# Outline

- Introducing Lexical Semantics
- Lexical Semantic Relations: Homonymy, Polysemy, Synonymy, Hyponymy
- WordNet: A Database of Lexical Relations

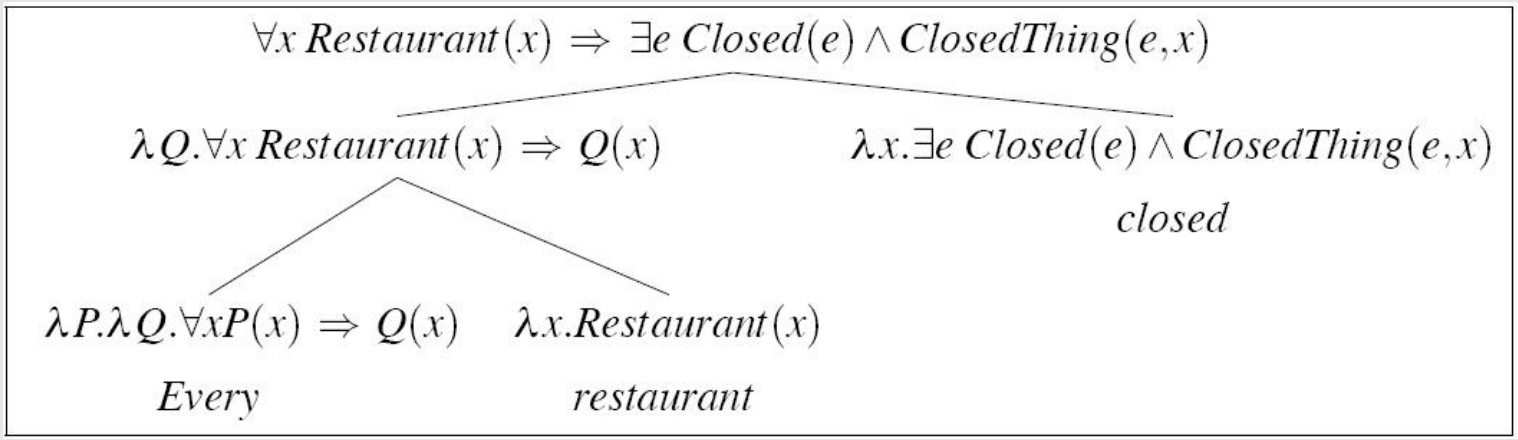
# Reading

- Reading:
  Chapter 19 of Jurafsky and Martin (2nd ed)
  From introduction to Section 19.3

- Further Background:
  D. A. Cruse. *Lexical Semantics*, Cambridge University Press, Cambridge, 1986.
  C. Kreidler. *Introducing English Semantics*, Routledge, London, 1998.

# Introducing Lexical Semantics

- Formal semantics deals with meaning representations for full sentences
  - ◇ aim is to construct meaning of syntactic wholes from parts (phrases, words)




- In this approach meanings of individual words are largely ignored

- Leads to view of lexicon as list of words without any systematic structure

# Introducing Lexical Semantics (cont)

- In fact there is rich structure
  - of semantic relations between words
  - internally between different meanings of individual words
- Study of this structure is called **lexical semantics**
- Useful to introduce notion of **lexeme** ('word' too ambiguous) – an individual entry in the **lexicon** (finite list of lexemes).
- Lexemes can be thought of (Jurafsky and Martin) as a pairing between
  - an orthographic and phonological form
  - a sense or meaning component

# Word Senses

- The meaning of a lemma can vary enormously given the context in which it appears
  - For example:
    I deposited the cheque at my bank
    Fishing from the river bank is prohibited
  - These are two separate lexemes with unrelated meanings that share an orthographic form – refer to these as, e.g. bank$^1$ and bank$^2$

- bank$^1$ and bank$^2$ are identical in both orthographic and phonological forms

- Pairs of lexemes like this are referred to as homonyms

# Homonyms, Homophones and Homographs

- Homonyms occur when pairs of lexemes have the same phonological and orthographic form.
- Words with distinct meanings which share phonological forms, but not orthographic forms are called homophones
  - ◇ e.g. wood-would, be-bee
- Words with distinct meanings which share orthographic forms, but not phonological forms are called homographs
  - ◇ e.g. bass(fish)-bass(guitar)

# Homonyms, Homophones and Homographs (contd)

- The application determines whether it is important to distinguish between homonyms, homophones or homographs

  Spelling correction
  - ◇ Distinguishing homophones may be important: whether-weather
  - ◇ Distinguishing imperfect homographs may be important: find-found (*Who found/founded the museum/treasure*)

  Speech recognition
  - ◇ Distinguishing homophones is important: need to be able to distinguish, e.g. to, two, too, their, there
  - ◇ Distinguishing homonyms may be important: for language modelling – if true homonyms are confused then n-gram probabilities will be incorrectly calculated (contexts for different homonyms conflated)

# Homonyms, Homophones and Homographs (contd)

## Text-to-speech

- Distinguishing homographs with distinct pronunciations important: bass/bass

## Information retrieval

- Distinguishing homographs may be important: users seeking information about bass(fish) are unlikely to be interested in documents about bass guitars
- Distinguishing homonyms may be important: users seeking information about bridge (card game) are unlikely to be interested in documents about bridge (structure spanning a river) construction

# Lexical Semantic Relations – Polysemy

- Homonyms are lexemes that have two distinct and unrelated meanings.
- Many lexemes have more than one related meaning – the phenomenon of multiple related meanings within a single lexeme is called polysemy
- For example:

  *While some banks furnish sperm only to married women, others are much less restrictive*

# Lexical Semantic Relations – Polysemy (cont)

- This use of bank is clearly not the bank$^2$ river bank meaning. However it is not the financial institution meaning either.
- bank has a range of sense related to repositories for biological entities (blood bank, egg bank, sperm bank etc.)
- Rather than create another lexeme bank$^3$ which would require asserting that this usage is unrelated to bank$^1$ (too strong) can say instead it is a different, *related* sense of the same lexeme
- Distinguishing homonymy from polysemy not easy – can use etymological evidence or can use intuitions about chance of relatedness
  - ◇ e.g. seems too much of a coincidence to believe uses of bank in repository sense are unrelated to financial institution sense.

# Lexical Semantic Relations – Polysemy (cont)

- Distinguishing multiple senses of a polysemous lexeme also difficult – domain of lexicographers.
- Use large corpus of examples + existing dictionary and try to account for all shades of meaning – can lead to more sense distinctions than ever required for any reasonable computational application (e.g. the Oxford concise dictionary has 30+ senses of go)
- Combine two separate uses of a lexeme into a single example using a conjunction (zeugma). For example:
  1. Which of those flights serve breakfast?
  2. Does Midwest Express serve Philadelphia?
  3. ?Does Midwest Express serve breakfast and Philadelphia?

  The oddness of 3. suggests that there are two related but distinct senses of serve here.

# Lexical Semantic Relations – Synonymy

- Two lexemes are **synonyms** if they have the same meaning.
  - ◇ e.g. cemetery/graveyard, car/automobile
- What does is it mean to have the same meaning?
- One answer: if one lexeme can be substituted for another in a sentence without changing the meaning or acceptability of the sentence.
- Example:

  How big is that plane?
  How large is that plane?

  Exchanging big and large and these examples has no effect $\Rightarrow$ big and large are synonyms
- Note synonymy does not require substitutability in *all* contexts – only in some.

  Sue became a kind of big sister to Timmy.
  ?Sue became a kind of large sister to Timmy.

# Lexical Semantic Relations – Metonymy

- Metonymy is a type of polysemy which occurs when there is a systematic relationship between the senses
- For example two closely related senses of bank are the *organisation* and *building* senses. Other words that share this sense distinction include school, university and hospital.

| | |
|---|---|
| Author (*Jane Austen wrote Emma*) | ↔ Works of Author (*I really love Jane Austen*) |
| Animal (*The chicken was domesticated in Asia*) | ↔ Meat (*The chicken was overcooked*) |
| Tree (*Plums have beautiful blossoms*) | ↔ Fruit (*I ate a preserved plum yesterday*) |

- More formally metonymy can be described as using one aspect of a concept to refer to other aspects of the entity, or the entity itself.

# Lexical Semantic Relations – Antonymy

- Antonyms are words with opposite meanings
  - ◇ long/short, big/little, fast/slow, cold/hot
- Difficult to formally define antonymy
- Different types exist:
  - ◇ Lexemes can be antonyms if they define a binary opposition or are at opposite end of some scale, e.g. long/short, fast/slow, big/little
  - ◇ Another group describes changes of direction or movement, e.g. rise/fall, up/down.
  - ◇ Can also distinguish between binary and non-binary antonyms

    Binary: left/right, dead/alive.
    Non-binary: old/young, wide/narrow
  - ◇ non-binary antonyms can be easily modified, e.g. very old, rather young, quite wide, extremely narrow

# Lexical Semantic Relations – Hyponymy

- One lexeme is a **hyponym** of another if it denotes a subclass of the other.

  E.g.: car is a hyponym of vehicle

- The more specific term (car) is the hyponym; the more general term (vehicle) is the **hypernym**.

| | | | | |
|---|---|---|---|---|
| **Superordinate** | vehicle | fruit | furniture | mammal |
| **Hyponym** | car | mango | chair | dog |

- To test for hyponymy use the schema:

  That is a x ⇒ That is a y
  That is a car ⇒ That is a vehicle

  If *x* is a hyponym of *y* then whenever the lefthand sentence in the schema is true the righthand sentence must be true as well.

# Ontologies and Taxonomies

- The concept of hyponymy is closely related to *ontologies* and *taxonomies*
  - ◇ Ontology: set of distinct objects characteristic of a domain or microworld
  - ◇ Taxonomy: arrangement of elements of an ontology into a tree-like class inclusion structure

  Consider: hound, mutt, puppy – all hyponyms of dog. However odd to construct a taxonomy from these pairs since the conceptual distinction underlying the subclass relation in each case is different in case (function: hound = breed of dog used for hunting; parentage: mutt = dog of mixed breed; age: puppy = young dog)

# WordNet: A Database of Lexical Relations

- WordNet is a large, freely available lexical database of English containing nouns, verbs, adjectives and adverbs grouped into sets of cognitive synonyms, called synsets, each expressing a distinct concept. See: http://wordnet.princeton.edu.
- Two entries are considered synonyms if they can be successfully substituted in some context.
- Sample synset: S: (n) chump, fool, gull, mark, patsy, fall guy, sucker, soft touch, mug (a person who is gullible and easy to take advantage of)

# WordNet: Lexical Semantic Relations for nouns

- Synsets are interlinked by means of lexical semantic relations.
- For nouns:
  - **hyponym** – from concepts to subtypes – meal $\rightarrow$ lunch
  - **hypernym** – from concepts to superordinates – breakfast $\rightarrow$ meal
  - **has-member** – from groups to their members – faculty $\rightarrow$ professor
  - **member-of** – from members to their groups – copilot $\rightarrow$ crew
  - **has-part** – from wholes to parts – table $\rightarrow$ leg
  - **part-of** – from parts to wholes – course $\rightarrow$ meal
  - **antonym** – opposites – leader $\rightarrow$ follower

# WordNet: Lexical Semantic Relations for verbs

- For verbs:
  - **hypernym** – from events to superordinate events – fly $\rightarrow$ travel
  - **troponym** – from events to their subtypes – walk $\rightarrow$ stroll
  - **entails** – from events to events they entail – snore $\rightarrow$ sleep
  - **antonym** – opposites – increase $\leftrightarrow$ decrease

# WordNet: Example

WordNet 3.0 entry for noun bass

The noun "bass" has 8 senses in WordNet.

1. bass[1] - (the lowest part of the musical range)
2. bass[2], bass part[1] - (the lowest part in polyphonic music)
3. bass[3], basso[1] - (an adult male singer with the lowest voice)
4. sea bass[1], bass[4] - (the lean flesh of a saltwater fish of the family Serranidae)
5. freshwater bass[1], bass[5] - (any of various North American freshwater fish with lean flesh (especially of the genus Micropterus))
6. bass[6], bass voice[1], basso[2] - (the lowest adult male singing voice)
7. bass[7] - (the member with the lowest range of a family of musical instruments)
8. bass[8] - (nontechnical name for any of numerous edible marine and freshwater spiny-finned fishes)

The adjective "bass" has 1 sense in WordNet.

1. bass[1], deep[6] - (having or denoting a low vocal or instrumental range)
   *"a deep voice"; "a bass voice is lower than a baritone voice"; "a bass clarinet"*

# WordNet: Example

The inherited hypernyms for two senses of bass

```
Sense 3
bass, basso --
(an adult male singer with the lowest voice)
=> singer, vocalist, vocalizer, vocaliser
   => musician, instrumentalist, player
      => performer, performing artist
         => entertainer
            => person, individual, someone...
               => organism, being
                  => living thing, animate thing,
                     => whole, unit
                        => object, physical object
                           => physical entity
                              => entity
               => causal agent, cause, causal agency
                  => physical entity
                     => entity

Sense 7
bass --
(the member with the lowest range of a family of
musical instruments)
=> musical instrument, instrument
   => device
      => instrumentality, instrumentation
         => artifact, artefact
            => whole, unit
               => object, physical object
                  => physical entity
                     => entity
```

# WordNet: Example

WordNet 3.0 entry for bill (nominal senses)

- S: (n) bill, measure (a statute in draft before it becomes law) *"they held a public hearing on the bill"*
- S: (n) bill, account, invoice (an itemized statement of money owed for goods shipped or services rendered) *"he paid his bill and left"*; *"send me an account of what I owe"*
- S: (n) bill, note, government note, bank bill, banker's bill, bank note, banknote, Federal Reserve note, greenback (a piece of paper money (especially one issued by a central bank)) *"he peeled off five one-thousand-zloty notes"*
- S: (n) bill (the entertainment offered at a public presentation)
- S: (n) circular, handbill, bill, broadside, broadsheet, flier, flyer, throwaway (an advertisement (usually printed on a page or in a leaflet) intended for wide distribution) *"he mailed the circular to all subscribers"*
- S: (n) poster, posting, placard, notice, bill, card (a sign posted in a public place as an advertisement) *"a poster advertised the coming attractions"*
- S: (n) bill (a list of particulars (as a playbill or bill of fare))
- S: (n) bill, billhook (a long-handled saw with a curved blade) *"he used a bill to prune branches off of the tree"*
- S: (n) bill, peak, eyeshade, visor, vizor (a brim that proj ects to the front to shade the eyes) *"he pulled down the bill of his cap and trudged ahead"*
- S: (n) beak, bill, neb, nib, pecker (horny projecting mouth of a bird)

# WordNet: Example (cont)

The inherited hypernyms of one sense of bill

- S: (n) beak, bill, neb, nib, pecker (horny projecting mouth of a bird)
  - * direct hyponym / full hyponym
  - * direct hypernym / **inherited hypernym** / sister term
    - S: (n) mouth (the externally visible part of the oral cavity on the face and the system of organs surrounding the opening) "she wiped lipstick from her mouth"
      - S: (n) orifice, opening, porta (an aperture or hole that opens into a bodily cavity) "the orifice into the aorta from the lower left chamber of the heart"
        - S: (n) passage,passageway (a path or channel or duct through or along which something may pass) "the nasal passages"
          - S: (n) structure, anatomical structure, complex body part, bodily structure, body structure (a particular complex anatomical part of a living thing) "he has good bone structure"
            - S: (n) body part (any part of an organism such as an organ or extremity)
              - S: (n) part, piece (a portion of a natural object) "they analyzed the river into three parts"; "he needed a piece of granite"
                - S: (n) thing (a separate and self-contained entity)
                  - S: (n) physical entity (an entity that has physical existence)
                    - S: (n) entity (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))

# WordNet Statistics (v3.0)

**Polysemy information**

| POS | Monosemous Words + Senses | Polysemous Words | Polysemous Senses |
|---|---|---|---|
| Noun | 101,863 | 15,935 | 44,449 |
| Verb | 6,277 | 5,252 | 18,770 |
| Adjective | 16,503 | 4,976 | 14,399 |
| Adverb | 3,748 | 733 | 1,832 |
| Totals | 128,391 | 26,896 | 79,450 |

| POS | Average Polysemy Inc Monosemous Words | Average Polysemy Exc Monosemous Words |
|---|---|---|
| Noun | 1.24 | 2.79 |
| Verb | 2.17 | 3.57 |
| Adjective | 1.40 | 2.71 |
| Adverb | 1.25 | 2.50 |

# Summary

- Lexical semantics is concerned with
  - ◇ common meaning relations that hold between words, e.g. synomymy, hyponomy/hyperonymy, antonymy, meronymy
  - ◇ multiple senses of single word forms – polysemy
- Traditional dictionaries provide catalogues of word senses.
- Electronic lexical databases such as WordNet provide not only word senses for individual words but a wide range of additional lexical semantic relational information.