP)*, or *Computational Linguistics (CL)*, is primarily concerned with the creation of computer programs that perform useful and interesting tasks with human languages (e.g., understanding, generation, learning).
- It is secondarily concerned with using computational metaphors to help us come to a better understanding of human language.
- The foundations of the field are in computer science (e.g., AI, theory), linguistics, mathematics and statistics, electrical engineering, and psychology.
- Studying NLP involves studying natural languages, formal representations, and algorithms for their manipulation.

Knowledge of Language

- Example dialogue from *2001: A Space Odyssey*

```
Dave: Open the pod bay doors, HAL.
HAL: I'm sorry Dave, I'm afraid I can't do that.
```

- To participate in such a conversation, HAL needs knowledge about many levels of language
  - words: producing contractions, plurals
  - syntax: questions versus statements, word order and grouping
  - semantics: meaning of words in isolation and compositionally
  - pragmatics: politeness and indirectness
  - discourse: between utterance references

Applications

What makes a computer application a language processing application?

- language processing applications require the use of knowledge of language

Little Applications

Little applications typically make use of only a small amount of a single kind of knowledge of language. Many of these types of applications are currently in use, but they are often nearly invisible.

- line breakers
- hyphenators
- spelling correctors
- OCR software
- grammar and style checkers

Big Applications

Big applications often make use of large amounts and varied kinds of knowledge of language.

- information retrieval and extraction
- question answering
- dialogue management
- text summarization
- machine translation
- HAL

Some Current Application Scenarios

Tutor students in areas such as physics

Access the web over the telephone

Speak to your appliances

Grade essays

Generate running commentary for robotic soccer

Generate weather reports in multiple languages

Demos

Dialogue Systems

- ELIZA (www-ai.ijs.si/eliza/eliza.html)

Question Answering

- AnswerBus (misshoover.si.umich.edu/∼zzheng/qa-new)

Summarization

- Newsblaster (www.cs.columbia.edu/nlp/newsblaster)
- NewsInEssence (www.newsinessence.com/nie.cgi)

Machine Translation

- Babelfish (babelfish.altavista.com)

Spoken Language Demos

Spoken Dialogue Systems

- JUPITER (www.sls.lcs.mit.edu/applications/jupiter.shtml)
  - 1-888-573-TALK
- CMU Communicator (fife.speech.cs.cmu.edu/Communicator)
  - 1-877-268-7526
- SpeechWorks (www.speechworks.com/demos/index.cfm)
  - see particular applications for phone numbers

Speech Synthesis

- Bell Labs Text-To-Speech (www.bell-labs.com/project/tts/)

HAL

Returning to our HAL example . . .

```
Dave: Open the pod bay doors, HAL.
HAL: I'm sorry Dave, I'm afraid I can't do that.
```

The knowledge that HAL needs to take part in this exchange can be broken down into somewhat discrete categories.

Each category can then be studied (and modeled computationally) in isolation. This, of course, leaves open the problem of how the categories interact.

Knowledge of Language

Phonetics and Phonology: speech sounds, their production, and the rule systems that govern their use

Morphology: words and their composition from more basic units

- *cat, cats*
- *friend, friendly*

Syntax: the structuring of words into legal larger phrases and sentences

- *The textbook for the NLP class is great.*
- *Jane met Mary*
- *Mary was met by Jane* (passive)

Semantics: the meaning of words and phrases

- **lexical semantics**: the study of the meanings of words

- **compositional semantics** how to combine words
- *river bank, financial bank*
- *met(Jan,Mary)*

Pragmatics: utterance interpretation in situational context

- *Do you know the time?*

Discourse: utterance interpretation in context of previous utterances

- *Sue took a trip to New York. She had a great time there.*

Deconstructing HAL

Recognizes speech and understands language

Decides how to respond and speaks reply

With personality

Recognizes the user's goals, adopts them, and helps to achieve them

Remembers the conversational history

Customizes interaction to different individuals

Learns from experience

Possesses vast knowledge, and is autonomous

Ambiguity

Why is the following sentence problematic?

- *I saw the woman with the telescope.*

Ambiguity, cont.

Why is the following sentence problematic?

- *I saw the woman with the telescope.*

It is syntactically ambiguous...

- *I saw (NP the woman with the telescope)*
- *I saw (NP the woman) (PP with the telescope)*

The categories of knowledge of language can be thought of as ambiguity-resolving components.

"I made her duck"

How many different interpretations does this sentence have?

- I cooked waterfowl for her
- I cooked waterfowl belonging to her
- I created the (plaster?) duck she owns
- I caused her to quickly lower her head or body
- . . .

What if the sentence was spoken? ("I" vs "eye")

How can the lexical, syntactic, and semantic ambiguities be resolved/disambiguated?

- tagging (part of speech) - is "duck" a noun or verb?
- parsing (syntactic structure) - is "her" part of the duck phrase?
- word sense disambiguation (lexical semantics), e.g., does "make" mean create or cook?

Disambiguation

Pick the most likely of n choices.

Tightly coupled processing

- decisions that can be made easily based on one kind of knowledge can inform the others

Loosely coupled processing

- do your best within a single kind of knowledge (maybe several)

More Examples

Models and Algorithms

Models

- formalisms to represent linguistic knowledge

Algorithms

- used to manipulate the representations to produce desired behavior

A small set of standard models and algorithms have proven to be effective

Some Example Models

State machines

- finite state automata, finite state transducers

Formal rule systems

- context free grammars, unification grammars, probabilistic grammars

Logic-based formalism

- first order predicate calculus, higher order logics

Models of uncertainty

- Bayesian probability theory (often augments other models).

Use of states and rules has roots in speech, morphology, syntax, while logic has roots in semantics, pragmatics, and discourse. But this is changing!

Algorithms

Almost all of the algorithms we'll study can be viewed as transducers or parsers, i.e., algorithms that accept an input and construct some structure based on that input.

Unfortunately, since language is ambiguous at all levels this is almost never simple. This leads us to employ algorithms that fall into two related categories:

- state space search
- dynamic programming

State Space Search

The states in a search represent a pairing of partial structures with pieces of input.

The goal is to arrive at the right/best structure after having processed all of the input.

As with most AI problems the spaces are too large and the criteria for "bestness" difficult to encode.

Dynamic Programming

Informally, in the course of searching we often run across or construct structures that must be present in all possible solutions.

Dynamic programming allows us to avoid recomputing such structures.

Have the students talking in the back row of this room ...

Can Machines Think?

The **Turing Test**: language as a test for intelligence

Three participants (2 humans, 1 computer)

- Human interrogator's goal: tell the machine and human apart
- Machine's goal: fool the interrogator into believing that a person is responding
- Other human's goal: help the interrogator reach his goal

A Brief NLP History

Four paradigms: stochastic, logic-based, nlu, discourse modelling

Empiricism and finite-state models return

Recent years: strong integration of different techniques, different areas (including speech and information retrieval)

For Next Time

HW 2

Next readings