

L90: Overview of Natural Language Processing

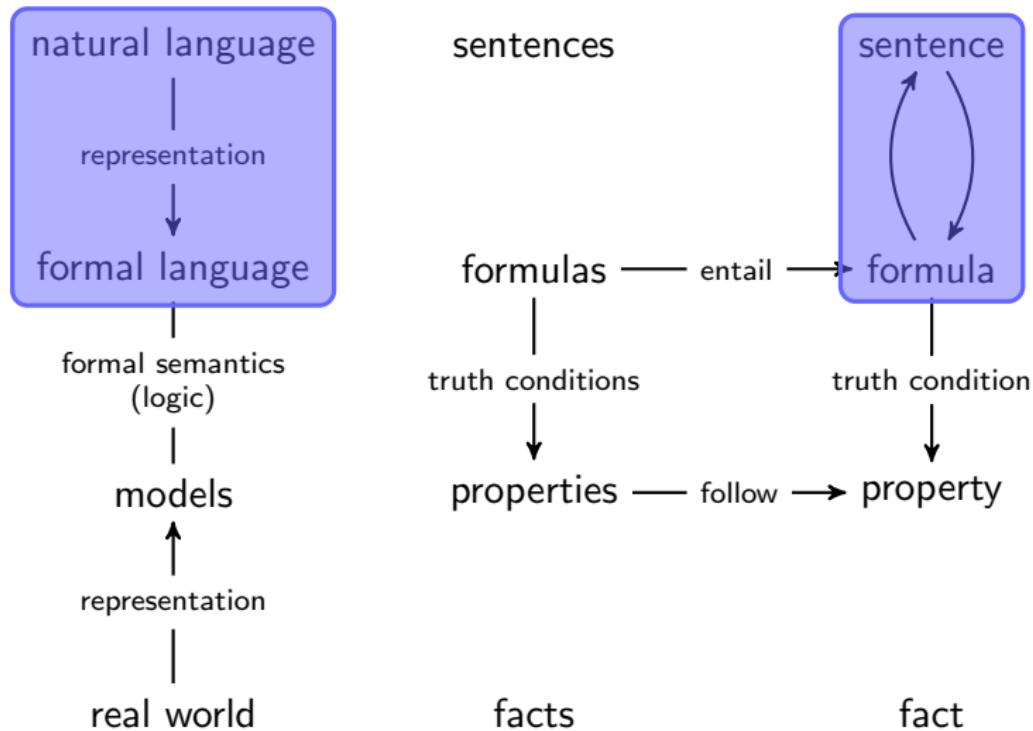
Lecture 10: Lexical Semantics

Weiwei Sun (materials mostly by Ann Copestake)

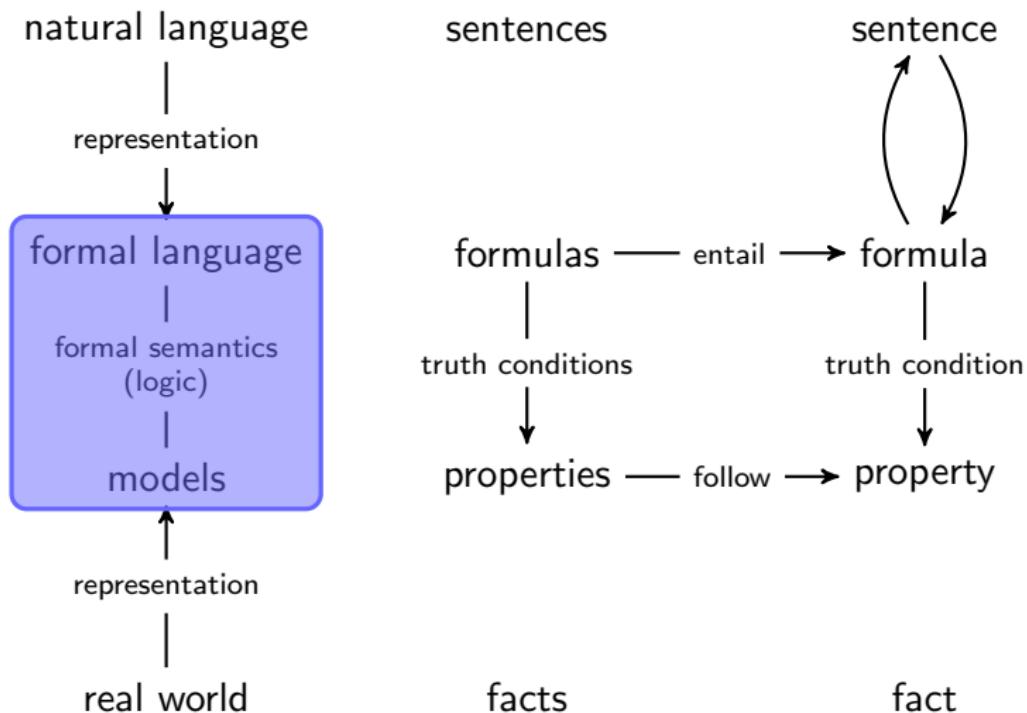
Department of Computer Science and Technology
University of Cambridge

Michaelmas 2020/21

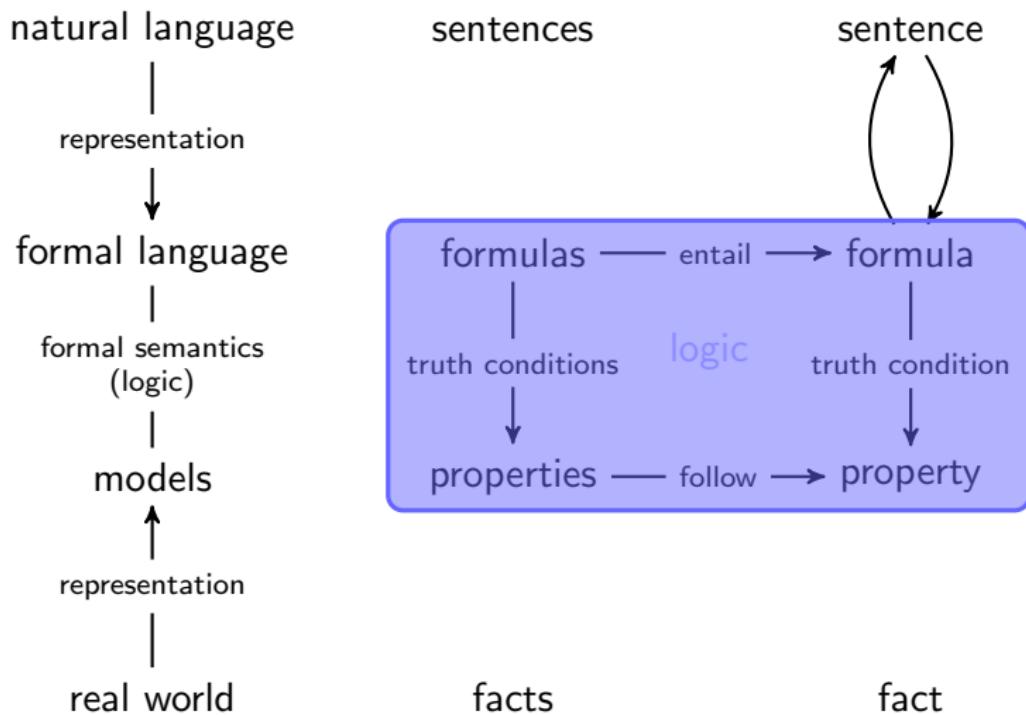
Recap: formal meaning representations



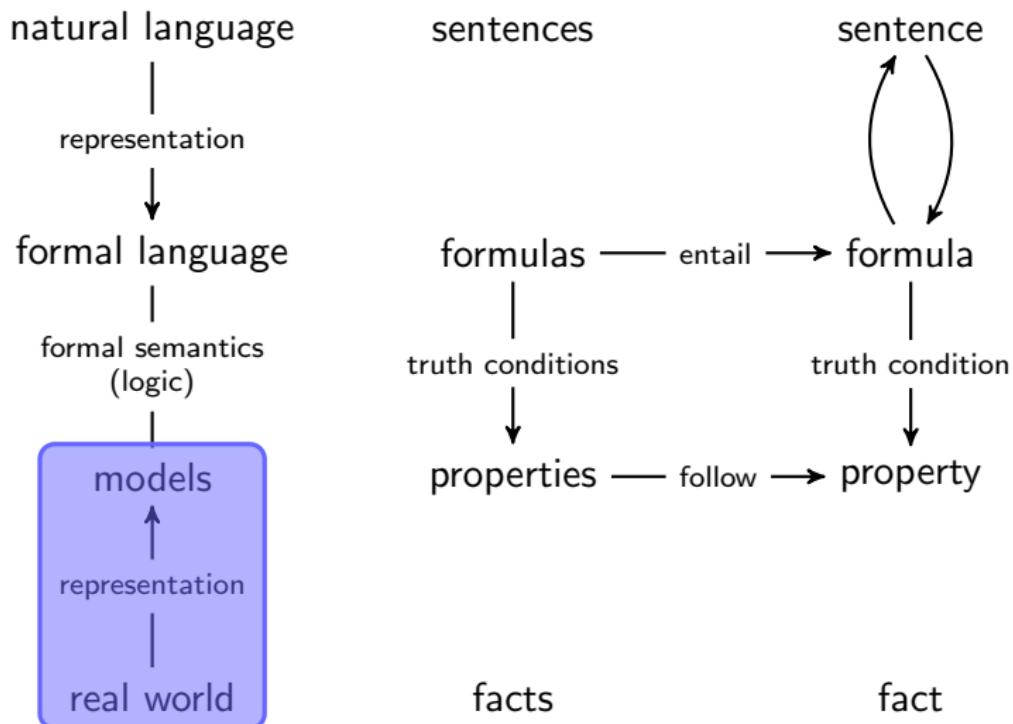
Recap: formal meaning representations



Recap: formal meaning representations



Recap: formal meaning representations





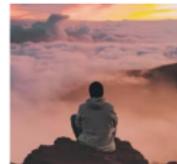
Life is Too Short...
becomingminimalist.com



Life (2017) - IMDb
imdb.com



Life (American TV series) - Wikipedia
en.wikipedia.org



9 Ways to Take Responsibility ...
thriveglobal.com



Pursue Meaning Instead of Happiness ...
medium.com



How Short Your Life REALLY Is - YouTube
m.youtube.com



Habits to Help You Balance a Creative Life
thriveglobal.com



What Will Life Be Like After the ...
knowledge.inead.edu



to Improve Your Personal Development ...
inc.com



Having a sense of meaning in life is ...
theconversation.com



The 10 Benefits of Knowing Your Purpos...
goalcast.com

search results by google

what is the meaning of life? — life'

Lecture 10: Lexical Semantics

1. Lexical semantics: semantic relations
2. Polysemy
3. Word sense disambiguation
4. Word sense induction

materials
mostly by
Ann Copestake

what makes soup, soup?

www.youtube.com/watch?v=Y1HVTNxwt7w&t=22s

What makes soup, soup? (1)

Formal semantics: **extension** — what words denote
e.g., soup': the set of all soups.



search results by google

What makes soup, soup? (1)

Formal semantics: **extension** — what words denote
e.g., soup': the set of all soups.



search results by google

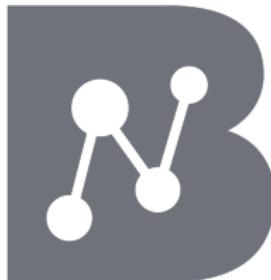
? if *a* and *b* designate the same object, there would be no difference.

? *Boris Johnson*=*Prime Minister*

A sign has both a reference and a “sense”

What makes soup, soup? (2)

- Limited domain: mapping to some knowledge base term(s).
- Knowledge base constrains possible meanings.
e.g. BabelNet (babelnet.org)



BabelNet

EN **soup** 🔊 🔊 💬

Liquid food especially of meat or fish or vegetable stock often containing pieces of solid food 🔊 WordNet

缺少定义

- Soup is a primarily liquid food, generally served warm or hot, that is made by combining ingredients such as meat and vegetables with stock, juice, water, or another liquid. 🔊 Wikipedia
- A liquidy food 🔊 Wikipedia (disambiguation)
- Primarily liquid food 🔊 Wikidata
- A cooked, liquid dish (made from meat or vegetables that are mixed with broth in a pot) that is often sold in tins. 🔊 OmegaWiki
- The liquid part of such a dish; the broth. 🔊 Wiktionary



Issues for broad coverage systems

- Boundary between lexical meaning and world knowledge.
- Representing lexical meaning.
- Acquiring representations.
- Polysemy and multiword expressions.

Approaches to lexical meaning

- Formal semantics: **extension** — what words denote
- Semantic primitives
 - e.g., *kill* means $\text{CAUSE}(\text{NOT}(\text{ALIVE}))$
- Meaning postulates:
 $\forall e_1, x, y [\text{kill}'(e_1, x, y) \rightarrow \exists e_2 [\text{cause}'(e_1, x, e_2) \wedge \text{die}'(e_2, y)]]$

Approaches to lexical meaning

- Formal semantics: **extension** — what words denote
- Semantic primitives
 - e.g., *kill* means $\text{CAUSE}(\text{NOT}(\text{ALIVE}))$
- Meaning postulates:
 $\forall e_1, x, y [\text{kill}'(e_1, x, y) \rightarrow \exists e_2 [\text{cause}'(e_1, x, e_2) \wedge \text{die}'(e_2, y)]]$
But ...

Approaches to lexical meaning

- Formal semantics: **extension** — what words denote
- Semantic primitives
 - e.g., *kill* means $\text{CAUSE}(\text{NOT}(\text{ALIVE}))$
- Meaning postulates:
 $\forall e_1, x, y [\text{kill}'(e_1, x, y) \rightarrow \exists e_2 [\text{cause}'(e_1, x, e_2) \wedge \text{die}'(e_2, y)]]$
But ...
- Distributional approaches (information vs knowledge)

Approaches to lexical meaning

- Formal semantics: **extension** — what words denote
- Semantic primitives
 - e.g., *kill* means $\text{CAUSE}(\text{NOT}(\text{ALIVE}))$
- Meaning postulates:
 $\forall e_1, x, y [\text{kill}'(e_1, x, y) \rightarrow \exists e_2 [\text{cause}'(e_1, x, e_2) \wedge \text{die}'(e_2, y)]]$
But ...
- Distributional approaches (information vs knowledge)
- Ontological relationships: informal or formal
this lecture (informal approaches).

Examples to think about

- tomato
- table
- thought
- democracy
- push
- sticky

Lexical Semantics: Semantic Relations

Taxonomic realtions

Hyponymy: IS-A

- (a sense of) *dog* is a hyponym of (a sense of) *animal*; *animal* is a *hypernym* of *dog*
- *dog* is more specific and belongs to a subclass of *animal*.
- *entailment*/IS-A: a sense *A* is a hyponym of a sense *B* if everything that is *A* is also *B*, and hence being an *A* entails being a *B*.
- hyponymy relationships form a *taxonomy*
- works best for concrete nouns

Taxonomic realtions

Hyponymy: IS-A

- (a sense of) *dog* is a hyponym of (a sense of) *animal*; *animal* is a *hypernym* of *dog*
- *dog* is more specific and belongs to a subclass of *animal*.
- *entailment*/IS-A: a sense *A* is a hyponym of a sense *B* if everything that is *A* is also *B*, and hence being an *A* entails being a *B*.
- hyponymy relationships form a *taxonomy*
- works best for concrete nouns

Some issues concerning hyponymy

- not useful for all words: *thought*, *democracy*, *push*, *sticky*?
- individuation differences: is *table* a hyponym of *furniture*?
- multiple inheritance: e.g., is *coin* a hyponym of both *metal* and *money*?
- what does the top of the hierarchy look like?

Other semantic relations

Classical relations

- **Meronymy:** PART-OF e.g., *arm* is a meronym of *body*, *steering wheel* is a meronym of *car* (piece vs part)
- **Synonymy** e.g., *aubergine/eggplant*.
- **Antonymy** e.g., *big/little*
- **Near-synonymy/similarity** e.g., *exciting/thrilling*
e.g., *slim/slender/thin/skinny*

The word *synonym* is commonly used to describe a relationship of approximate or rough synonymy.

- *craft, skill*
- *apple, fruit*

WordNet

wordnet.princeton.edu

- large-scale, open source resource for English
- wordnets being built for other languages, e.g. Open Multilingual Wordnet (compling.hss.ntu.edu.sg/omw)
- hand-constructed
- organized into *synsets*: synonym sets (near-synonyms)

- large-scale, open source resource for English
- wordnets being built for other languages, e.g. Open Multilingual Wordnet (compling.hss.ntu.edu.sg/omw)
- hand-constructed
- organized into *synsets*: synonym sets (near-synonyms)

Overview of adj red

S: (adj) **red**, reddish, ruddy, blood-red, carmine, cerise, cherry, cherry-red, crimson, ruby, ruby-red, scarlet (of a color at the end of the color spectrum (next to orange); resembling the color of blood or cherries or tomatoes or rubies)

- similar to
 - S: (adj) chromatic (being or having or characterized by hue)
- derivationally related form
- antonym
 - W: (adj) achromatic [Indirect via chromatic] (having no hue) “neutral colors like black or white”

WordNet labels each synset with a lexicographic category/*supersenses*.

Category	Example	Category	Example	Category	Example
ACT	<i>service</i>	GROUP	<i>place</i>	PLANT	<i>tree</i>
ANIMAL	<i>dog</i>	LOCATION	<i>area</i>	POSSESSION	<i>price</i>
ARTIFACT	<i>car</i>	MOTIVE	<i>reason</i>	PROCESS	<i>process</i>
ATTRIBUTE	<i>quality</i>	NATURAL EVENT	<i>experience</i>	QUANTITY	<i>amount</i>
BODY	<i>hair</i>	NATURAL OBJECT	<i>flower</i>	RELATION	<i>portion</i>
COGNITION	<i>way</i>	OTHER	<i>stuff</i>	SHAPE	<i>square</i>
COMMUNICATION	<i>review</i>	PERSON	<i>people</i>	STATE	<i>pain</i>
FEELING	<i>discomfort</i>	PHENOMENON	<i>result</i>	SUBSTANCE	<i>oil</i>
FOOD	<i>food</i>			TIME	<i>day</i>

Hyponymy in WordNet

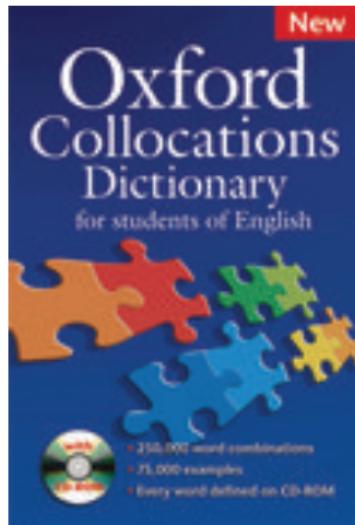
search wordnet

Using hyponymy

- Semantic classification: e.g., for named entity recognition.
e.g., *JJ Thomson Avenue* is a place.
- RTE style inference: *find / discover*
- Word sense disambiguation
- Query expansion in search

Collocation

- two or more words that occur together more often than expected by chance (informal description — there are others)
- some collocations are multi-word expressions



Polysemy

Polysemy

homonymy: unrelated word senses.

bank (raised land) vs *bank* (financial institution)

bank (financial institution) vs *bank* (the building belonging to a financial institution):

(1) The bank is on the corner of Nassau and Witherspoon.

BUILDING ↔ ORGANIZATION

bank (N) (raised land) vs bank (V) (to create some raised land): *regular polysemy*. Compare *pile*, *heap* etc

Related but distinct senses

- No clearcut distinctions.
- Dictionaries are not consistent.

Word Sense Disambiguation

Word sense disambiguation

- selecting the correct sense for a word in a context.
- challenges: inventory of potential word senses, datasets
- needed for many applications, problematic for large domains.
- *lexical sample task*: to disambiguate a small pre-selected set of words.
simple supervised classification approaches work very well.
- *all-words task*: to disambiguate every word in the text. similar to part-of-speech tagging.
- SemCor: a subset of the Brown Corpus; over 226,036 words; manually tagged with WordNet senses; all-words task

Assumes that we have a standard set of word senses (e.g., WordNet)

Sense	Supersense	Target Word in Context
bass ⁴	FOOD	... fish as Pacific salmon and striped <i>bass</i> and...
bass ⁷	ARTIFACT	... play <i>bass</i> because he doesn't have to solo...

Aspects of WSD

- baseline: *most frequent sense*
frequency: e.g., *diet*: the food sense (or senses) is much more frequent than the parliament sense (*Diet of Worms*)
- *one sense per discourse*: a word appearing multiple times in a discourse often appears with the same sense.
- collocations: e.g. *striped bass* (the fish) vs *bass guitar*: syntactically related or in a window of words (latter sometimes called 'cooccurrence').
Generally 'one sense per collocation'.
- selectional restrictions/preferences (e.g., *Kim eats bass*, must refer to fish)

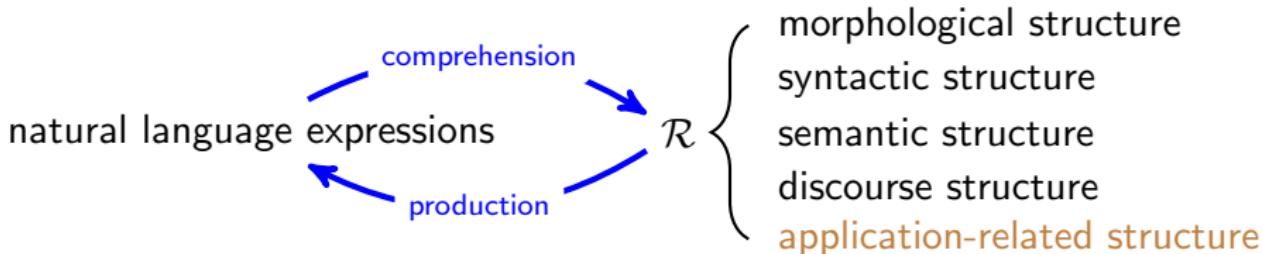
WSD techniques

- supervised learning: cf. POS tagging from lecture 3 with neural encoders and classifiers (lecture 6–8).
- sense-tagged corpora are difficult to construct; algorithms need far more data than POS tagging
- contextual embeddings + nearest-neighbor $v_s = \frac{1}{n} \sum_i c_i$
- Feature-based algorithms for WSD are extremely simple and function almost as well as contextual language model algorithms.
 - part-of-speech tags
 - collocation features of words or N -grams
 - weighted average of embeddings of all words in a window

WSD techniques

- supervised learning: cf. POS tagging from lecture 3 with neural encoders and classifiers (lecture 6–8).
- sense-tagged corpora are difficult to construct; algorithms need far more data than POS tagging
- contextual embeddings + nearest-neighbor $v_s = \frac{1}{n} \sum_i c_i$
- Feature-based algorithms for WSD are extremely simple and function almost as well as contextual language model algorithms.
 - part-of-speech tags
 - collocation features of words or N -grams
 - weighted average of embeddings of all words in a window
- *unsupervised* learning: learn from $\{x^{(1)}, x^{(2)}, \dots, x^{(m)}\}$
 - machine readable dictionaries (MRDs): e.g., look at overlap with words in definitions and example sentences
 - selectional preferences: don't work very well by themselves, useful in combination with other techniques

Word Sense Induction



How can we get proper representations?

word sense induction automatically create the set of “senses” of each word.

grammar induction automatically create the tree of each sentence.

semantic role induction automatically create predicate–argument links of words.

Clustering

- For each token w_i of word w in a corpus, compute a context vector c .
- Use a clustering algorithm to cluster these word-token context vectors c into a predefined number of groups or clusters. Each cluster defines a sense of w .
- Compute the vector centroid of each cluster. Each vector centroid s_j is a sense vector representing that sense of w .

Lexical meaning: what doesn't work

- meaning of tomato is tomato' or TOMATO
- meaning postulates
- dictionary definition: good dictionary definition allows reader with some familiarity with a concept to identify it

tomato: mildly acid red or yellow pulpy fruit eaten as a vegetable

Unanswered questions

- how far does distributional semantics get us?
- grounding often claimed for systems combining vision and language: is this enough?
- are virtual worlds a possible basis for grounding?

Readings

- Ann's notes
- D Jurafsky and J Martin. *Speech and Language Processing*
Chapter 19. web.stanford.edu/~jurafsky/slp3/19.pdf