6.863J Natural Language Processing
Lecture 15: Word semantics - Working with Wordnet

Robert C. Berwick
berwick@ai.mit.edu

The Menu Bar
• Administrivia:
• Lab 4 due April 9 (Weds.);
  • Start w/ final projects, unless there are objections
• Agenda:
• Working with Wordnet
  • What’s Wordnet
  • What can we do with it?
  • Solving some reasoning problems:
    • Mending a torn dress
    • Enjoying a movie; What’s a shelf?
    • Implementing EVCA and Wordnet together
Wordnet motivation

But people have persistent problem. When they look up a word, especially a commonly used word, they often find a dozen or more different meanings. What the dictionary does not make clear are the contexts in which each of these different meanings would be understood. So we know what kind of information is required, but we have not yet learned how to provide it to a computer.

(G. Miller, U.S./Japan Joint Workshop on Electronic Dictionaries and Language Technologies January 23--25, 1993.)

What’s Wordnet?

• Psychological motivation
• Nouns, verbs, adjectives organized into (fairly) distinct networks of
• Synonym Sets (synsets)
• Each synset = 1 concept
• Supposedly intersubstitutable within synset (“synonomy”)
Practical motivation

• What’s not in a dictionary?

• Take example, like tree – “large, woody perennial plant with a distinct trunk”

• What info is missing?

Psychological motivation

• Why these categories?
• Words association: first word thought of drawn from difft syntactic categories
• Modal response – same as probe: noun probes elicit nouns 79% of the time; verbs, v’s, 43%; adjs, adjs 65%
• Not just contiguity (since that fails)
• “Middle level” descriptions for nouns
Psychological motivation

• Where do categories come from??

• How do we ‘carve up’ nature at its joints?

Synonymy

• Two entries synonyms if the can be substituted in some context

• set of synonyms = Synset
  • {chump, fish, fool, gull, mark, patsy, fall guy, sucker, schlemiel, shlemiel, soft touch, mug}
  • “easy to take advantage of”
  • A concept that has been lexicalized
Basic stats

<table>
<thead>
<tr>
<th>POS</th>
<th>Unique Strings</th>
<th>Synsets</th>
<th>Word-Sense Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>109195</td>
<td>75804</td>
<td>134716</td>
</tr>
<tr>
<td>Verb</td>
<td>11088</td>
<td>13214</td>
<td>24169</td>
</tr>
<tr>
<td>Adjective</td>
<td>21460</td>
<td>18576</td>
<td>31184</td>
</tr>
<tr>
<td>Adverb</td>
<td>4607</td>
<td>3629</td>
<td>5748</td>
</tr>
<tr>
<td>Totals</td>
<td>146350</td>
<td>111223</td>
<td>195817</td>
</tr>
</tbody>
</table>

Example synset

cigarette
Add relations... then stir

• Concepts related through (possibly iterated) applications of basic relation:
  1. Is-a relation (Hypernym): from concept to superordinate (denoted → )
     e.g., breakfast → meal
     This is unidirectional
  2. Meronomy ("part of")
  3. Antonymy (opposite)

Note: multiple inheritance;
No unique ‘top’ concept: dirt has top concept entity; while gossip has act
Is-a merges function/nonfunctional isa relations
Wordnet Relations

cigarette, butt, fag, coffin nail

↓

smoke ←−−−−→ smoke
down

↓

 tobacco

↑

plant product

street drug

↑

substance

artifact ←−−−−→ create

↑

physical object ← verbs of perception

Wordnet Relations

<table>
<thead>
<tr>
<th>X hyp Y</th>
<th>Y hypernym of X</th>
<th>x-repair, y-improve</th>
</tr>
</thead>
<tbody>
<tr>
<td>X ent Y</td>
<td>X entails Y</td>
<td>x-breathe, y-inhale</td>
</tr>
<tr>
<td>X sim Y</td>
<td>(adj) Y similar X</td>
<td>x-achromatic, y-white</td>
</tr>
<tr>
<td>X cs Y</td>
<td>Y is a cause of X</td>
<td>x-anesthetize, y-sleep</td>
</tr>
<tr>
<td>X vgp Y</td>
<td>(verbs) Y similar X</td>
<td>x-behave, y-pretend</td>
</tr>
<tr>
<td>X ant Y</td>
<td>X, Y antonyms</td>
<td>x-present, y-absent</td>
</tr>
<tr>
<td>X sa Y</td>
<td>X, see also Y</td>
<td>x-breathe, y-breathe out</td>
</tr>
<tr>
<td>X ppl Y</td>
<td>X participle of Y</td>
<td>x-applied, y-apply</td>
</tr>
<tr>
<td>X per y</td>
<td>X pertains to Y</td>
<td>x-abaxial, y-axial</td>
</tr>
</tbody>
</table>
### Noun relations

<table>
<thead>
<tr>
<th>Relation</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypernym</td>
<td>From concepts to superordinates</td>
<td>breakfast → meal</td>
</tr>
<tr>
<td>Hyponym</td>
<td>From concepts to subtypes</td>
<td>meal → lunch</td>
</tr>
<tr>
<td>Has-Member</td>
<td>From groups to their members</td>
<td>faculty → professor</td>
</tr>
<tr>
<td>Member-Of</td>
<td>From members to their groups</td>
<td>copilot → crew</td>
</tr>
<tr>
<td>Has-Part</td>
<td>From wholes to parts</td>
<td>table → leg</td>
</tr>
<tr>
<td>Part-Of</td>
<td>From parts to wholes</td>
<td>course → meal</td>
</tr>
<tr>
<td>Antonym</td>
<td>Opposites</td>
<td>leader → follower</td>
</tr>
</tbody>
</table>

### Verb relations

<table>
<thead>
<tr>
<th>Relation</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypernym</td>
<td>From events to superordinate events</td>
<td>fly → travel</td>
</tr>
<tr>
<td>Troponym</td>
<td>From events to their subtypes</td>
<td>walk → stroll</td>
</tr>
<tr>
<td>Entails</td>
<td>From events to the events they entail</td>
<td>snore → sleep</td>
</tr>
<tr>
<td>Antonym</td>
<td>Opposites</td>
<td>increase ↔ decrease</td>
</tr>
</tbody>
</table>
Wordnet: why?

- Can draw inferences about some events
- We’ll give 3 case studies...

- Consider first:
  1. Susan mended the torn dress
  2. Susan mended the red dress

- Mend refers to some action, resulting in a change of state for direct objects

The inferential questions

- For 1: After the dress is mended, is it still torn?
- For 2: After the dress is mended, is it still red?
- Call this ‘semantic opposition’, e.g.:
  - The woman on the boat jumped into the river
  - The prisoner escaped from prison
“event template” idea doesn’t help

• \([x \text{ cause } [\text{become } y \text{ <mended}>] ]\)

And more examples

• The plumber fixed every leaky faucet
• The plumber fixed every blue faucet
• Mary fixed the flat tire
• The mother comforted the crying child
• John painted the white house blue
• Mary rescued the drowning man

These are all examples of the famous...
Frame Problem in AI!

- Which things remain the same in a changing world?
- “except for things explicitly known to change, everything else remains the same”

Yet other examples

- Not just change of state verbs:
  - John brushed the dirty carpet
  - John brushed the dirty carpet clean
- (seep, wipe, broom, paint,...)
Wordnet solution

- Use transitivity of hypernym relation

- Given adjective & change of state verb:
  - Compute shortest path between them in Wordnet
  - If antonym exists on the path, then cancel the adjective
  - Otherwise, the adjective still prevails
- A test of transitivity, and Wordnet

Example

- mend vs. tear
  - Repair is in same synset
  - Break and bust#1 are in same synset
  - Bust#1 and Bust#3 both verbs of contact
  - Bust#3 in same synset as tear
  - So chain looks like this:
This path or bust

Can’t get there from here

- 5 other ways (longer chains) between mend and tear (one w/o antonym):
  1. Mend and fix in same synset
  2. Fix1 and Fix3 in synsets related by verb change
  3. Fix3 is an instance of attach
  4. Attach1 and attach3 in synsets related by verb contact
  5. Attach3 instance of touch
  6. Touch1 and touch3 in synsets related by verb contact
  7. Touch see also touch down
  8. Touch down instance of land
  9. Land and shoot down in same synset
  10. Shoot down1 and shoot down2 in synsets related by verb of motion
  11. Shoot down2 and tear in same synset
OK, how does this work?

• Do BFS on nodes - unidirectional
• Bi-directional would work even better
• 11 test examples 9 work, 2 don’t

Mary rescued...

Parsing: Mary rescued the drowning man
drown/v (200329171,1)
instance of eliminate/v (200328742,1)
instance of destroy/v (20114042,1)
instance of unmake/v (201113462,1)
and make/v (20113245,2) are antonyms
and make/v (201185771,4) in synsets related by verb.creation
instance of direct/v (201651432,1)
instance of deal/v (201658906,2)
and deal/v (201619807,1) in synsets related by verb.social
instance of deport/v (201716569,4)
and deport/v (201706176,3) in synsets related by verb.social
and deliver/v (201706176,2) in same synset
and deliver/v (201739567,2) in synsets related by verb.social
and rescue/v (201739567,1) in same synset
### Test cases

<table>
<thead>
<tr>
<th>Pair</th>
<th>Chain len</th>
<th>Semantic opposition</th>
<th>Search size</th>
</tr>
</thead>
<tbody>
<tr>
<td>mend-tear</td>
<td>5</td>
<td>Yes</td>
<td>1261</td>
</tr>
<tr>
<td>mend-red</td>
<td>-</td>
<td>No</td>
<td>11974</td>
</tr>
<tr>
<td>fix-leaky</td>
<td>5</td>
<td>Yes</td>
<td>12167</td>
</tr>
<tr>
<td>fix-blue</td>
<td>11</td>
<td>No</td>
<td>14553</td>
</tr>
<tr>
<td>mix-powdered</td>
<td>6</td>
<td>Yes</td>
<td>11931</td>
</tr>
<tr>
<td>comfort-crying</td>
<td>9</td>
<td>Yes</td>
<td>11359</td>
</tr>
<tr>
<td>blue-white</td>
<td>-</td>
<td>No</td>
<td>24432</td>
</tr>
<tr>
<td>rescue-drowning</td>
<td>13</td>
<td>Yes</td>
<td>9142</td>
</tr>
<tr>
<td>clean-dirty</td>
<td>1</td>
<td>Yes</td>
<td>61</td>
</tr>
<tr>
<td>fill-empty</td>
<td>1</td>
<td>Yes</td>
<td>48</td>
</tr>
</tbody>
</table>

### Why the failures - analysis

- Can we reduce length threshold below 11? No...
- Why does the color system fail?
Color system in Wordnet

argent blue-black charcoal gray hueless neutral white

achromatic

antonym

chromatic

amber azure blue brown dun green red ... yellow

1. *white* and *achromatic* (3003677747.2) in same synset
2. *achromatic* (300364634.1) and *chromatic* (300355823.1) are antonyms
3. *chromatic* (300355823.1) and *blue* are similar

Wordnet defects for semantic inference

- Shortest path/threshold only work if length of chain inversely correlated with reliability
- Semantic opposition not always encoded – how to do this?
Application 2: logical metonomy: telic (functional/purposive) roles distinguished

- What’s that?

\[(1)\]
\[\begin{align*}
\text{a.} & \quad \text{John began the novel (reading/writing)} \\
\text{b.} & \quad \text{The author began the unfinished novel back in 1962 (writing)}
\end{align*}\]

- Begin can have Agent role, for the writer, or it can be a Telic role (function), for the reader
- Problem is how to define ‘context’ here

Application 2: Metonomy

- John began the novel (reading/writing)
- Context can alter: He really enjoyed your book (reading)
- My dog eats anything
- He really enjoyed your book (eating)

- !John enjoyed the rock
- !! John enjoyed the door
Using Wordnet here

- Wordnet can pick out contexts in which NPs represent events, relative to classes (types)

- Point is: locus of variation is not lexical structure, but in more general ontology (we shall return to this point later)

‘Enjoy’ has lots of purposes...

a. Mary enjoyed seeing the garden  
b. Mary enjoyed inspecting the garden  
c. Mary enjoyed visiting the garden  
d. Mary enjoyed strolling through the garden  
e. Mary enjoyed rollerblading in the garden  
f. Mary enjoyed sitting in the garden  
g. Mary enjoyed dozing in the garden

How can we recover these – distinguish between agent and telic (function) roles?
Context can be subtle

- He really enjoyed your book (reading)
- My dog eats anything
- He really enjoyed your book (eating)

2 contextual function search rules using Wordnet

**Principle of Specificity**: Prefer $R_i$ to $R_j$ in the sequence
\[
\begin{array}{c}
R_i \\
\downarrow \\
C \\
\rightarrow \cdots \\
\rightarrow C_i \\
\rightarrow \cdots \\
\rightarrow C_j \\
\rightarrow \cdots \\
\downarrow \\
R_j
\end{array}
\]

**Principle of Locality**: Plausibility of $R_i$ scales with $m$ and inversely with $l$ in
\[
\begin{array}{c}
R_i \\
\downarrow \\
C \\
\rightarrow \cdots \\
\rightarrow C_i \\
\rightarrow \cdots \\
\rightarrow C_T \\
\leftarrow l \\
\leftarrow m
\end{array}
\]
Test verb

a. EXP enjoy NP
b. EXP_i enjoy [PRO_i [V(ing) NP]]

Mary enjoyed the cigarette (*smoking*)

Wordnet

cigarette, butt, fag, coffin nail

↓

smoke ~~~~ smoke

↓

tobacco

plant product

street drug

↓

substance

artifact ~~~~ create

↓

physical object ~ verbs of perception
Check

Mary enjoyed the cigarette (*smoking*)

Given the hypernym hierarchy in (13), *smoke* (*PRO*, cigarette) is the strongly preferred interpretation since the concept *smoke* is highly specific (*l* small) and distant from general concepts *artifact* and *physical object* (*m* large).

Links for sonata

a. Mary enjoyed the sonata (*listening to/playing*)
b. Mary began the sonata (*playing/composing*)
For begin

\[ music \rightarrow art \rightarrow creation \rightarrow artifact \]

\[ compose, write \rightarrow create \]

Door...

!!John enjoyed the door
Specifically, a door can function both as an entrance (enter) and a barrier (block) to an enclosure. However, the telic verb block has form block(\text{door}, \text{ENCLOSURE}) which is incompatible with the prototype V(\text{PRO}, \text{door}), thus ruling out block. Similar reasoning applies to enter(\text{PRO}, \text{ENCLOSURE}). At the other end of the hierarchy, the canonical events associated with physical object are predicted to be implausible ($l$ large, $m$ small).
Enjoy rock

John enjoyed the rock
  rock
  natural object
  physical object ← perception verbs

Enjoy wine

Mary enjoyed the wine (drinking)
  wine
    alcoholic beverage
      drink
        drink
          food
            liquid
              food
                substance
                  physical object

6.863J/9.611J Lecture 15 Sp03
Where do classes come from?

What’s a natural word?

- NO: NALL = ‘not all’
- NO: green and an hour long
- Nonconstituent: John ate pizza and Mary, bread = “Mary bread”
A tree grows in Brooklyn

Keil’s predicability tree and conceptual ‘naturalness’
Bleaching – noun meaning bleached

a. John boxed the present
b. John put the present in a <BOX>
c. John boxed the present in a gift box
d. # John boxed the present in a brown paper bag

a. Mary buttered the piece of toast
b. Mary put <BUTTER> on the piece of toast
c. Mary buttered the toast with margarine/unsalted butter
d. # Mary buttered the toast with marmalade/onions

More examples

a. Peter shelved a book
b. Peter shelved a book on the windowsill/mantelpiece/table/stand
c. # Peter shelved a book on the hall/spike/ceiling/floor/balcony

a. Sue breaded the fish
b. Sue breaded the fish with breadcrumbs/shredded coconut/crushed almonds
c. # Sue breaded the fish with marmalade/butter/treacle/ice

a. x put y on <SHELF>
b. x put y on z & shelf-like-object(z)

a. # x put <BREAD> on y
b. x put crumbs of <BREAD> on y
c. x put crumbs of z on y
Wordnet hypothesis

Denominial root \( Y \) may be bleached using \( X \) if
a. \( X \) is a hyponym\(^+\) of \( Y \), or
b. \( Z \) is a functional hypernym\(^+\) of \( Y \), and \( X \) is a hyponym\(^+\) of \( Z \)

Path from shelf to windowsill

(a)

(b)

6.863|9.611j Lecture 15 Sp03
More about shelf

![Diagram showing the relationship between shelf, table, support, and other related terms.]

Blanket

![Diagram showing the relationship between blanket, clothing, covering, and other related terms.]

6.863J/9.611J Lecture 15 Sp03
What information is in the lexicon?

/shelf/ vs. /put/ vs. /butter/

/shelf/
Bob shelved the book on the windowsill.

/put/
Bob put the book on the shelf.
Bob put the book with the others.

/butter/
He buttered the bread with margarine.

/put/, /shelf/ imposes $P_{LOCATION}$ on arguments
/put/, /butter/ imposes $P_{LOCATION}$ on arguments
Hypothesis 1: **Lexicon** Contains Selection Criteria

/shelf/ has \( p_{LOCATION} \) selection in lexicon \((=p_{LOCATION} =d(et) \ \text{v})\)

Also: /shelf/ is \( n_{LOCATION} \)

/butter/ has \( p_{LOCATION} \) selection in lexicon \((=p_{LOCATION} =d(et) \ \text{v})\)

Also: /butter/ is \( n_{LOCATION} \)

So then the **Lexicon** cannot derive:

* 1. Bob shelved the windowsill with the book.
* 2. Bob buttered the margarine onto the bread.

Information about butter and shelf – where is it located?

---

Hypothesis 1 Problem

Problem: How does **Lexicon** acquire the following:

<table>
<thead>
<tr>
<th>/shelf/</th>
<th>/butter/</th>
<th>/shovel/</th>
<th>/pencil/</th>
<th>/mop/</th>
<th>/email/</th>
<th>etc.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( p_{LOCATION} \) \( =d(et) \ \text{v} \)

\( n_{LOCATION} \)

\( p_{LOCATION} \) \( =d(et) \ \text{v} \)

\( n_{LOCATION} \)

\( p_{LOCATION} \) \( =d(et) \ \text{v} \)

\( n_{LOCATION} \)

\( p_{LOCATION} \) \( =d(et) \ \text{v} \)

\( n_{LOCATION} \)

\( n_{INST-COMM} \) \( =p_{HAVE} \ \text{v} \)

\( n_{INST-COMM} \) \( =p_{DEST} \ \text{v} \)

\( n_{INST-COMM} \) \( =p_{SOURCE} \ \text{v} \)

\( n_{INST-COMM} \) \( =p_{DEST} \ \text{v} \)

\( n_{INST-COMM} \) \( =p_{SOURCE} \ \text{v} \)

etc.

Solution 1: Solve the above problem

Solution 2: Push problem OUT of **Lexicon** and INTO Encyclopedia
**Solution 2:** Push problem OUT of Lexicon and INTO Encyclopedia

Encyclopedia, not lexicon, is source of 'Oddness' of:

# (1) Bob shelved the windowsill with the book.
# (2) Bob buttered the margarine onto the bread.

Lexicon is NOT:

<table>
<thead>
<tr>
<th>Lexicon</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>shelf</td>
<td>/shelf/ =p LOCATION =d(et) V</td>
</tr>
<tr>
<td></td>
<td>/butter/ =d +k LOCATION</td>
</tr>
<tr>
<td>into</td>
<td>/int/ =d +case LOCATION</td>
</tr>
<tr>
<td></td>
<td>/with/ =d +case LOCATION</td>
</tr>
</tbody>
</table>

But instead:

<table>
<thead>
<tr>
<th>Lexicon</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>shelf</td>
<td>/shelf/ =p =d V</td>
</tr>
<tr>
<td></td>
<td>/butter/ =d +case p</td>
</tr>
<tr>
<td>into</td>
<td>/int/ =d +case p</td>
</tr>
<tr>
<td></td>
<td>/with/ =d +case p</td>
</tr>
</tbody>
</table>

Thus insofar as the lexicon is concerned, (1) and (2) are GRAMMATICAL.

---

**Encyclopedia vs. Lexicon**

Lexicon does NOT hold real-world knowledge, only:

<table>
<thead>
<tr>
<th>ROOT</th>
<th>Lexicon</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>arrive</td>
<td>+v, +DP, +cause</td>
<td>John arrived. The arrival of John</td>
</tr>
<tr>
<td>big</td>
<td>-v, +DP</td>
<td>The big X</td>
</tr>
<tr>
<td>open</td>
<td>+v, +DP, +cause</td>
<td>John opened X. X opened.</td>
</tr>
<tr>
<td>destroy</td>
<td>+v, +DP, +cause</td>
<td>John destroyed X. John's destruction of X</td>
</tr>
</tbody>
</table>

Encyclopedia holds knowledge 'rejecting' the following GRAMMATICAL sentences:

# John thought the book to Mary        # John's growth of tomatoes
# Sue walked in an hour                 # Bob shelved the window sill with the book.
# Bob buttered the margarine onto the bread.
2 Language Acquisition Problems:
Lexicon vs Encyclopedia

<table>
<thead>
<tr>
<th>ROOT</th>
<th>LEXICON ENTRIES</th>
<th>ENCYCLOPEDIA ENTRIES</th>
</tr>
</thead>
<tbody>
<tr>
<td>/shelf/</td>
<td>n_r = p = d V_casser</td>
<td>p_location = d V</td>
</tr>
<tr>
<td>/butter/</td>
<td>n_r = =d V_casser</td>
<td>=d V</td>
</tr>
<tr>
<td>/into/</td>
<td>=d +k p</td>
<td>=d +k p</td>
</tr>
<tr>
<td>/with/</td>
<td>=d +k p</td>
<td>=d +k p</td>
</tr>
</tbody>
</table>

**LEXICON ACQUISITION:**
How do LEXICAL roots get assigned to feature set?

**ENCYCLOPEDIA ACQUISITION:**
How do ENCYCLOPEDIA roots get assigned to feature set?

Distributed Semantics

Encyclopedic roots distributed *everywhere* in mental architecture. Lexicon roots are convenient abstractions for encyclopedic roots so syntax operates autonomously:

- v -subj, +comp p +subj, +comp
- n -subj, -comp a +subj, -comp

/x/ meaning X is placeholder for experiences of X. What 'experiences' of X are requires theorizing about attention, iconic bottlenecks, etc.

/butter/ is understood as plocatum and nlocatum because BUTTER memories contain a relatively higher frequency of plocatum over plocation primitives activated. /shelf/ is the opposite.

When /x/ has "two entries", there is a bimodal distribution of utterances associated with two primitives firing, neither more 'defining' or 'core' than the other. Try /color/, /open/, /ache/, /see/, /look/, /remember/, /forget/, /think/.

Interface conditions of encyclopedia to roots to these modules
Predicate-arguments to thematic roles

- Use linking rules
- These say whether, e.g, Subject is the agent...
- Is there a theory for this?
- How do we build this knowledge?

Predicate-argument structures for lose

lose1 (Agent: animate, Patient: physical-object)

lose2 (Agent: animate, Patient: competition)

Agent  <=>  subj
Patient  <=>  obj
Machine Translation Lexical Choice - Word Sense Disambiguation

Iraq lost the battle.
Ilakuka centwey ciessta.
[Iraq] [battle] [lost].

John lost his computer.
John-i computer-lul ilepelyessta.
[John] [computer] [misplaced].

Word sense disambiguation with Source Language Semantic Class Constraints (co-occurrence patterns)

\[\text{lose1} (\text{Agent, Patient: competition}) \iff \text{ciessta}\]

\[\text{lose2} (\text{Agent, Patient: physobj}) \iff \text{ilepelyessta}\]
Is there enough data?

- Break

Levin classes (3100 verbs)

- 47 top level classes, 150 second and third level
- Based on pairs of syntactic frames.
  
  - John broke the jar. / Jars break easily. / The jar broke.
  - John cut the bread. / Bread cuts easily. / *The bread cut.
  - John hit the wall. / *Walls hit easily. / *The wall hit.

- Reflect underlying semantic components
  contact, directed motion, exertion of force, change of state

- Synonyms, syntactic patterns, relations
Another alternation example

- Another example: Causative/inchoative
  - The window broke
  - John broke the window
  - The rabbit suddenly appeared
  - *The magician appeared the rabbit

- Benefactive:
  - Sue carved a toy out of wood for Hansel
  - Sue carved hansel a toy out of wood
  - Sue carved some wood into a toy for Hansel
  - *Sue carved Hansel some wood into a toy

- Middle formation:
  - The whale frightens easily
  - *The whale sees easily

Alternations:

- Sue broke the vase/ The vase broke (change-of-state)
- The vase broke easily
- Conative: *Sue broke at the vase

- Bill cut the bread/ *The bread cut (change-of-state, no “telic” endpoint)
- The bread cut easily
- Bill cut at the bread

- Mary touched the cat / *The cat touched
- *The cat touched easily (no change-of-state)
- *Mary touched at the cat

- Joe kicked the tire / *The tire kicked
- *The tire kicked easily
- Joe kicked at the tire

Alternations can be lang-specific: “break” is a causative/inchoative in English, but not Italian.
Lexical Gaps: English to Chinese

<table>
<thead>
<tr>
<th>English</th>
<th>Chinese</th>
</tr>
</thead>
<tbody>
<tr>
<td>break</td>
<td>da po - irregular pieces</td>
</tr>
<tr>
<td>smash</td>
<td>da sui - small pieces</td>
</tr>
<tr>
<td>shatter</td>
<td>pie duan - line segments</td>
</tr>
<tr>
<td>snap</td>
<td></td>
</tr>
</tbody>
</table>
Intersective Levin classes

So we want...
Thematic Roles

• E w,x,y,z Giving (x) ^ Giver(w,x) ^ Givee(z, x) ^ Given(y,x)
• E w,x,z Breaking (x) ^ Breaker(w,x) ^ Broken(z,x)
• A set of roles:
  • agent, experiencer, force, theme, result, content, instrument, beneficiary, source, goal,...

  The dog ate the cheeseburger.
  What is cheeseburger?
  The sniper shot his victim with a rifle.
  What is rifle?

Schank's Conceptual Dependency

• Eleven predicate primitives represent all predicates
• Objects decomposed into primitive categories and modifiers
• But few predicates result in very complex representations of simple things

  Ex,y Atrans(x) ^ Actor(x,John) ^ Object(x,Book) ^ To(x,Mary) ^ Ptrans(y) ^ Actor(y,John) ^ Object(y,Book) ^ To(y,Mary)

  John caused Mary to die vs. John killed Mary
Selection via sortal hierarchy

- John ate a clam
- They served clams

- “logical” form: $\exists x, y, e[\text{eat}(e) \& \text{eater}(e, y) \& \text{eaten}(e, x) \& \text{john}(y) \& \text{clam}(x) \& \text{past}(e)]$

- So...

Sortal hierarchy (‘ontology’)

Entity
  \[\text{thing} \quad \text{being} \quad \text{state} \]
  \[\text{food} \quad \text{implement} \]
Selection via sortal hierarchy

1. eater([Eating],[Being])
2. eat([Eating])
3. eaten([Eating],[Food])
4. server([Serving],[Being])
5. serve$_1$([Serving])
6. served([Serving],[Food])
7. john([Person])
8. they([Person])
9. mussel$_1$([Food])
10. mussel$_2$([Creature])

But...

• Which airlines serve Denver?

• You ate glass on an empty stomach

• Metonymy: What airlines fly to Boston?
But how can we/computer learn this?

- Two parts: pred-arg linking to thematic roles – which verbs do what
- Selectional restrictions

pour vs. fill

- Different linking entails semantic difference - when in Object position, the Goal seems "affected" in a way not so in the PP
- Fill: Cause X to become full of Y by means of causing Y to be in X
- Pour: Cause X to go in a downward stream into Y
- Fill has two events: a state change (the glass) and a location change (the water)
- Pour has one event: location change
- The Main-change argument gets Old-Info structure and main event status. Main event of Fill: state change of glass
The Problem of Ambiguity

Possible Hypotheses

- Rabbit (whole object)
- Animal (superordinate)
- Flopsie (individual)
- Furry (property)
- Ear (part)
- Walk by (activity)
- Undetached rabbit parts ......

"Gavagai!"
Two Bootstrapping Proposals

- Children use syntactic cues to verb meaning (Gleitman 1990)
- Children use (verb) meaning to figure out how its arguments are realized in the syntax of the language (Pinker 1989)

Semantic Bootstrapping (Pinker 1984)

Semantic Bootstrapping involves the pairing of a situational context with some syntactic pattern.

- Kids learn syntax by first learning the semantic argument structure of the verb.
  - SWIM = one participant (the “swimmer”)
  - EAT = two participants (“eater”, “eatee”)
  - TAKE = two/three participants (“taker”, “takee”, and “person taken from”...)
Gleitman: Not So Fast, Pinker...

Temporal ambiguity  Situation ambiguity  Mental unobservable!

... more than just real-world observation...

Syntactic Bootstrapping
(Landau and Gleitman 1986, Naigles 1990)

Syntactic frames provide evidence for meaning:

\[ H_1: \text{arm wheel} \]
\[ H_2: \text{cause to squat} \]

/\text{X and Y are gorping}/
/\text{Look, gorping}/
/\text{X is gorping Y}/
Verbs Classes Grouped by Cause Feature

<table>
<thead>
<tr>
<th>Feature</th>
<th>Verb Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>Externally Caused (touch, load)</td>
</tr>
<tr>
<td></td>
<td>F1: He touched the glass.</td>
</tr>
<tr>
<td></td>
<td>* F0: The glass touched.</td>
</tr>
<tr>
<td>H0</td>
<td>Internally Caused (laugh, glimmer)</td>
</tr>
<tr>
<td></td>
<td>* F1: He laughed the child.</td>
</tr>
<tr>
<td></td>
<td>F0: He laughed.</td>
</tr>
<tr>
<td>H*</td>
<td>Externally Causable (open, break)</td>
</tr>
<tr>
<td></td>
<td>F1: He opened the door.</td>
</tr>
<tr>
<td></td>
<td>F0: The door opened.</td>
</tr>
</tbody>
</table>

Hypothesis space $H$ Evidence $x$ in $X = \{0, 1\}$ Hi in $H$

One-shot learning

within a Bayesian framework.

Syntactic Evidence Semantic Evidence | Evidence $x$

Linguistic Theory $H = \{H_1, H_2, \ldots\}$

Prior: $p(H_i)$

Likelihood $p(x|H_i)$

Bayesian Acquisition Device

Acquired Lexicon

$/(seb/\text{means} Posterior:/) p(H_i|x)$

$p(H_i|x) = \frac{p(x|H_i)p(H_i)}{p(x)}$
Learning Value of Verbs Cause Feature

**Syntactic Evidence:**
/He glipped the balloon/
x = F1

**Syntactic Theory:**
H = {H₁, H₀, H₅}

Prior:
p(H₁) = .333

Likelihood
p(x|Hᵢ)

x = F0  x = F1
H₁ .05  .95
H₀ .95  .05
H₅ .50  .50

**Acquired Lexicon**

p(H₁|x) = p(x|H₁)p(H₁) / p(H₁)

p(H₁|x=F1) = (.95)(.33)

p(H₀|x=F1) = (.05+.95+.50)(.33)

p(H₅|x=F1) = .333

Syntactic Evidence X:
/He glipped the balloon/
/X gorped Y/ /X gorped Y/
/X sebbed Y/ /Y sebbed/
/X meefed Y/ /Y meefed/
/Y foomed/

**Syntactic Theory:**
H = {H₁, H₀, H₅}

Prior p(Hᵢ)

Likelihood p(x|Hᵢ)

Lexicon:  Evidence X  p(H₁|X)  p(H₀|X)  p(H₅|X)
/glip/  F1  .633  .033  .333
/gorp/  F1, F1  .781  .002  .217
/seb/  F1, F0  .137  .137  .724
/meef/  F1², F0  .712  5e-6  .288
/foom/  F0⁶  2e-8  979  .021
Bayesian Learning at the Syntax-Semantics Interface

**Syntactic Evidence**
/X is gorping Y into Z/
/X is pilking Z with Y/
/Look! jebbing!/

**Semantic Evidence**
Person pours water into a glass, filling it
Person pours water into a glass, filling it
Person pours water into a glass, filling it

**Linguistic Theory**
H={H₁, H₂, ...}  
Prior: p(Hᵢ)  
Likelihood p(x|Hᵢ)

**Acquired Lexicon**

|   | p(POUR|x) | p(FILL|x) | p(MOVE|x) |
|---|-----------|-----------|-----------|
| gorp | .880      | .000      | .101      |
| pilk | .001      | .989      | .000      |
| jeb  | .463      | .463      | .005      |

How to get ‘real semantics’ in?