Chapter NLP:V

V. Semantics

- Semantic Phenomena
- Symbolic Semantics
- Distributional Semantics
- Compositional Semantics
- Frame Semantics

Lexical Semantics

Definition [OxfordRE Linguistics]

Lexical semantics describes the relation between meaning and form of words:

Semasiology Which meaning can be assumed from word form. Semantic relations describe difference in meaning with identical form.

Polysemy:	`mouse`	animal *vs.* input device
Specialization:	`corn`	wheat *vs.* oats
Generalization:	`moon`	of earth *vs.* any satellite
Metaphor:	`desktop`	of a desk *vs.* on screen

Onomasiology Which relations exist between the concepts.
Which forms exists for the concepts.

Lexical Relations:	`mouse` *vs.* `rodent`
Lexical Fields	
Frames	
Distributional relations	
. . .	

Lexical Semantics

Word Senses

Meaning is usually represented discretely as a word sense.

- Senses are often denominated by superscript ^{1}.
- The informal description of a sense is it's gloss,
- Senses can be modeled explixitly with their lexical relations or by componential analysis, or intrinsically via distributional semantics.

Sense description by gloss:

Sense	Gloss
Bank^{1}:	financial institution that accepts, deposits, and lends money.
Bank^{2}:	sloping land (especially the slope beside a body of water).
Red^{1}:	the color of blood or a ruby.
Blood^{1}:	the red liquid that circulates in the vains of animals.

Lexical Semantics

Lexical Relations (selection)

Polysemy	Same lexeme, different sense. The semantic relations are subtypes of polysemy.
Synonym	Different sense but similar meaning. `couch` $\longleftrightarrow$ `sofa` `big` $\longleftrightarrow$ `large`
Antonymy	Opposite meaning. `long` $\longleftrightarrow$ `short`

Hyponymy/Hypernomy	One sense is less/more specific. Also called IS-A
	`car` $\overrightarrow{\text{Hyponym}}$ `vehicle`
	`car` $\overrightarrow{\text{Hypernym}}$ `ID.3`

Meronym/Holonym	The part-whole relation.
	`wheel` $\overrightarrow{\text{Meronym}}$ `car`
	`car` $\overrightarrow{\text{Holonym}}$ `wheel`

Relations are defined over senses, not lexemes:

Synonymous	$\texttt{big}^{1}$ `plane`	$\longleftrightarrow$	$\texttt{large}^{1}$ `plane`
Not synonymous	$\texttt{big}^{2}$ `sister`	$\longleftrightarrow$	$\texttt{large}^{1}$ `sister`

Lexical Semantics

WordNet

The largest English database of word senses is WordNet. [Fellbaum, 1998]

- WordNet has entries for lemmas.
- An entry has 1 or more synsets: sets of near-synonymous **senses**. Synsets represent concepts of meaning.
- Synsets and lemmas are split by word class: Noun, Verb, Adjective, Adveb.
- Each synset has a topic (supersense), assigned from a per word class fixed set.

Entry for the lemma `bass (Noun)`:

Synset	Topic	Gloss
bass^1	attribute	the lowest part of the musical range
bass^2	animal	edible marine and freshwater spiny-finned fishes
sea bass, bass^2	food	the lean flesh of a saltwater fish of the family Serranidae
$\text{bass},^1$ bass part	communication	part the lowest part in polyphonic music
$\text{bass},^1$ basso	person	an adult male singer with the lowest voice

Lexical Semantics
WordNet

The largest English database of word senses is WordNet. [Fellbaum, 1998]

- WordNet has entries for lemmas.
- An entry has 1 or more synsets: sets of near-synonymous **senses**. Synsets represent concepts of meaning.
- Synsets and lemmas are split by word class: Noun, Verb, Adjective, Adveb.
- Each synset has a topic (supersense), assigned from a per word class fixed set.

Entry for the lemma `ride (Verb)`:

Synset	**Supersense**	**Gloss**
sit, ride	motion	sit and travel on the back of animal
sit, ride	motion	be carried or travel on or in a vehicle
tease, ride, rally, bait, rag, twit, tantalize, razz, taunt, cod	communication	harass with persistent criticism or carping
ride	stative	continue undisturbed and without interference ?Let it ride?

Remarks:

- WordNet synsets are separated by word class and do not overlap:
  - Nouns: 117,798 lemmas (avg. 1.23 senses)
  - Verbs: 11,529 lemmas (avg. 2.16 senses)
  - Adjectives: 22,479 lemmas
  - Adverbs: 4,481 lemmas
- In WordNet, semantic relations are encoded as one lexical relation (Polysemy) with 3 additional subrelations:
  - Constructional/structured polysemy: Same sense entry refers to different entitires

    `(The) Times` printed paper *vs.* the news contained in it *vs.* the organization
  - Sense extension polysemy: Derives a new synset from an old sense

    `chicken` animal *vs.* meat of animal
  - Homonymy: Same sense, very different meaning

    `bank` river bank *vs.* financial bank

Lexical Semantics

Word Sense Disambiguation

Word Sense Disambiguation (WSD) is the task of assigning to each word in a text the correct sense from a sense lexicon.

- WSD is similar to tagging, but more difficult.
  There are several classes for each word.
- To disambiguate a small set of words, classification works.
- Using the most frequent sense every time is a stong baseline. Disambiguate medical terms in lab reports
- There are several datasets (semantic concordance) where each word is annotated with its sense.
  - SEMCOR (Brown Corpus, English), SENSEVAL-3 (English), [Vossen et al., 2011] (Dutch), [Heinrich et al., 2012] (German)
  - Annotated are word class and sense id for open-class words.

Example annotations from SEMCOR:

You will find$^{9}_{v}$ that avocado$^{1}_{n}$ is$^{1}_{v}$ unlike$^{1}_{j}$ other$^{1}_{j}$ fruit$^{1}_{n}$ you have ever$^{1}_{r}$ tasted$^{2}_{v}$

Lexical Semantics

Word Sense Disambiguation: Lesk

Idea: The context of a word should overlap with the words in the gloss of its sense.

- Does not need training data.
- Easy to apply to new or low-resource languages.
- Glosses can easily be extended with (annotated) examples.

Sense	**Gloss**
bank^1	A financial institution that accepts deposits and channels the money into lending activities
bank^2	sloping land (especially the slope beside a body of water)

w_{i-4}	w_{i-3}	w_{i-2}	w_{i-1}	w_i	w_{i+1}	w_{i+2}	w_{i+3}	w_{i+4}	w_{i+5}
We	guarantee	that	your	bank	deposits	will	cover	future	costs.

Lexical Semantics

Word Sense Disambiguation: Lesk

SIMPLIFIED LESK:

For the word w_i in a sequence $w_1, \ldots, w_{i-k}, \ldots, w_i, \ldots, w_{i+k}, \ldots, w_n$ with window size k and glosses $G_{w_i} = \{g_{w_i,1}, \ldots, g_{w_i,j}\}$:

1. Remove stopwords from and lemmatize the context window $v_i := (w_{i-k}, \ldots, w_{i-1}, w_{i+1}, \ldots, w_{i+k})$ and all glosses $g_j \in G_{w_i}$.
2. Vectorize v_i and all $g_j \in G_{w_i}$.
3. Disambiguate w_i by the lowest cosine between context and gloss vectors.

- LESK can be improved using tf·idf-weighted vectors or any other (semantic) similarity measure.
- Gloss vectors can be pre-computed.

Lexical Semantics

Word Sense Disambiguation: Classification

Idea: Classify the sense with sliding-window features (cf. sequence tagging).

Example features:

1. Words (lemmas/stems) in the context window
2. Part-of-speech tags for each word in the window
3. n-grams
4. Weighted average of the word embeddings

Example:

w_{i-4}	w_{i-3}	w_{i-2}	w_{i-1}	w_i	w_{i+1}	w_{i+2}	w_{i+3}	w_{i+4}	w_{i+5}
`We`	`guarantee`	`that`	`your`	`bank`	`deposits`	`will`	`cover`	`future`	`costs.`

Features for w_i with $k = 2$:

w_{i-2}	POS_{i-2}	w_{i-1}	POS_{i-1}	w_{i+1}	POS_{i+1}	w_{i+2}	POS_{i+2}	w_{i-2}^{i-1}	w_{i+1}^{i+2}
that	IN	your	PRP	deposits	NN	will	MD	that your	deposits will

Remarks:

- The state of the art for WSD uses contextualized word embeddings.
- Word embeddings are identical for each sense of the same word form. They are not intrinsically useful for WSD.
- However, since word vector spaces embed semantic similarity, they can solve tasks like lexical substitution with a simple nearest neighbor search.
- A large language model like BERT produces contextualized word embeddings more or less as a by-product. Here, the same lexeme can have different vectors, depending on its context words. BERT solves WSD extremely well if there are vectors for each sense.
- `Word Sense Induction` tries to create lexicons like WordNet automatically by clustering the embedding space. This produces a synset collection with context vectors for each sense (the mean vector of each cluster) in an unsupervised fashion.

Lexical Semantics

Lexical Substitution

Lexical substitution tasks are subtask of WSD.

- Classic lexical subsitution looks for one or more semantically similar replacement for certain words.

  `My favorite thing about her is her straightforward honesty`

  $\rightarrow$ `My favorite thing about her is her sincere/genuine/frank honesty`

- Lexical simplification looks for a easier to understand byt semantically similar replacement.

  `John composed these verses` $\rightarrow$ `John wrote these poems`

- Lexical substitution often uses the same techiniques as WSD but does not require a lexicon.

Lexical Semantics

Multi-Word Expressions

Multi-word expressions (MWE) can function as singular lexical units.

MWE semantics can be

1. compositional,

   `driving instructor`

   `argumentation quality assessment`

2. idiomatic,

   `vice versa`

   `kick the bucket`

3. or in-between.

   `Long time no see`