Lexical Semantics, Distributions, Predicate-Argument Structure, and Frame Semantic Parsing

11-711 Algorithms for NLP
2 December 2014
(With thanks to Noah Smith)
11-711 Course Context

- Previous semantics lectures discussed composing meanings of parts to produce the correct global sentence meaning
  - *The mailman bit my dog.*
- The “atomic units” of meaning have come from the lexical entries for words
- The meanings of words have been overly simplified (as in FOL): atomic objects in a set-theoretic model
First, review: what is a “word”?

• Some Asian languages have obvious issues:
  利比亚“全国过渡委员会”执行委员会主席凯卜
  22日在首都的黎波里公布“过渡政府”内阁名单，
  宣告过渡政府正式成立。

• But German too: Noun-noun compounds:
  Gesundheitsversicherungsgesellschaften

• Spanish clitics: Darmelo

• Even English has issues, to a small degree:
  Gregg and Bob’s house
What is a “word”?

• Some Asian languages have obvious issues:
  利比亚“全国过渡委员会”执行委员会主席凯卜22日在首都的黎波里公布“过渡政府”内阁名单，宣告过渡政府正式成立。

• But German too: Noun-noun compounds:
  Gesundheits-versicherungs-gesellschaften (health insurance companies)

• Spanish clitics: Darmelo

• Even English has issues, to a small degree:
  Gregg and Bob’s house
What is a “word”?  

• Some Asian languages have obvious issues:  
  利比亚“全国过渡委员会”执行委员会主席凯卜22日在首都的黎波里公布“过渡政府”内阁名单，宣告过渡政府正式成立。

• But German too: Noun-noun compounds:
  Gesundheitsversicherungsgesellschaften

• Spanish clitics: Dar-me-lo (To give me it)

• Even English has issues, to a small degree:
  Gregg and Bob’s house
Again, what’s a Word?

• Lemma (base form)
  – “animals” becomes animal
  – “brought” becomes bring

• Sometimes more than one word
  – “throws up” → throw up
  – “New York” → New York
Word Sense

• Instead, a bank can hold the investments in a custodial account in the client’s name.
• But as agriculture burgeons on the east bank, the river will shrink even more.
• While some banks furnish sperm only to married women, others are much less restrictive.
• The bank is near the corner of Forbes and Murray.
Four Meanings of “Bank”

• **Synonyms:**
  - bank$^1$ = “financial institution”
  - bank$^2$ = “sloping mound”
  - bank$^3$ = “biological repository”
  - bank$^4$ = “building where a bank$^1$ does its business”

• The connections between these different **senses** vary from practically none (**homonymy**) to related (**polysemy**).
  - The relationship between the senses bank$^4$ and bank$^1$ is called **metonymy**.
Antonyms

• White/black, tall/short, skinny/American, …
• But different dimensions possible:
  – White/Black vs. White/Colorful
  – Often culturally determined
• Partly interesting because automatic methods have trouble separating these from synonyms
  – Same *semantic field*
How Many Senses?

• This is a hard question, due to vagueness.
• Considerations:
  – Truth conditions (serve meat / serve time)
  – Syntactic behavior (serve meat / serve as senator)
  – Zeugma test:
    • #Does United serve breakfast and Pittsburgh?
    • ??She poaches elephants and pears.
Ambiguity vs. Vagueness

• **Lexical ambiguity:** *My wife has two kids* (children or goats?)

• **Vs. Vagueness:** 1 sense, but indefinite: *horse* (*mare, colt, filly, stallion, ...*) vs. *kid*:
  – *I have two horses and George has three*
  – *I have two kids and George has three*

• Verbs too: *I ran last year and George did too*

• **Vs. Reference:** *I, here, the dog* not considered ambiguous in the same way
Related Phenomena

- Homophones (*would/wood, two/too/to*)
  - *Mary, merry, marry* in some dialects, not others
- Homographs (*bass/bass*)
Word Senses and Dictionaries

**sentence** |ˈsentns|
noun
1 a set of words that is complete in itself, typically containing a subject and predicate, conveying a statement, question, exclamation, or command, and consisting of a main clause and sometimes one or more subordinate clauses.
- Logic a series of signs or symbols expressing a proposition in an artificial or logical language.
2 the punishment assigned to a defendant found guilty by a court: her husband is serving a three-year sentence for fraud.
- the punishment fixed by law for a particular offense: slander of an official carried an eight-year prison sentence.

verb [trans.]
declare the punishment decided for (an offender): ten army officers were sentenced to death.

PHRASES
**under sentence of** having been condemned to: he was under sentence of death.

**ORIGIN** Middle English (in the senses [way of thinking, opinion, court's declaration of punishment,] and [gist of a piece of writing]): via Old French from Latin sententia 'opinion,' from sentire 'feel, be of the opinion.'

---

**statement** |ˈstætəmənt|
noun
a definite or clear expression of something in speech or writing: do you agree with this statement? this is correct as a **statement of fact**.
- an official account of facts, views, or plans, esp. one for release to the media: the officials issued a joint statement calling for negotiations.
- a formal account of events given by a witness, defendant, or other party to the police or in a court of law: she made a statement to the police.
- a document setting out items of debit and credit between a bank or other organization and a customer.
- the expression of an idea or opinion through something other than words: their humorous kitschiness makes a statement of serious wealth.
- Music the occurrence of a musical idea or motive within a composition: a carefully structured musical and dramatic progression from the first statement of this theme.
**expression** | /ɪkˈspres·ʃən |
---
noun
1. the process of making known one's thoughts or feelings: *his views found expression in his moral sermons* | *she accepted his expressions of sympathy.*
- the conveying of opinions publicly without interference by the government: *the right to freedom of expression.*
- the look on someone's face that conveys a particular emotion: *a sad expression.*
- the ability to put an emotion into words: *envious beyond expression.*
- a word or phrase, esp. an idiomatic one, used to convey an idea: *no place is the expression “garbage in, garbage out” any truer.*
- the style or phrasing of written or spoken words: *subtlety of expression.*
- the conveying of feeling in the face or voice, in a work of art, or in the performance of a piece of music: *eyes empty of expression* | *their instruments have a rich variety of expression.*
- Mathematics a collection of symbols that jointly express a quantity: *the expression for the circumference of a circle is 2πr.*
- Genetics the appearance in a phenotype of a characteristic or effect attributed to a particular gene.
- (also *gene expression*) Genetics the process by which possession of a gene leads to the appearance in the phenotype of the corresponding character.
2. the production of something, esp. by pressing or squeezing it out: *essential oils obtained by distillation or expression.*
Ontologies

• For NLP, databases of word senses are typically organized by lexical relations such as hypernym (IS-A) into a DAG
• This has been worked on for quite a while
• Aristotle’s classes (about 330 BC)
  – substance (physical objects)
  – quantity (e.g., numbers)
  – quality (e.g., being red)
  – Others: relation, place, time, position, state, action, affection
Word senses in WordNet3.0

The noun “bass” has 8 senses in WordNet.
1. bass\textsuperscript{1} - (the lowest part of the musical range)
2. bass\textsuperscript{2}, bass part\textsuperscript{1} - (the lowest part in polyphonic music)
3. bass\textsuperscript{3}, basso\textsuperscript{1} - (an adult male singer with the lowest voice)
4. sea bass\textsuperscript{1}, bass\textsuperscript{4} - (the lean flesh of a saltwater fish of the family Serranidae)
5. freshwater bass\textsuperscript{1}, bass\textsuperscript{5} - (any of various North American freshwater fish with lean flesh (especially of the genus Micropterus))
6. bass\textsuperscript{6}, bass voice\textsuperscript{1}, basso\textsuperscript{2} - (the lowest adult male singing voice)
7. bass\textsuperscript{7} - (the member with the lowest range of a family of musical instruments)
8. bass\textsuperscript{8} - (nontechnical name for any of numerous edible marine and freshwater spiny-finned fishes)

The adjective “bass” has 1 sense in WordNet.
1. bass\textsuperscript{1}, deep\textsuperscript{6} - (having or denoting a low vocal or instrumental range)
   “a deep voice”; “a bass voice is lower than a baritone voice”; “a bass clarinet”
Synsets

- (bass6, bass-voice1, basso2)
- (bass1, deep6)  (Adjective)

- (chump1, fool2, gull1, mark9, patsy1,
  fall guy1, sucker1, soft touch1, mug2)
“Rough” Synonymy

- Jonathan Safran Foer’s *Everything is Illuminated*

AN OVERTURE TO THE COMMENCEMENT OF A VERY RIGID JOURNEY

My legal name is Alexander Perchov. But all of my many friends dub me Alex, because that is a more flaccid-to-utter version of my legal name. Mother dubs me Alexi-stop-sleening-me!, because I am always sleening her. If you want to know why I am always sleening her, it is because I am always elsewhere with friends, and disseminating so much currency, and performing so many things that can spleen a mother. Father used to dub me Shapka, for the fur hat I would don even in the summer month. He ceased dubbing me that because I ordered him to cease dubbing me that. It sounded boyish to me, and I have always thought of myself as very potent and generative. I have many many girls, believe me, and they all have a different name for me. One dubs me Baby, not because I am a baby, but because she attends to me. Another dubs me All Night. Do you want to know why? I have a girl who dubs me Currency, because I disseminate so much currency around her. She licks my chops for it. I have a miniature brother who dubs me Alli. I do not dig this name very much, but I dig him very much, so OK, I permit him to dub me Alli. As for his name, it is Little Igor, but Father dubs him Clumsy One, because he is always promenading into things. It was only four days previous that he made his eye blue from a mismanagement with a brick wall. If you’re wondering what my bitch’s name is, it is Sammy Davis, Junior, Junior. She has this name because Sammy Davis, Junior was Grandfather’s beloved singer, and the bitch is his, not mine, because I am not the one who thinks he is blind.

As for me, I was sired in 1977, the same year as the hero of this story. In truth, my life has been very ordinary. As I mentioned before, I do
# Noun relations in WordNet3.0

<table>
<thead>
<tr>
<th>Relation</th>
<th>Also Called</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyponym</td>
<td>Subordinate</td>
<td>From concepts to subtypes</td>
<td>meal¹ → lunch¹</td>
</tr>
<tr>
<td>Instance Hyponym</td>
<td>Instance</td>
<td>From instances to their concepts</td>
<td>Austen¹ → author¹</td>
</tr>
<tr>
<td>Member Meronym</td>
<td>Has-Member</td>
<td>From groups to their members</td>
<td>composer¹ → Bach¹</td>
</tr>
<tr>
<td>Member Holonym</td>
<td>Member-Of</td>
<td>From members to their groups</td>
<td>faculty² → professor¹</td>
</tr>
<tr>
<td>Part Meronym</td>
<td>Has-Part</td>
<td>From wholes to parts</td>
<td>copilot¹ → crew¹</td>
</tr>
<tr>
<td>Part Holonym</td>
<td>Part-Of</td>
<td>From parts to wholes</td>
<td>table² → leg³</td>
</tr>
<tr>
<td>Substance Meronym</td>
<td></td>
<td>From substances to their subparts</td>
<td>course⁷ → meal¹</td>
</tr>
<tr>
<td>Substance Holonym</td>
<td></td>
<td>From parts of substances to wholes</td>
<td>water¹ → oxygen¹</td>
</tr>
<tr>
<td>Antonym</td>
<td></td>
<td>Semantic opposition between lemmas</td>
<td>gin¹ → martini¹</td>
</tr>
<tr>
<td>Derivationally Related Form</td>
<td></td>
<td>Lemmas w/same morphological root</td>
<td>leader¹ ↔ follower¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>destruction¹ ↔ destroy¹</td>
</tr>
</tbody>
</table>
Sense 3
bass, basso --
(an adult male singer with the lowest voice)
=> singer, vocalist, vocalizer, vocaliser
  => musician, instrumentalist, player
  => performer, performing artist
  => entertainer
  => person, individual, someone...
    => organism, being
      => living thing, animate thing,
      => whole, unit
      => object, physical object
        => physical entity
          => entity
          => causal agent, cause, causal agency
            => physical entity
              => entity

Sense 7
bass --
(the member with the lowest range of a family of musical instruments)
=> musical instrument, instrument
  => device
    => instrumentality, instrumentation
      => artifact, artefact
      => whole, unit
      => object, physical object
        => physical entity
Is a hamburger food?

Sense 1
hamburger, beefburger --
(a fried cake of minced beef served on a bun)
=> sandwich
  => snack food
  => dish
    => nutriment, nourishment, nutrition...
    => food, nutrient
    => substance
    => matter
    => physical entity
    => entity
Verb relations in WordNet3.0

<table>
<thead>
<tr>
<th>Relation</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyponym</td>
<td>From events to superordinate events</td>
<td>fly(^9) \rightarrow\ travel^{5}\</td>
</tr>
<tr>
<td>Troponym</td>
<td>From events to subordinate event (often via specific manner)</td>
<td>walk(^1) \rightarrow\ stroll(^1)</td>
</tr>
<tr>
<td>Entails</td>
<td>From verbs (events) to the verbs (events) they entail</td>
<td>snore(^1) \rightarrow\ sleep(^1)</td>
</tr>
<tr>
<td>Antonym</td>
<td>Semantic opposition between lemmas</td>
<td>increase(^1) \leftrightarrow\ decrease(^1)</td>
</tr>
<tr>
<td>Derivationally</td>
<td>Lemmas with same morphological root</td>
<td>destroy(^1) \leftrightarrow\ destruction(^1)</td>
</tr>
<tr>
<td>Related Form</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

• Not nearly as much information as nouns
Word similarity

• Human language words seem to have real-valued semantic distance (vs. logical objects)

• Two main approaches:
  – Thesaurus-based methods
    • E.g., WordNet-based
  – Distributional methods
    • Distributional “semantics”, vector “semantics”
    • More empirical, but affected by more than semantic similarity ("word relatedness")
Human-subject Word Associations

**Stimulus: wall**

Number of different answers: 39
Total count of all answers: 98
BRICK 16 0.16
STONE 9 0.09
PAPER 7 0.07
GAME 5 0.05
BLANK 4 0.04
BRICKS 4 0.04
FENCE 4 0.04
FLOWER 4 0.04
BERLIN 3 0.03
CEILING 3 0.03
HIGH 3 0.03
STREET 3 0.03
...

**Stimulus: giraffe**

Number of different answers: 26
Total count of all answers: 98
NECK 33 0.34
ANIMAL 9 0.09
ZOO 9 0.09
LONG 7 0.07
TALL 7 0.07
SPOTS 5 0.05
LONG NECK 4 0.04
AFRICA 3 0.03
ELEPHANT 2 0.02
HIPPOPOTAMUS 2 0.02
LEGS 2 0.02
...

From Edinburgh Word Association Thesaurus, [http://www.eat.rl.ac.uk/](http://www.eat.rl.ac.uk/)
Thesaurus-based Word Similarity

• Simplest approach: path length
Better approach: weighted links

- Use corpus stats to get probabilities of nodes
- Refinement: use info content of LCS:
  \[2 \times \log P(\text{g.f.})/(\log P(\text{hill}) + \log P(\text{coast})) = 0.59\]
Distributional Word Similarity

• Determine similarity of words by their *distribution* in a corpus
  – “You shall know a word by the company it keeps!” (Firth 1957)

• E.g.: 100k *dimension* vector, “1” if word occurs within “2 lines”:

<table>
<thead>
<tr>
<th></th>
<th>arts</th>
<th>boil</th>
<th>data</th>
<th>function</th>
<th>large</th>
<th>sugar</th>
<th>summarized</th>
<th>water</th>
</tr>
</thead>
<tbody>
<tr>
<td>apricot</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>pineapple</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>digital</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>information</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

• “Who is my neighbor?” Which functions?
Who is my neighbor?

• Linear window? 1-500 words wide. Or whole document. Remove *stop words*?

• Use dependency-parse relations? More expensive, but maybe better relatedness.

<table>
<thead>
<tr>
<th>cell</th>
<th>subj-of, absorb</th>
<th>subj-of, adapt</th>
<th>subj-of, behave</th>
<th>pobj-of, inside</th>
<th>pobj-of, into</th>
<th>nmod-of, abnormality</th>
<th>nmod-of, anemia</th>
<th>nmod-of, architecture</th>
<th>obj-of, attack</th>
<th>obj-of, call</th>
<th>obj-of, come from</th>
<th>obj-of, decorate</th>
<th>nmod, bacteria</th>
<th>nmod, body</th>
<th>nmod, bone marrow</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>16</td>
<td>30</td>
<td>3</td>
<td>8</td>
<td>1</td>
<td>6</td>
<td>11</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Weights vs. just counting

- Weight the counts by the *a priori* chance of co-occurrence
- Pointwise Mutual Information (PMI)
- Objects of *drink*:

<table>
<thead>
<tr>
<th>Object</th>
<th>Count</th>
<th>PMI Assoc</th>
<th>Object</th>
<th>Count</th>
<th>PMI Assoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>bunch beer</td>
<td>2</td>
<td>12.34</td>
<td>wine</td>
<td>2</td>
<td>9.34</td>
</tr>
<tr>
<td>tea</td>
<td>2</td>
<td>11.75</td>
<td>water</td>
<td>7</td>
<td>7.65</td>
</tr>
<tr>
<td>Pepsi</td>
<td>2</td>
<td>11.75</td>
<td>anything</td>
<td>3</td>
<td>5.15</td>
</tr>
<tr>
<td>champagne</td>
<td>4</td>
<td>11.75</td>
<td>much</td>
<td>3</td>
<td>5.15</td>
</tr>
<tr>
<td>liquid</td>
<td>2</td>
<td>10.53</td>
<td>it</td>
<td>3</td>
<td>1.25</td>
</tr>
<tr>
<td>beer</td>
<td>5</td>
<td>10.20</td>
<td>&lt;SOME AMOUNT&gt;</td>
<td>2</td>
<td>1.22</td>
</tr>
</tbody>
</table>
Distance between vectors

• Compare sparse high-dimensional vectors
  – Normalize for vector length
• Just use vector cosine?
• Several other functions come from IR community
Lots of functions to choose from

\[
\begin{align*}
\text{assoc}_{\text{prob}}(w, f) &= P(f|w) & (20.35) \\
\text{assoc}_{\text{PMI}}(w, f) &= \log_2 \frac{P(w,f)}{P(w)P(f)} & (20.38) \\
\text{assoc}_{\text{Lin}}(w, f) &= \log_2 \frac{P(w,f)}{P(w)P(r|w)P(w'|w)} & (20.39) \\
\text{assoc}_{\text{t-test}}(w, f) &= \frac{P(w,f)-P(w)P(f)}{\sqrt{P(f)P(w)}} & (20.41)
\end{align*}
\]

\[
\begin{align*}
\text{sim}_{\text{cosine}}(\vec{v}, \vec{w}) &= \frac{\vec{v} \cdot \vec{w}}{||\vec{v}|| ||\vec{w}||} = \frac{\sum_{i=1}^{N} v_i w_i}{\sqrt{\sum_{i=1}^{N} v_i^2} \sqrt{\sum_{i=1}^{N} w_i^2}} & (20.47) \\
\text{sim}_{\text{Jaccard}}(\vec{v}, \vec{w}) &= \frac{\sum_{i=1}^{N} \min(v_i, w_i)}{\sum_{i=1}^{N} \max(v_i, w_i)} & (20.48) \\
\text{sim}_{\text{Dice}}(\vec{v}, \vec{w}) &= \frac{2 \times \sum_{i=1}^{N} \min(v_i, w_i)}{\sum_{i=1}^{N} (v_i + w_i)} & (20.49) \\
\text{sim}_{\text{JS}}(\vec{v}||\vec{w}) &= D(\vec{v}|\frac{\vec{v}+\vec{w}}{2}) + D(\vec{w}|\frac{\vec{v}+\vec{w}}{2}) & (20.52)
\end{align*}
\]
Distributionally Similar Words

<table>
<thead>
<tr>
<th>Rum</th>
<th>Write</th>
<th>Ancient</th>
<th>Mathematics</th>
</tr>
</thead>
<tbody>
<tr>
<td>vodka</td>
<td>read</td>
<td>old</td>
<td>physics</td>
</tr>
<tr>
<td>cognac</td>
<td>speak</td>
<td>modern</td>
<td>biology</td>
</tr>
<tr>
<td>brandy</td>
<td>present</td>
<td>traditional</td>
<td>geology</td>
</tr>
<tr>
<td>whisky</td>
<td>receive</td>
<td>medieval</td>
<td>sociology</td>
</tr>
<tr>
<td>liquor</td>
<td>call</td>
<td>historic</td>
<td>psychology</td>
</tr>
<tr>
<td>detergent</td>
<td>release</td>
<td>famous</td>
<td>anthropology</td>
</tr>
<tr>
<td>cola</td>
<td>sign</td>
<td>original</td>
<td>astronomy</td>
</tr>
<tr>
<td>gin</td>
<td>offer</td>
<td>entire</td>
<td>arithmetic</td>
</tr>
<tr>
<td>lemonade</td>
<td>know</td>
<td>main</td>
<td>geography</td>
</tr>
<tr>
<td>cocoa</td>
<td>accept</td>
<td>indian</td>
<td>theology</td>
</tr>
<tr>
<td>chocolate</td>
<td>decide</td>
<td>various</td>
<td>hebrew</td>
</tr>
<tr>
<td>scotch</td>
<td>issue</td>
<td>single</td>
<td>economics</td>
</tr>
<tr>
<td>noodle</td>
<td>prepare</td>
<td>african</td>
<td>chemistry</td>
</tr>
<tr>
<td>tequila</td>
<td>consider</td>
<td>japanese</td>
<td>scripture</td>
</tr>
<tr>
<td>juice</td>
<td>publish</td>
<td>giant</td>
<td>biotechnology</td>
</tr>
</tbody>
</table>

(from an implementation of the method described in Lin. 1998. Automatic Retrieval and Clustering of Similar Words. COLING-ACL. Trained on newswire text.)
Human-subject Word Associations

**Stimulus:** *wall*

Number of different answers: 39
Total count of all answers: 98
BRICK 16 0.16
STONE 9 0.09
PAPER 7 0.07
GAME 5 0.05
BLANK 4 0.04
BRICKS 4 0.04
FENCE 4 0.04
FLOWER 4 0.04
BERLIN 3 0.03
CEILING 3 0.03
HIGH 3 0.03
STREET 3 0.03
...

**Stimulus:** *giraffe*

Number of different answers: 26
Total count of all answers: 98
NECK 33 0.34
ANIMAL 9 0.09
ZOO 9 0.09
LONG 7 0.07
TALL 7 0.07
SPOTS 5 0.05
LONG NECK 4 0.04
AFRICA 3 0.03
ELEPHANT 2 0.02
HIPPOPOTAMUS 2 0.02
LEGS 2 0.02
...

From Edinburgh Word Association Thesaurus, [http://www.eat.rl.ac.uk/](http://www.eat.rl.ac.uk/)
Recent events (2013-now)

- RNNs (Recurrent Neural Networks) as another way to get feature vectors
  - Hidden weights accumulate fuzzy info on words in the neighborhood
  - The set of hidden weights is used as the vector!
- Composition by multiplying (etc.)
  - Mikolov et al (2103): “king – man + woman = queen”(!?)
  - CCG with vectors as NP semantics, matrices as verb semantics(!?)
Event participants: A story

- Noah built an ark out of gopher wood.
- He loaded two of every animal onto the ark.
- Noah piloted the ark through stormy weather.
- When the skies cleared, all rejoiced.
- The sun came out and everyone on the ark celebrated.
Similar sentences

• Noah piloted an ark.
• Noah sailed an ark.
• Noah commanded an ark.
• Noah built an ark.
Syntactic constraints

• Noah piloted an ark. *Noah piloted.
• Noah sailed an ark. Noah sailed.
• Noah commanded an ark. *Noah commanded.
• Noah built an ark. *Noah built.
Semantic constraints

- Noah piloted an ark **through stormy seas**.
- Noah sailed an ark **through stormy seas**.
- Noah commanded an ark **through stormy seas**.
- #Noah built an ark **through stormy seas**.
- #An ark built Noah.
Rough paraphrases

• Noah built an ark out of gopher wood.
• An ark was built by Noah. It was made from gopher wood.
• Noah constructed an ark with wood from a gopher tree.
• Using gopher wood, Noah managed to put together an ark.
• Noah built an ark.
• ...
Predicates

• Noah **built** an ark out of gopher wood.
• An ark was **built** by Noah. It was **made** from gopher wood.
• Noah **constructed** an ark with wood from a gopher tree.
• Using gopher wood, Noah managed to **put together** an ark.
Predicates + Arguments

- Noah built an ark out of gopher wood.
- An ark was built by Noah. It was made from gopher wood.
- Noah constructed an ark with wood from a gopher tree.
- Using gopher wood, Noah managed to put together an ark.
Predicates + Arguments

- **Noah built** an ark out of gopher wood.
- He loaded two of every animal onto the ark.
- Noah piloted the ark through stormy weather.
- When the skies cleared, all rejoiced.
- The sun came out and everyone on the ark celebrated.
Breaking, Eating, Opening

• John broke the window.
• The window broke.
• John is always breaking things.

• We ate dinner.
• We already ate.
• The pies were eaten up quickly.

• Open up!
• Someone left the door open.
• John opens the window at night.
Breaking, Eating, Opening

- John broke the window.
- The window broke.
- John is always breaking things.

- We ate dinner.
- We already ate.
- The pies were eaten up quickly.

- Open up!
- Someone left the door open.
- John opens the window at night.
Can We Generalize?

• **Thematic roles** describe general patterns of participants in generic events.
• This gives us a kind of shallow, partial semantic representation.
• First proposed by Panini, before 400 BC!
# Thematic Roles

<table>
<thead>
<tr>
<th>Role</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agent</td>
<td>Volitional causer of the event</td>
<td>The waiter spilled the soup.</td>
</tr>
<tr>
<td>Force</td>
<td>Non-volitional causer of the event</td>
<td>The wind blew the leaves around.</td>
</tr>
<tr>
<td>Experiencer</td>
<td></td>
<td>Mary has a headache.</td>
</tr>
<tr>
<td>Theme</td>
<td>Most directly affected participant</td>
<td>Mary swallowed the pill.</td>
</tr>
<tr>
<td>Result</td>
<td>End-product of an event</td>
<td>We constructed a new building.</td>
</tr>
<tr>
<td>Content</td>
<td>Proposition of a propositional event</td>
<td>Mary knows you hate her.</td>
</tr>
<tr>
<td>Instrument</td>
<td></td>
<td>You shot her with a pistol.</td>
</tr>
<tr>
<td>Beneficiary</td>
<td></td>
<td>I made you a reservation.</td>
</tr>
<tr>
<td>Source</td>
<td>Origin of a transferred thing</td>
<td>I flew in from Pittsburgh.</td>
</tr>
<tr>
<td>Goal</td>
<td>Destination of a transferred thing</td>
<td>Go to hell!</td>
</tr>
</tbody>
</table>
Thematic Grid or Case Frame

• We can enumerate the possible sets of thematic roles of a particular verb.

• Example: break
  – agent/subject, theme/object
  – agent/subject, theme/object, instrument/PP_{with}
  – instrument/subject, theme/object
  – theme/subject

• Syntactic layout is not fully determined from the semantic roles (e.g., “give”)!
The Trouble With Thematic Roles

• They are not formally defined.
• They are overly general.
• “agent verb theme with instrument” and “instrument verb theme” ...
  – The cook opened the jar with the new gadget.
    → The new gadget opened the jar.
  – Susan ate the sliced banana with a fork.
    → #The fork ate the sliced banana.
Two Datasets

• Proposition Bank (PropBank): verb-specific thematic roles
• FrameNet: “frame”-specific thematic roles
Proposition Bank (PropBank)

- A set of **verb-sense-specific** “frames” with informal English glosses describing the roles
- Conventions for labeling optional modifier roles
- Penn Treebank, labeled with those verb-sense-specific semantic roles.
“Agree” in PropBank

- arg0: agreeer
- arg1: proposition
- arg2: other entity agreeing
- The group agreed it wouldn’t make an offer.
- Usually John agrees with Mary on everything
“Fall (move downward)” in PropBank

- arg1: logical subject, patient, thing falling
- arg2: extent, amount fallen
- arg3: starting point
- arg4: ending point
- argM-loc: medium
- Sales fell to $251.2 million from $278.8 million.
- The average junk bond fell by 4.2%.
- The meteor fell through the atmosphere, crashing into Cambridge.
FrameNet

- FrameNet is similar, but abstracts from specific verbs, so that semantic **frames** are first-class citizens.
- For example, there is a single frame called `change_position_on_a_scale`. 
Oil rose in price by 2%
It has increased to having them 1 day a month.
Microsoft shares fell to 7 5/8.
Colon cancer incidence fell by 50% among men.
FrameNet

• Frames are not just for verbs!
• **Verbs**: advance, climb, decline, decrease, diminish, dip, double, drop, dwindle, edge, explode, fall, fluctuate, gain, grow, increase, jump, move, mushroom, plummet, reach, rise, rocket, shift, skyrocket, slide, soar, swell, swing, triple, tumble
• **Nouns**: decline, decrease, escalation, explosion, fall, fluctuation, gain, growth, hike, increase, rise, shift, tumble
• **Adverb**: increasingly
FrameNet

• Includes inheritance and causation relationships among frames.
• Examples included, but little fully-annotated corpus data.
SemLink

• It would be really useful if these different resources were interconnected in a useful way.
• SemLink project is (was?) trying to do that
• Unified Verb Index (UVI) connects
  – PropBank
  – VerbNet
  – FrameNet
  – WordNet/OntoNotes
Semantic Role Labeling

- What are the arguments of each predicate (usually, the verb)? (predicate-argument structures)
- Somewhere between syntactic parsing and full-fledged compositional semantics.
Semantic Role Labeling

• Input: sentence
• Output: segmentation into roles, with labels

• Example:
  • \([\text{arg}_0 \text{ The Examiner}] \text{ issued } [\text{arg}_1 \text{ a special edition}] [\text{arg}_M \text{-tmp yesterday}]\)
Semantic Role Labeling: How It Works

• First, parse.

• For each predicate word in the parse:
  – For each node in the parse:
    • **Classify** the node with respect to the predicate.
Yet Another Classification Problem!

• As before, there are many techniques (e.g., Naïve Bayes)
• Key: what features?
Features for Semantic Role Labeling

• What is the predicate?
• Phrase type of the constituent
• Head word of the constituent, its POS
• Path in the parse tree from the constituent to the predicate
• Active or passive
• Is the phrase before or after the predicate?
• Subcategorization (≈ grammar rule) of the predicate
Feature example

• Example sentence:
  - [arg0 The Examiner] issued [arg1 a special edition] [argM-tmp yesterday]

• Arg0 features:
  issued, NP, Examiner, NNP, path, active, before, VP->VBD NP PP
Figure 20.16: Parse tree for a PropBank sentence, showing the PropBank argument labels. The dotted line shows the path feature NP ↑ S ↓ VP ↓ VBD for ARG0, the NP-SBJ constituent *The San Francisco Examiner*. 
Additional Issues

• Initial filtering of non-arguments
• Using chunking or partial parsing instead of full parsing
• Enforcing consistency (e.g., non-overlap)
• Phrasal verbs, support verbs/light verbs
  – *take a nap*: verb *take* is syntactic head of VP, but predicate is *napping*, not *taking*
SEMAFØR

• A FrameNet-based semantic role labeling system developed within Noah’s research group
  ‣ It uses a dependency parser (the MST Parser) for preprocessing
  ‣ Identifies and disambiguates predicates; then identifies and disambiguates each predicate’s arguments
Questions?
Noun compounds

- A very flexible (*productive*) syntactic structure in English
  - The *noun noun* pattern is easily applied to name new concepts (*Web browser*) and to disambiguate known concepts (*fire truck*)
  - Can also combine two NPs: incumbent protection plan, [*undergraduate* [*computer science*] [*lecture course*]]
  - Sometimes creates ambiguity, esp. in writing where there is no phonological stress: *Spanish teacher*
  - People are creative about interpreting even nonsensical compounds

- Also present in many other languages, sometimes with special morphology
  - German is infamous for loving to merge words into compounds. e.g. *Fremdsprachenkenntnisse*, ‘knowledge of foreign languages’
Noun compounds

• SemEval 2007 task: **Classification of Semantic Relations between Nominals**
  ‣ *7 predefined relation types*
  1. Cause-Effect: flu virus
  2. Instrument-User: laser printer
  3. Product-Producer: honeybee
  4. Origin-Entity: rye whiskey
  5. Purpose-Tool: soup pot
  6. Part-Whole: car wheel
  7. Content-Container: apple basket

• [http://nlp.cs.swarthmore.edu/semeval/tasks/task04/description.shtml](http://nlp.cs.swarthmore.edu/semeval/tasks/task04/description.shtml)
Noun compounds

- SemEval 2010 task: **Noun compound interpretation using paraphrasing verbs**
  - A dataset was compiled in which subjects were presented with a noun compound and asked to provide a verb describing the relationship
  - *nut bread* elicited: contain(21); include(10); be made with(9); have(8); be made from(5); use(3); be made using(3); feature(2); be filled with(2); taste like(2); be made of(2); come from(2); consist of(2); hold(1); be composed of(1); be blended with(1); be created out of(1); encapsulate(1); diffuse(1); be created with(1); be flavored with(1)

Thesaurus/dictionary-based similarity measures

\[
\begin{align*}
\text{sim}_{\text{path}}(c_1, c_2) &= -\log \text{pathlen}(c_1, c_2) \\
\text{sim}_{\text{Resnik}}(c_1, c_2) &= -\log P(\text{LCS}(c_1, c_2)) \\
\text{sim}_{\text{Lin}}(c_1, c_2) &= \frac{2 \times \log P(\text{LCS}(c_1, c_2))}{\log P(c_1) + \log P(c_2)} \\
\text{sim}_{\text{jc}}(c_1, c_2) &= \frac{1}{2 \times \log P(\text{LCS}(c_1, c_2)) - (\log P(c_1) + \log P(c_2))} \\
\text{sim}_{\text{eLesk}}(c_1, c_2) &= \sum_{r, q \in \text{RELS}} \text{overlap}(\text{gloss}(r(c_1)), \text{gloss}(q(c_2)))
\end{align*}
\]
Euclidean$(\vec{a}, \vec{b}) = L_2(\vec{a}, \vec{b})$

Manhattan$(\vec{a}, \vec{b}) = L_1(\vec{a}, \vec{b})$