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

11-711 Algorithms for NLP
13 November 2018
(With thanks to Noah Smith and Lori Levin)

## Semantics so far in course

- 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

## 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.

## 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

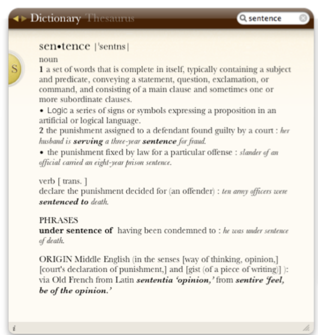
## How Many Senses?

- 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.*

## Related Phenomena

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

## Word Senses and Dictionaries

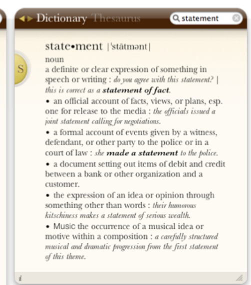

Dictionary Thesaurus sentence

sen•tence |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.*

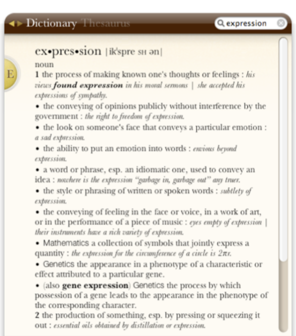
PHRASES
**under sentence of** having been condemned to : *he was under sentence of death.*

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.'**


Dictionary Thesaurus statement

state•ment |ˈstātmənt|
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.*

## Word Senses and Dictionaries


Dictionary Thesaurus expression

ex•pres•sion |ikˈspre sн ən|
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 : *nowhere 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$^1$ - (the lowest part of the musical range)
2. bass$^2$, bass part$^1$ - (the lowest part in polyphonic music)
3. bass$^3$, basso$^1$ - (an adult male singer with the lowest voice)
4. sea bass$^1$, bass$^4$ - (the lean flesh of a saltwater fish of the family Serranidae)
5. freshwater bass$^1$, bass$^5$ - (any of various North American freshwater fish with lean flesh (especially of the genus Micropterus))
6. bass$^6$, bass voice$^1$, basso$^2$ - (the lowest adult male singing voice)
7. bass$^7$ - (the member with the lowest range of a family of musical instruments)
8. bass$^8$ - (nontechnical name for any of numerous edible marine and freshwater spiny-finned fishes)

The adjective "bass" has 1 sense in WordNet.

1. bass$^1$, deep$^6$ - (having or denoting a low vocal or instrumental range)
   *"a deep voice"; "a bass voice is lower than a baritone voice"; "a bass clarinet"*

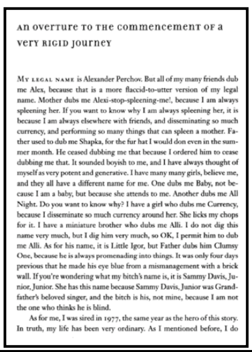
# Synsets

- (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-spleening-me!, because I am always spleening her. If you want to know why I am always spleening 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

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

```
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
```

# Is a hamburger food?

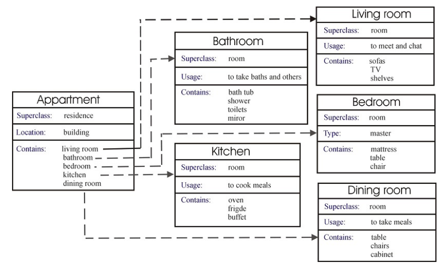
```
Sense 1
hamburger, beefburger --
(a fried cake of minced beef served on a bun)
=> sandwich
   => snack food
      => dish
         => nutriment, nourishment, nutrition...
            => food, nutrient
               => substance
                  => matter
                     => physical entity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

## Verb relations in WordNet3.0

| Relation | Definition | Example |
|---|---|---|
| Hypernym | From events to superordinate events | *fly*$^9$ → *travel*$^5$ |
| Troponym | From events to subordinate event (often via specific manner) | *walk*$^1$ → *stroll*$^1$ |
| Entails | From verbs (events) to the verbs (events) they entail | *snore*$^1$ → *sleep*$^1$ |
| Antonym | Semantic opposition between lemmas | *increase*$^1$ ⟺ *decrease*$^1$ |
| Derivationally Related Form | Lemmas with same morphological root | *destroy*$^1$ ⟺ *destruction*$^1$ |

- Not nearly as much information as nouns

## Frame based Knowledge Rep.




- Organize relations around concepts
- Equivalent to (or weaker than) FOPC
  - *Image from futurehumanevolution.com*

## Still no "real" semantics?

- Semantic primitives:
  Kill(x,y) = CAUSE(x, BECOME(NOT(ALIVE(y))))
  Open(x,y) = CAUSE(x, BECOME(OPEN(y)))
- Conceptual Dependency: PTRANS,ATRANS,…
  The waiter brought Mary the check
  PTRANS(x)∧ACTOR(x,Waiter)∧(OBJECT(x,Check)
  ∧TO(x,Mary)
  ∧ATRANS(y)∧ACTOR(y,Waiter)∧(OBJECT(y,Check)
  ∧TO(y,Mary)

## 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
…

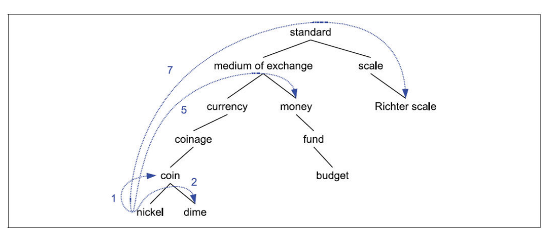
**Stimulus:** ***giraffe***

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/

## Thesaurus-based Word Similarity

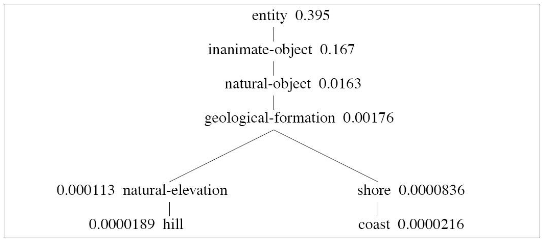
- Simplest approach: path length



## Better approach: weighted links

- Use corpus stats to get probabilities of nodes
- Refinement: use info content of LCS:

  2*logP(g.f.)/(logP(hill) + logP(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":

| | arts | boil | data | function | large | sugar | summarized | water |
|---|---|---|---|---|---|---|---|---|
| **apricot** | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |
| **pineapple** | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |
| **digital** | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 |
| **information** | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 |

- "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.

| | *subj-of*, absorb | *subj-of*, adapt | *subj-of*, behave | ... | *pobj-of*, inside | *pobj-of*, into | ... | *nmod-of*, abnormality | *nmod-of*, anemia | *nmod-of*, architecture | ... | *obj-of*, attack | *obj-of*, call | *obj-of*, come from | *obj-of*, decorate | ... | *nmod*, bacteria | *nmod*, body | *nmod*, bone marrow |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| cell | 1 | 1 | 1 | | 16 | 30 | | 3 | 8 | 1 | | 6 | 11 | 3 | 2 | | 3 | 2 | 2 |

## Weights vs. just counting

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

| Object | Count | PMI Assoc | Object | Count | PMI Assoc |
|---|---|---|---|---|---|
| bunch beer | 2 | 12.34 | wine | 2 | 9.34 |
| tea | 2 | 11.75 | water | 7 | 7.65 |
| Pepsi | 2 | 11.75 | anything | 3 | 5.15 |
| champagne | 4 | 11.75 | much | 3 | 5.15 |
| liquid | 2 | 10.53 | it | 3 | 1.25 |
| beer | 5 | 10.20 | <SOME AMOUNT> | 2 | 1.22 |

## 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

$$\text{assoc}_{\text{prob}}(w,f) = P(f|w) \qquad (20.35)$$

$$\text{assoc}_{\text{PMI}}(w,f) = \log_2 \frac{P(w,f)}{P(w)P(f)} \qquad (20.38)$$

$$\text{assoc}_{\text{Lin}}(w,f) = \log_2 \frac{P(w,f)}{P(w)P(r|w)P(w'|w)} \qquad (20.39)$$

$$\text{assoc}_{\text{t-test}}(w,f) = \frac{P(w,f)-P(w)P(f)}{\sqrt{P(f)P(w)}} \qquad (20.41)$$

$$\text{sim}_{\text{cosine}}(\vec{v},\vec{w}) = \frac{\vec{v}\cdot\vec{w}}{|\vec{v}||\vec{w}|} = \frac{\sum_{i=1}^{N} v_i \times w_i}{\sqrt{\sum_{i=1}^{N} v_i^2}\sqrt{\sum_{i=1}^{N} w_i^2}} \qquad (20.47)$$

$$\text{sim}_{\text{Jaccard}}(\vec{v},\vec{w}) = \frac{\sum_{i=1}^{N} \min(v_i,w_i)}{\sum_{i=1}^{N} \max(v_i,w_i)} \qquad (20.48)$$

$$\text{sim}_{\text{Dice}}(\vec{v},\vec{w}) = \frac{2\times\sum_{i=1}^{N} \min(v_i,w_i)}{\sum_{i=1}^{N}(v_i+w_i)} \qquad (20.49)$$

$$\text{sim}_{\text{JS}}(\vec{v}||\vec{w}) = D(\vec{v}|\tfrac{\vec{v}+\vec{w}}{2}) + D(\vec{w}|\tfrac{\vec{v}+\vec{w}}{2}) \qquad (20.52)$$

## Distributionally Similar Words

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

31

(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

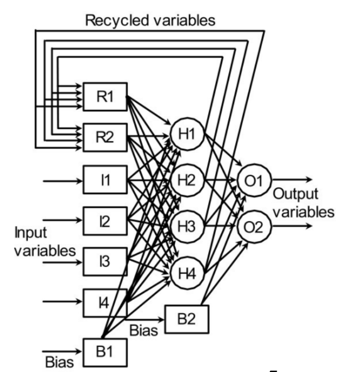
NECK 33 0.34
ANIMAL 9 0.09
ZOO 9 0.09
LONG 7 0.07
TALL 7 0.07
SPOTS 5 0.05
LONG NECK 4 0.04
AFRICA 3 0.03
ELEPHANT 2 0.02
HIPPOPOTAMUS 2 0.02
LEGS 2 0.02
...

From Edinburgh Word Association Thesaurus, http://www.eat.rl.ac.uk/

## 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!

## RNNs




*From openi.nlm.nih.gov*

## 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 (2013): "king – man + woman = queen"(!?)
  - CCG with vectors as NP semantics, matrices as verb semantics(!?)

## Semantic Cases/Thematic Roles

- Developed in late 1960's and 1970's
- Postulate a limited set of abstract semantic relationships between a verb & its arguments: *thematic roles* or *case roles*
- In some sense, part of the verb's semantics

36 Semantic Processing [2]

## Problem: Mismatch between FOPC and linguistic arguments

- *John broke the window with a hammer.*
  - Broke(j,w,h)
- *The hammer broke the window.*
  - Broke(h,w)
- *The window broke.*
  - Broke(w)

- Relationship between 1st argument and the predicate is implicit, inaccessible to the system

## 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.<br>• The window broke.<br>• John is always breaking things. | breaker,<br>broken thing,<br>breaking frequency? |
| • We ate dinner.<br>• We already ate.<br>• The pies were eaten up quickly. | eater,<br>eaten thing,<br>eating speed? |
| • Open up!<br>• Someone left the door open.<br>• John opens the window at night. | opener,<br>opened thing,<br>opening time? |

## Thematic Role example

- *John broke the window with the hammer*
- *John*: AGENT role
  *window*: THEME role
  *hammer*: INSTRUMENT role
- Extend LF notation to use semantic roles



## Thematic Roles

- Is there a precise way to define meaning of AGENT, THEME, etc.?
- By definition:
  - "The AGENT is an instigator of the action described by the sentence."
- Testing via sentence rewrite:
  - *John <u>intentionally</u> broke the window*
  - **The hammer <u>intentionally</u> broke the window*



## Thematic Roles [2]

- THEME
  - Describes the primary object undergoing some change or being acted upon
  - For transitive verb X, "what was Xed?"
  - *The gray eagle saw the mouse*
    "What was seen?" (A: the mouse)

## 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

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

## Verb Subcategorization

Verbs have sets of allowed args. Could have many sets of VP rules. Instead, have a SUBCAT feature, marking sets of allowed arguments:

- +none -- Jack laughed
- +np -- Jack found a key
- +np+np -- Jack gave Sue the paper
- +vp:inf -- Jack wants to fly
- +np+vp:inf -- Jack told the man to go
- +vp:ing -- Jack keeps hoping for the best
- +np+vp:ing -- Jack caught Sam looking at his desk
- +np+vp:base -- Jack watched Sam look at his desk
- +np+pp:to -- Jack gave the key to the man
- +pp:loc -- Jack is at the store
- +np+pp:loc -- Jack put the box in the corner
- +pp:mot -- Jack went to the store
- +np+pp:mot -- Jack took the hat to the party
- +adjp -- Jack is happy
- +np+adjp -- Jack kept the dinner hot
- +sthat -- Jack believed that the world was flat
- +sfor -- Jack hoped for the man to win a prize

50-100 possible ***frames*** for English; a single verb can have several.
*(Notation from James Allen “Natural Language Understanding”)*

## Thematic Grid or Case Frame

- Example: break
  - The child broke the vase. < agent theme >
    subj obj
  - The child broke the vase with a hammer.
    < agent theme instr >
    subj obj PP
  - The hammer broke the vase. < theme instr >
    obj subj
  - The vase broke. < theme >
    subj

## Thematic Grid or Case Frame

- Example: break
  - The child broke the vase. < agent theme >
    subj obj
  - The child broke the vase with a hammer.
    < agent theme instr >
    subj obj PP
  - The hammer broke the vase. < theme instr >
    obj subj
  - The vase broke. < theme >
    subj

The Thematic Grid or Case Frame shows
- How many arguments the verb has
- What roles the arguments have
- Where to find each argument
  - For example, you can find the agent in the subject position

## Diathesis Alternation:

a change in the number of arguments or the grammatical relations associated with each argument

| | | | |
|---|---|---|---|
| Chris gave a book to Dana. | < agent<br>subj | theme<br>obj | goal ><br>PP |
| A book was given to Dana by Chris. | < agent<br>PP | theme<br>subj | goal ><br>PP |
| Chris gave Dana a book. | < agent<br>subj | theme<br>obj2 | goal ><br>obj |
| Dana was given a book by Chris. | < agent<br>PP | theme<br>obj | goal ><br>subj |

## 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

- These are lexicons containing case frames/thematic grids for each verb.

## 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 is labeled with those verb-sense-specific semantic roles.

## "Agree" in PropBank

- arg0: agreer
- 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**.

## change_position_on_a_scale

| Core Roles | |
|---|---|
| ATTRIBUTE | The ATTRIBUTE is a scalar property that the ITEM possesses. |
| DIFFERENCE | The distance by which an ITEM changes its position on the scale. |
| FINAL_STATE | A description that presents the ITEM's state after the change in the ATTRIBUTE's value as an independent predication. |
| FINAL_VALUE | The position on the scale where the Item ends up. |
| INITIAL_STATE | A description that presents the ITEM's state before the change in the ATTRIBUTE's value as an independent predication. |
| INITIAL_VALUE | The initial position on the scale from which the ITEM moves away. |
| ITEM | The entity that has a position on the scale. |
| VALUE_RANGE | A portion of the scale, typically identified by its end points, along which the values of the ATTRIBUTE fluctuate. |
| **Some Non-Core Roles** | |
| DURATION | The length of time over which the change takes place. |
| SPEED | The rate of change of the VALUE. |
| GROUP | The GROUP in which an ITEM changes the value of an ATTRIBUTE in a specified way. |

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.

**Many words, not just verbs, share the same frame:**

**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

## Conversely, one word has many frames

Example: rise

- **Change-position-on-a-scale**: Oil ROSE in price by two percent.
- **Change-posture**: a protagonist changes the overall position or posture of a body.
  - Source: starting point of the change of posture.
  - Charles ROSE from his armchair.
- **Get-up**: A Protagonist leaves the place where they have slept, their Bed, to begin or resume domestic, professional, or other activities. Getting up is distinct from Waking up, which is concerned only with the transition from the sleeping state to a wakeful state.
  - I ROSE from bed, threw on a pair of camouflage shorts and drove my little Toyota Corolla to a construction clearing a few miles away.
- **Motion-directional**: In this frame a Theme moves in a certain Direction which is often determined by gravity or other natural, physical forces. The Theme is not necessarily a self-mover.
  - The balloon ROSE upward.
- **Sidereal-appearance:** An Astronomical_entity comes into view above the horizon as part of a regular, periodic process of (apparent) motion of theAstronomical_entity across the sky. In the case of the sun, the appearance begins the day.
  - At the time of the new moon, the moon RISES at about the same time the sun rises, and it sets at about the same time the sun sets.

  Each day the sun's RISE offers us a new day.

## 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

- Input: sentence
- Output: for each predicate*, labeled spans identifying each of its arguments.
- Example:

  [$_{agent}$ The batter] hit [$_{patient}$ the ball] [$_{time}$ yesterday]

- Somewhere between syntactic parsing and full-fledged compositional semantics.

*Predicates are sometimes identified in the input, sometimes not.

## But wait. How is this different from dependency parsing?

- Semantic role labeling
  - [$_{agent}$ The batter] hit [$_{patient}$ the ball] [$_{time}$ yesterday]
- Dependency parsing
  - [$_{subj}$ The batter] hit [$_{obj}$ the ball] [$_{mod}$ yesterday]

## But wait. How is this different from dependency parsing?

- Semantic role labeling
  - [$_{agent}$ The batter] hit [$_{patient}$ the ball] [$_{time}$ yesterday]
- Dependency parsing
  - [$_{subj}$ The batter] hit [$_{obj}$ the ball] [$_{mod}$ yesterday]

1. These are not the same task.
2. Semantic role labeling is much harder.

## Subject vs agent

- Subject is a grammatical relation
- Agent is a semantic role

- In English, a subject has these properties
  - It comes before the verb
  - If it is a pronoun, it is in nominative case (in a finite clause)
    - I/he/she/we/they hit the ball.
    - *Me/him/her/us/them hit the ball.
  - If the verb is in present tense, it agrees with the subject
    - She/he/it hits the ball.
    - I/we/they hit the ball.
    - *She/he/it hit the ball.
    - *I/we/they hits the ball.
    - I hit the ball.
    - I hit the balls.

## Subject vs agent

- In the most typical sentences (for some definition of "typical"), the agent is the subject:
  - The batter hit the ball.
  - Chris opened the door.
  - The teacher gave books to the students.

- Sometimes the agent is not the subject:
  - The ball was hit by the batter.
  - The balls were hit by the batter.

- Sometimes the subject is not the agent:
  - The door opened.
  - The key opened the door.
  - The students were given books.
  - Books were given to the students.

## Semantic Role Labeling

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

- Example from book:
- [$_{arg0}$ The Examiner] issued [$_{arg1}$ a special edition] [$_{argM\text{-}tmp}$ 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] [$_{argM\text{-}tmp}$ yesterday]
- Arg0 features:
  issued, NP, Examiner, NNP, ***path***, active, before, VP->VBD NP PP

## Example

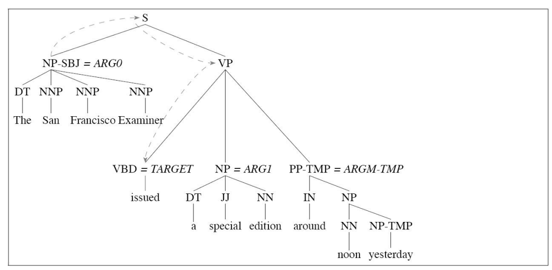



Figure 20.16: Parse tree for a PropBank sentence, showing the PropBank argument labels. The dotted line shows the **path** feature NP ↑ S ↓ VP ↓ VBD for ARG0, the NP-SBJ constituent *The San Francisco Examiner*.

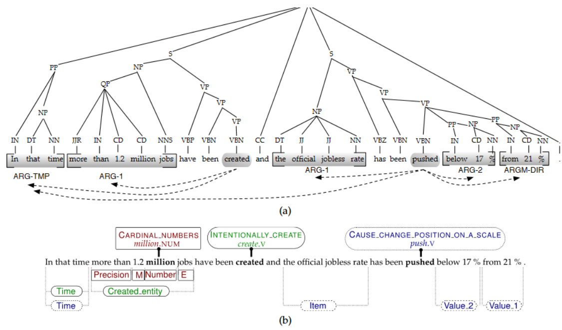
## Additional Issues

- Initial filtering of non-arguments
- Using chunking or partial parsing instead of full parsing
- Enforcing consistency (e.g., non-overlap, only one arg0)
- Phrasal verbs, support verbs/light verbs
  - *take a nap*: verb *take* is syntactic head of VP, but predicate is **napping**, not **taking**

## Two datasets, two systems

- Example from book uses PropBank
- Locally-developed system SEMAFOR works on SemEval problem, based on FrameNet

## PropBank vs FrameNet




## Shallow approaches to deep problems

- For many problems:
  - Shallow approaches much easier to develop
    - As in, *possible at all* for unlimited vocabularies
  - Not wonderful performance yet
    - Sometimes claimed to help a particular system, but often doesn't seem to help
  - Definitions are not crisp
    - There clearly is *something* there, but the granularity of the distinctions very problematic
- Deep Learning will fix everything?

## Questions?

## Similarities to WSD

- Pick correct choice from N ambiguous possibilities
- Definitions are not crisp
- Need to pick a labelling scheme, corpus
  - Choices have big effect on performance, usefulness

## Shallow approaches to deep problems

- For both WSD and SRL:
  - Shallow approaches much easier to develop
    - As in, *possible at all* for unlimited vocabularies
  - Not wonderful performance yet
    - Sometimes claimed to help a particular system, but often doesn't seem to help
  - Definitions are not crisp
    - There clearly is *something* there, but the granularity of the distinctions very problematic
- Deep Learning will fix everything?

## SEMAFOR

- 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
  - Trained on frame-annotated corpora from SemEval 2007/2010 tasks. Domains: weapons reports, travel guides, news, Sherlock Holmes stories.

## 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

## 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)
- http://semeval2.fbk.eu/semeval2.php?location=tasks#T12

## Thesaurus/dictionary-based similarity measures

$$
\begin{aligned}
\text{sim}_{\text{path}}(c_1,c_2) &= -\log \text{pathlen}(c_1,c_2)\\
\text{sim}_{\text{Resnik}}(c_1,c_2) &= -\log P(\text{LCS}(c_1,c_2))\\
\text{sim}_{\text{Lin}}(c_1,c_2) &= \frac{2\times\log P(LCS(c_1,c_2))}{\log P(c_1)+\log P(c_2)}\\
\text{sim}_{\text{jc}}(c_1,c_2) &= \frac{1}{2\times\log P(\text{LCS}(c_1,c_2))-(\log P(c_1)+\log P(c_2))}\\
\text{sim}_{\text{eLesk}}(c_1,c_2) &= \sum_{r,q\in\text{RELS}} \text{overlap}(\text{gloss}(r(c_1)),\text{gloss}(q(c_2)))
\end{aligned}
$$