6.863J Natural Language Processing
Lecture 21: the meaning of it all: Lexical semantics

Instructor: Robert C. Berwick
berwick@ai.mit.edu

The Menu Bar

• Administrivia:
  • Lab 4(a&b) out – due April 30th
  • Agenda:
    How to do things with words...
Review: what’s right, what’s wrong

- Montague grammar
  - Meaning of whole is combination of the sum of its parts
  - Rule of syntax ⇔ rule of semantics
- Today & next time - goals:
  - Understand the failures of the above from:
    - Historical perspective
    - Open problems

No “Unified Theory of Semantics”*

Different goals → different semantic theories:
- Syntactician: why do different words appear in different constructions?
- Semanticist: what is an adequate meaning representation of a vocabulary item?
- Lexicographer: what are all the things we know about a word’s meaning?
- IR Engineers: what is the meaning abstraction of a piece of text?
- Roboticist: how can the robot appear to understand me?
- Child Dev Psych (Vocab + Grammar)
- Historical linguist

*(Is there a unified theory for chemistry? physics?)
Uncertainty in terms*

- Grammar: How much semantics should be in it?
- Grammaticality: Is a semantically anomalous sentence ungrammatical?
  - He gave the book to John.
  - He thought the book to John.
- Grammatical category: What are their essences?
- Word Meaning: What is a meaning representation?
- Concepts: How are they related to words?
  - How is what we know about TIGER related to /tiger/?

*No one knows the answer! Wait: When did science know the definition of an atom, electron, proton, …?

Formal Semantics

Montague 1970, Partee 1974 - YOUR Lab 4
Given grammar, mechanical procedures derive semantic representation:

(1) Thematic roles
   EXISTS e | Agent(x) & Theme(y) & Eat(e, x, y)
(2) Lexical conceptual structures
   CAUSE([[Thing i]], GO([[Thing j]], IN-MOUTH-OF([[Thing i]]))

Do you believe these things really work??

*No one knows the answer! Wait: When did science know the definition of an atom, electron, proton, …?
Lambda calculus Semantics

• /Bob put the book on the shelf/
  (cause :agent (bob) :effect (go :theme (book) :path (path :oper (on) :terminal+ (shelf)) :tense past))

• /What did Bob put on the shelf/
  (cause :agent (bob) :effect (go :theme (? (what)) :path (path :oper (on) :terminal+ (shelf)) :tense past))

• /What did Bob put the book on/
  (query :event (cause :agent (bob) :effect (go :theme (book) :path (path :oper (on) :terminal+ (? (what))) :tense past))

• /Where did Bob put the book/
  (query :event (cause :agent (bob) :effect (go :theme (book) :path (path :oper () :terminal+ (? (where))) :tense past))

• */.../
Subcategorization
(Chomsky 1965)

- Verbs have classes:
  John ate a tomato.
  A tomato was eaten.
  John resembled a tomato.
  * A tomato was resembled.
  You have seen this in your labs!

- Use features:
  +animate, -passive, +male, +human, ...
  - If we allow +human, then do we allow +good-to-eat-with-chicken?
  - Wait: where are the restrictions on these features?
  - Major problem: Blank check on features

Subcategorization
(Levin 1993)

- +motion +contact –effect
- Hit, touch, break, cut classes

Any notion that rules apply blindly without paying attention to "semantics" is pure wishful thinking. The question is how much attention.
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
Lexical semantics Primitives

Jackendoff 1983, 1990
Goal: Not just syntax, but cognition
If we postulate a CAUSE primitive in /melt/
Entailment
Map to syntax: linking rules
Present day: Levin and Rappoport 1998

Entries in the lexicon are Root Lexical Conceptual Structures, RLCS

*-marked positions are unified with other compatible RLCSs during composition.

Roll:

```
           GO
          /    \   
*theme   *path   ROLL
       /   \    \   
  (theme)  AT  (goal)
 /     \  /   \  
thing   *place
  \    /  \    /  
thing at
```

6.863J/9.611J SP04 Lecture 21
Lexical Conceptual Structure

- Each node contains:
  - Primitive:
    - CLOSED CLASS: GO, STAY, BE, ON, IN, AT...
    - OPEN CLASS: JOHN, RUN-INGLY, ...
  - Field: Analogy to motion/position in Localist approach: LOCATIONAL, POSSESSIONAL, TEMPORAL, ...
  - Type: EVENT, STATE, PATH, POSITION, MANNER...

Constituency vs. Dependency

- Constituency Tree
  - pieces of syntax constructed out of smaller pieces of syntax
  - rules apply to constituents:
    - [Several smelly fish] [bothered Nigel]
    - *Several smelly [fish bothered] Nigel

- Dependency Tree
  - dependents are children of heads
  - subjects are included as children of verbs
The Fulton County Grand Jury said Friday an investigation of Atlanta’s recent primary election produced no evidence that any irregularities took place.
The ball rolled towards Beth.

Recursively compose the children, then assign the composed children to *-marked positions in the current RLCS. This yields a Composed LCS, CLCS.

Possible Words

- No language has a word that means X
  - If it has /kick/, then it doesn’t have /skick/
- If true, this says something about the internal language
- But it is hard to actually do this research:
  - E.g. English doesn’t have a word that means “make sing” Maybe it’s a universal.
  - But XXX language does.
- So, can’t be too certain too quickly
Language of Thought (Fodor 1975)

- Children acquiring a language are mapping words onto internal language.
- This internal language cannot be induced on the basis of language learning.
- Are the lexical semantics primitives the LOT?

Structural vs. Content Meaning Component

- Verbs in a class share structural component.
- Verbs in a class distinguished by content component.
Structural vs. Content Meaning

- Verbs in a class share structural component
- Verbs in a class are distinguished by content component

- \( \text{cause} \)
- \( \text{go-ident} \)
- \( \text{thing}_1 \)
- \( \text{thing}_2 \)
- \( \text{toward-ident} \)
- \( \text{at-ident} \)
- \( \text{buttered} \)
- \( \text{carpeted} \)
- \( \text{feathered} \)
- \( \text{saddled} \)
- \( \text{salted} \)

Structural vs. Content Meaning

- \( \text{go-loc} \)
- \( \text{thing}_2 \)
- \( \text{from-loc} \)
- \( \text{to-loc} \)
- \( \text{amble-ingly} \)
- \( \text{lope-ingly} \)
- \( \text{skitter-ingly} \)
- \( \text{zoom-ingly} \)
Common objections

Definition = Structure + Plus X, for unknown X
- Consider paint, water, butter, ...:
  - She painted a house, he watered a plant, he buttered bread
- Claim: Structure is “put N on X” (Hale&Keyser 2003)
- Plus X: (story about putting)

Undefinable primitives:
1. Thematic Roles: Agent, Patient, Goal, ...
   Remedy: Reduce them and define them (Dowty 1991)
   Remedy: Define/derive them structurally (Hale & Keyser 2003)
2. Lexical Semantic Primitives: CAUSE, GO, BE, HAVE, ...
   Remedy: Decompose them even more (Jackendoff 1991, 1996)

What information is in the lexicon?
Hypothesis 1: **Lexicon** Contains Selection Criteria

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

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

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

Also: /butter/ is \( n_{\text{LOCATUM}} \)

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:

- `/shelf/` =pLOCATION =d V
- `/butter/` =pLOCATION =d V
- `/shovel/` =pINST-MOT =pLOCATION =d V
- `/pencil/` =pINST-MOT =pLOCATION =d V
- `/mop/` =pINST-REM =pLOCATION =d V
- `/email/` =pINST-COMM =pPERSON =d V

etc.

Solution 1: Solve the above problem

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:

- `/shelf/` =pLOCATION =d V
- `/butter/` =d +k pLOCATUM
- `/into/` =d +case pLOCATION
- `/with/` =d +case pLOCATUM

But instead:

- `/shelf/` =p =d V
- `/butter/` =d +case p
- `/into/` =d +case p
- `/with/` =d +case p

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 windowsill with the book.
#. Bob buttered the margarine onto the bread.

2 Language Acquisition Problems: Lexicon vs Encyclopedia

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

ENCYCLOPEDIA ACQUISITION:
How do ENCYCLOPEDIA roots get assigned to feature set?
WordNet
(Miller et al 1998)

- Widely used in computational linguistics
- Dictionary-like definitions organized by links:
  - Nouns: X is a kind-of/part-of Y
  - Verbs: X involves doing Y
    - Also with common syntactic frames
  - Other than the above, no conceptual structure, no meaning postulates
-Enumerates lists of senses, does not relate these senses

Senses

- How many senses per a word? WordNet examples:
  - bank – 10 noun senses, 8 verb senses
  - have – 1 noun sense, 19 verb senses
  - smoke – 8 noun sense, 2 verb senses
- Are these different senses? How are they structurally related?
  - relating them structurally requires conceptual metalanguage
Meaning isn’t (always) at the Word Level

pick up, throw up, turn on does NOT have picking, throwing, turning (at least not directly)

Antidisestablishmentarism
(morphosemantics theory very poor)

And there is pragmatics (too large a topic)

Words appear in a very wide variety of constructions

He sirened her down.
The car sirened its way to NY.
She sirened Bill the message.

... Fantasy:
VP => V142 PP
V142 -> siren

A more flexible approach needed!
Idioms / Constructions

- Are idioms to be stored in the lexicon?
  - Examples:
    - Kicked the bucket, Paint the town red
    - Spic-and-span, kit and kaboodle
    - What’s X doing Y? The X-er, The Y-er
  - H1: Yes
    - BUT then: how do you treat Tense, agreement, ...
  - H2: No
    - BUT then: then where is “meaning” stored?
      - Answer: the encyclopedia
      - But that is a non-answer

Failure to Compose

- Defeasability:
  - He climbed the mountain vs He climbed down the mountain
  - Red hair vs red Porsche
    - Does this work? Red(x) & Hair(x)
    - Meaning of RED in context > outside context?
  - Former friend
    - Does this work? Former(x) & Friend(x)
  - Good knife vs good book vs good life
    - Does this work? Good(x) & Knife(x)
    - Good knives cut well, Good books ..., Good lives ...
Metaphoric Meanings

- “No silver bullet on 9/11”
- “My surgeon is a butcher” vs “My butcher is a surgeon”
- “Don’t get high on Montague grammar”
- Appears way more often than you think.

Meanings are highly private

Before they become adults, children think:
/uncle/ is a friendly middle-aged man
/island/ is a beachy area with palm trees
/two/ is some small number greater than one
and not anything like

Blind children’s LOOK
What can one do?

Let’s ...

Show why everyone is wrong (Fodor 1998)

Summarize corpora statistically

P(V142|D) = .011
P(V143|D) = .004
P(V144|D) = .0014

Promise: Helps parsing.
Unpromise:
(1) Why parse?
(2) This is a mere redescription
• I thought the book to Mary.

Promise: NTH, -
Unpromise: Has atypical ideas on what it means to “have” a concept
Let’s …

**Build robots**

Collect knowledge from people

- Promise: Machine Learning used to get /apple/ associated to RED, ...
- Unpromise: only as good as your concept metalanguage, which is sensorimotor by nature. Reading minds is much harder.

**Track Computers**

- Promise: If machines could understand what is collected, Plus-X goes away.
- Unpromise: (1) IF (2) Data without a theory.

Let’s

- Figure out how children learn
**The Problem of Ambiguity**

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

“Gavagai!”

**Constraints Guide Learning**

(Brown 1957, Rosch, Markman, Clark, Baldwin, Gleitman, Landau, ...)

- Syntactic Cues
  
  But also:
  
  - "Basic Level" Constraint
  - Whole Object Constraint
  - Mutual Exclusivity
  - Eye gaze-referent
  - Shape bias
  - Theory of mind

Look! He’s sebbing!  Look! A seb!

Look, some seb!
Look! Glipping!

Look! He is glipping water into the glass!

Look! He is glipping the glass with water!

/glip/ means POUR

/glip/ means FILL

**KEY HUMAN COMPETENCE:**

- Fast-mapping of words from syntax & semantics

---

Semantic Bootstrapping: Distribution of *scenes* determines word-concept mapping

(Pinker 1989)

Syntactic Bootstrapping: Distribution of *syntactic frames* disambiguates

(Gleitman 1990, Naigles 1990, Fisher et al 1994, ...)

---

6.863J/9.611J SP04 Lecture 21
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”...)
... more than just real-world observation...

Syntactic Bootstrapping
(Landau and Gleitman 1986, Naigles 1990)

Syntactic frames provide evidence for meaning:

\( H_1: \) arm wheel

\( /X \text{ and } Y \text{ are gorping}/ \)

\( /\text{Look, gorping}/ \)

\( /X \text{ is gorping } Y/ \)

\( H_2: \) cause to squat

\( /X \text{ is gorping } Y/ \)

Temporal ambiguity
Situation ambiguity
Mental unobservable!
One-shot learning

within a Bayesian framework.

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

Bayesian Learning at the Syntax-Semantics Interface

Syntactic Evidence  Semantic Evidence  | Evidence x

Linguistic Theory  \[ H = \{H_1, H_2, ...\} \]
Prior: \( p(H_i) \)
 Likelihood \( p(x|H_i) \)

BAYESIAN Acquisition Device

Acquired Lexicon

\( /\text{seb}/ \text{ means Posterior: } \)

\( p(H_i|x) \)

Syntactic Evidence

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

Semantic Evidence

pour-fill: (G001, W110)
pour-fill: (G001, W110)
pour-fill: (G001, W110)

Syntactic Theory  \[ \rightarrow \]

Bayesian Language Acquisition Device

Semantic Theory

Acquired Lexicon

\( /\text{gorp}/ \text{ means POUR} \)
\( /\text{pilk}/ \text{ means FILL} \)
\( /\text{jebi}/ \text{ means POUR or FILL} \)
Verbs Classes Grouped by Cause

<table>
<thead>
<tr>
<th>Feature</th>
<th>Hi Verb Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</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
Hᵢ in H
Evidence x in X = {0, 1}

Learning Value of Verb’s Cause Feature

<table>
<thead>
<tr>
<th>Feature</th>
<th>Acquired Lexicon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntactic Theory:</td>
<td>/He glipped the balloon/</td>
</tr>
<tr>
<td>H = {H₁, H₀, H*}</td>
<td>x = F₁</td>
</tr>
<tr>
<td>Prior:</td>
<td>p(H₁) = .333</td>
</tr>
<tr>
<td>Likelihood</td>
<td>p(x</td>
</tr>
<tr>
<td>x = F₀</td>
<td>H₁: .05, H₀: .95, H*: .50</td>
</tr>
<tr>
<td>x = F₁</td>
<td>H₁: .95, H₀: .05, H*: .50</td>
</tr>
<tr>
<td>p(Hᵢ</td>
<td>x) = p(x</td>
</tr>
<tr>
<td>p(H₀</td>
<td>x=F₁) = .033</td>
</tr>
<tr>
<td>p(H*</td>
<td>x=F₁) = .333</td>
</tr>
</tbody>
</table>

Posterior p(Hᵢ|x) = p(x|Hᵢ)p(Hᵢ)/p(H) = (.95)(.33)/(.95+.95+.50)(.33)
Example 3: How to Learn Semantics from Syntax

fg Verb Class

- Verbs of “Terminal” Coincidence (Figure)
  - F1: He sprayed water into the glass.
  - * F0: He sprayed the glass with water.
    (pour, spill ... – Figure "manner of motion")

- Verbs of “Central” Coincidence (Ground)
  - * F1: He filled water into the glass.
  - F0: He filled the glass with water.
    (cover, ... – Ground “change of state”)

- Alternating Verbs
  - F1: He loaded the wagon with hay.
  - F0: He loaded the hay onto the wagon.

Syntactic bootstrapping (Gleitman) - distribution of syntactic frames aids learning

If 25 /X glipped Y into Z/ => /glip/ like /pour/ (fg:[1])
If 25 /X glipped Y with Z/ => /glip/ like /fill/ (fg:[0])
If 12 /X glipped Y into Z/ + 13 /X glipped Y with Z/ => /glip/ like /load/ (fg:[*])

Language Acquisition Device

Acquired Syntactic Knowledge

| Lexicon: | X       | p(H_1|X) | p(H_0|X) | p(H_*=X) |
|----------|---------|---------|---------|----------|
| /glip/   | F1      | .633    | .033    | .333     |
| /gorp/   | F1^2    | .781    | .002    | .002     |
| /seb/    | F0,F1   | .137    | .137    | .724     |
| /meef/   | F0,F1^5 | .712    | 5e-6    | .288     |
| /foom/   | F0^6    | 2e-8    | .979    | .021     |

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/
Example 2: Learning Verb Semantics (ILP)

Verb meanings are logic programs: (Siskind 1996)

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)

Verb Logic Program
/lower/ 1-1*101**-11*
/raise/ 1-1*011**-11*
/rise/ 0-1*01***
/fall/ 0-1*10***

Hypothesis space \( H \): All LPs
Evidence \( X \): Bit Vector Examples
(e.g. 1-1010100-110)

Learning Problem: \( p(H_i|X) \)
(Inverting a ROM -- c.f.Yip & Sussman 1997)

Semantic Bootstrapping - distribution of semantic evidence results in verb acq.

Semantic Evidence:
/Look! Glipping!/ X1=000
/Look! Gorping!/ X2=000,001
/Look! Sebbing!/ X3=000,000,000
/Look! Meefing!/ X4=000,101,010,111,000

Language Acquisition Device

Acquired Semantic Knowledge

| Lexicon | p(H_{grip}|X) | p(H_{gorp}|X) | p(H_{seb}|X) | p(H_{meef}|X) |
|---------|--------------|--------------|--------------|--------------|
| grip    | .30          | .15          | .07          | .03          |
| gorp    | .00          | .64          | .16          | .04          |
| seb     | .70          | .09          | .01          | .001         |
| meef    | .00          | .00          | .00          | 1.0          |

6.8633/9.6113 SP04 Lecture 21
Example 3: Semantic Agreement

**Syntax-Semantic Theory:**

<table>
<thead>
<tr>
<th>Probe</th>
<th>Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>/spray/</td>
<td>/water into glass/</td>
</tr>
<tr>
<td>fg: [1] ↔ fg: [1]</td>
<td></td>
</tr>
<tr>
<td>/spray/</td>
<td>/glass with water/</td>
</tr>
<tr>
<td>fg: [1] ↔ fg: [0]</td>
<td></td>
</tr>
<tr>
<td>/fill/</td>
<td>/water into glass/</td>
</tr>
<tr>
<td>fg: [0] ↔ fg: [1]</td>
<td></td>
</tr>
<tr>
<td>/fill/</td>
<td>/glass with water/</td>
</tr>
<tr>
<td>fg: [0] ↔ fg: [0]</td>
<td></td>
</tr>
<tr>
<td>/load/</td>
<td>/hay into wagon/</td>
</tr>
<tr>
<td>fg: [*] ↔ fg: [1]</td>
<td></td>
</tr>
<tr>
<td>/load/</td>
<td>/wagon with hay/</td>
</tr>
<tr>
<td>fg: [*] ↔ fg: [0]</td>
<td></td>
</tr>
</tbody>
</table>

Hypothesis space $H = \{H_1, H_0, H_\star\}$

Priors $p(P) = .25$

Likelihood $p(P, G)$

<table>
<thead>
<tr>
<th>$G$</th>
<th>$P=0$</th>
<th>$P=1$</th>
<th>$P=\star$</th>
<th>$P=*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.165</td>
<td>.0025</td>
<td>.1225</td>
<td>.0825</td>
</tr>
<tr>
<td>1</td>
<td>.0025</td>
<td>.165</td>
<td>.1225</td>
<td>.0825</td>
</tr>
<tr>
<td>*</td>
<td>.0025</td>
<td>.0025</td>
<td>.0025</td>
<td>.0025</td>
</tr>
<tr>
<td>-</td>
<td>.08</td>
<td>.08</td>
<td>.0025</td>
<td>.08</td>
</tr>
</tbody>
</table>

Example 3: Syntactic Bootstrapping - distribution of syntactic frames results in verb acq.

**Syntactic Evidence:**

4 /X glipped Y with Z/
4 /X pilked Y/  
2 /X jirged Y/, 2 /X jirged Y/  
23 /X meefed Y with Z/, 1 /X meefed Y into Z/  

**Syntactic Theory:**

Hypothesis space $H = \{H_1, H_0, H_\star, H_\star\}$

Priors $p(P)$

Likelihood $p(P, G)$

**Language Acquisition Device**

**Acquired Linguistic Knowledge**

| Lexicon | $p(H_0|X)$ | $p(H_1|X)$ | $p(H_\star|X)$ | $p(H_\star|X)$ |
|---------|------------|------------|---------------|---------------|
| /glip/  | .732       | .00        | .222          | .046          |
| /pilk/  | .319       | .319       | .000          | .361          |
| /jirg/  | .789       | .000       | .000          | .210          |
| /meef/  | .998       | .000       | .002          | .000          |
Example 2: Learning Semantic Features

**Semantic “Theory”:** (3 bits)

- **Hypothesis space** $H$: 27 LPs
  - $q$:
    - 0: 000, 001, 010, 011, 100, 101, 110, 111
    - 1: 00*, 01*, 10*, 11*, 0*0, 0*1, 1*0, 1*1
    - 2: 0**, 1**, *0*, *1*, **0, **1
    - 3: ***

- **Bayesian Language Acquisition Device**

**Semantic Evidence:**

- /Look! Glipping!/ $X=000$
- /Look! Gorping!/ $X=000,001$
- /Look! Sebbing!/ $X=000,000,000$
- /Look! Meefing!/ $X=000,101,011,111,000$

**Acquired Semantic Knowledge**

- Lexicon: $p(H_{000}|X) p(H_{00*}|X) p(H_{0**}|X) p(H_{***}|X)$

**Example 3: Semantic and Syntactic Bootstrapping**

**Syntax-Semantic Theory:**

- Hypothesis space $H_{fg} = \{H_p, H_o, H_s\}$

**Verb $V$**

- /pour/ /water into glass/
  - $fg: [1] \leftrightarrow fg: [1]$

- /pour/ /glass with water/
  - $fg: [0] \leftrightarrow fg: [1]$

- /fill/ /water into glass/
  - $fg: [0] \leftrightarrow fg: [0]$

- /load/ /hay into wagon/
  - $fg: [*] \leftrightarrow fg: [1]$

**Complement $C$**

- /water into/ glass/
- /glass with/ water/
- /into/ fg: [1]
- /with/ fg: [0]
- /load/ with hay/
- /wagon with/ hay/
- /load/ fg: [0]
Semantic Evidence

Description

pour-fill (G001, W110) X pours W into G, filling G
splash-fill (G001, W120) X splashes W into G, filling G
spray-fill (G001, W130) X splash-es W into G, filling G
pour-empty (G002, W110) X pours W from G, emptying G
pour-none (G000, W130) X pours W into G

Syntactic Evidence

Attention Features

/X Ved water into glass/ W 1 * *
/X Ved glass with water/ G 0 * *
/Look, Ving!/ * * *

H

Features

POUR 1 1 *
SPLASH 1 2 *
SPRAY 1 3 *
FILL 0 * 1
EMPTY 0 * 2
MOVE 1 * *

Bayesian Language Acquisition Device

Acquired Lexicon

GRAND FINALE

Syntactic-Semantic Evidence X

pour-fill /X glipped water into glass/ .889
pour-fill /X glipped glass with water/ .000 .000 .990 .009 .0001
pour-fill /Glipping!/ .468 .004 .004 .468 .004 .049
pour-empty/X glipped water from glass/.998 .000 .000 .000 .000 .002
pour-none/X glipped water/

Acquired Lexical Knowledge p(H|X)

pour spray splash fill empty move

.008 .008 .000 .000 .093
.000 .000 .990 .009 .0001
.004 .004 .468 .004 .049
.998 .000 .000 .000 .000 .002

6.8633/9.6113 SP04 Lecture 21
Semantic Evidence Description
pour-fill (G001, W110) Person pours water into a glass, filling it
splash-fill (G001, W120) Person splashes water into a glass, filling it
spray-fill (G001, W130) Person splashes water into a glass, filling it
pour-empty (G002, W110) Person pours water from a glass, emptying it
pour-none (G000, W130) Person pours some water into a glass

Syntactic Evidence Attention Features
/X Ved water into glass/ W 1 - -
/X Ved glass with water/ G 0 - -
/Look, Ving!/ - - - -

Concept Features
Hpour 11-
Hsplash 12-
Hspray 13-
Hempty 0-1
Hfill 0-2
Hmove 1--

Acquired Lexical Knowledge p(Hi|X)

Scene-Utterance Evidence X | pour | spray | splash | fill | empty | move
---|---|---|---|---|---|---
pour-fill /X glipped water into glass/ | .880 | .010 | .010 | .000 | .000 | .101
pour-fill /X glipped glass with water/ | .000 | .000 | .000 | .989 | .011 | .0001
pour-fill /Glipping!/ | .463 | .006 | .006 | .463 | .005 | .058
none /X glipped water into glass/ | .246 | .246 | .246 | .004 | .004 | .254
none /X glipped glass with water/ | .007 | .007 | .007 | .485 | .485 | .007
none /Glipping!/ | .166 | .166 | .166 | .166 | .166 | .170
pour-fill /Glipping!/ | .998 | .000 | .000 | .000 | .000 | .002
pour-empty/X glipped water from glass/ | .061 | .066 | .066 | .000 | .000 | .806
pour-none/X glipped water/ | .061 | .066 | .066 | .000 | .000 | .806

GRAND FINALE
WHY IT WORKS

Fast-mapping of words from syntactic and semantic evidence can be modeled within a Bayesian framework.

- **Syntactic Evidence** (frames)
- **Semantic Evidence** (scenes)

**Linguistic Theory** → **Bayesian Language Acquisition Device**

**Evidence x**

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

**Acquired Lexicon**

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