

TTIC 31190: Natural Language Processing

Kevin Gimpel
Spring 2018

Lecture 3: Word Sense and Lexical Semantics

- All materials are posted on the course website:
`ttic.uchicago.edu/~kgimpel/teaching/31190-s18/index.html`
- Assignment 1 due 6:00 pm on Wed., April 11th

Roadmap

- words, morphology, lexical semantics
- text classification
- simple neural methods for NLP
- language modeling and word embeddings
- recurrent/recursive/convolutional networks in NLP
- sequence labeling, HMMs, dynamic programming
- syntax and syntactic parsing
- semantics, compositionality, semantic parsing
- machine translation and other NLP tasks

Words

- types and tokens
- morphology
- **distributional word vectors**
- word sense and lexical semantics

Counting Context Words

sugar, a sliced lemon, a tablespoonful of
 enjoyment. Cautiously she sampled her first
 well suited to programming on the digital
 for the purpose of gathering data and

apricot
pineapple
computer.
information

preserve or jam, a pinch each of,
 and another fruit whose taste she likened
 In finding the optimal R-stage policy from
 necessary for the study authorized in the

| | aardvark | computer | data | pinch | result | sugar | ... |
|-------------|----------|----------|------|-------|--------|-------|-----|
| apricot | 0 | 0 | 0 | 1 | 0 | 1 | ... |
| pineapple | 0 | 0 | 0 | 1 | 0 | 1 | ... |
| digital | 0 | 2 | 1 | 0 | 1 | 0 | ... |
| information | 0 | 1 | 6 | 0 | 4 | 0 | ... |
| ... | | | | | | | |

Handling Sentence Boundaries

`<s> these symptoms do not imply sickness . </s>`

- pad sentences with `<s>` at the beginning and `</s>` at the end
- these are context words just like any others
- when context window goes outside the sentence, you can use a single padding symbol or enough to fill the window
 - (I used a single padding symbol at each end)

Cosine Similarity

- dot product divided by vector length product

$$\frac{\mathbf{u}^\top \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}$$

- cosine of angle between the vectors

Context words of “cooked” with largest counts

| | | | |
|-----|------|----|--------|
| 123 | , | 13 | as |
| 92 | and | 12 | for |
| 79 | the | 12 | food |
| 71 | . | 11 | which |
| 68 | <s> | 11 | that |
| 66 | </s> | 11 | meat |
| 53 | in | 11 | can |
| 39 | a | 11 | by |
| 38 | is | 10 | when |
| 35 | of | 9 | rice |
| 30 | with | 9 | raw |
| 28 | are | 9 | beef |
| 25 | to | 7 | they |
| 23 | or | 7 | their |
| 23 | it | 7 | on |
| 20 | (| 7 | not |
| 19 | be | 7 | from |
| 15 |) | 6 | leaves |
| 14 | " | 6 | has |

Pointwise Mutual Information (PMI)

- do two events x and y co-occur more often than if they were independent?

$$\text{pmi}(x; y) = \log_2 \frac{p(x, y)}{p(x)p(y)}$$

- replace raw counts with pmi scores

Context words of “cooked” with highest PMIs

| | | | |
|---------|------------|---------|-----------|
| 9.30533 | beef | 7.66406 | chili |
| 8.88418 | shrimp | 7.56264 | rice |
| 8.63397 | potatoes | 7.56167 | soup |
| 8.61946 | ate | 7.45315 | flour |
| 8.56584 | dishes | 7.43874 | steamed |
| 8.50945 | eaten | 7.43715 | crushed |
| 8.4931 | beans | 7.41193 | meals |
| 8.33137 | texture | 7.39793 | digest |
| 8.29489 | vegetables | 7.39175 | rockies |
| 8.25088 | soda | 7.34773 | ramsay |
| 8.20831 | meat | 7.33211 | honey |
| 8.15708 | sauce | 7.32253 | toxicity |
| 8.08345 | consuming | 7.29057 | cared |
| 7.9532 | cuisine | 7.28626 | tomatoes |
| 7.94043 | raw | 7.27912 | boiling |
| 7.78435 | curry | 7.27769 | dal |
| 7.7563 | juice | 7.27485 | citrus |
| 7.74444 | vegetable | 7.25649 | doncaster |

How should we evaluate word vectors?

WordSim353

(Finkelstein et al., 2002)

| word pair | | similarity |
|-----------|-----------|------------|
| journey | voyage | |
| king | queen | |
| computer | software | |
| law | lawyer | |
| forest | graveyard | |
| rooster | voyage | |

WordSim353

(Finkelstein et al., 2002)

Instructions:

Assign a numerical similarity score between 0 and 10
(0 = words are totally unrelated,
10 = words are VERY closely related).

| | | |
|----------|-----------|--|
| computer | software | |
| law | lawyer | |
| forest | graveyard | |
| rooster | voyage | |

WordSim353

(Finkelstein et al., 2002)

Instructions:

Assign a numerical similarity score between 0 and 10
(0 = words are totally unrelated,
10 = words are VERY closely related).

When estimating similarity of antonyms, consider them "similar" (i.e., belonging to the same domain or representing features of the same concept), **rather than "dissimilar"**.

| | | |
|---------|-----------|--|
| forest | graveyard | |
| rooster | voyage | |

WordSim353

(Finkelstein et al., 2002)

| word pair | | similarity |
|-----------|-----------|------------|
| journey | voyage | 9.3 |
| king | queen | 8.6 |
| computer | software | 8.5 |
| law | lawyer | 8.4 |
| forest | graveyard | 1.9 |
| rooster | voyage | 0.6 |

SimLex-999

(Hill et al., 2014)

| word pair | | similarity |
|-----------|---------|------------|
| insane | crazy | 9.6 |
| attorney | lawyer | 9.4 |
| author | creator | 8.0 |
| diet | apple | 1.2 |
| new | ancient | 0.2 |

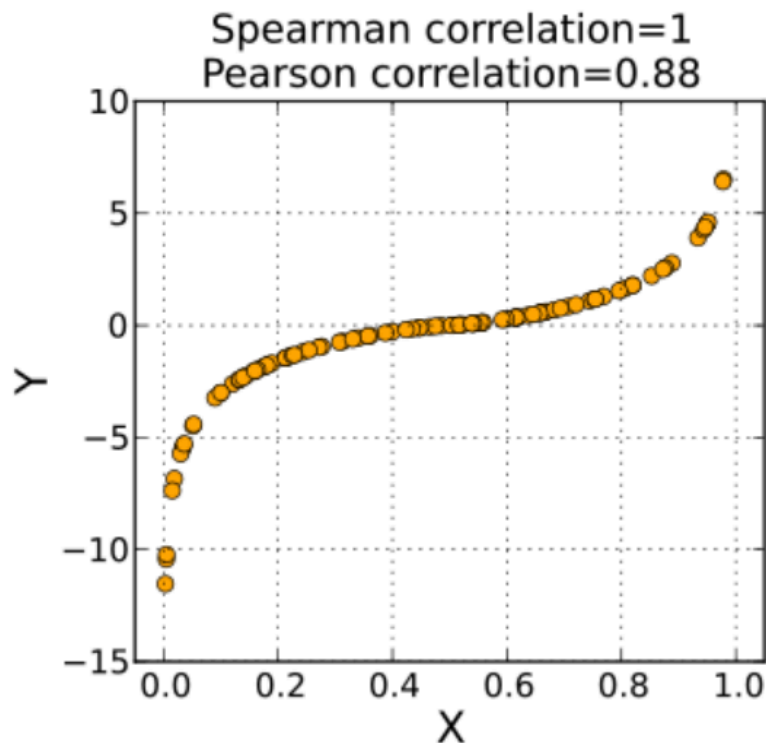
measures **paraphrastic** similarity:

two words are “similar” if they have similar meanings

- there are many word similarity datasets
- some focus on topical **relatedness**, others focus on similarity in **meaning**
- in assignment 1, you will evaluate your word vectors using MEN (relatedness) and SimLex-999 (meaning)

Evaluation Metrics for Word Similarity

- Spearman rank correlation coefficient
- measures correlation between two variables:
 - variable 1: human-annotated similarities for word pairs
 - variable 2: cosine similarities computed with your word vectors for the same word pairs



Sparse versus dense vectors

- so far, our vectors are
 - **long** (length = 25,000)
 - **sparse** (mostly zero)
- why might we want to reduce vector dimensionality?

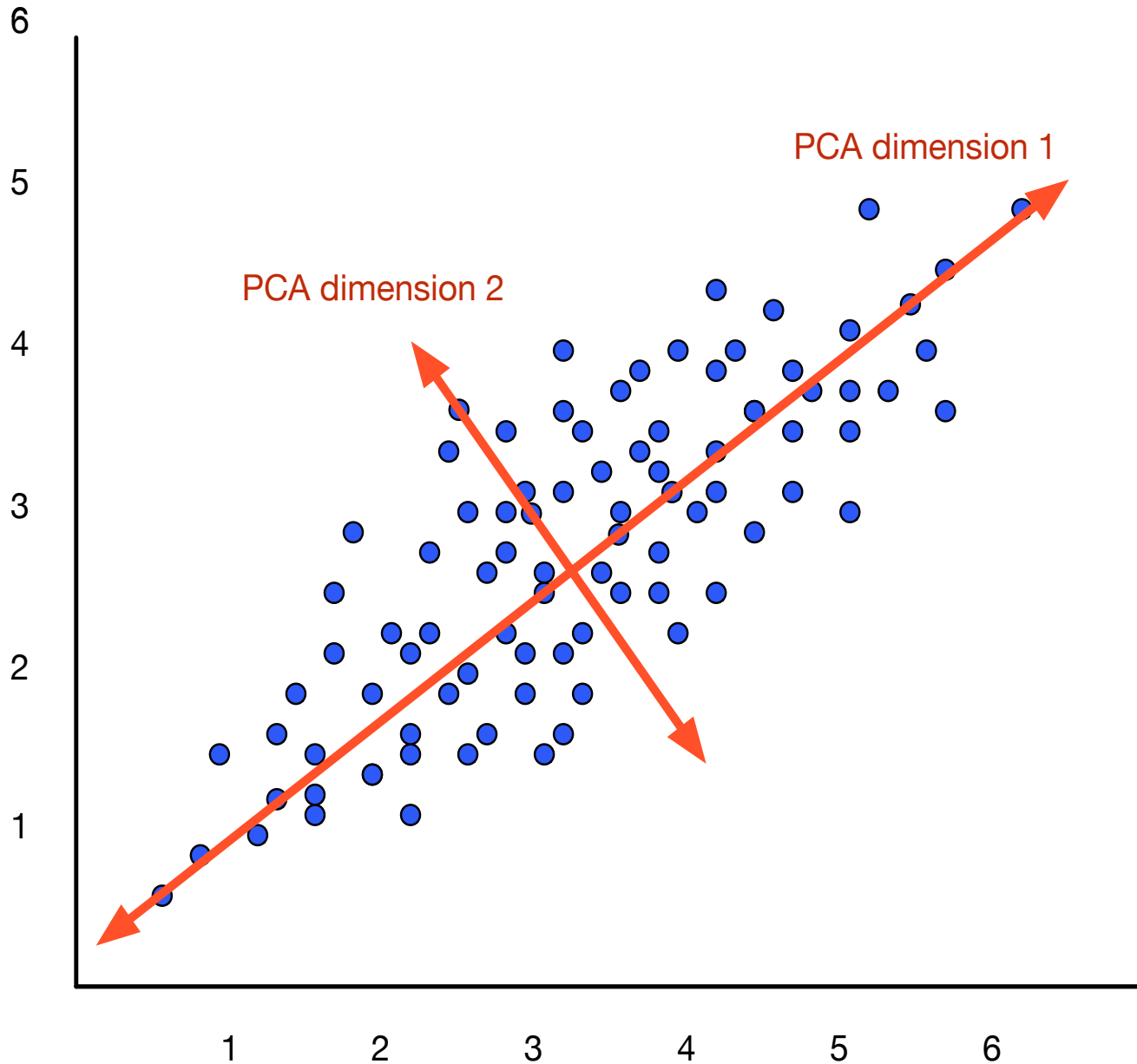
Why reduce dimensionality?

- short vectors may be easier to use as features (fewer weights to tune)
- reducing dimensionality may better handle variability in natural language due to synonymy:
 - *car* and *automobile* are synonyms, but are distinct dimensions
 - fails to capture similarity between a word with *car* as a neighbor and one with *automobile* as a neighbor

Dimensionality Reduction: Intuition

- approximate an N -dimensional dataset using fewer dimensions:
 - rotate axes into a new space
 - in which first dimension captures most variance in original dataset
- many such (related) methods:
 - principal component analysis (PCA)
 - factor analysis
 - singular value decomposition (SVD)

Dimensionality reduction



SVD embeddings versus sparse vectors

- dense SVD embeddings sometimes work better than sparse PMI vectors at tasks (like word similarity)
 - denoising: low-order dimensions may represent unimportant information
 - truncation may help the models generalize better to unseen data
 - smaller number of dimensions may make it easier for classifiers to effectively assign weights to dimensions for the task
 - dense models may do better at capturing higher order co-occurrence

Words

- types and tokens
- morphology
- distributional word vectors
- word sense and lexical semantics

Word Sense Ambiguity

- many words have multiple meanings

Word Sense Ambiguity



Word Sense Ambiguity



credit: A. Zwicky

Terminology: lemma and wordform

- **lemma**
 - words with same lemma have same stem, part of speech, rough semantics
- **wordform**
 - inflected word as it appears in text

| wordform | lemma |
|----------|--------|
| banks | bank |
| sung | sing |
| duermes | dormir |

Lemmas have senses

- one lemma *bank* can have many meanings:

sense 1: ...a *bank*₁ can hold the investments in a custodial account

sense 2: ...as agriculture burgeons on the east *bank*₂ the river will shrink even more

- **sense** (or **word sense**)
 - a discrete representation of an aspect of a word's meaning
- the lemma *bank* here has two senses

- two ways to categorize the patterns of multiple meanings of words:
 - **homonymy**: the multiple meanings are unrelated (coincidental?)
 - **polysemy**: the multiple meanings are related

Homonymy

homonyms: words that share a form but have unrelated, distinct meanings:

- *bank*₁: financial institution *bank*₂: sloping land
- *bat*₁: club for hitting a ball *bat*₂: nocturnal flying mammal

homographs: same spelling, different meanings

bank/bank, bat/bat

homophones: same pronunciation, different meanings

write/right, piece/peace

Homonymy causes problems for NLP

- information retrieval
 - query: *bat care*
- machine translation
 - *bat*: *murciélago* (animal) or *bate* (for baseball)
- text-to-speech
 - *bass* (stringed instrument) vs. *bass* (fish)

Polysemy

1: *The **bank** was constructed in 1875 out of local red brick.*

2: *I withdrew the money from the **bank**.*

- are these the same sense?
 - sense 2: “a financial institution”
 - sense 1: “the building belonging to a financial institution”
- a **polysemous** word has **related** meanings
 - most non-rare words have multiple related meanings

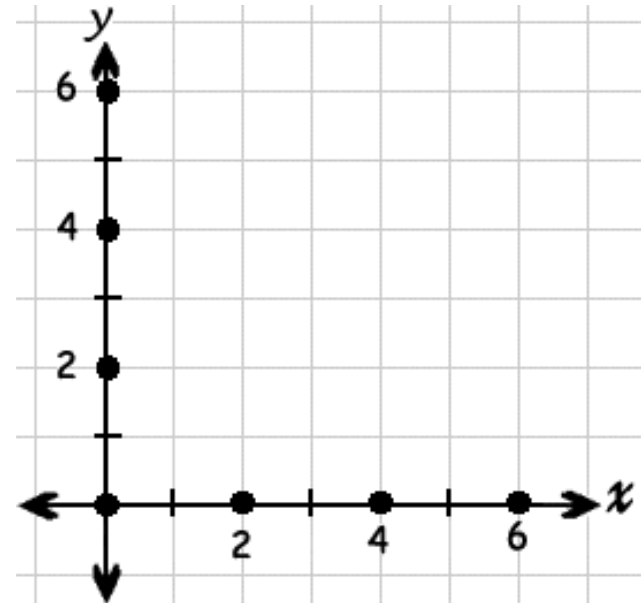
Homonymy or Polysemy?

axes

an edge tool with a heavy bladed head mounted across a handle



a fixed reference line for the measurement of coordinates



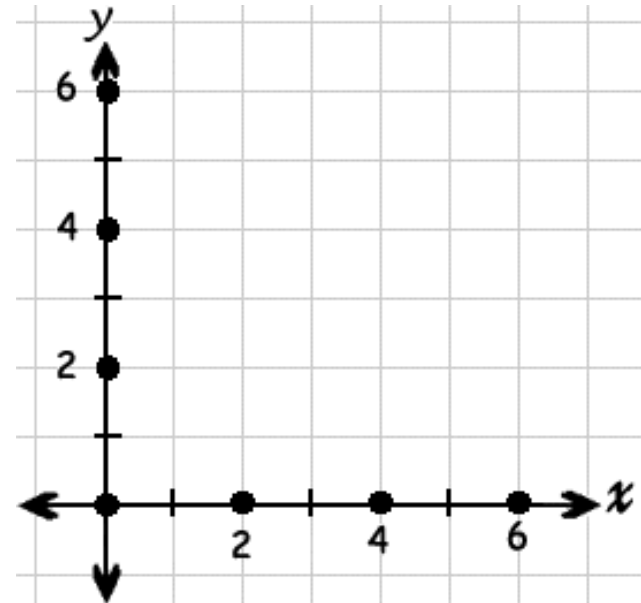
Homonymy or Polysemy?

axes

an edge tool with a heavy
bladed head mounted
across a handle



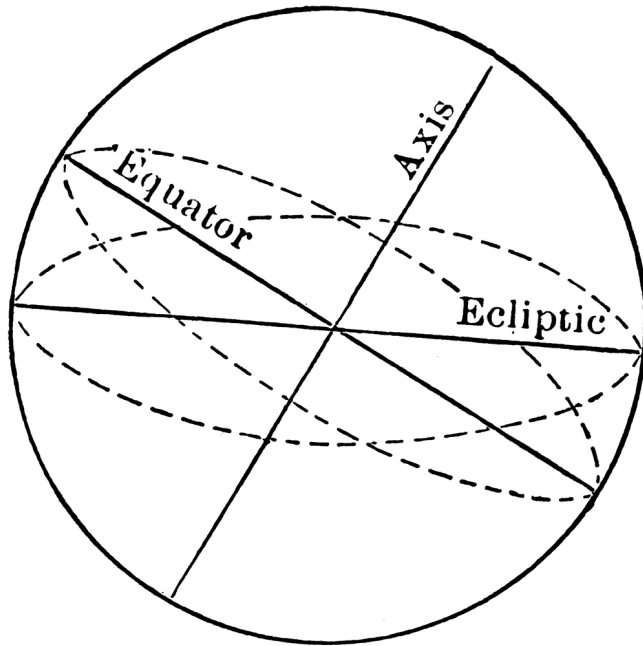
a fixed reference line for
the measurement of
coordinates



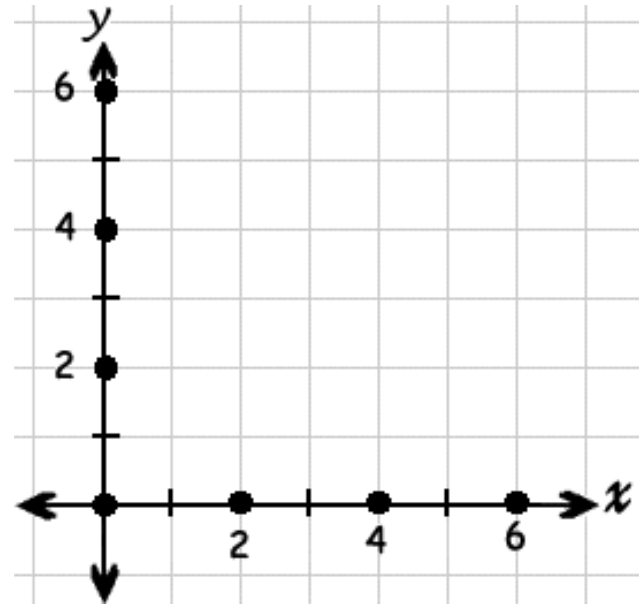
Homonymy or Polysemy?

axes

an imaginary line about which a body rotates



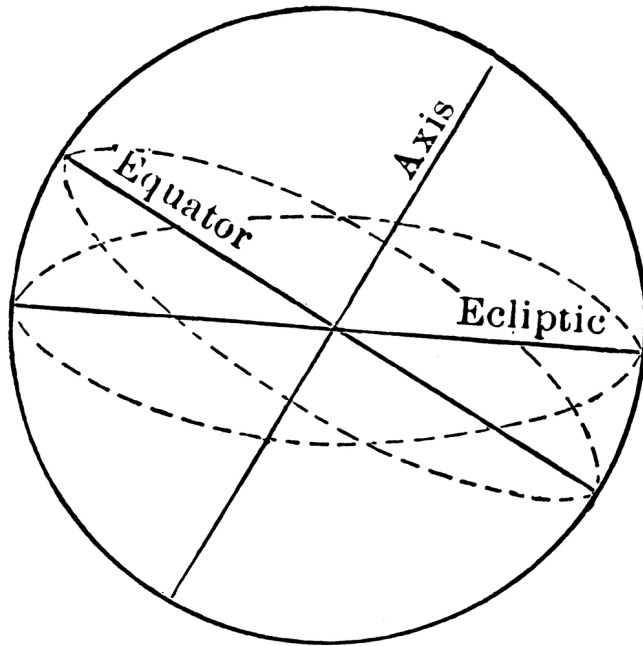
a fixed reference line for the measurement of coordinates



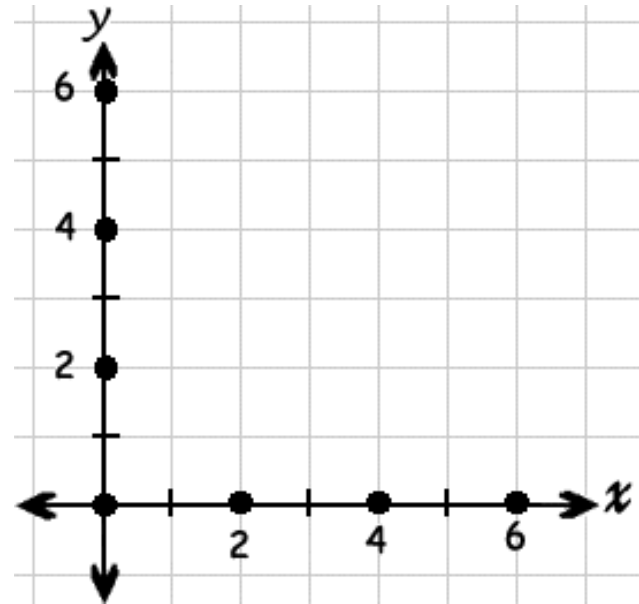
Homonymy or **Polysemy**?

axes

an imaginary line about which a body rotates



a fixed reference line for the measurement of coordinates

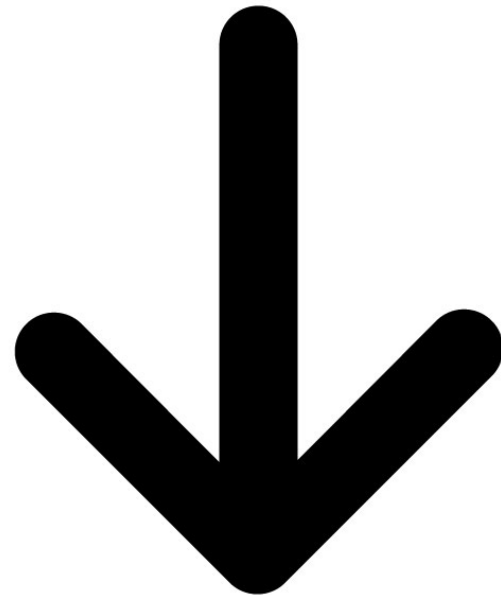
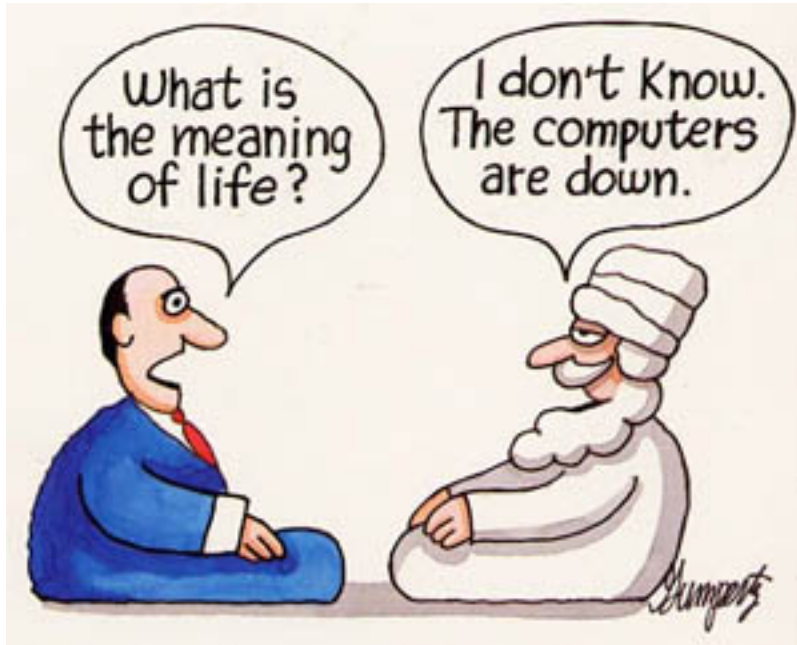


Homonymy or Polysemy?

down

in an inactive or inoperative state

being or moving lower in position or less in some value

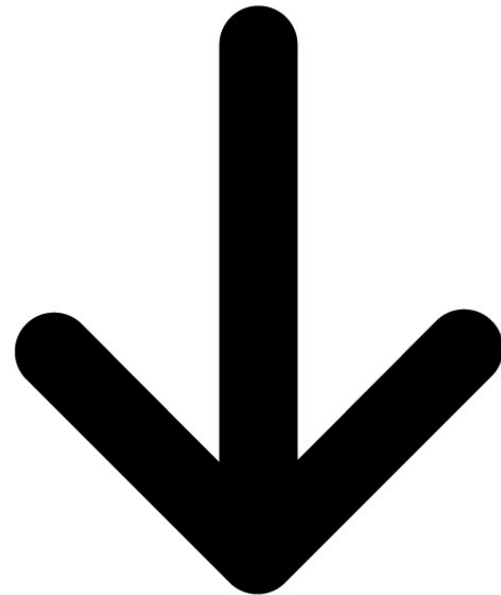
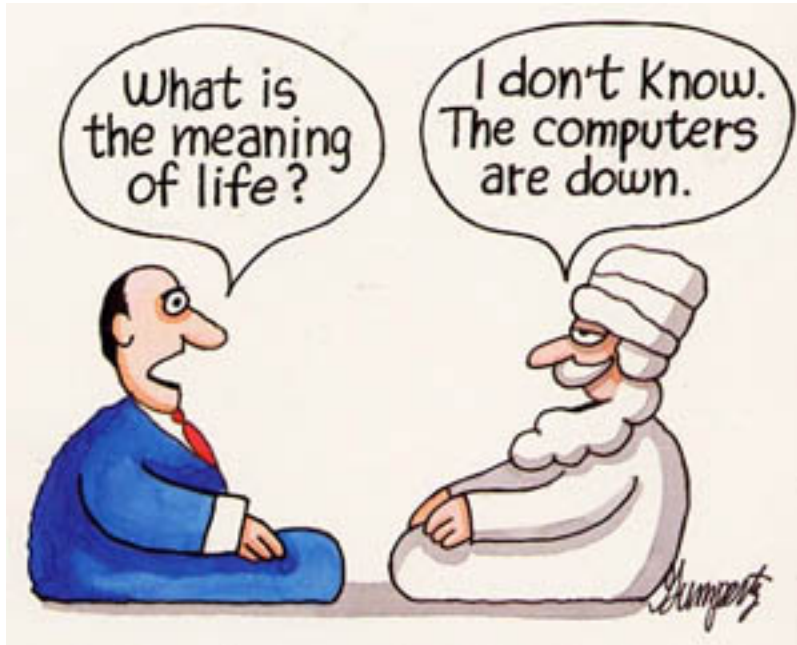


Homonymy or **Polysemy**?

down

in an inactive or inoperative state

being or moving lower in position or less in some value



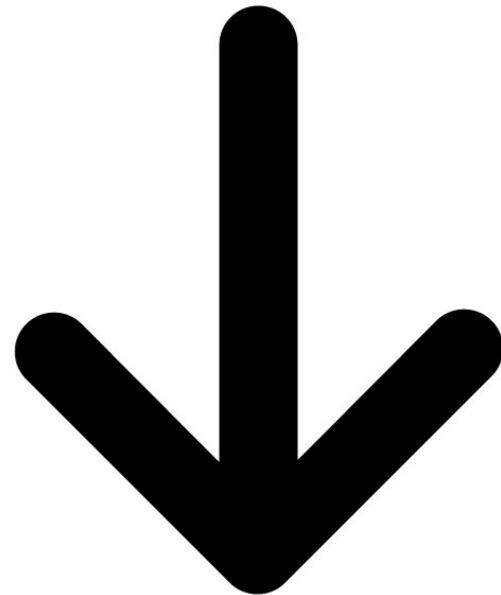
Homonymy or Polysemy?

down

soft fine feathers



being or moving lower in position or less in some value



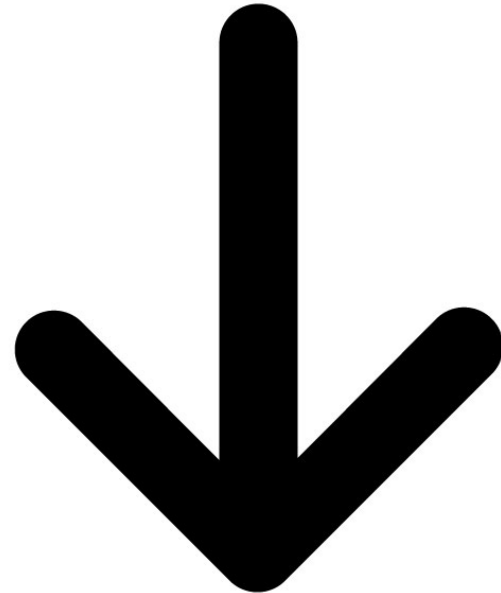
Homonymy or Polysemy?

down

soft fine feathers



being or moving lower in position or less in some value



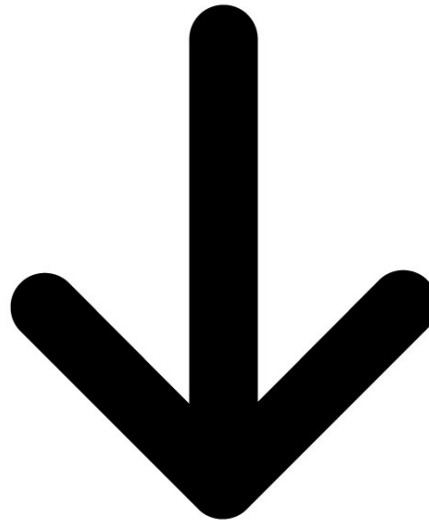
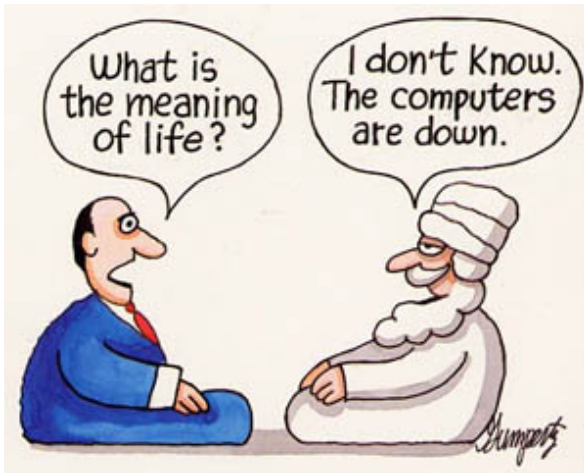
Homonymy or Polysemy?

down

in an inactive or inoperative state

being or moving lower in position or less in some value

unhappy



Metonymy or Systematic Polysemy: A systematic relationship between senses

- lots of types of polysemy are systematic
 - *school, university, hospital*
 - all can mean the institution or the building

- a systematic relationship:

– *building* ↔ *organization*

- other such kinds of systematic polysemy:

Author (Jane Austen wrote Emma) ↔ *Works of Author* (I love Jane Austen)

Tree (Plums have beautiful blossoms) ↔ *Fruit* (I ate a preserved plum)

How do we know when a word has more than one sense?

- “zeugma” test: two senses of *serve*?
 - *Which flights **serve** breakfast?*
 - *Does Lufthansa **serve** Philadelphia?*
 - *?Does Lufthansa serve breakfast and Philadelphia?*
- since this conjunction sounds weird, we say that these are **two different senses of *serve***

Synonyms

- words with same meaning in some or all contexts:
 - *filbert / hazelnut*
 - *couch / sofa*
 - *big / large*
 - *water / H₂O*
- two lexemes are synonyms if they can be substituted for each other in all situations

Synonyms

- few (or no) examples of perfect synonymy
 - even if many aspects of meaning are identical
 - still may not preserve the acceptability based on notions of politeness, slang, register, genre, etc.
- examples:
 - *water / H₂O*
 - *big / large*
 - *brave / courageous*

Synonymy is a relation between **senses** rather than words

- consider the words *big* and *large*
- are they synonyms?
 - How ***big*** is that plane?
 - Would I be flying on a ***large*** or small plane?
- how about here:
 - Miss Nelson became a kind of ***big*** sister to Benjamin.
 - ?Miss Nelson became a kind of ***large*** sister to Benjamin.
- why?
 - *big* has a sense that means being older or grown up
 - *large* lacks this sense

Antonyms

- senses that are opposites with respect to one feature of meaning
- otherwise, they are very similar!
dark/light short/long fast/slow rise/fall
hot/cold up/down in/out
- more formally, antonyms can
 - define a binary opposition or be at opposite ends of a scale
 - *long/short, fast/slow*
 - be **reversives**:
 - *rise/fall, up/down*

Hyponymy and Hypernymy

- one sense is a **hyponym** of another if the first sense is more specific, denoting a subclass of the other
 - *car* is a hyponym of *vehicle*
 - *mango* is a hyponym of *fruit*
- conversely: **hypernym** (“hyper is super”)
 - *vehicle* is a hypernym of *car*
 - *fruit* is a hypernym of *mango*

Meronymy/Holonymy

- part-whole relation
 - *wheel* is a **meronym** of *car*
 - *car* is a **holonym** of *wheel*

WordNet 3.0

- hierarchically organized lexical database
- on-line thesaurus + aspects of a dictionary
 - some languages available or under development: Arabic, Finnish, German, Portuguese...

| Category | Unique Strings |
|-----------|----------------|
| Noun | 117,798 |
| Verb | 11,529 |
| Adjective | 22,479 |
| Adverb | 4,481 |

Senses of *bass* in WordNet

Noun

- **S: (n) bass** (the lowest part of the musical range)
- **S: (n) bass, [bass part](#)** (the lowest part in polyphonic music)
- **S: (n) bass, [basso](#)** (an adult male singer with the lowest voice)
- **S: (n) [sea bass](#), bass** (the lean flesh of a saltwater fish of the family Serranidae)
- **S: (n) [freshwater bass](#), bass** (any of various North American freshwater fish with lean flesh (especially of the genus *Micropterus*))
- **S: (n) bass, [bass voice](#), [basso](#)** (the lowest adult male singing voice)
- **S: (n) bass** (the member with the lowest range of a family of musical instruments)
- **S: (n) bass** (nontechnical name for any of numerous edible marine and freshwater spiny-finned fishes)

Adjective

- **S: (adj) bass, [deep](#)** (having or denoting a low vocal or instrumental range) "*a deep voice*"; "*a bass voice is lower than a baritone voice*"; "*a bass clarinet*"

How is “sense” defined in WordNet?

- **synset (synonym set)**: set of near-synonyms; instantiates a sense or concept, with a **gloss**
- example: *chump* as a noun with gloss:
“a person who is gullible and easy to take advantage of”
- this sense of *chump* is shared by 9 words:
chump¹, fool², gull¹, mark⁹, patsy¹, fall guy¹, sucker¹, soft touch¹, mug²
- each of **these** senses have this same gloss
 - (not **every** sense; sense 2 of *gull* is the aquatic bird)

Noun

- **S: (n)** [fool](#), [sap](#), [saphead](#), [muggins](#), [tomfool](#) (a person who lacks good judgment)
- **S: (n)** [chump](#), [fool](#), [gull](#), [mark](#), [patsy](#), [fall guy](#), [sucker](#), [soft touch](#), [mug](#) (a person who is gullible and easy to take advantage of)
- **S: (n)** [jester](#), [fool](#), [motley fool](#) (a professional clown employed to entertain a king or nobleman in the Middle Ages)

ambiguity

- one form, multiple meanings → split form
 - the three senses of *fool* belong to different synsets

variability

- multiple forms, one meaning → merge forms
 - each synset contains senses of several different words

WordNet Hypernym Hierarchy for *bass*

- (n) **bass**, **basso** (an adult male singer with the lowest voice)
 - **direct hypernym** / **inherited hypernym** / **sister term**
 - **S: (n) singer, vocalist, vocalizer, vocaliser** (a person who sings)
 - **S: (n) musician, instrumentalist, player** (someone who plays a musical instrument (as a profession))
 - **S: (n) performer, performing artist** (an entertainer who performs a dramatic or musical work for an audience)
 - **S: (n) entertainer** (a person who tries to please or amuse)
 - **S: (n) person, individual, someone, somebody, mortal, soul** (a human being) "*there was too much to do for one person to do*"
 - **S: (n) organism, being** (a living thing that has (or can develop) the ability to act or function independently)
 - **S: (n) living thing, animate thing** (a living (or once living) entity)
 - **S: (n) whole, unit** (an assemblage of parts that is regarded as a single entity) "*what part compared to the whole?*"; "*the team is a unit*"
 - **S: (n) object, physical object** (a tangible and visible entity; an entity that has a shadow) "*it was full of rackets, balls and other objects*"
 - **S: (n) physical entity** (an entity that has physical existence)
 - **S: (n) entity** (that which is perceived or known or inferred to have a distinct existence (living or nonliving))

Supersenses: top level hypernyms in hierarchy

(counts from Schneider & Smith's Streusel corpus)

Noun

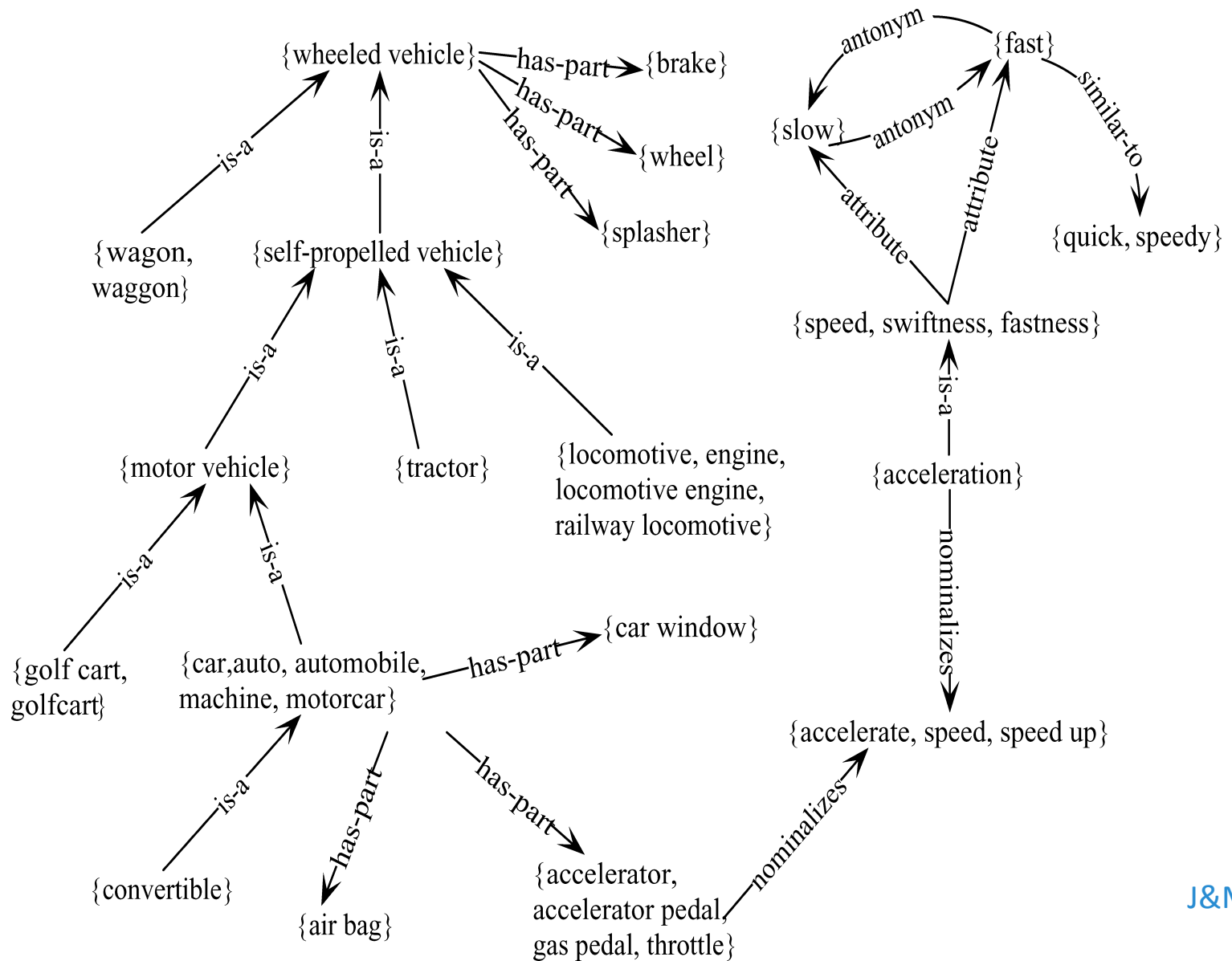
Verb

| | | |
|------------|------|-------------------|
| GROUP | 1469 | <i>place</i> |
| PERSON | 1202 | <i>people</i> |
| ARTIFACT | 971 | <i>car</i> |
| COGNITION | 771 | <i>way</i> |
| FOOD | 766 | <i>food</i> |
| ACT | 700 | <i>service</i> |
| LOCATION | 638 | <i>area</i> |
| TIME | 530 | <i>day</i> |
| EVENT | 431 | <i>experience</i> |
| COMMUNIC.* | 417 | <i>review</i> |
| POSSESSION | 339 | <i>price</i> |
| ATTRIBUTE | 205 | <i>quality</i> |
| QUANTITY | 102 | <i>amount</i> |
| ANIMAL | 88 | <i>dog</i> |

| | | |
|--------------|----|-------------------|
| BODY | 87 | <i>hair</i> |
| STATE | 56 | <i>pain</i> |
| NATURAL OBJ. | 54 | <i>flower</i> |
| RELATION | 35 | <i>portion</i> |
| SUBSTANCE | 34 | <i>oil</i> |
| FEELING | 34 | <i>discomfort</i> |
| PROCESS | 28 | <i>process</i> |
| MOTIVE | 25 | <i>reason</i> |
| PHENOMENON | 23 | <i>result</i> |
| SHAPE | 6 | <i>square</i> |
| PLANT | 5 | <i>tree</i> |
| OTHER | 2 | <i>stuff</i> |

| | | |
|-------------|------|-------------------|
| STATIVE | 2922 | <i>is</i> |
| COGNITION | 1093 | <i>know</i> |
| COMMUNIC.* | 974 | <i>recommend</i> |
| SOCIAL | 944 | <i>use</i> |
| MOTION | 602 | <i>go</i> |
| POSSESSION | 309 | <i>pay</i> |
| CHANGE | 274 | <i>fix</i> |
| EMOTION | 249 | <i>love</i> |
| PERCEPTION | 143 | <i>see</i> |
| CONSUMPTION | 93 | <i>have</i> |
| BODY | 82 | <i>get...done</i> |
| CREATION | 64 | <i>cook</i> |
| CONTACT | 46 | <i>put</i> |
| COMPETITION | 11 | <i>win</i> |
| WEATHER | 0 | — |

WordNet: Viewed as a graph





is a (hyponym/hypernym/meronym/holonym) of





is a (hyponym/hypernym/meronym/**holonym**) of



piano₁



is a

(hyponym/hypernym)

of

instrument₁



piano₁



is a

(**hyponym**/hypernym)

of

instrument₁



Word Sense Disambiguation (WSD)

- given:
 - a word in context
 - a fixed inventory of potential word senses
- decide which sense of the word this is
- why? machine translation, question answering, sentiment analysis, text-to-speech
- what set of senses?
 - English-to-Spanish machine translation: set of Spanish translations
 - text-to-speech: homographs like *bass* and *bow*
 - in general: the senses in a thesaurus like WordNet

Two Variants of WSD Task

- lexical sample task
 - small pre-selected set of target words (*line, plant, bass*)
 - inventory of senses for each word
 - supervised learning: train a classifier for **each** word
- all-words task
 - every word in an entire text
 - a lexicon with senses for each word
 - data sparseness: can't train word-specific classifiers

8 Senses of *bass* in WordNet

Noun

- **S: (n) bass** (the lowest part of the musical range)
- **S: (n) bass, bass part** (the lowest part in polyphonic music)
- **S: (n) bass, basso (an adult male singer with the lowest voice)**
- **S: (n) sea bass, bass** (the lean flesh of a saltwater fish of the family Serranidae)
- **S: (n) freshwater bass, bass** (any of various North American freshwater fish with lean flesh (especially of the genus *Micropterus*))
- **S: (n) bass, bass voice, basso** (the lowest adult male singing voice)
- **S: (n) bass** (the member with the lowest range of a family of musical instruments)
- **S: (n) bass** (nontechnical name for any of numerous edible marine and freshwater spiny-finned fishes)

Inventory of Sense Tags for *bass*

| WordNet Sense | Spanish Translation | Roget Category | Target Word in Context |
|-------------------|---------------------|----------------|---|
| bass ⁴ | lubina | FISH/INSECT | ... fish as Pacific salmon and striped bass and... |
| bass ⁴ | lubina | FISH/INSECT | ... produce filets of smoked bass or sturgeon... |
| bass ⁷ | bajo | MUSIC | ... exciting jazz bass player since Ray Brown... |
| bass ⁷ | bajo | MUSIC | ... play bass because he doesn't have to solo... |

WSD Evaluation and Baselines

- best evaluation: **extrinsic (“task-based”)**
 - embed WSD in a task and see if it helps!
- **intrinsic** evaluation often done for convenience
- strong baseline: most frequent sense

Most Frequent Sense

- WordNet senses are ordered by frequency
- most frequent is first
- sense frequencies come from *SemCor* corpus

| Freq | Synset | Gloss |
|------|--|---|
| 338 | plant ¹ , works, industrial plant | buildings for carrying on industrial labor |
| 207 | plant ² , flora, plant life | a living organism lacking the power of locomotion |
| 2 | plant ³ | something planted secretly for discovery by another |
| 0 | plant ⁴ | an actor situated in the audience whose acting is rehearsed but seems spontaneous to the audience |

Performance Ceiling

- human inter-annotator agreement
 - compare annotations of two humans on same data, given same tagging guidelines
- human agreements on all-words corpora with WordNet style senses: 75%-80%

Training Data for WSD

- **semantic concordance**: corpus in which each open-class word is labeled with a sense from a specific dictionary/thesaurus
 - SemCor: 234,000 words from Brown Corpus, manually tagged with WordNet senses
 - SENSEVAL-3 competition corpora: 2081 tagged word tokens

Features for WSD?

Features for WSD?

Intuition from Warren Weaver (1955):

“If one examines the words in a book, one at a time as through an opaque mask with a hole in it one word wide, then it is obviously impossible to determine... the meaning of the words...



But if one lengthens the slit in the opaque mask, until one can see not only the central word in question but also say N words on either side, then if N is large enough one can unambiguously decide the meaning of the central word...

‘What minimum value of N will... lead to the correct choice of meaning for the central word?’”

Example

- using a window of +/- 3 from the target:

*An **electric guitar** and **bass** player stand off to one side not really part of the scene*

Semi-Supervised Learning

problem: supervised learning requires large hand-built resources

what if you don't have much training data?

solution: bootstrapping

generalize from a very small hand-labeled seed set

Bootstrapping

- “one sense per collocation” heuristic:
 - a word reoccurring in collocation with the same word will almost surely have the same sense
- For *bass*:
 - *play* occurs with the music sense of *bass*
 - *fish* occurs with the fish sense of *bass*

Sentences extracted using *fish* and *play*

We need more good teachers – right now, there are only a half a dozen who can **play** the free **bass** with ease.

An electric guitar and **bass player** stand off to one side, not really part of the scene, just as a sort of nod to gringo expectations perhaps.

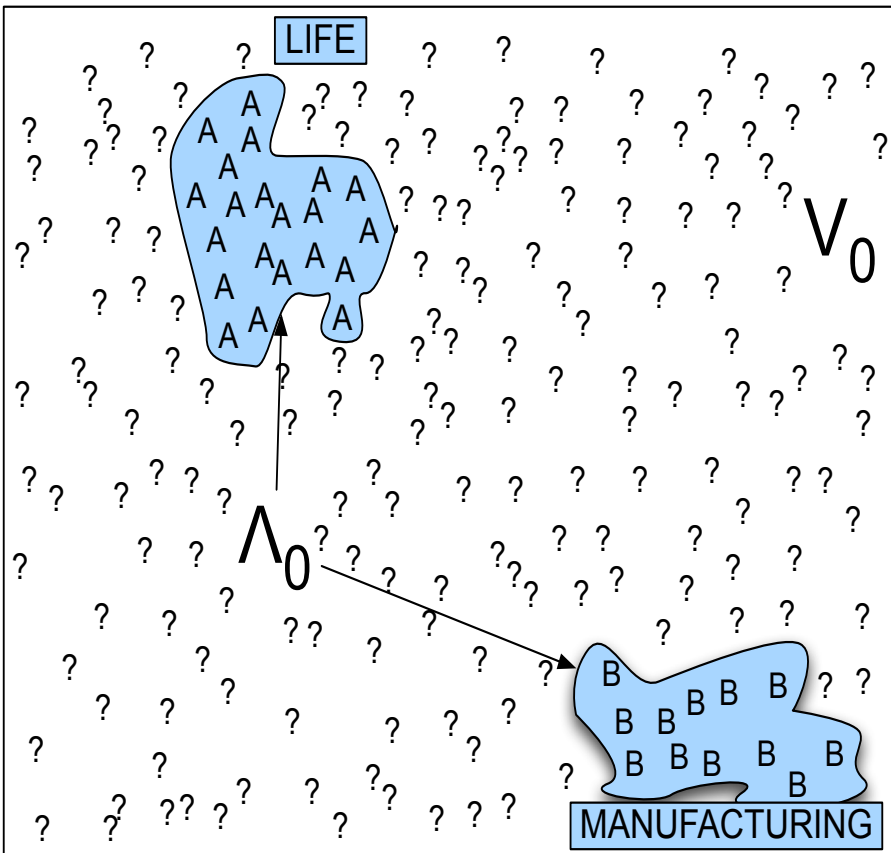
The researchers said the worms spend part of their life cycle in such **fish** as Pacific salmon and striped **bass** and Pacific rockfish or snapper.

And it all started when **fishermen** decided the striped **bass** in Lake Mead were too skinny.

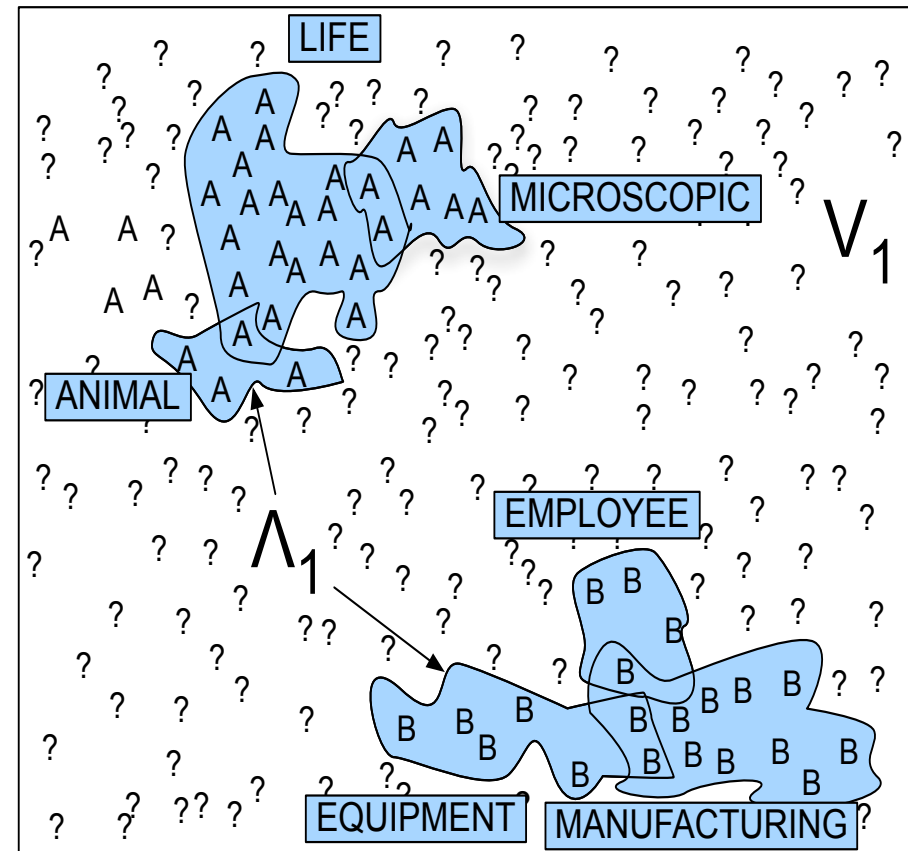
Bootstrapping

- “one sense per collocation” heuristic:
 - a word reoccurring in collocation with the same word will almost surely have the same sense
- “one sense per discourse” heuristic:
 - sense of a word is highly consistent within a document (Yarowsky, 1995)
 - especially topic-specific words

Stages in Yarowsky bootstrapping algorithm for *plant*



(a)



(b)

Exceptions

- Exceptions to one sense per collocation?
 - wedding band
- Exceptions to one sense per discourse?
 - “I’m going to rest for the rest of the day.”
 - “Last year was his last year.”
 - “Those plants generate so much pollution that no plants grow within a hundred feet.”

Summary

- word sense disambiguation: choosing correct sense in context
- applications: MT, QA, etc.
- main intuition:
 - lots of information in a word's context
 - simple algorithms based on word counts can be surprisingly good