The Neural Basis of Thought and Language

Final Review Session
Administrivia

- Final in class Thursday, May 8th
  - Be there on time!
  - Format:
    - closed books, closed notes
    - short answers, no blue books
- A9 due tonight
- Final paper due on Monday, May 12
Resources

- Textbook!
- Class slides
- Section slides
- Joe Makin's class notes from 2006
  - on notes page
Overview

• Bailey Model
  – feature structures
  – Bayesian model merging

• Bayes Nets

• KARMA
  – X-schema, frames
  – aspect
  – event-structure metaphor
  – inference

• FrameNet
  – frames
  – image schemas

• Reinforcement Learning

• ECG
  – SemSpecs
  – parsing
  – constructions
  – learning algorithm
Bailey’s VerbLearn Model

- 3 Levels of representation
  1. cognitive: words, concepts
  2. computational: f-structs, x-schemas
  3. connectionist: structured models, learning rules
- Input: labeled hand motions (f-structs)
- learning:
  1. the correct number of senses for each verb
  2. the relevant features in each sense, and
  3. the probability distributions on each included feature
- execution: perform a hand motion based on a label
Bayes Nets

- Probability
- Bayes' Rule / Product Rule
  - $P(x,y) = P(x) \cdot P(y|x)$
  - $= P(y) \cdot P(x|y)$
  - $P(x|y) = P(x) \cdot P(y|x) / P(y)$
- Write factored distribution $P(x,y,z,\ldots) = P(x) \cdot P(y|x) \ldots$
- Infer distributions over variables given evidence
  - variable elimination (by summation: $P(x) = \sum_y P(x,y)$)
- Temporal Bayes' Nets
Event Structure Metaphor

- States are Locations
- Changes are Movements
- Causes are Forces
- Causation is Forced Movement
- Actions are Self-propelled Movements
- Purposes are Destinations
- Means are Paths
- Difficulties are Impediments to Motion
- External Events are Large, Moving Objects
- Long-term, Purposeful Activities are Journeys
Ego Moving versus Time Moving

Fig. 1. (a) Schematic of the ego-moving schema used to organize events in time. (b) Schematic of the time-moving schema used to organize events in time.
## Results

<table>
<thead>
<tr>
<th></th>
<th>Meeting is Monday</th>
<th>Meeting is Friday</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRIME</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ego Moving</td>
<td>26.7%</td>
<td>73.3%</td>
</tr>
<tr>
<td>Object Moving</td>
<td>69.2%</td>
<td>30.8%</td>
</tr>
</tbody>
</table>
KARMA

- DBN to represent target domain knowledge
- Metaphor maps link target and source domain
- X-schema to represent source domain knowledge
X-Schemas

• Active representation

• Has hierarchical actions
  – defined by network structure

• Actions have structure (e.g. ready, iterating, ongoing, failed, complete)
  – defined by network structure

• Properly-designed nets will be goal-directed
  – take best actions to reach goal, given current context
  – related to “reinforcement learning”
Reinforcement Learning

- unsupervised learning
- learn behaviors
- reward
  - discounts
- use estimated future value

\[
V(s) = \max_a Q(s, a) \\
Q(s, a) = E[r(s, a) + \gamma V(s')] 
\]
Reinforcement Learning

• Learning methods
  – Value iteration
  – Q-learning

• Biology
  – dopamine = reward difference
    • only for reward, not punishment
  – non-exponential discounting
    • preference switching
Language

- Grammar
  - Syntax
- Semantics
- Metaphor
- Simulation
- Unification
Grammar

• A grammar is a set of rules defining a formal language
• a common example is Context-Free Grammar
  
  \[ \alpha \rightarrow \beta \]

\[ \forall \ \alpha: \text{single non-terminal} \]

• \( \beta \) : any combination of terminals and non-terminals

S \rightarrow NP \ VP
NP \rightarrow \text{Det Noun} \ | \ \text{ProperNoun}
VP \rightarrow \text{Verb NP} \ | \ \text{Verb PP}
PP \rightarrow \text{Preposition NP}

Noun \rightarrow \text{kiwi} \ | \ \text{orange} \ | \ \text{store}
ProperNoun \rightarrow \text{Pat} \ | \ \text{I}
Det \rightarrow \text{a} \ | \ \text{an} \ | \ \text{the}
Verb \rightarrow \text{ate} \ | \ \text{went} \ | \ \text{shop}
Preposition \rightarrow \text{to} \ | \ \text{at}
Sentence generation: *Pat ate the kiwi*

- start from S and apply any applicable rules
- forward expansion

```
S → NP VP
NP → Det Noun | ProperNoun
VP → Verb NP | Verb PP
PP → Preposition NP
Noun → kiwi | orange | store
ProperNoun → Pat | I
Det → a | an | the
Verb → ate | went | shop
Preposition → to | at
```
Unification Grammar

- Basic idea: capture these agreement features for each non-terminal in feature structures

\[
\begin{align*}
\text{I} & \left( \text{agreement} \left\{ \text{number: SG}, \text{person: 1st} \right\} \right) \\
\text{Went} & \left( \text{agreement} \right) \\
\text{Pat} & \left( \text{agreement} \left\{ \text{number: SG}, \text{person: 3rd} \right\} \right) \\
\text{Shop} & \left( \text{agreement} \left\{ \text{number: }, \text{person: 1st} \right\} \right)
\end{align*}
\]

Enforce constraints on these features using unification rules

\[
\begin{align*}
\text{VP} & \rightarrow \text{Verb NP} \\
\text{VP.agreement} & \leftrightarrow \text{Verb.agreement} \\
\text{S} & \rightarrow \text{NP VP} \\
\text{NP.agreement} & \leftrightarrow \text{VP.agreement}
\end{align*}
\]
Poverty and Opulence

• Poverty of the stimulus
  – Coined to suggest how little information children have to learn from

• Opulence of the substrate
  – Opulence = "richness"
  – Coined in response to suggest how much background information children have
“Harry walked into the café.”
The INTO construction

construction INTO
subcase of Spatial-Relation
form
  self_f.orth ← “into”
meaning: Trajectory-Landmark
  evokes Container as cont
  evokes Source-Path-Goal as spg
trajector ← spg.trajector
landmark ← cont
cont.interior ← spg.goal
cont.exterior ← spg.source
The Spatial-Phrase construction

construction SPATIAL-PHRASE

constructional
constituents
sr : Spatial-Relation
lm : Ref-Expr

form
sr_f before lm_f

meaning
sr_m · landmark ↔ lm_m
The Directed-Motion construction

construction DIRECTED-MOTION

constructional

constituents

a : Ref-Exp
m: Motion-Verb
p : Spatial-Phrase

form

af before mf
mf before pf

meaning

evokes Directed-Motion as dm
selfm.scene m dm
dm.agent m am
dm.motion m m
dm.path m pm

schema Directed-Motion

roles

agent : Entity
motion : Motion
path : SPG
Do not forget the SemSpec!
What exactly is simulation?

- Belief update and/or X-schema execution
Learning-Analysis Cycle (Chang, 2004)

1. Learner passes input (Utterance + Situation) and current grammar to Analyzer.


3. Learner updates grammar:
   a. Hypothesize new map.
   b. Reorganize grammar (merge or compose).
   c. Reinforce (based on usage).
Three ways to get new constructions

- Relational mapping
  - throw the ball

- Merging
  - throw the block
  - throwing the ball

- Composing
  - throw the ball
  - ball off
  - you throw the ball off

\[ \text{THROW} < \text{BALL} \]
\[ \text{THROW} < \text{OBJECT} \]
\[ \text{THROW} < \text{BALL} < \text{OFF} \]
Grammar merging

• How can we measure description length?
  – complicated rules are bad
  – lots of rules are bad
    • measure “derivation length”
  – $\alpha \times \text{size(rules)} + \text{derivationCost(rules, sentences)}$
How do you learn...

the meanings of spatial relations,

the meanings of verbs,

the metaphors, and

the constructions?
How do you learn...

the meanings of spatial relations,

the meanings of verbs,

the metaphors, and

the constructions?

That’s the Regier model.
How do you learn…

the meanings of spatial relations,

the meanings of verbs,

the metaphors, and

the constructions?

That's Bailey's model
How do you learn…

the meanings of spatial relations,

the meanings of verbs,

the metaphors, and

the constructions?

conflation hypothesis
( primary metaphors)
How do you learn...

the meanings of spatial relations,

the meanings of verbs,

the metaphors, and

the constructions?

construction learning
further questions?