Words, Senses, and WordNet

Bonan Min

bonanmin@gmail.com

Some slides are based on class materials from Thien Huu Nguyen

Announcement

Office hours: Tuesdays/Thursdays before class, by appointment

Location: Halligan 206 (knock on the door if it’s closed☺)

Words and Senses

Until now we have manipulated structures based on words

But if we are really interested in the meaning of sentences, we must consider the *senses* of words

- most words have several senses
- frequently several words share a common sense
- both are important for information extraction

A word sense is a representation of one aspect of a word's meaning.

Word Senses

I'm going to the bank

- bank_1 = "financial institution"
- bank_2 = "sloping mound"
- bank_3 = "biological repository"
- bank_4 = "building where a bank_1 does its business"

Word Senses

serve

Verb

- S: (v) **serve**, function (serve a purpose, role, or function) *"The tree stump serves as a table"; "The female students served as a control group"; "This table would serve very well"; "His freedom served him well"; "The table functions as a desk"*
- S: (v) **serve** (do duty or hold offices; serve in a specific function) *"He served as head of the department for three years"; "She served in Congress for two terms"*
- S: (v) **serve** (contribute or conduce to) *"The scandal served to increase his popularity"*
- S: (v) service, **serve** (be used by; as of a utility) *"The sewage plant served the neighboring communities"; "The garage served to shelter his horses"*
- S: (v) **serve**, help (help to some food; help with food or drink) *"I served him three times, and after that he helped himself"*
- S: (v) **serve**, serve up, dish out, dish up, dish (provide (usually but not necessarily food)) *"We serve meals for the homeless"; "She dished out the soup at 8 P.M."; "The entertainers served up a lively show"*
- S: (v) **serve** (devote (part of) one's life or efforts to, as of countries, institutions, or ideas) *"She served the art of music"; "He served the church"; "serve the country"*
- S: (v) **serve**, serve well (promote, benefit, or be useful or beneficial to) *"Art serves commerce"; "Their interests are served"; "The lake serves recreation"; "The President's wisdom has served the country well"*
- S: (v) **serve**, do (spend time in prison or in a labor camp) *"He did six years for embezzlement"*
- S: (v) **serve**, attend to, wait on, attend, assist (work for or be a servant to) *"May I serve you?"; "She attends the old lady in the wheelchair"; "Can you wait on our table, please?"; "Is a salesperson assisting you?"; "The minister served the King for many years"*
- S: (v) **serve**, process, swear out (deliver a warrant or summons to someone) *"He was processed by the sheriff"*
- S: (v) suffice, do, answer, **serve** (be sufficient; be adequate, either in quality or quantity) *"A few words would answer"; "This car suits my purpose well"; "Will $100 do?"; "A \`B' grade doesn't suffice to get me into medical school"; "Nothing else will serve"*
- S: (v) **serve** (do military service) *"She served in Vietnam"; "My sons never served, because they are short-sighted"*
- S: (v) **serve**, service (mate with) *"male animals serve the females for breeding purposes"*
- S: (v) **serve** (put the ball into play) *"It was Agassi's turn to serve"*

Polysemy vs Homophony

Polysemy refers to phenomenon that one and the same word acquires different, though obviously related, meanings, often with respect to particular contexts.

- The bank raised its interest rates yesterday.
- The store is next to the newly constructed bank.
- The bank appeared first in Italy in the Renaissance.

Homophony refers to cases in which two words “accidentally” have the same phonological form

- Mary walked along the bank of the river.
- HarborBank is the richest bank in the city.

Zeugma (/ˈzo͞ogmə/)

Conjunction ("yoke") of antagonistic readings; one test for whether word senses are distinct (often used intentionally to either confuse the reader or inspire them to think more deeply)

The storm sank my boat.

The storm sank my dreams.

The storm sank my boat and my dreams.

All over Ireland the farmers grew potatoes, barley, and bored.

Relationships Between Senses

Synonym

- Two senses of different words are synonyms of each other if their meaning is nearly identical
- Two words are never exactly the same in their meaning, distribution of use, dialect or other contexts in which they're licensed.
- Synonyms can be exchanged for each other without changing the truth conditions of a sentence.

couch	sofa
filbert	hazelnut
car	automobile
fair	impartial
fair	pale

Relationships Between Senses

Synonym

- Synonymy holds between word senses, not words

- How big is that plane?
- Would I be flying on a large or small plane?

- Miss Nelson, for instance, became a kind of big sister to Benjamin
- ?Miss Nelson, for instance, became a kind of large sister to Benjamin

Relationships Between Senses

Antonym (anotonymy)

- Two senses of different words are antonymous of each other if their meaning is nearly opposite

- All aspects of meaning are nearly identical between antonyms, except one (very much like synonyms in this respect)

long	short	both describe length
big	little	both describe size
fast	slow	both describe speed
cold	hot	both describe temperature
dark	light	both describe luminescence

Relationships Between Senses

Hyponymy

- Sense A is a hyponym of sense B if A is a subclass of B
- Formally, entailment: for entity x, $A(x) \Rightarrow B(x)$
- Hyponymy is generally transitive

hypo = “under”
(e.g., hypothermia)

hyponym/subordinate	hypernym/superordinate
car	vehicle
mango	fruit
chair	furniture
dog	mammal
mammal	animal

Relationships Between Senses

Meronymy

- Part-whole relations. A meronym is a part of a holonym.

meronym	holonym
leg	chair
wheel	car
car	automobile

WordNet

A large-scale database of lexical relations

Organized as graph whose nodes are synsets
(synonym sets)

- Each synset consists of 1 or more word senses which are considered synonymous
- Fine-grained senses

Primary relation: hyponym / hypernym

Available on Web

- Along with [foreign-language Wordnets]

Relations in WordNet

Relation	Also Called	Definition	Example
Hypernym	Superordinate	From concepts to superordinates	*breakfast*^{1} $\rightarrow$ *meal*^{1}
Hyponym	Subordinate	From concepts to subtypes	*meal*^{1} $\rightarrow$ *lunch*^{1}
Instance Hypernym	Instance	From instances to their concepts	*Austen*^{1} $\rightarrow$ *author*^{1}
Instance Hyponym	Has-Instance	From concepts to concept instances	*composer*^{1} $\rightarrow$ *Bach*^{1}
Member Meronym	Has-Member	From groups to their members	*faculty*^{2} $\rightarrow$ *professor*^{1}
Member Holonym	Member-Of	From members to their groups	*copilot*^{1} $\rightarrow$ *crew*^{1}
Part Meronym	Has-Part	From wholes to parts	*table*^{2} $\rightarrow$ *leg*^{3}
Part Holonym	Part-Of	From parts to wholes	*course*^{7} $\rightarrow$ *meal*^{1}
Substance Meronym		From substances to their subparts	*water*^{1} $\rightarrow$ *oxygen*^{1}
Substance Holonym		From parts of substances to wholes	*gin*^{1} $\rightarrow$ *martini*^{1}
Antonym		Semantic opposition between lemmas	*leader*^{1} $\Longleftrightarrow$ *follower*^{1}
Derivationally Related Form		Lemmas w/same morphological root	*destruction*^{1} $\Longleftrightarrow$ *destroy*^{1}

Figure 17.2 Noun relations in WordNet.

Synsets in WordNet

synset	gloss
mark, grade, score	a number or letter indicating quality
scratch, scrape, scar, mark	an indication of damage
bell ringer, bull's eye, mark, home run	something that exactly succeeds in achieving its goal
chump, fool, gull, mark, patsy, fall guy, sucker, soft touch, mug	a person who is gullible and easy to take advantage of
mark, stigma, brand, stain	a symbol of disgrace or infamy

Synsets in WordNet

- S: (n) victim, dupe (a person who is tricked or swindled)
  - S: (n) person, individual, someone, somebody, mortal, soul (a human being) *"there was too much 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) *"how big is that part compared to the whole?"; "the team is a unit"*
          - S: (n) object, physical object (a tangible and visible entity; an entity that can cast 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 its own distinct existence (living or nonliving))

Hypernyms of {chump, fool, gull, patsy, fall guy, sucker, soft touch, mug} synset

WordNet

WordNet encodes human-judged measures of similarity. Learn distributed representations of words that respect WordNet similarities (Faruqui et al. 2015)

By indexing word senses, we can build annotated resources on top of it for word sense disambiguation (WSD).

Semcor: 200K+ words from Brown corpus tagged with Wordnet senses.

original	It urged that the city take steps to remedy this problem
lemma sense	It urge^{1} that the city^{2} take^{1} step^{1} to remedy^{1} this problem^{2}
synset number	It urge$^{2:32:00}$ that the city$^{1:15:01}$ take$^{2:41:04}$ step$^{1:04:02}$ to remedy$^{2:30:00}$ this problem$^{1:10:00}$

http://web.eecs.umich.edu/~mihalcea/downloads/semcor/semcor3.0.tar.gz

“All-word” Word Sense Disambiguation

“Only$_{\text{only1}}$ a relative$_{\text{relative1}}$ handful$_{\text{handful1}}$ of such$_{\text{such0}}$ reports$_{\text{report3}}$ was received$_{\text{receive2}}$”

For all content words in a sentence, resolve each token to its sense in an fixed sense inventory (e.g., WordNet).

Methods:

- Dictionary methods (Lesk)
- Supervised (machine learning)
- Sem-supervised (boostrapping)

Dictionary Methods

Predict the sense for a given token that has the highest overlap between the token's context and sense's dictionary gloss.

The boat washed up on the river bank.

bank^1	Gloss:	a financial institution that accepts deposits and channels the money into lending activities
	Examples:	"he cashed a check at the bank", "that bank holds the mortgage on my home"
bank^2	Gloss:	sloping land (especially the slope beside a body of water)
	Examples:	"they pulled the canoe up on the bank", "he sat on the bank of the river and watched the currents"

Lesk Algorithm

function SIMPLIFIED LESK(*word, sentence*) **returns** best sense of *word*

best-sense ← most frequent sense for *word*
max-overlap ← 0
context ← set of words in *sentence*
for each *sense* **in** senses of *word* **do**
  signature ← set of words in the gloss and examples of *sense*
  overlap ← COMPUTEOVERLAP(*signature, context*)
  if *overlap* > *max-overlap* **then**
    max-overlap ← *overlap*
    best-sense ← *sense*
end
return(*best-sense*)

Extension (Basile et al. 2014): measure similarity between gloss g = $\{g_1, \ldots g_G\}$ and context c = $\{c_1, \ldots, c_C\}$ as cosine similarity between sum of distributed representations.

Supervised WSD

We have labeled training data; let’s learn from it

- Decision trees (Yarowsky 1994)
- Naive Bayes, log-linear classifiers, support vector machines (Zhong and Ng 2010)
- Bidirectional LSTM (Raganato et al. 2017)

Typical features

- Collocational: words in specific positions before/after the target word to be disambiguated (e.g., one word before and after)
- Bag-of-words: words in window around target (without encoding specific position)
- part of speech tagging, lemmatization, syntactic parsing (headwords, dependency relations)

feature
w_{i-1} = fish
w_{i-2}= fish
w_{i+1} = fish
w_{i+2} = fish
word in context = fish
…

Can we apply Naïve Bayes for this?
What’s the problem?

Supervised WSD

	Dev	Test Datasets				Concatenation of All Test Datasets				
	SE07	**SE2**	**SE3**	**SE13**	**SE15**	**Nouns**	**Verbs**	**Adj.**	**Adv.**	**All**
BLSTM	61.8	71.4	68.8	65.6	69.2	70.2	56.3	75.2	**84.4**	68.9
BLSTM + att.	62.4	71.4	**70.2**	66.4	70.8	71.0	**58.4**	75.2	83.5	69.7
BLSTM + att. + LEX	63.7	**72.0**	69.4	66.4	**72.4**	**71.6**	57.1	**75.6**	83.2	**69.9**
BLSTM + att. + LEX + POS	**64.8**	**72.0**	69.1	**66.9**	71.5	71.5	57.5	75.0	83.8	**69.9**
Seq2Seq	60.9	68.5	67.9	65.3	67.0	68.7	54.5	74.0	81.2	67.3
Seq2Seq + att.	62.9	69.9	69.6	65.6	67.7	69.5	57.2	74.5	81.8	68.4
Seq2Seq + att. + LEX	64.6	70.6	67.8	66.5	68.7	70.4	55.7	73.3	82.9	68.5
Seq2Seq + att. + LEX + POS	63.1	70.1	68.5	66.5	69.2	70.1	55.2	75.1	84.4	68.6
IMS	61.3	70.9	69.3	65.3	69.5	70.5	55.8	75.6	82.9	68.9
IMS+emb	**62.6**	**72.2**	**70.4**	65.9	71.5	**71.9**	56.6	**75.9**	**84.7**	**70.1**
Context2Vec	61.3	71.8	69.1	65.6	**71.9**	71.2	**57.4**	75.2	82.7	69.6
Lesk_{ext}+emb	⋆56.7	63.0	63.7	66.2	64.6	70.0	51.1	51.7	80.6	64.2
UKB_{gloss} w2w	42.9	63.5	55.4	⋆62.9	63.3	64.9	41.4	69.5	69.7	61.1
Babelfy	51.6	⋆67.0	63.5	**66.4**	70.3	68.9	50.7	73.2	79.8	66.4
MFS	54.5	65.6	⋆66.0	63.8	⋆67.1	67.7	49.8	73.1	80.5	65.5

Raganato et al. 2017

Supervised vs. Semi-supervised

Problem: training some classifiers (such as WSD) needs lots of labeled data

- supervised learners: all data labeled

Alternative: semi-supervised learners

- some labeled data (“seed”) + lots of unlabeled data

Bootstrapping: A Semi-supervised Learner

Basic idea of bootstrapping:

start with a small set of labeled seeds L and a large set of unlabeled examples U

repeat

train classifier C on L

apply C to U

identify examples with most confident labels; remove them from U and add them (with labels) to L

Bootstrapping WSD

Premises:

one sense per discourse (document)

one sense per collocation

Example

“bass” as fish or musical term

Example

bass catch bass	bass play bass
catch bass	play bass

Example

label initial examples

bass
fish

catch bass

bass
music

play bass

catch bass

play bass

Example

label other instances in same document

bass
fish

catch bass
fish

bass
music

play bass
music

catch bass

play bass

Example

learn collocations: catch … → fish; play … → music

bass fish catch bass fish	bass music play bass music
catch bass	play bass

Example

label other instances of collocations

bass
fish

catch bass
fish

bass
music

play bass
music

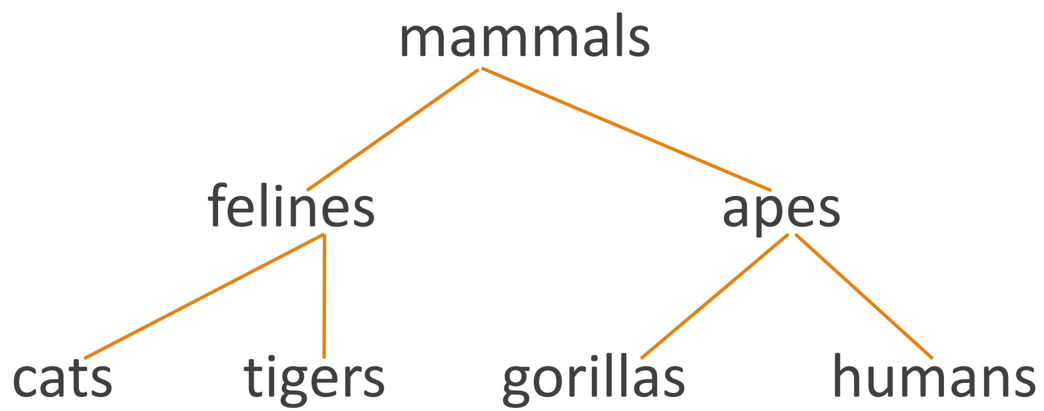
catch bass
fish

play bass
music

Using WordNet

Simplest measures of semantic similarity based on WordNet: path length:

longer path → less similar



- Variants: Wu and Palmer (1994), Leacok and Chodorow (1998)

Using WordNet

Path length ignores differences in degrees of generalization in different hyponym relations:

mammals

cats people

a cat's view of the world (cats and people are similar)

Information Content

$P(c)$ = probability that a word in a corpus is an instance of the concept (matches the synset c or one of its hyponyms) (computed based on a corpus)

Information content of a concept

$$IC(c) = -\log P(c)$$

If $LCS(c_1, c_2)$ is the *lowest common subsumer* of c_1 and c_2, the IC distance between c_1 and c_2 is $IC(c_1) + IC(c_2) - 2\, IC(LCS(c_1, c_2))$

Variants: Resnik Similarity, Jiang-Conrath Similarity

http://www.nltk.org/howto/wordnet.html