SI 630

Natural Language Processing: Algorithms and People

Lecture 12: Semantics April 7, 2021

Some slides adapted from David Bamman, Dan Jurafsky, and Ray Mooney

Semantics

Word representation

…	…
the	0
a	0
an	0
for	0
in	0
on	0
dog	1
cat	0
…	…

dog is a point in V-dimensional space

dog

4.1
-0.9

dog is a point in 2-dimensional space

Polysemy

the movie is so bad , in fact , that it retains that ridiculous tarzan call that was so tirelessly mocked in last summer's comedy " george of the jungle .

一词多义

the performances are bad .

Polysemy 一词多义 一词多词性

	Sense 1	Sense 2
bad	Of poor quality or little worth.	slang (orig. U.S.). Formidable, good.
sick	Suffering from illness of any kind; ill, unwell, ailing.	slang (now esp. Skateboarding and Surfing). Excellent, impressive; risky.
awesome	Inspiring awe; appalling, dreadful, weird.	Expressing enthusiastic approval: great, excellent, highly impressive; fantastic.
wicked	Bad in moral character, disposition, or conduct	Excellent, splendid; remarkable. slang (orig. U.S.).

Oxford English Dictionary

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

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

“line”

Noun

- S: (n) **line** (a formation of people or things one beside another) *"the line of soldiers advanced with their bayonets fixed"; "they were arrayed in line of battle"; "the cast stood in line for the curtain call"*
- S: (n) **line** (a mark that is long relative to its width) *"He drew a line on the chart"*
- S: (n) **line** (a formation of people or things one behind another) *"the line stretched clear around the corner"; "you must wait in a long line at the checkout counter"*
- S: (n) **line** (a length (straight or curved) without breadth or thickness; the trace of a moving point)
- S: (n) **line** (text consisting of a row of words written across a page or computer screen) *"the letter consisted of three short lines"; "there are six lines in every stanza"*
- S: (n) **line** (a single frequency (or very narrow band) of radiation in a spectrum)
- S: (n) **line** (a fortified position (especially one marking the most forward position of troops)) *"they attacked the enemy's line"*
- S: (n) argumentation, logical argument, argument, line of reasoning, **line** (a course of reasoning aimed at demonstrating a truth or falsehood; the methodical process of logical reasoning) *"I can't follow your line of reasoning"*
- S: (n) cable, **line**, transmission line (a conductor for transmitting electrical or optical signals or electric power)
- S: (n) course, **line** (a connected series of events or actions or developments) *"the government took a firm course"; "historians can only point out those lines for which evidence is available"*
- S: (n) **line** (a spatial location defined by a real or imaginary unidimensional extent)
- S: (n) wrinkle, furrow, crease, crinkle, seam, **line** (a slight depression or fold in the smoothness of a surface) *"his face has many lines"; "ironing gets rid o most wrinkles"*
- S: (n) pipeline, **line** (a pipe used to transport liquids or gases) *"a pipeline runs from the wells to the seaport"*
- S: (n) **line**, railway line, rail line (the road consisting of railroad track and roadbed)
- S: (n) telephone line, phone line, telephone circuit, subscriber line, **line** (a

Word senses

- They rarely serve red meat
- He served as U.S. ambassador to Norway
- He might have served his time.

Jurafsky & Martin (2008)

Zeugma

- Conjunction ("yoke") of antagonistic readings; one test for whether word senses are distinct.

- Which flights serve breakfast?
- Does Midwest Express serve Philadelphia?

同一句句子中serve有两个不同的含义

- Does Midwest Express serve breakfast and Philadelphia?

Aside: Zeugma is often a key component of humor!

可以代表语气等隐状态

- I like my X like I like my Y: Z
- A successful joke often requires yoking Z with two divergent meanings that are each applicable
  - rich
  - thick
  - hung
  - bitter
  - bound in leather
  - capable of making my heart race

- S: (adj) **rich** (possessing material wealth) *"her father is extremely rich"; "many fond hopes are pinned on rich uncles"*
- S: (adj) **rich** (having an abundant supply of desirable qualities or substances (especially natural resources)) *"blessed with a land rich in minerals"; "rich in ideas"; "rich with cultural interest"*
- S: (adj) **rich** (of great worth or quality) *"a rich collection of antiques"*
- S: (adj) fat, fertile, productive, **rich** (marked by great fruitfulness) *"fertile farmland"; "a fat land"; "a productive vineyard"; "rich soil"*
- S: (adj) deep, **rich** (strong; intense) *"deep purple"; "a rich red"*
- S: (adj) **rich** (very productive) *"rich seams of coal"*
- S: (adj) **rich** (high in mineral content; having a high proportion of fuel to air) *"a rich vein of copper"; "a rich gas mixture"*
- S: (adj) **rich** (suggestive of or characterized by great expense) *"a rich display"*
- S: (adj) **rich** (containing plenty of fat, or eggs, or sugar) *"rich desserts"; "they kept gorging on rich foods"*
- S: (adj) full-bodied, racy, **rich**, robust (marked by richness and fullness of flavor) *"a rich ruby port"; "full-bodied wines"; "a robust claret"; "the robust flavor of fresh-brewed coffee"*
- S: (adj) **rich** (pleasantly full and mellow) *"a rich tenor voice"*
- S: (adj) ample, copious, plenteous, plentiful, **rich** (affording an abundant supply) *"had ample food for the party"; "copious provisions"; "food is plentiful"; "a plenteous grape harvest"; "a rich supply"*

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

Synonymy

Synonymy holds between word senses, not words

big和large只有在特定语境下才是近义词

- 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

Antonymy

在特定语义下是反义词

- Two senses of different words are synonyms 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

Hyponymy

hypo = “under”
(e.g., hypothermia)

- Sense A is a hyponym of sense B if A is a subclass of B

- Formally, entailment: for entity x, A(x) ⇒ B(x)

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

transitive

Meronymy

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

部分-整体关系

meronym	holonym
leg	chair
wheel	car
car	automobile

WordNet

- Lexical database for nouns, verbs and adjectives/ adverbs.

- Each word sense is arranged in a synset (category of near-synonyms) and each synset is related to others in terms of their sense relations.

Relations

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

Figure 17.2 Noun relations in WordNet.

Synsets

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

- 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, mark, patsy, fall guy, sucker, soft touch, mug} sunset

WordNet + Distributional Semantics

- 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.

Semcor

- Semcor: 200K+ words from Brown corpus tagged with Wordnet senses.
    - http://web.eecs.umich.edu/~mihalcea/downloads/semcor/semcor3.0.tar.gz

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$^{1:10:00}$

Word Sense Disambiguation

can be preprocess of word2vec，each sense is embedded with a vector

“All-word” WSD

“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).

Word Sense Disambiguation (WSD)

- Dictionary methods (Lesk)
- Supervised (machine learning)
- Semi-supervised (Bootstrapping)

Dictionary methods

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

看单词每个sense和周围context的重叠率

Dictionary methods

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"

"The boat washed up on the the river bank."

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)
```

Lesk Algorithm

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

$$\cos\left(\sum_{i=1}^{G} g_i, \sum_{i=1}^{C} c_i\right)$$

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)

Supervised WSD

- Collocational: words in specific positions before/after the target word to be disambiguation
- Bag-of-words: words in window around target (without encoding specific position)

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

Supervised learning

"I got money from the bank to buy a bass"

- Pre-processing: part of speech tagging, lemmatization, syntactic parsing (headwords, dependency relations)

- Collocations:
  - Token 1 word to the left, 1 word to the right
- All words within window of n tokens

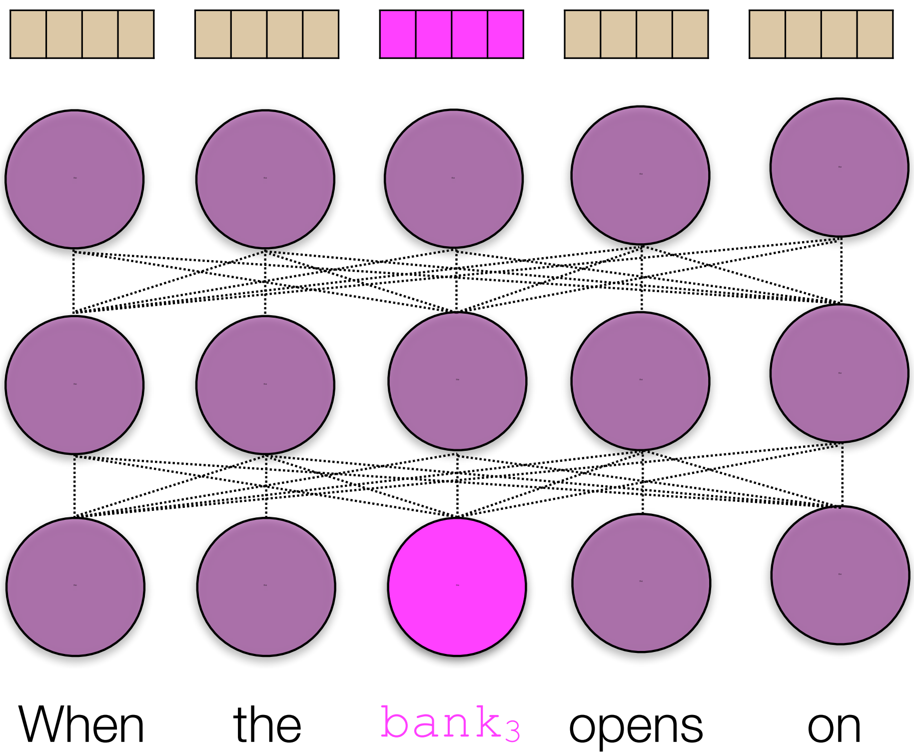
BLSTM提升有限

	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

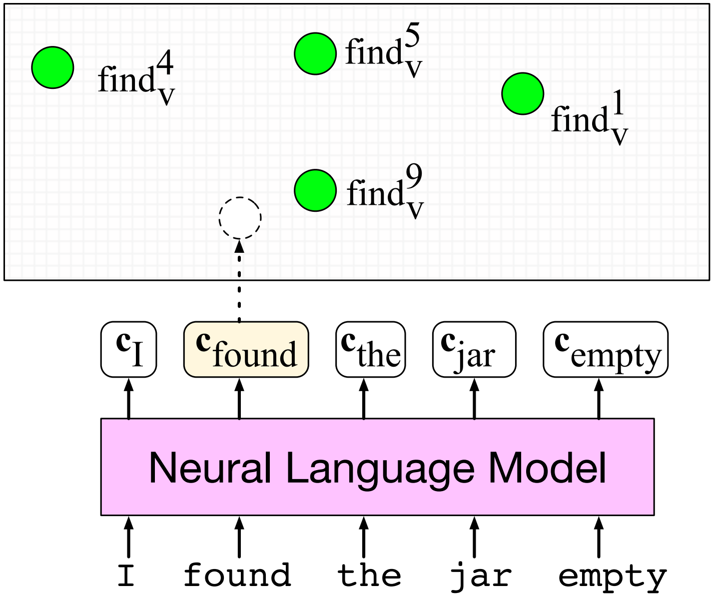
WSD with Neural Language Models

Training: Use contextual embeddings on sense-tagged data to learn what each sense "looks like"



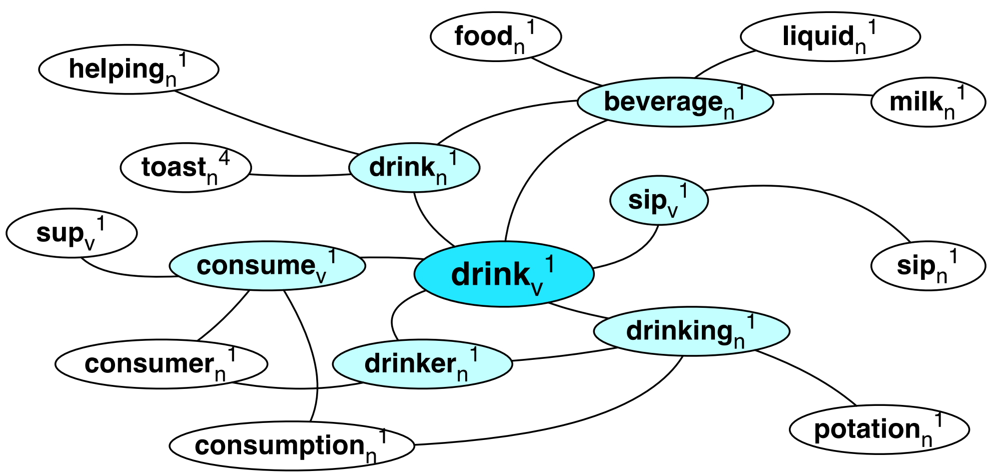
WSD with Neural Language Models

During testing, generate a contextual embedding for the target word and choose the nearest neighbor of the sense-tagged contextual embeddings from training



Graph-based WSD methods

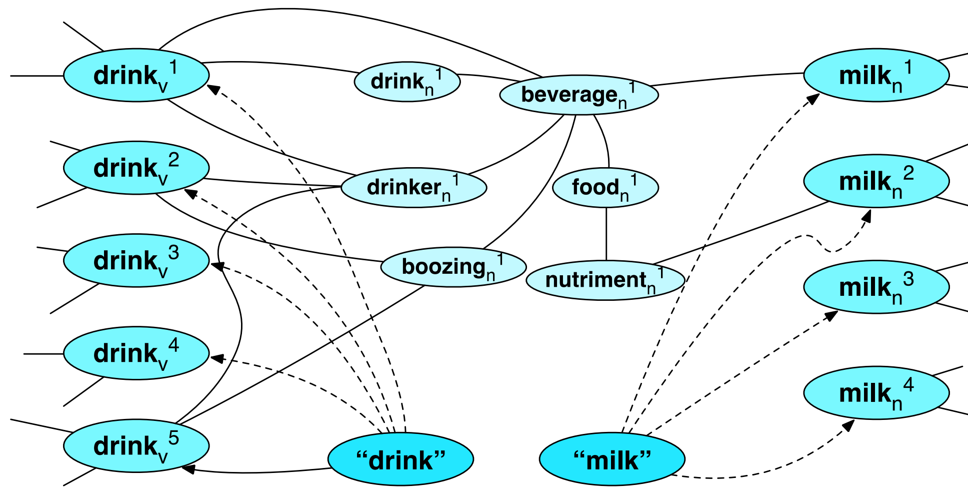
- WordNet can be viewed as a knowledge graph
  - senses are nodes
  - relations (hypernymy, meronymy) are edges
  - Also add edge between word and unambiguous gloss words



How to use the graph for WSD

- Insert target word and words in its sentential context into the graph, with directed edges to their senses

- "She drank some milk"
- Now choose the most central sense
- Add some probability to "drink" and "milk" and compute node with highest PageRank



One sense per discourse

- If a word appears multiple times in a document, it's usually with the same sense. (Gale et al. 1992)
    - Articles about financial banks don't use talk about river banks.

Semi-supervised WSD

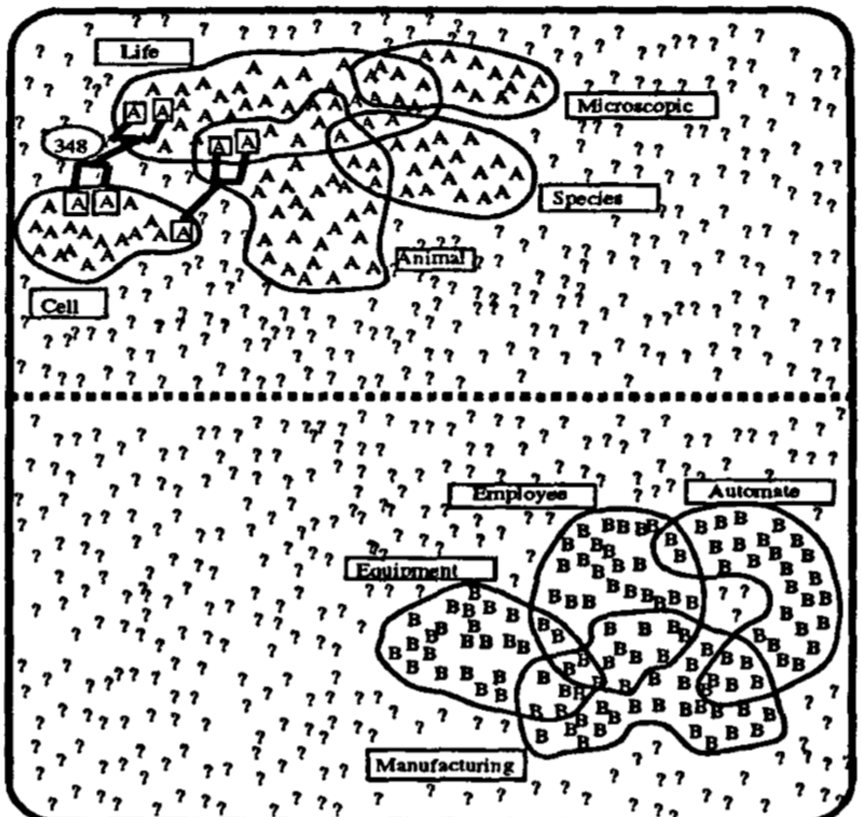
1. Produce seeds (dictionary definitions, single defining collocate, or label common collocates)

2. Repeat until convergence:

    1. Train supervised classifier on labeled examples

    2. Label all examples, and keep labels for high-confidence instances

Semi-supervised WSD



"Plant"

A = SENSE-A training example
B = SENSE-B training example
? = currently unclassified training example
Life = Set of training examples containing the collocation "life".

Evaluation

- Annotated data; cross-validation.
  - Semcor
  - Ontonotes
- Semeval/Senseval competitions

Hyponymy

	… Entity
… Entity	Artifact
Animal	Instrumentality
Vertebrate	Conveyance
Mammal	Wheeled vehicle
Ungulate	Self-propelled vehicle
Equine	Motor vehicle
Horse	Car

NOUNS			
SUPERSENSE	NOUNS DENOTING	SUPERSENSE	NOUNS DENOTING
act	acts or actions	object	natural objects (not man-made)
animal	animals	quantity	quantities and units of measure
artifact	man-made objects	phenomenon	natural phenomena
attribute	attributes of people and objects	plant	plants
body	body parts	possession	possession and transfer of possession
cognition	cognitive processes and contents	process	natural processes
communication	communicative processes and contents	person	people
event	natural events	relation	relations between people or things or ideas
feeling	feelings and emotions	shape	two and three dimensional shapes
food	foods and drinks	state	stable states of affairs
group	groupings of people or objects	substance	substances
location	spatial position	time	time and temporal relations
motive	goals	Tops	abstract terms for unique beginners
VERBS			
SUPERSENSE	VERBS OF	SUPERSENSE	VERBS OF
body	grooming, dressing and bodily care	emotion	feeling
change	size, temperature change, intensifying	motion	walking, flying, swimming
cognition	thinking, judging, analyzing, doubting	perception	seeing, hearing, feeling
communication	telling, asking, ordering, singing	possession	buying, selling, owning
competition	fighting, athletic activities	social	political and social activities and events
consumption	eating and drinking	stative	being, having, spatial relations
contact	touching, hitting, tying, digging	weather	raining, snowing, thawing, thundering
creation	sewing, baking, painting, performing		

Ciarmita and Altun 2003

Supersense tagging

Hyponymy标注

artifact artifact motion time group

The station wagons arrived at noon, a long shining line

motion location location

that coursed through the west campus.

Noun supersenses (Ciarmita and Altun 2003)

Supersense tagging

- Ciarameta and Altun (2006). Trained on data from Semcor (Miller et al. 1993); Brown corpus annotated with WordNet synset labels
- Token-level predictor – each instance of a word has its own supersense tag.
- Maximum-entropy Markov Model (MEMM) trained with averaged perceptron. Features for: word token identity, part-of-speech tag, word shape, previous label + supersense for most frequent synset for word.
- In-domain accuracy: 77.1 F score (cf. 66 F MFS baseline)

Data

- Semcor: 200K+ words tagged with Wordnet senses.
  http://www.cse.unt.edu/~rada/downloads.html#semcor
- WordNet
  https://wordnet.princeton.edu/wordnet/download/
- CROWN
  https://github.com/davidjurgens/crown

Frame and Role Semantics

使用Frame把语义结构化,并推测context

What do you know about this situation?

The travelers spent a few hours on land________

the ground

Frame semantics aim to capture this type of background information that readers naturally infer

Frame semantics

- A semantic frame is a coherent structure of related concepts
- Alternative view: a frame is a data-structure representing a stereotyped situation (Minsky, 1975).
- Assumes the meaning of a word cannot be understood without access to all the knowledge that relates to that word.

Sally gave Bill $100 for (a) tip

ransom

allowance

refund

honorarium

bounty

tuition

retainer

bonus

rent

fare

child support

bus money

salary

reward

alimony

根据名词可以推测Sally和Bill的context信息

Representation of meaning

- Constants name individual entities in the world
- Relations are sets of entities
- Variables refer to entities that have not yet been specified
- Quantifiers bind variables.
    - ∃ (existential quantifier)
    - ∀ (universal quantifier)

Representation of meaning

Pat likes Sal

- Constants: Pat, Sal
- Relations: likes(x,y)

- The denotation ⟦likes⟧ = the ordered set of entities for whom the relation is true
- likes(Pat, Sal) = true
- ⟦likes⟧ = {(Pat, Sal), (…, …)}

Representation of meaning

- Quantifiers bind variables.
  - ∃ (existential quantifier)
  - ∀ (universal quantifier)

- Order matters!

  For anyone X, must speak one kind of language Y

  - $\forall x \exists y$ speaks(x,y)
  - $\exists y \forall x$ speaks(x,y)

  There is a language Y such everyone speaks

Representation of meaning

- Relations: likes(x,y) is scoped over two variables
- We can represent the partial representation of meaning with lambda expressions:

λx.likes(x,Sal)

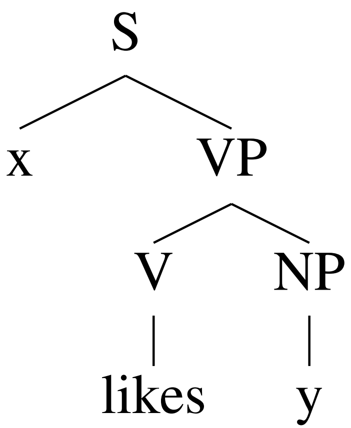
Expect one other argument to complete the meaning of this relation

Church (1940)

Representation of meaning

λy.λx.likes(x,y)

Lambda expressions let us tie semantics explicitly to phrases (subtrees in syntax)



Representation of meaning

- Relations: likes(x,y) is scoped over two variables
- We can represent the partial representation of meaning with lambda expressions:

λx.likes(x,Sal)

Expect one other argument to complete the meaning of this relation

Church (1940)

Shallow semantics

∃e,y.EVENT(e)
∧ CAUSER-OF-ACTION(e,Sasha)
∧ RECIPIENT-OF-ACTION(e,y)
∧ "window"(y)

∃e,y.BREAKING-EVENT(e)
∧ BREAKER(e,Sasha)
∧ BROKEN-THING(e,y)
∧ WINDOW(e,y)

∃e,y.EVENT(e)
∧ CAUSER-OF-ACTION(e,Pat)
∧ RECIPIENT-OF-ACTION(e,y)
∧ "door"(y)

∃e,y.OPENING-EVENT(e)
∧ OPENER(e,Pat)
∧ OPENED-THING(e,y)
∧ DOOR(e,y)

These roles have a lot in common: direct causal responsibility for the events, have volition, often animate

Event semantics

Yesterday, Pat gives Sal a book reluctantly

$\exists x.\text{book}(x) \wedge \textsc{give}(\text{Pat, Sal, x, yesterday, reluctantly})$

- One option: extend the arity of the relation (require more arguments)
- But that's not great because we need a separate predicate for every possible combination of arguments (even those that aren't required).

Eisenstein 2017

Event semantics

In model-theoretic semantics, each of these has some denotation in the world model.

Example: WINDOW has a identifier in some knowledge base (e.g., Freebase) uniquely identifying its properties.

$\exists e,y.$BREAKING-EVENT(e)
$\wedge$ BREAKER$(e,Sasha)$
$\wedge$ BROKEN-THING(e,y)
$\wedge$ WINDOW(e,y)

$\exists e,y.$OPENING-EVENT(e)
$\wedge$ OPENER(e,Pat)
$\wedge$ OPENED-THING(e,y)
$\wedge$ DOOR(e,y)

Event semantics

This requires a comprehensive representation of the world

$\exists e,y.$BREAKING-EVENT(e)
$\wedge$ BREAKER$(e,Sasha)$
$\wedge$ BROKEN-THING(e,y)
$\wedge$ WINDOW(e,y)

$\exists e,y.$OPENING-EVENT(e)
$\wedge$ OPENER(e,Pat)
$\wedge$ OPENED-THING(e,y)
$\wedge$ DOOR(e,y)

SLP3

Shallow semantics

Agent: Sasha
Theme: window

$\exists e,y.\text{BREAKING-EVENT}(e)$
$\wedge\ \text{BREAKER}(e,\text{Sasha})$
$\wedge\ \text{BROKEN-THING}(e,y)$
$\wedge\ \text{WINDOW}(e,y)$

Agent: Pat
Theme: door

$\exists e,y.\text{OPENING-EVENT}(e)$
$\wedge\ \text{OPENER}(e,\text{Pat})$
$\wedge\ \text{OPENED-THING}(e,y)$
$\wedge\ \text{DOOR}(e,y)$

Thematic roles

Thematic roles capture the semantic commonality among arguments for different relations (predicates)

- John broke the window
- The window was broken by John

Different syntactic roles, but the same thematic role.

SLP3

Thematic roles

Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event
Content	The proposition or content of a propositional event
Instrument	An instrument used in an event
Beneficiary	The beneficiary of an event
Source	The origin of the object of a transfer event
Goal	The destination of an object of a transfer event

- John broke the window
- The window was broken by John
- John broke the window with a rock
- The rock broke the window
- The window broke

SLP3

Thematic roles

Agent	*The waiter* spilled the soup. waiter本质是agent
Experiencer	John has *a headache*.
Force	*The wind* blows debris from the mall into our yards.
Theme	Only after Benjamin Franklin broke *the ice*...
Result	The city built *a regulation-size baseball diamond*...
Content	Mona asked *"You met Mary Ann at a supermarket?"*
Instrument	He poached catfish, stunning them with *a shocking device*...
Beneficiary	Whenever Ann makes hotel reservations for *her boss*...
Source	I flew in from *Boston*.
Goal	I drove to *Portland*.

SLP3

Thematic roles

The thematic roles for verbs generally are predictable by the syntactic position of the argument (specific to each verb class). Some allow for consistent alternations:

Agent Theme Goal

Doris gave the book to Cary

Agent Goal Theme

Doris gave Cary the book

Thematic roles capture common roles across different verbs

The student wrote the program in python using PyCharm

The musician played a sonata on the guitar

The lawyer delivered a closing argument

- student, musician, and lawyer all have agency, are acting of their own volition, and directly cause the event to transpire
- program, sonata, and argument all are things that are the results of their events

Thematic Roles

Thematic Role	Definition
Agent	The volitional causer of an event
Experiencer	The experiencer of an event
Force	The non-volitional causer of the event
Theme	The participant most directly affected by an event
Result	The end product of an event
Content	The proposition or content of a propositional event
Instrument	An instrument used in an event
Beneficiary	The beneficiary of an event
Source	The origin of the object of a transfer event
Goal	The destination of an object of a transfer event

The student wrote the program in python using PyCharm

Speech & Language Processing (3ed.) Table 20.1

Thematic roles

- Thematic roles are very useful but different to formally define AGENT, THEME, etc.
- At the same time, they may be too coarse for some applications.

Thematic roles

Intermediary instruments can be subjects

- The cook opened the jar with the new gadget
- The new gadget opened the jar

- Shelly ate the sliced banana with a fork
- *The fork ate the sliced banana

Enabling instruments cannot

Levin and Rappaport Hovav 2005; SLP3

Coarsening: Proto-roles

- Proto-roles = generalized thematic roles
- Proto-agent: causing an event, having volition wrt event, moving, acting with intention
- Proto-patient: change of state, causally affected by event)

Propbank

- Sentences from the Penn Treebank annotated with proto-roles, along with lexical entries for each sense of a verb identifying the specific meaning of each proto-role for that verb sense.

https://propbank.github.io

Propbank

(22.11) **agree.01**

Arg0: Agreer

Arg1: Proposition

Arg2: Other entity agreeing

Ex1: [$_{\text{Arg0}}$ The group] *agreed* [$_{\text{Arg1}}$ it wouldn't make an offer].

Ex2: [$_{\text{ArgM-TMP}}$ Usually] [$_{\text{Arg0}}$ John] *agrees* [$_{\text{Arg2}}$ with Mary] [$_{\text{Arg1}}$ on everything].

(22.12) **fall.01**

Arg1: Logical subject, patient, thing falling

Arg2: Extent, amount fallen

Arg3: start point

Arg4: end point, end state of arg1

Ex1: [$_{\text{Arg1}}$ Sales] *fell* [$_{\text{Arg4}}$ to \$25 million] [$_{\text{Arg3}}$ from \$27 million].

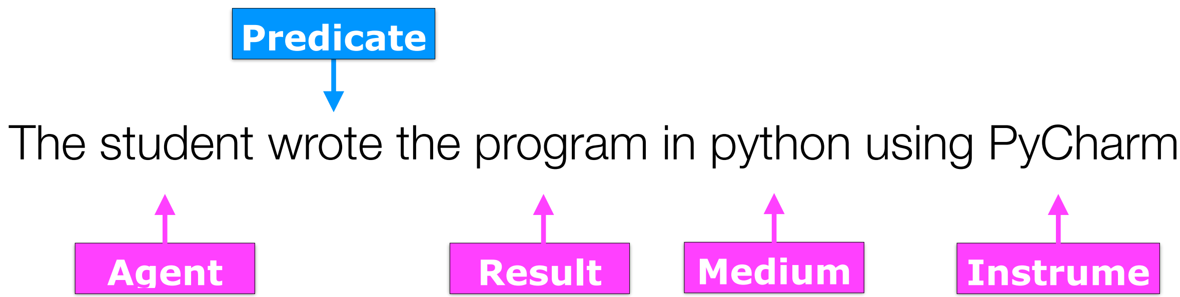
Ex2: [$_{\text{Arg1}}$ The average junk bond] *fell* [$_{\text{Arg2}}$ by 4.2%].

Propbank

- Verb-specific argument structures lets us map the commonalities among the different surface forms
  - [$_{Arg0}$ Big Fruit Co.] increased [$_{Arg1}$ the price of bananas].
  - [$_{Arg1}$ The price of bananas] was increased again [$_{Arg0}$ by Big Fruit Co.]
  - [$_{Arg1}$ The price of bananas] increased [$_{Arg2}$ 5%].

- [$_{\text{Arg1}}$ The price of bananas] increased [$_{\text{Arg2}}$ 5%].
- [$_{\text{Arg1}}$ The price of bananas] rose [$_{\text{Arg2}}$ 5%].
- There has been a [$_{\text{Arg2}}$ 5%] rise [$_{\text{Arg1}}$ in the price of bananas].

How to label roles?



- Treat it as a supervised learning task!
- Example using data from PropBank

PropBank

- Extra annotation of the Penn TreeBank, which was originally designed for parsing
- Each verb has its own roles, but these loosely align:
    - Arg0: Proto-Agent: volitional, cause, movers, sentience…
    - Arg1: Proto-Patient: thing that is affected, changes, stationary
    - Arg2: benefactive, instrument, attribute, or end state (mostly)
    - Arg3: starting point, benefactive, instrument, or attribute (mostly)
    - Arg4: …

PropBank-labeled sentence for increase

- Arg0: Cause of the increase (proto-agent!)
- Arg1: Thing that is increased (proto-patient!)
- Arg2: Amount of increase
- Arg3: Starting point prior to increase
- Arg4: Ending point after increase

Arg0 Arg1 Arg2
[The student] increased [their grade] [5%]

Arg1 Arg2
[The assignment 4 grade] increased by [5%]

Arg1 Arg3 Arg2
[The grade] was increased from [90] by [5 points]

FrameNet and Semantic Role Labeling

FrameNet

- Propbank maps argument structure for individual verb senses
- FrameNet maps argument structure for frames, which are evoked by a lexical unit (typically a verb)

https://framenet.icsi.berkeley.edu/fndrupal/framenet_data

Frames

AI

- Schank and Abelson 1975, 1977
- Minksky 1974

Linguistics

- Fillmore 1975, 1982, Tannen 1979

Cognitive Psychology

- Rumelhart 1975, 1980

Sociology

- Goffman 1975

Media Studies

- Entman 1993

Frames

- "A frame is a data-structure for representing a stereotyped situation" (Minsky 1975)

- By the term 'frame' I have in mind any system of concepts related in such a way that to understand any one of them you have to understand the whole structure in which it fits; when one of the things in such a structured is introduced … all of the others are automatically made available.'' (Fillmore 1982)

Who did what to whom?

purchase frame

Thing bought

- John bought the car at the dealership
- The car was bought by John
- John's purchase of the car
- The sale of the car cleared their inventory.

PropNet compared to FrameNet

- PropBank reflects the structure and types of arguments for an individual verb sense
  - increase.v: prop只从语义出发
    - Arg0: Cause of the increase
    - Arg1: Thing that is increased
    - Arg2: Amount of increase
    - …
- FrameNet reflects the structure and types of arguments for frames, which are signaled by a lexical unit
  - cooking_event
    - cook
    - produced_food
    - heating_instrument
    - …

https://framenet.icsi.berkeley.edu/fndrupal/framenet_data

Semantic Frame

APPLY_HEAT

- Lexical units: words that signal the frame's presence

bake.v, barbecue.v, blanch.v, boil.v, braise.v, broil.v, brown.v, char.v, coddle.v, cook.v, deep fry.v, fry.v, grill.v, microwave.v, parboil.v, plank.v, poach.v, roast.v, saute.v, scald.v, sear.v, simmer.v, singe.v, steam.v, steep.v, stew.v, toast.v

- Core Frame Elements:

Cook	The Cook applies heat to the Food.
Food	Food is the entity to which heat is applied by the Cook.
Heating instrument	The entity that directly supplies heat to the Foo
Container	The Container holds the Food to which heat is applied.
Temperature setting	The Temperature_setting of the Heating_instrument for the Food.

Semantic Frame

DESTROY

- Lexical units:

annihilate.v, annihilation.n, blast.v, blow up.v, demolish.v, demolition.n, destroy.v, destruction.n, destructive.a, devastate.v, devastation.n, dismantle.v, dismantlement.n, lay waste.v, level.v, obliterate.v, obliteration.n, raze.v, ruin.v, take out.v, unmake.v, vaporize.v

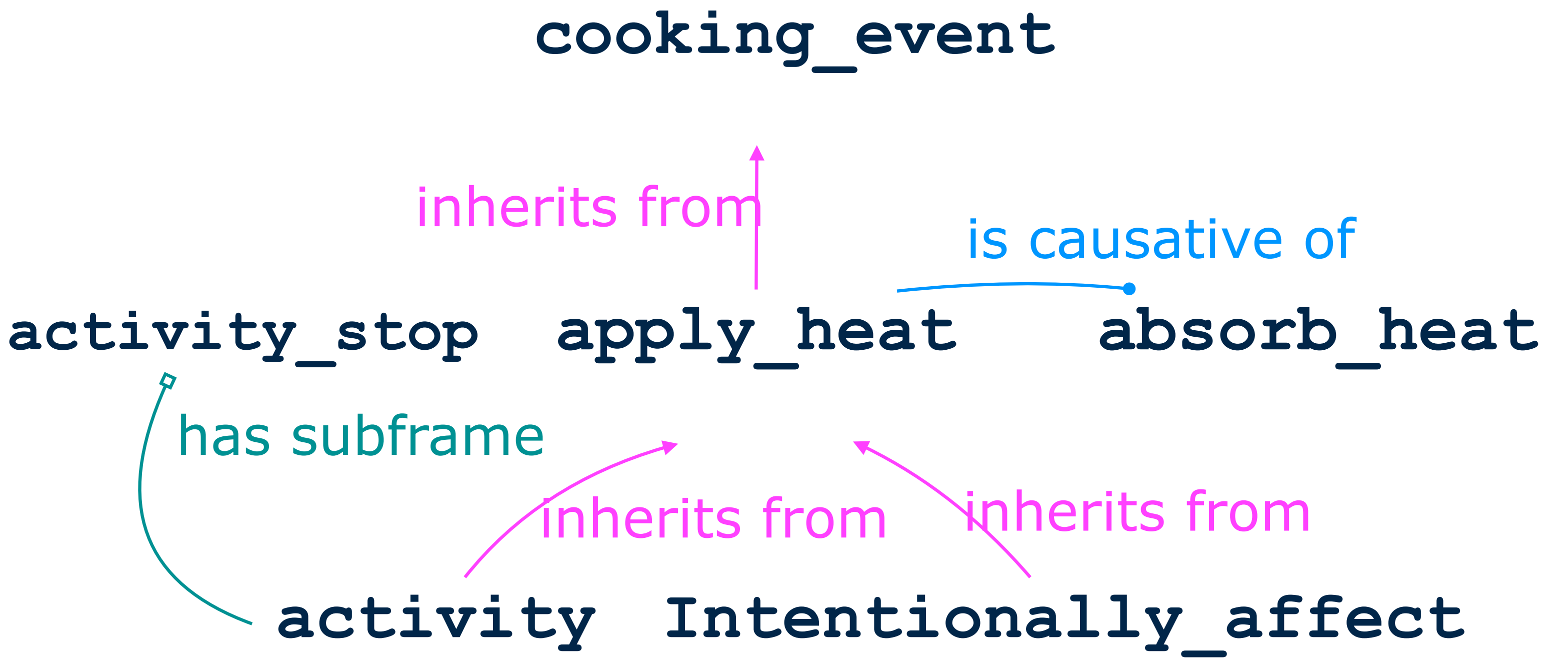
- Core Frame Elements:

Cause	The event or entity which is responsible for the destruction of the Patient.
Destroyer	The conscious entity, generally a person, that performs the intentional action that results in the Patient's destruction.
Patient	The entity which is destroyed by the Destroyer.

Lexical units aren't just verbs

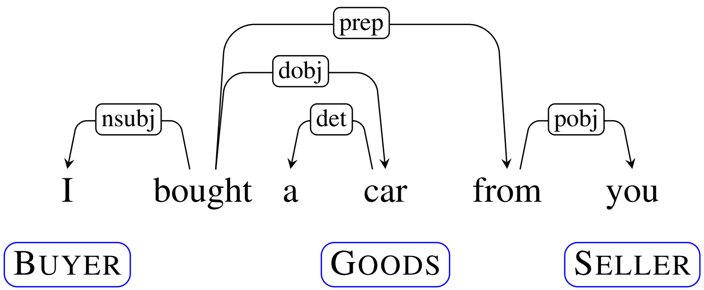
- For the "change in position on a scale" frame, example units…
  - verbs: increase, advance, jump, slide
  - nouns: escalation, increase, growth, shift
  - adjectives: accelerated, growing
  - adverbs: increasingly

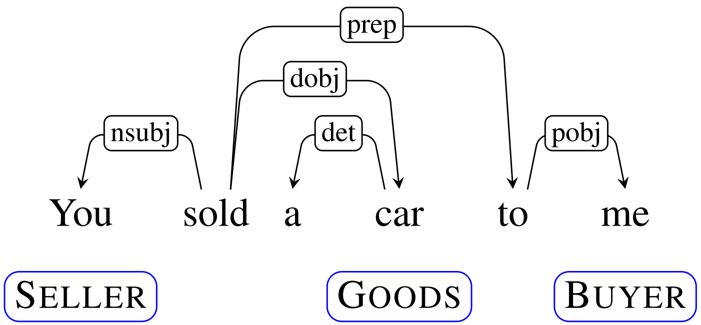
Frames can be related as well



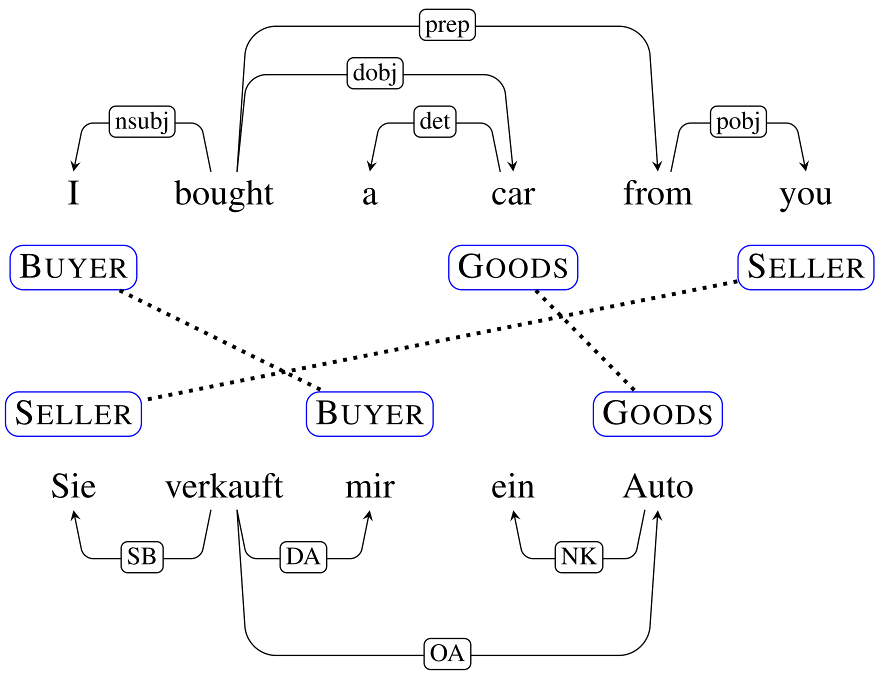
Semantic representations

Two different perspectives on a commercial transaction





Multilingual frames



Multilingual frames

- French
- Chinese
- Brazilian Portuguese
- German
- Spanish
- Japanese
- Swedish
- Korean

https://framenet.icsi.berkeley.edu/fndrupal/framenets_in_other_languages

Semantic role labeling

- Input: a sentence
- Output:
  - A list of predicates, each containing:
    - a label (e.g., Framenet frame)
    - a span
    - a set of arguments, each containing:
      - a label (thematic role, FrameNet role)
      - a span

Smith 2017

Semantic role labeling

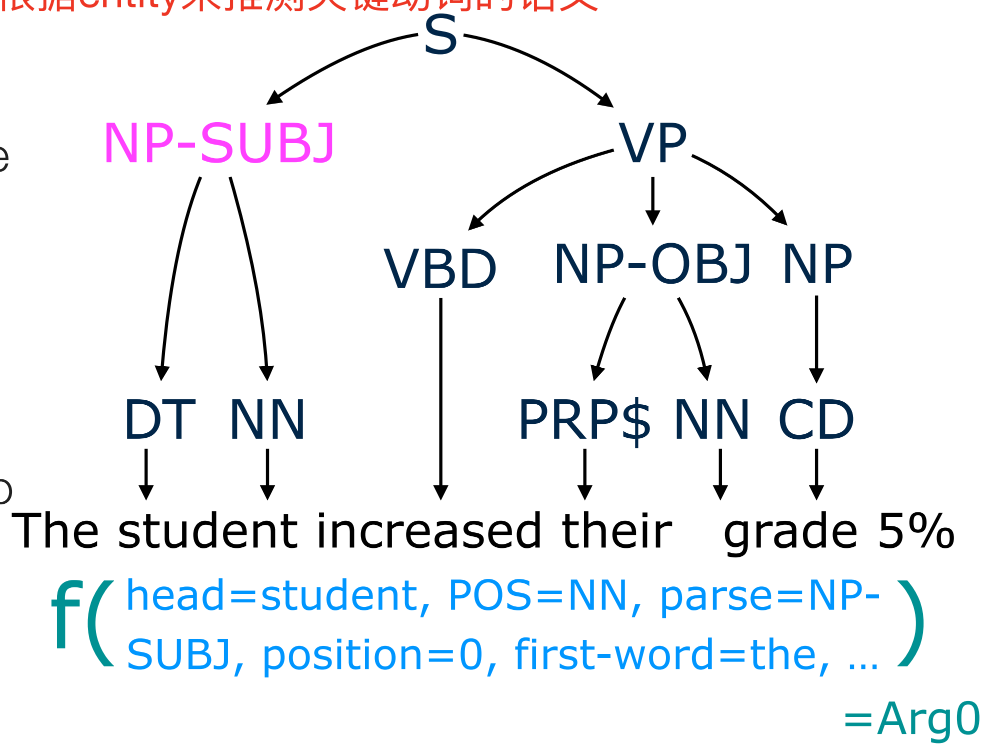
FrameNet	[You]	can't	[blame]	[the program]	[for being unable to identify it]
	COGNIZER		TARGET	EVALUEE	REASON

PropBank	[The San Francisco Examiner]	issued	[a special edition]	[yesterday]
	ARG0	TARGET	ARG1	ARGM-TMP

Labeling sentences with frame roles

看到关键动词就用frame解析,再根据entity来推测关键动词的语义

- General algorithm:
  - Constituency parse the sentence
  - For each predicate (recursively)
    - Featurize the sentence relative to the predicate
    - Predict the role (if any) of the predicate



Semantic Role Labeling is Challenging!

- Too many predicates to classify all—pruning is needed!
- Many types of arguments makes leveraging cross-frame labels difficult
  - Simpler schemes like Proto-Roles have a huge advantage here
- Training data is growing but still limited
  - PropBank annotations are available in many languages!
- Neural approaches are growing in usage—but still data-limited!

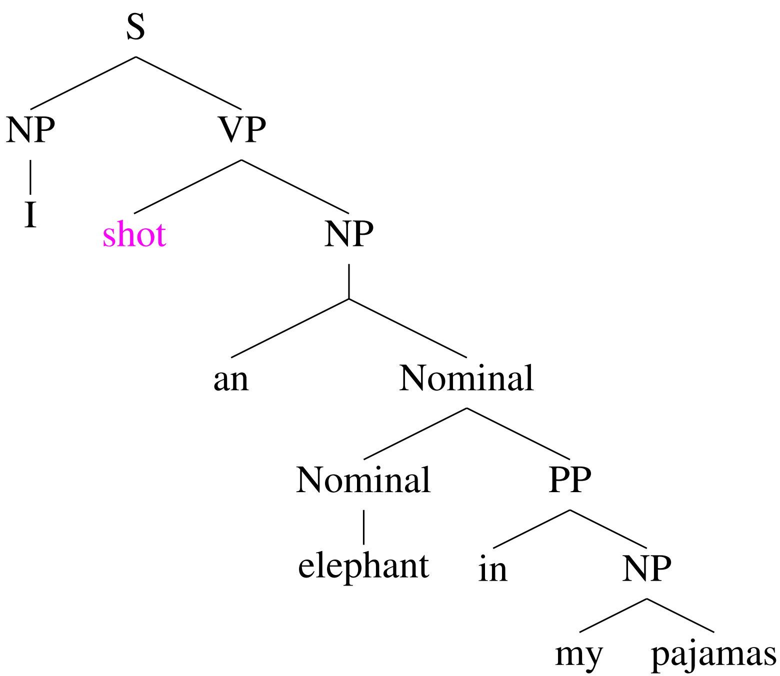
Semantic role labeling

function SEMANTICROLELABEL(*words*) **returns** labeled tree

```
parse ← PARSE(words)
for each predicate in parse do
    for each node in parse do
        featurevector ← EXTRACTFEATURES(node, predicate, parse)
        CLASSIFYNODE(node, featurevector, parse)
```


Gildea and Jurafsky 2002; SLP3

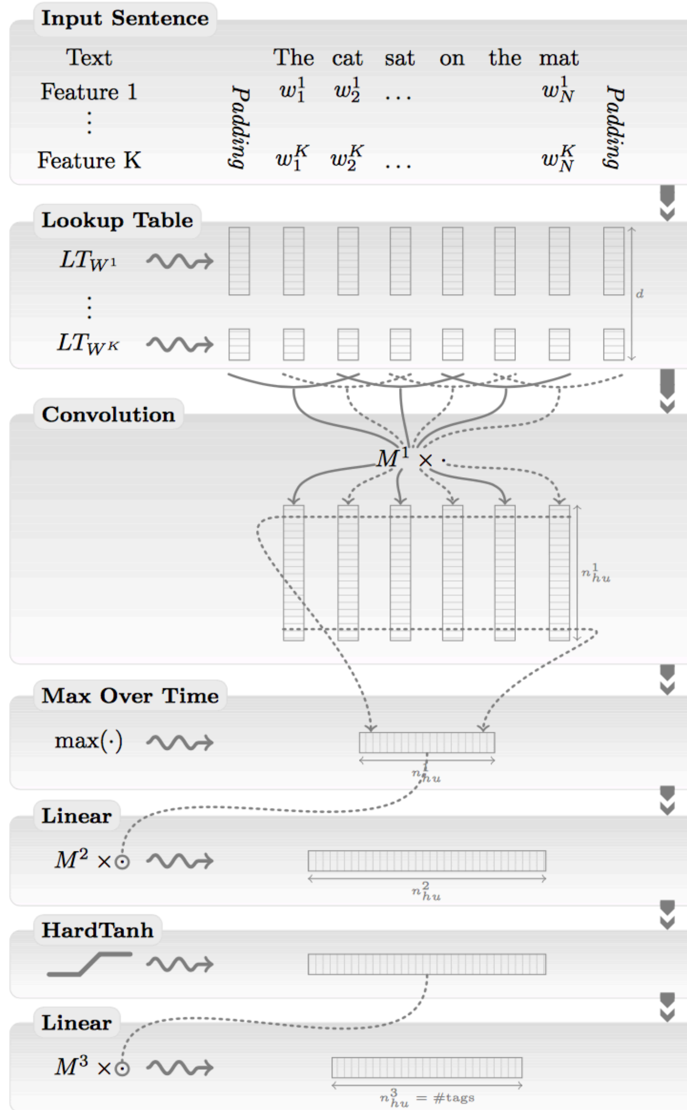
Semantic role labeling



feature
predicate: shot
phrase type = NP
headword of phrase = elephant
path = NP↑S↓VP
voice of verb = active
voice of verb = passive
phrase before verb?
first/last words of phrase

Semantic role labeling

Collobert et al. (2011), Natural Language Processing (Almost) from Scratch



Semantic role labeling

- Sentence-level constraints:
  - Arguments can't overlap
  - For a given predicate, typically only one argument of each type (e.g., ARG0, BUYER)
- Approximate joint decoding (Das et al. 2010)
- Constrained optimization (e.g., ILP)

Data

- CCGBank [through UCB Library]
  http://groups.inf.ed.ac.uk/ccg/ccgbank.html

- PropBank
  https://propbank.github.io

- FrameNet
  https://framenet.icsi.berkeley.edu/fndrupal/framenet_data

Semantic Parsing

Semantic Parsing

- “Semantic Parsing” is, ironically, a semantically ambiguous term
  - Semantic role labeling
  - Finding generic relations in text
  - Transforming a natural language sentence into its meaning representation

Semantic Parsing

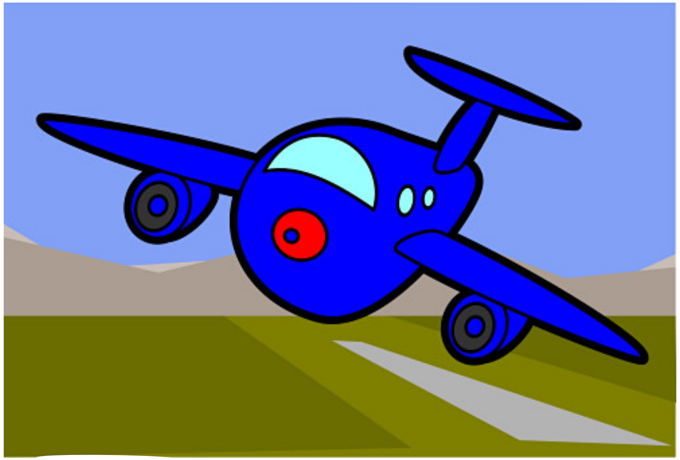
- "Semantic Parsing" is, ironically, a semantically ambiguous term
    - Semantic role labeling
    - Finding generic relations in text
    - → Transforming a natural language sentence into its meaning representation

Semantic Parsing

- Semantic Parsing: Transforming natural language (NL) sentences into computer executable complete meaning representations (MRs) for domain-specific applications
- Realistic semantic parsing currently entails domain dependence
- Example application domains
  - ATIS: Air Travel Information Service
  - CLang: Robocup Coach Language
  - Geoquery: A Database Query Application

ATIS: Air Travel Information Service

- Interface to an air travel database [Price, 1990]
- Widely-used benchmark for spoken language understanding



May I see all the flights from Cleveland to Dallas?

Semantic Parsing

Air-Transportation
Show: (Flight-Number)
Origin: (City "Cleveland")
Destination: (City "Dallas")

Query

NA 1439,
TQ 23,
…

Geoquery: A Database Query Application

Which rivers run through the states bordering Texas?

Semantic Parsing

answer(traverse(next_to(stateid('texas'))))

Query

Answer

Arkansas, Canadian, Cimarron, Gila, Mississippi, Rio Grande …

Meaning Representation Languages

- Meaning representation language (MRL) for an application is assumed to be present
- MRL is designed by the creators of the application to suit the application's needs independent of natural language
- CLang was designed by RoboCup community to send formal coaching instructions to simulated players
- Geoquery's MRL was based on the Prolog database
- SQL as a representation of grounded question intent for text-to-SQL

Engineering Motivation for Semantic Parsing

- Applications of domain-dependent semantic parsing
  - Natural language interfaces to computing systems
  - Communication with robots in natural language
  - Personalized software assistants
  - Question-answering systems
- Machine learning makes developing semantic parsers for specific applications more tractable
- Training corpora can be easily developed by tagging natural-language glosses with formal statements

Distinctions from Other NLP Tasks

- Shallow semantic processing
  - Information extraction
  - Semantic role labeling
- Intermediate linguistic representations
  - Part-of-speech tagging
  - Syntactic parsing
  - Semantic role labeling
- Output meant for humans
  - Question answering
  - Summarization
  - Machine translation

What you should know

- Words can have different meanings, known as senses which relate to other senses in a variety of ways
- WSD is a challenging task (why?) with many approaches
- Frame semantics describe common sense knowledge about events and the roles of entities in those events
    - Semantic Role Labeling determines which words/ phrases are particular roles in a frame
- Semantic Parsing turns natural next into a semantic representation that can be evaluated (think text-to-SQL)