CS474 Natural Language Processing

- Today
  - Tiny intro to lexical semantic analysis
- Next lectures
  - Word sense disambiguation

Critique Guidelines

- <=1 page, typed (single space)
- The purpose of a critique is not to summarize the paper; rather you should choose one or two points about the work that you found interesting.
- Examples of questions that you might address are:
  - What are the strengths and limitations of its approach?
  - Is the evaluation fair? Does it achieve it support the stated goals of the paper?
  - Does the method described seem mature enough to use in real applications? Why or why not? What applications seem particularly amenable to this approach?
  - What good ideas does the problem formulation, the solution, the approach or the research method contain that could be applied elsewhere?
  - What would be good follow-on projects and why?

- Are the paper's underlying assumptions valid?
- Did the paper provide a clear enough and detailed enough description of the proposed methods for you to be able to implement them? If not, where is additional clarification or detail needed?

- Avoid unsupported value judgments, like "I liked..." or "I disagreed with..." If you make judgments of this sort, explain why you liked or disagreed with the point you describe.
- Be sure to distinguish comments about the writing of the paper from comment about the technical content of the work.
Semantic analysis

- Assigning meanings to linguistic utterances
- **Compositional semantics**: we can derive the meaning of the whole sentence from the meanings of the parts.
  - Max ate a green apple.
- Relies on knowing:
  - the meaning of individual words
  - how the meanings of individual words combine to form the meaning of groups of words
  - how it all fits in with syntactic analysis

Caveats

- Problems with a compositional approach
  - a former congressman
  - a toy elephant
  - kicked the bucket

Introduction to lexical semantics

- Lexical semantics is the study of
  - the systematic meaning-related connections among words and
  - the internal meaning-related structure of each word
- Lexeme
  - an individual entry in the lexicon
  - a pairing of a particular orthographic and phonological form with some form of symbolic meaning representation
- Sense: the lexeme’s meaning component
- Lexicon: a finite list of lexemes

Dictionary entries

- right  *adj.* located nearer the right hand esp. being on the right when facing the same direction as the observer.
- left  *adj.* located nearer to this side of the body than the right.
- red  *n.* the color of blood or a ruby.
- blood  *n.* the red liquid that circulates in the heart, arteries and veins of animals.
Lexical semantic relations: homonymy

- Homonyms: words that have the same form and unrelated meanings
  - Instead, a bank\(^1\) can hold the investments in a custodial account in the client's name.
  - But as agriculture burgeons on the east bank\(^2\), the river will shrink even more.
- Homophones: distinct lexemes with a shared pronunciation
  - E.g. would and wood, see and sea.
- Homographs: identical orthographic forms, different pronunciations, and unrelated meanings
  - The expert angler from Dora, Mo., was fly-casting for bass rather than the traditional trout.
  - The curtain rises to the sound of angry dogs baying and ominous bass chords sounding.

Why do these distinctions matter?

- One type or another is more likely to affect specific NLP applications.
  - Spelling correction?
  - Speech recognition?
  - Text-to-speech?
  - Information retrieval?

Lexical semantic relations: polysemy

- Polysemy: the phenomenon of multiple related meanings within a single lexeme
  - Example: While some banks furnish blood only to hospitals, others are much less restrictive.
  - New sense, e.g. bank\(^3\)?
  - Polysemy allows us to associate a lexeme with a set of related senses.
- Distinguishing homonymy from polysemy is not always easy. Decision is based on:
  - Etymology: history of the lexemes in question
  - Intuition of native speakers

Polysemous lexemes

- For any given single lexeme we would like to be able to answer the following questions:
  - What distinct senses does it have?
  - How are these senses related?
  - How can they be reliably distinguished?
- Answers dictate how well semantic analyzers, search engines, NL generators, and MT systems perform their tasks.
Polysemous lexemes

- For any given single lexeme we would like to be able to answer the following questions:
  - What distinct senses does it have?
    » generally rely on lexicographers
  - How are these senses related?
    » relatively little work in this area
  - How can they be reliably distinguished?
    » this is the task of word sense disambiguation

Word sense disambiguation

- Given a fixed set of senses associated with a lexical item, determine which of them applies to a particular instance of the lexical item

Two fundamental approaches
- WSD occurs during semantic analysis as a side-effect of the elimination of ill-formed semantic representations
- Stand-alone approach
  » WSD is performed independent of, and prior to, compositional semantic analysis
  » Makes minimal assumptions about what information will be available from other NLP processes
  » Applicable in large-scale practical applications