6.863J Natural Language Processing
Lecture 14: Word semantics I

Robert C. Berwick
berwick@ai.mit.edu
The Menu Bar

• Administrivia:
• Lab 4 due April 9
• Agenda:
• Lexical semantics: the meanings of words: how hard can it be?
• Tense and time (if there’s time)
Word sense

- The benevolent alien race that visits earth.
- Their great book is entitled How to Serve Humans
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)

\[ \text{lose}_1(\text{Agent, Patient}: \text{competition}) \iff \text{ciessta} \]

\[ \text{lose}_2(\text{Agent, Patient}: \text{physobj}) \iff \text{ilepelyessta} \]
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.
Break Levin class - *Change-of-state*

Diagram:
- break
  - chip
  - crack
  - crash
  - crush
  - fracture
  - rip
  - tear
  - split
  - splinter
  - smash
  - shatter
Lexical Gaps: English to Chinese

break → ?

smash

dausible - irregular pieces

dau sui - small pieces

shatter

snap → pie duan - line segments
Intersective Levin classes

- "Cut" Verbs: scrape, (clip) (snip), scratch, (chip) cut, (slash) saw, hack, hew
- "Split" Verbs: pull, (draw), (kick) (yank), tug, shove, push
- "Carry" Verbs
- "Push/Pull" Verbs
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

\[
\text{Ex, } x \ A\text{trans}(x) \wedge \text{Actor}(x, \text{John}) \wedge \\
\text{Object}(x, \text{Book}) \wedge \text{To}(x, \text{Mary}) \wedge \text{Ptrans}(y) \wedge \\
\text{Actor}(y, \text{John}) \wedge \text{Object}(y, \text{Book}) \wedge \text{To}(y, \text{Mary})
\]

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
  - thing
  - being
  - state
    - food
    - implement
Selection via sortal hierarchy

1. eater([Eating],[Being])
2. eat([Eating])
3. eaten([Eating],[Food])
4. server([Serving],[Being])
5. serve₁([Serving])
6. served([Serving],[Food])
7. john([Person])
8. they([Person])
9. mussel₁([Food])
10. mussel₂([Creature])
But...

- Which airlines serve Denver?
- You ate glass on an empty stomach
- Metonomy: 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
Look! He’s sebbing!

Look! A seb!

Look, some seb!

/seb/ means MIXING

/seb/ means BOWL

/seb/ means STUFF

KEY HUMAN COMPETENCE:

One-shot integration of syntax & semantics
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...

THE INFORMATION GIVEN BY THE REAL WORLD

Go show Granny what you’re doing. She’ll think you push the truck sooo well!

Time

GO SHOW THINK PUSH

Temporal ambiguity

CAN SYNTAX OVERRIDE "SALIENCES" IN THE SCENE?
(Fisher, Hall, Rakowitz, and Gleitman, 1991)

GIVE.
TAKE

CHASE
FLEE

PUSH
FALL

Situation ambiguity

Mental unobservable!

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

6.863J/9.611J Lecture 14 Sp03
Syntactic Bootstrapping
(Landau and Gleitman 1986, Naigles 1990)

Syntactic frames provide evidence for meaning:

\[ H_1: \text{arm wheel} \]

\[ H_2: \text{cause to squat} \]

/X and Y are gorping!/
/Look, gorping!/
/X is gorping Y!/
Verbs Classes Grouped by Cause Feature

\[ H_i \quad \text{Verb Class} \]
\[ H_1 \quad \text{Externally Caused (\textit{touch, load})} \]
\[ F1: \text{He touched the glass.} \]
\[ * F0: \text{The glass touched.} \]

\[ H_0 \quad \text{Internally Caused (\textit{laugh, glimmer})} \]
\[ * F1: \text{He laughed the child.} \]
\[ F0: \text{He laughed.} \]

\[ H^* \quad \text{Externally Causable (\textit{open, break})} \]
\[ F1: \text{He opened the door.} \]
\[ F0: \text{The door opened.} \]

Hypothesis space \( H \)
Evidence \( x \) in \( X = \{0, 1\} \)
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

\[
p(H_i|x) = \frac{p(x|H_i)p(H_i)}{p(x)}
\]

Acquired Lexicon  
(\( /seb/ \) means Posterior: )  
\( p(H_i|x) \)
Learning Value of Verbs Cause Feature

Syntactic Theory: 
\[ H = \{ H_1, H_0, H_* \} \]

Prior:
\[ p(H_i) = .333 \]

Likelihood
\[ p(x|H_i) \]

\begin{align*}
  x = F0 & \quad x = F1 \\
  H_1 & \quad .05 \quad .95 \\
  H_0 & \quad .95 \quad .05 \\
  H_* & \quad .50 \quad .50 \\
\end{align*}

Syntactic Evidence:
\[ /He glipped the balloon/ \]
\[ x = F1 \]

Acquired Lexicon

Posterior \( p(H_i|x) \)
\[ p(H_1|x=F1) = .633 \\
 p(H_0|x=F1) = .033 \\
 p(H_*|x=F1) = .333 \]

\[
p(H_i|x) = \frac{p(x|H_i)p(H_i)}{p(H_i)}
\]

\[
p(H_i|x=F1) = \frac{(0.95)(0.33)}{(0.05+0.95+0.50)(0.33)}
\]
**Syntactic Evidence X:**
/He glipped the balloon/
/X gorped Y/, /X gorped Y/
/X sebbed Y/, /Y sebbed/
/X meefed Y\(^5\), /Y meefed/
/Y foomed/\(^6\)

**Syntactic Theory:**
\(H = \{H_1, H_0, H_*\}\)

- Prior \(p(H_i)\)
- Likelihood \(p(x|H_i)\)

**Bayesian Language Acquisition Device**

**Acquired Syntactic Knowledge**

| Lexicon  | Evidence X | \(p(H_1|X)\) | \(p(H_0|X)\) | \(p(H_*/X)\) |
|----------|------------|----------------|----------------|----------------|
| /glip/   | \(F1\)     | .633           | .033           | .333           |
| /gorp/   | \(F1, F1\) | .781           | .002           | .217           |
| /seb/    | \(F1, F0\) | .137           | .137           | .724           |
| /meef/   | \(F1^5, F0\) | .712       | 5e-6           | .288           |
| /foom/   | \(F0^6\)   | 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(Hᵢ|x)

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

Bayesian Language Acquisition Device
How to get ‘real semantics’ in?
Verb meanings are logic programs (LPs):

**General:** cause(e)

**One args x:** move(x), rotate(x), move-dn(x), move-up(x)
supported(x), liquid(x), container(x)

**Two args x,y:** contact(x,y), support(x,y), attach(x,y)
(if cause(e)=1)

<table>
<thead>
<tr>
<th>Verb</th>
<th>Logic Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>/lower/</td>
<td>1 1<em>101** 11</em></td>
</tr>
<tr>
<td>/raise/</td>
<td>1 1<em>011** 11</em></td>
</tr>
<tr>
<td>/rise/</td>
<td>0 1<em>01</em>**</td>
</tr>
<tr>
<td>/fall/</td>
<td>0 1<em>10</em>**</td>
</tr>
</tbody>
</table>

**Hypothesis space H:** All LPs

**Evidence X:** Bit Vector Examples
(e.g. 1 1010100 110)
Learning Semantic Features

**Semantic “Theory”:** (3 bits)

Hypothesis space $\mathbf{H}$: 27 LPs

$q$  $H_i$
0    000, 001, 010, 011
     100, 101, 110, 111
1    00*, 01*, 10*, 11*
     0*0, 0*1, 1*0, 1*1
     *00, *01, *10, *11
2    0**, 1**, *0*, *1*, **0, **1
3    ***

Prior $p(H_i) = 1/27$

Likelihood $p(x|H_i) = \{2^{-q} \text{ if } x \in H_i\}$

- $p(x=000|H_{000}) = 1$
- $p(x=000|H_{00*}) = 0.5$
- $p(x=000|H_{0**}) = 0.25$
- $p(x=000|H_{***}) = 0.125$

**Semantic Evidence:**

- /Look! Glipping!/ $X_1 = 000$
- /Look! Gorping!/ $X_2 = 000, 001$
- /Look! Sebbing!/ $X_3 = 000, 000, 000$
- /Look! Meefing!/ $X_4 = 000, 101, 010, 111, 000$

**Acquired Semantic Knowledge**

**Lexicon:**

- $p(H_{000}|X)$
  - /glip/ 0.30
  - /gorp/ 0.00
  - /seb/ 0.70
  - /mee/ 0.00

- $p(H_{00*}|X)$
  - /glip/ 0.15
  - /gorp/ 0.64
  - /seb/ 0.09
  - /mee/ 0.00

- $p(H_{0**}|X)$
  - /glip/ 0.07
  - /gorp/ 0.16
  - /seb/ 0.01
  - /mee/ 0.00

- $p(H_{***}|X)$
  - /glip/ 0.03
  - /gorp/ 0.04
  - /seb/ 0.001
  - /mee/ 1.0

Bayesian Language Acquisition Device
But... what are the possible arguments?

- Predicate-arguments can be complicated... can we crank it out?
- Argument structure is syntax
- There are no specialized mechanisms of ‘thematic role assignment’
- Everything is really predication
Hale-Keyser: arguments are syntax
The basic form

```
spec
  comp
```
H & K: The framework

• There are only three places a verb argument can come from
  • The complement or specifier of a “basic” lexical item
  • An external “addition”
  • As for “basic lexical items” there are four types: N, V, A, P
  • Why so few thematic roles? Because so few basic lexical items (entity/instance, event, state, relation)
N, V, A, P

- N takes no arguments
- V are predicational, and take one argument, a complement.
- P are relational, and take two arguments
- A are predicational, and take one argument, but require some help; thus an A is always the complement of a verb, which then projects for an external arg.
Hale-Keyser Incorporation

• 4 Fundamental Primitives Yield Different Argument Structures
HK Allows Us to Discard Thematic Roles

• Agent, Patient, Theme, Instrument, Goal, ...
  ... derived from positions in structural configurations.
• V-N:  
  • V-A
  • V-P

6.863J/9.611J Lecture 14 Sp03
What can N get us?

• Intransitive verbs:

Nouns cannot project arguments. A noun (run, laugh, play, cough, snore, burp) incorporates into the verb. An external argument is adjoined to \( v \). Thus, rather than having cognate N and V copies in the lexicon, verbs are derived by syntax.
Unergatives vs.
Simple Transitives

• Unergatives: no external agent The child laughed
  • [NP [v [V+N (N)]]]

• No verbs like *The clown laughed the child / *The alfalfa sneezed the colt (The N complement to V has incorporated, where would the “object NP” reside?)
  • [NP [V+N (N) NP?]]

• Simple transitive (non-creation) The clown made the child laugh
  • [NP [v [NP [V+N (N)]]]]
  • Extensions: get+A (I got drunk, I got Josh drunk)
    • But not for get+N (I got the measles, *I got Josh the measles)
Explaining Gaps in the Lexicon

• *It cowed a calf, *It dusted the horses blind, *It machined the wine into bottles (cf. The cow had a calf, the dust made the horses blind, the machines put the wine into bottles)

• The above items would be the result of the external subject incorporating into the verb, which is ruled out by the syntax elsewhere (items raise & incorporate up, but not down)

• If all “denominal” verbs are the result of incorporation of the complement to the V head, rather than unconstrained “category change”, these non-verbs are predicted
V: Verbs of Creation: The simple case

- bake a cake, make trouble, build a house, have puppies

- V has a complement NP(=DP). External argument is projected and adjoined to v.
P gives put-type Verbs

- The P frame has a specifier and complement. The whole P-complex is a verb complement. An external argument is projected and adjoined.
P gives locatum-type verbs

- With a bare N as the PP complement, the N conflates with the P, which conflates with the V, giving saddled the horse, boxed the gift, roofed the house (all have P-meaning)
Picture

V

horse, gift, house

P

Saddle, box, roof

6.863J/9.611J Lecture 14 Sp03
Implementation

(define-verb-class "PUT VERBS: put verbs (Section 9.1)"
 "putting entity at some location (but not to or from)"
 '(arrange immerse install lodge mount place position put set 
   situate sling stash stow)
 (list '(* the water put into a bowl))
 '((+ he put the water into the bowl)
   (vp ()
     (v* (v put (feature CAUSE)))
     (pp (n the water)
       (p* (p into (feature MOVELOCATION))
         (n a bowl)))))))

6.863J/9.611J Lecture 14 Sp03
Argument Structure: The Moral

• No specialized mechanism of “thematic role assignment”. Everything is predication.

• Do these mechanisms of derived verbs happen in the syntax with everything else, or “prior to lexical insertion”, e.g. “in the lexicon”? What do you think? Should this distinction matter?