Outline

- Challenges
- Distributional Semantics
- Word Sense
- Semantic Role Labeling
The Challenge of Designing Semantic Representations

• Q: What is semantics?
  A: The study of meaning

• Q: What is meaning?
  A: …
We know it when we see it

• These sentences/phrases all have the same meaning:

  • XYZ corporation bought the stock.
  • The stock was bought by XYZ corporation.
  • The purchase of the stock by XYZ corporation...
  • The stock purchase by XYZ corporation...
But how to formally define it?
Example Representations

Sentence: “I have a car”

Logic Formula

\[ \exists e,y \, \text{Having}(e) \land \text{Haver}(e, \text{Speaker}) \land \text{HadThing}(e,y) \land \text{Car}(y) \]

Graph Representation

Key-Value Records

Car

\[ \uparrow \text{POSS-BY} \]

Speaker

Haver

Had-Thing

Having

Haver: Speaker

HadThing: Car
Example Representations

Sentence: “I have a car”

“Ich habe ein Auto” As translation in another language
There’s no single agreed-upon representation that works in all cases

- Different emphases:
  - Words or Sentences
  - Syntax-Semantics interface, Logical Inference, etc.

- Different aims:
  - “Deep (and narrow)” vs “Shallow (and broad)”
  - e.g. Show me all flights from BWI to NRT.
    - Do we link to actual flight records?
    - Or general concept of flying machines?
Outline

- Challenges
- Distributional Semantics
- Word Sense
- Semantic Role Labeling
Distributional Semantics

What information is needed for a good representation?

You shall know a word by the company it keeps. – J. R. Firth (linguist)

- Distributional Semantics: a word’s meaning is based on its positional distribution in text
Learning Distributional Semantics from large text dataset

- Your pet dog is so cute
- Your pet cat is so cute
- The dog ate my homework
- The cat ate my homework

$\text{neighbor(dog)}$ overlaps-with $\text{neighbor(cats)}$

so $\text{meaning(dog)}$ is-similar-to $\text{meaning(cats)}$
Word2Vec implements Distribution Semantics

From: Mikolov, Tomas; et al. (2013). "Efficient Estimation of Word Representations in Vector Space"
Latent Semantic Analysis (LSA) also implements Distributional Semantics

Document-Term Matrix

<table>
<thead>
<tr>
<th></th>
<th>pet</th>
<th>dog</th>
<th>is</th>
<th>cat</th>
<th>the</th>
<th>ate</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>S2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>S3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>S4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

S1: ... *pet dog is* ...
S2: ... *pet cat is* ...
S3: The *dog ate* ...
S4: The *cat ate* ...

Pet-peeve: fundamentally, neural approaches aren’t so different from classical LSA. They just use more GPUs!
Advantages of Distributional Semantics

• Do you know what’s a bar-ba-loot?

• What is a more likely sentence?

  1. Bar-ba-loots like to eat fruits

  2. The pirate ship Bar-ba-loots looted Barbados

• What if I tell you: $\text{vector}(\text{Bar-ba-loots}) \sim \text{vector}(\text{bear})$
Advantages of Distributional Semantics

- Similarity metric between vectors allow processing of related words

Bar-ba-loots from Dr. Seuss’ children books
Outline

- Challenges
- Distributional Semantics
  - Word Sense
- Semantic Role Labeling
Word Sense

• Senses of the word **bank**:  
  • The bank\(^1\) is investing in CDO’s. 
  • The river bank\(^2\) is flooding. 
  • The food bank\(^3\) is providing free meals.  
  • The bank\(^4\) is at the corner of 1st and Main St. 

• bank\(^1\) & bank\(^2\) are **homonyms** (coincidentally same sound/orthography, otherwise unrelated in meaning) 

• bank\(^1\) & bank\(^3\) exhibit **polysemy** (related meaning: “a repository for stuff”) 

• bank\(^4\) shows there’s a relationship between BUILDING and INSTITUTION
WordNet

**Hypernym:**
- dog is-a-kind-of canine
- to butt is-a-kind-of to hit

**Meronym:**
- window is-part-of building

**Entailment:**
- to sleep is entailed by to snore

*Note: Relation is defined in terms of synsets, not words (as this simplified slide might suggest)*
• When senses of two different words are similar, we say they’re synonyms
  
  • e.g. couch & sofa; bank & repository

• Instead of talking about two words being synonyms, we talk of synset as set of senses that are similar

  • e.g. couch\(^1\) & sofa\(^1\); bank\(^1\) & repository\(^2\)
Word to search for: serve

Noun

- **S: (n)** serve#1, service#12 ((sports) a stroke that puts the ball in play) "his powerful serves won the game"

Verb

- (55) **S: (v)** serve#1, function#2 (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"
- (36) **S: (v)** serve#2 (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"
- (24) **S: (v)** serve#3 (contribute or conduce to) "The scandal served to increase his popularity"
- (23) **S: (v)** service#1, serve#4 (be used by; as of a utility) "The sewage plant served the neighboring communities"; "The garage served to shelter his horses"
- (21) **S: (v)** serve#5, help#5 (help to some food; help with food or drink) "I served the guests"
Word Sense Disambiguation (WSD)

- Input Sentence: He served as Secretary of State
- Task: What sense is served?
  - serve¹, serve², serve³, serve⁴?
Supervised methods for WSD

Training Data: Sentences with Sense Labels

Extract features from neighboring words & from WordNet descriptions, then train

Classifier

Try to capture selectional preferences:

Example:
1. I hate washing dishes
2. I can stir-fry some simple dishes

There is little ambiguity in dishes\(^1\) (a physical plate) vs dishes\(^2\) (a particular food item, like Chicken Fried Rice) for us because:
- washing and stir-fry “select” for different kinds of objects
Semi-Supervised methods for WSD: Yarowsky Algorithm

Small Training Data

Unlabelled Data

Unlabelled Data

Extract features from neighboring words & from WordNet descriptions, then train

Classifier v0

Label

Retrain

Classifier v1

Label

Retrain

Classifier v2
Outline

- Challenges
- Distributional Semantics
- Word Sense
- Semantic Role Labeling
Who did what to whom

• Jerry hit Tom with a hammer today
• Today, Tom was hit by Jerry with a hammer
• With a hammer, Tom hit Jerry today
Thematic Roles

- **Agent**: volitional causer for an event. “The waiter spilled the soup”
- **Experiencer**: the experiencer of an event. “John has a headache”
- **Theme**: the participant most directly effected by the event. “Fred threw the rock”
- **Result**: the end product of an event. “The government has built a stadium.”
- **Force**: the non-volitioner cause of an event. “The wind blew things away”
- **Instrument**: an instrument used in an event
Alternatives to thematic roles

- It’s difficult to create a standard set of roles

Solutions:

- Fewer roles: each role is more general —> PropBank
- More roles: define roles specific to each group of predicate —> FrameNet
PropBank roles

• Each verb sense has numbered argument:
  • Arg0: Proto-Agent
  • Arg1: Proto-Patient
  • Arg2-Arg5: depends on the verb sense, includes benefactive, instrument, attribute, end state, ..
  • ArgM-: modifiers or adjuncts
• These are annotated on top of a syntactic parse
PropBank Example

• increase.01 “go up incrementally”
  • Arg0: Proto-Agent - causer of increase
  • Arg1: Proto-Patient - thing increased
  • Arg2: Amount increased
  • Arg3: Start point; Arg4: End point

• Now we can see these sentences have similar meanings
  • [Arg0 The shop] increased [Arg1 the price] [Arg3 today]
  • [Arg1 The price] increased [Arg2 10%] [Arg0 by the shop]
FrameNet

- These different forms of "increase" are related:
  - [Arg1 The price] increased [Arg2 10%]
  - [Arg1 The price] rose
  - There has been a [Arg2 10%] rise [Arg1 in the price]

- Let's define a frame: change_position_on_a_scale, with elements like item's attribute, initial/final value, difference
  - Verbs evoking this frame: rise, increase, jump, grow,…
  - Nouns evoking this frame: rise, increase, growth, escalation,…
  - [Item The price] increased [Difference 10%]
  - There has been a [Difference 10%] rise [Item in the price]
Semantic Role Labeling Task Formulation (there are variants)

- Determine the semantic role of constituents of a sentence, given the predicate

Task 1: Identification

Task 2: Classification

PropBank Example
Semantic Role Labeling Task Formulation (there are variants)

- Determine the semantic role of constituents of a sentence, given the predicate

**Task 1: Identification**
Are these spans possible arguments for the predicate “issued”? Yes/No

**Task 2: Classification**

PropBank Example
Semantic Role Labeling Task Formulation (there are variants)

- Determine the semantic role of constituents of a sentence, given the predicate

**Task 1: Identification**

Are these spans possible arguments for the predicate “issued”? Yes/No

**Task 2: Classification**

Given Task 1 span results, what is the label? Arg0? Arg1? Arg2? ArgM?
Models for SRL (1)

- Log-linear classifier for each task

- Features include:
  - headword/POS of constituent
  - voice (active or passive)
  - grammar rule (subcategorization) for predicate VP->VBD NP PP
  - named entity type (The San Francisco Examiner = ORGANIZATION)
  - path in tree from constituent to predicate: NP↑S↓VP↓VBD

- Apply classifier to each node on tree
Models for SRL (2)

- Formulate as sequence labeling (BIO encoding)

He et. al. (2017) Deep SRL. What works and what’s next
Summary

• Challenges: Many ways to define “semantics”

• Distributional Semantics: “You shall know a word by the company it keeps.” e.g. Word2Vec

• Word Sense: Synsets & WordNet relations

• Semantic Role Labeling:
  • Who did what to whom.
  • Thematic roles, PropBank, FrameNet