Orientation

This course divides roughly into three parts:

1. Words and sequences
2. Syntactic structure
3. Meaning (of words, sentences, and discourse)

Today we start part 3.

(We also have a bit more to say about some things from parts 1 and 2, will intersperse these topics.)

Meaning

• The grand goal of artificial intelligence
  – machines that do not mindlessly process data
  – ... but that ultimately understand its meaning

• But how do we know if we succeeded?

Eliza

A famous computer program from 1969 shows people can be easily fooled into thinking that machines have some deep understanding.

young woman: Men are all alike.
eliza: In what way?
young woman: They’re always bugging us about something specific or other.
eliza: Can you think of a specific example?
young woman: Well, my boyfriend made me come here.
eliza: Your boyfriend made you come here?
young woman: He says I’m depressed much of the time.
eliza: I am sorry to hear you are depressed.
young woman: It’s true. I’m unhappy.
eliza: Do you think coming here will help you not to be unhappy?

Online demo: http://nlp-addiction.com/eliza/
What is meaning? What is understanding?

• These are deep philosophical questions
• NLP usually takes a more pragmatic view: can the computer behave as though it understands (in order to do what we want)?
  – Dialogue systems (e.g., Eliza)
  – Machine translation
  – Question answering
• What issues will we face in building such systems?

A Concrete Goal

• We would like to build
  – a machine that answers questions in natural language.
  – may have access to knowledge bases
  – may have access to vast quantities of English text
• Basically, a smarter Google
• This is typically called Question Answering

Semantics

• To build our QA system we will need to deal with issues in semantics, i.e., meaning.
• Lexical semantics: the meanings of individual words (next few lectures)
• Sentential semantics: how word meanings combine (after that)
• Consider some examples to highlight problems in lexical semantics

Example Question

• Question
  When was Barack Obama born?
• Text available to the machine
  Barack Obama was born on August 4, 1961
• This is easy.
  – just phrase a Google query properly:
    "Barack Obama was born on *"
  – syntactic rules that convert questions into statements are straight-forward
Example Question (2)

- Question
  What plants are native to Scotland?

- Text available to the machine
  A new chemical plant was opened in Scotland.

- What is hard?
  - words may have different meanings (senses)
  - we need to be able to disambiguate between them

Example Question (3)

- Question
  Where did Theresa May go on vacation?

- Text available to the machine
  Theresa May spent her holiday in Cornwall

- What is hard?
  - words may have the same meaning (synonyms)
  - we need to be able to match them

Example Question (4)

- Question
  Which animals love to swim?

- Text available to the machine
  Polar bears love to swim in the freezing waters of the Arctic.

- What is hard?
  - words can refer to a subset (hyponym) or superset (hypernym) of the concept referred to by another word
  - we need to have database of such A is-a B relationships, called an ontology

Example Question (5)

- Question
  What is a good way to remove wine stains?

- Text available to the machine
  Salt is a great way to eliminate wine stains

- What is hard?
  - words may be related in other ways, including similarity and gradation
  - we need to be able to recognize these to give appropriate responses
Example Question (6)

Question
Did Poland reduce its carbon emissions since 1989?

Text available to the machine
Due to the collapse of the industrial sector after the end of communism in 1989, all countries in Central Europe saw a fall in carbon emissions.
Poland is a country in Central Europe.

What is hard?
– we need to do inference
– a problem for sentential, not lexical, semantics

Word Sense Ambiguity

Not all problems can be solved by WordNet alone.

Two completely different words can be spelled the same (homonyms): I put my money in the bank. vs. He rested at the bank of the river.
You can do it! vs. She bought a can of soda.

More generally, words can have multiple (related or unrelated) senses (polysemses)

Polysemous words often fall into (semi-)predictable patterns: see next slides (from Hugh Rabagliati in PPLS).

WordNet

Some of these problems can be solved with a good ontology.

WordNet (for English: see http://wordnet.princeton.edu/) is a hand-built ontology containing 117,000 synsets: sets of synonymous words.

Synsets are connected by relations such as
– hyponym/hypernym (IS-A: chair-furniture)
– meronym (PART-WHOLE: leg-chair)
– antonym (OPPOSITES: good-bad)

globalwordnet.org now lists wordnets in over 50 languages (but variable size/quality/licensing)

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Participating Senses</th>
<th>Example Sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animal for fur</td>
<td>Mink, chinchilla, rabbit, beaver, raccoon*, alpaca*, crocodile*</td>
<td>The mink drank some water / She likes to wear mink</td>
</tr>
<tr>
<td>Animal/Object</td>
<td>Chicken, sheep, pig, snake, star*, rat*, doll*</td>
<td>The chicken drank some water / He is a chicken</td>
</tr>
<tr>
<td>for personality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Animal for meat</td>
<td>Chicken, lamb, fish, shrimp, salmon*, rabbit*, lobster*</td>
<td>The chicken drank some water / The chicken is tasty</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Artifact for</td>
<td>Shower, bath, sauna, baseball,</td>
<td>The shower was leaking / The shower was relaxing</td>
</tr>
<tr>
<td>activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body part for</td>
<td>Arm, leg, hand, face, back*, head*, foot*, shoulder*, lip*,</td>
<td>John’s arm was tired / The arm was reupholstered</td>
</tr>
<tr>
<td>object part</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Building for people</td>
<td>Church, factory, school, airplane,</td>
<td>The church was built 20 years ago / The church sang a song</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complement Coercion</td>
<td>Begin, start, finish, try</td>
<td>John began reading the book / John began the book</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Container for contents</td>
<td>Bottle, can, pot, pan, bowl*, plate*, box*, bucket*</td>
<td>The bottle is made of steel / He drank half of the bottle</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word for question</td>
<td>Price, weight, speed</td>
<td>The price of the coffee was low / John asked the price of the coffee</td>
</tr>
</tbody>
</table>
### How many senses?

- How many senses does the word *interest* have?
  - She pays 3% *interest* on the loan.
  - He showed a lot of *interest* in the painting.
  - Microsoft purchased a controlling *interest* in Google.
  - It is in the national *interest* to invade the Bahamas.
  - I only have your best *interest* in mind.
  - Playing chess is one of my *interests*.
  - Business *interests* lobbied for the legislation.

- Are these seven different senses? Four? Three?

- Also note: distinction between polysemy and homonymy not always clear!

### WordNet senses for *interest*

**S1:** a sense of concern with and curiosity about someone or something, Synonym: involvement

**S2:** the power of attracting or holding one’s interest (because it is unusual or exciting etc.), Synonym: interestingness

**S3:** a reason for wanting something done, Synonym: sake

**S4:** a fixed charge for borrowing money; usually a percentage of the amount borrowed

**S5:** a diversion that occupies one’s time and thoughts (usually pleasantly), Synonyms: pastime, pursuit

**S6:** a right or legal share of something; a financial involvement with something, Synonym: stake

**S7:** (usu. plural) a social group whose members control some field of activity and who have common aims, Synonym: interest group
Polysemy in WordNet

- Polysemous words are part of multiple synsets
- This is why relationships are defined between synsets, not words
- On average,
  - nouns have 1.24 senses (2.79 if excluding monosemous words)
  - verbs have 2.17 senses (3.57 if excluding monosemous words)
- Is Wordnet too fine grained?

Stats from: http://wordnet.princeton.edu/wordnet/man/wnstats.7WN.html

Sharon Goldwater Word senses and relations 20

Different sense = different translation

- Another way to define senses: if occurrences of the word have different translations, these indicate different sense
  - Example interest translated into German
    - Zins: financial charge paid for load (Wordnet sense 4)
    - Anteil: stake in a company (Wordnet sense 6)
    - Interesse: all other senses
- Other examples might have distinct words in English but a polysemous word in German.

Sharon Goldwater Word senses and relations 21

Word sense disambiguation (WSD)

- For many applications, we would like to disambiguate senses
  - we may be only interested in one sense
  - searching for chemical plant on the web, we do not want to know about chemicals in bananas
- Task: Given a polysemous word, find the sense in a given context
- Popular topic, data driven methods perform well

WSD as classification

- Given word token in context, which sense (class) is it?
  - Just train a classifier, if we have sense-labeled training data:
    - She pays 3% interest/INTEREST-MONEY on the loan.
    - He showed a lot of interest/INTEREST-CURIOSITY in the painting.
    - Playing chess is one of my interests/INTEREST-HOBBY.
- SensEval and later SemEval competitions provide such data
  - held every 1-3 years since 1998
  - provide annotated corpora in many languages for WSD and other semantic tasks

Sharon Goldwater Word senses and relations 22
Other sources of training data

- Pseudo-words: create artificial corpus by conflating words
  - Example: replace all occurrences of banana and door with new word bananadoor to get training data like:
    
    She ordered a bananadoor/BANANA milkshake.
    The bananadoor/DOOR opened quietly.
    He looked out the bananadoor/DOOR.

  - Tells us how well system might work in theory, but not on real words.

- Multilingual parallel corpora (translated texts aligned by sentence)
  - translation indicates sense

What kind of classifier?

- Naive Bayes (see Lecture 10)
- Maximum entropy model (see next lecture)
- Decision lists (see J&M, 20.2.2)
- Decision trees (see any ML textbook)

Naive Bayes for WSD

- Naive Bayes requires estimates of:
  - The prior probability of each class (sense)
  - The probability of each feature given each class

- These can be estimated from the training data.

- But what features to use? (Same question for other classifiers!)

Simple features

- Directly neighboring words
  - interest paid
  - rising interest
  - lifelong interest
  - interest rate
  - interest piqued

- Any content words in a 50 word window
  - pastime
  - financial
  - lobbied
  - pursued
More features

• Syntactically related words
• Syntactic role in sense
• Topic of the text
• Part-of-speech tag, surrounding part-of-speech tags

Of course, with Naive Bayes we have the usual problem with correlated features... stay tuned for next lecture.

Evaluation

• Extrinsic: test as part of IR, QA, or MT system
• Intrinsic: evaluate classification accuracy or precision/recall against gold-standard senses
• Baseline: choose the most frequent sense (sometimes hard to beat)

Issues with WSD

• Not always clear how fine-grained the gold-standard should be
• Classifiers must be trained separately for each word
  – Hard to learn anything for infrequent or unseen words
  – Requires new annotations for each new word
  – Motivates unsupervised and semi-supervised methods (see J&M 20.5, 20.10)

Semantic roles

• Often we want to know who did what to whom?
• But the same event and participants can have different syntactic realizations:
  Sandy broke the glass. vs. The glass was broken by Sandy.
  She gave the boy a book. vs. She gave a book to the boy.
• Instead of focusing on syntax, consider the semantic roles (also called thematic roles) defined by each event.
Commonly used thematic roles

<table>
<thead>
<tr>
<th>Role</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agent</td>
<td><em>The boy kicked his toy</em></td>
</tr>
<tr>
<td>Theme</td>
<td><em>The boy kicked his toy</em></td>
</tr>
<tr>
<td>Experiencer</td>
<td><em>The boy felt sad</em></td>
</tr>
<tr>
<td>Result</td>
<td><em>The girl built a shelf with power tools</em></td>
</tr>
<tr>
<td>Instrument</td>
<td><em>The girl built a shelf with power tools</em></td>
</tr>
<tr>
<td>Source</td>
<td><em>She came from home</em></td>
</tr>
</tbody>
</table>

• J&M give definitions and additional roles

Issues with thematic roles

• No universally agreed-upon set of roles

• Items with the “same” role (e.g., Instrument) may not behave quite the same
  
  Sandy opened the door with a key  The key opened the door  Sandy ate the salad with a fork  *The fork ate the salad

• The two main NLP resources for thematic roles avoid these problems by defining very fine-grained roles:
  – Specific to individual verbs only (PropBank)
  – Specific to small groups of verbs (FrameNet)

Semantic role labeling

• The NLP task of identifying which words/phrases play which roles in an event.

• Supervised techniques similar to other classification tasks:
  – Training data from FrameNet or PropBank
  – Features are mostly related to syntactic structure and the particular words involved
  – Use one of many standard classifiers from machine learning

• Current research focuses on reducing the need for training data (e.g., to work on non-English languages)

Summary

• Aspects of lexical semantics:
  – Word senses, and methods for disambiguating.
  – Lexical semantic relationships, like synonymy, hyponymy, and meronymy.
  – Semantic roles: the roles of each argument in an event.

• Resources that provide annotated data for lexical semantics:
  – WordNet (senses, relations)
  – Propbank, FrameNet (semantic roles)