Starting from Scratch in Semantic Role Labeling

Dan Roth

Department of Computer Science
University of Illinois at Urbana-Champaign
Connecting Language to the World

Can I get a coffee with sugar and no milk

Semantic Parser

MAKE(COFFEE,SUGAR=YES,MILK=NO)

Great!

Can we rely on this interaction to provide supervision (and, eventually, recover meaning)?

- How to recover meaning from text?
  - Annotate with meaning representation; use (standard) “example based” ML
    - Teacher needs deep understanding of the learning agent
    - Annotation burden; not scalable.
  - Instructable computing
    - Natural communication between teacher/agent
Scenarios I: Understanding Instructions \[\text{IJCAI'11, ACL'13, MLJ'13}\]

- Understanding Games’ Instructions
  
  A top card can be moved to the tableau if it has a different color than the color of the top tableau card, and the cards have successive values.

- An automated learner reads natural instructions and “understands” them by interacting with the game
  - Agnostic of agent's internal representations
  - Contrasts with traditional 'example-based' ML

A series of works: Dan Goldwasser’s Thesis
Starting from Scratch in Semantic Role Labeling

Michael Connor            Yael Gertner            Cynthia Fisher

How do children begin to understand sentences?

- Topid rivvo den marplox.
Language Acquisition

“The gubish fib fribs the nurt nit”

The dog chases the cat?
The cat is quick?
The cat and the dog are playing?
It’s a nice day today?

The cat flees the dog?

Traditional view: knowledge of meaning drives learning of words and syntax
Language Acquisition

“The gubish **dog** fribs the nurt **cat**”

The dog chases the cat?
The cat is quick?
The cat and the dog are playing?
Its a nice day today?

The cat flees the dog?
The dog is excited?
Language Acquisition

(Agent) (Patient)
“The gubish **dog** fribs the nurt **cat**”

The dog chases the cat?
The cat is quick?
The cat and the dog are playing?

The cat flees the dog?
The dog is excited?

It's a nice day today?

- What if we know or predict abstract semantics of sentence?
  - Who does what to whom?
Language Acquisition

(Agent) (Patient)
“The gubish dog fribs the nurt cat”

The dog chases the cat?
The cat is quick? The cat flees the dog?
The cat and the dog are playing? The dog is excited?
It's a nice day today?

- Neither syntax or semantics are obvious from sentence or scene
- Only by using limited knowledge of both can we learn
  - Even simple structural representations (such as identity of nouns) can help
The ambiguity of situations as evidence for sentence meaning

The situation

![Image of a child playing with toys]

The sentence?

You can put the blue one there.
The blue one goes there.
Try the blue one.

It’s my turn.
She wants me to put this here.

The gap between the speaker’s likely communicative goal and her words is the unknown unknown.
The language-world mapping problem

“the language”
[Topid rivvo den marplox.]

“the world”
Observe how words are distributed across situations

Smur! Rivvo della frowler.

Topid rivvo den marplox.

Blert dor marplox, arno.

Marplox dorinda blicket.
Structure-mapping: A proposed starting point for syntactic bootstrapping

- Children can learn the meanings of some *nouns* via cross-situational observation alone [Fisher 1996, Gillette, Gleitman, Gleitman, & Lederer, 1999; Snedeker & Gleitman, 2005]

- But how do they learn the meaning of verbs?
  - Sentences comprehension is grounded by the acquisition of an initial set of concrete nouns
  - These nouns yields a *skeletal sentence structure* — candidate arguments; cue to its semantic predicate—argument structure.
  - Represent sentence in an abstract form that permits generalization to new verbs

  ![Example Sentences](example_sentences.png)

- Nouns identified

  - [Topid rivvo den marplox.]
Syntactic Bootstrapping Makes Three main claims:

1. **Structure-mapping**: Syntactic bootstrapping begins with an unlearned bias toward one-to-one mapping between nouns in sentences and semantic arguments of predicate terms (Fisher et al., 1994; Gertner & Fisher, 2012; Gillette et al., 1999; Yuan, Fisher & Snedeker, 2012)

2. **Early abstraction**: Learners are biased toward abstract representations of language experience (Gertner, Fisher & Eisengart, 2006; Pinker, 1989; Thothathiri & Snedeker, 2008)

A fourth corollary:

- (4) 'Real' syntactic bootstrapping: Children can learn which words are verbs by tracking their syntactic argument-taking behavior in sentences.

Hey, she pushed her. 
Will you push me on the swing? 
John pushed the cat off the sofa ...
**Strong Predictions** [Gertner & Fisher, 2006]

- Test 21 month olds on assigning arguments with novel verbs
- How order of nouns influences interpretation: Transitive & Intransitive

Agent-first: The boy and the girl are daxing!
Transitive: The boy is daxing the girl!
Agent-last: The girl and the boy are daxing!

Error disappears by 25 months
preferential looking paradigm
BabySRL

- Develop a machine learning model to support psycholinguistic theories of syntactic bootstrapping in early stages of language acquisition

- Develop Semantic Role Labeling System (BabySRL) to experiment with theories of early language acquisition
  - SRL as minimal level language understanding
  - Determine who does what to whom.

- **Realistic Computational model** for Syntactic Bootstrapping
  - Verbs meanings are learned via their syntactic argument-taking roles
  - Semantic feedback to improve **syntactic & meaning representation**

- Inputs and knowledge sources
  - Only those we can defend children have access to
Verb Meaning = Semantic Role Labeling (SRL)

- A semantic analysis of the sentences at the level of who does what to whom
- For each verb in the sentence:
  - SRL system tries to identify all constituents that fill a semantic role
  - and to assign them roles (agent, patient, goal, etc.)
- Inspired by our SRL program (Punyakanok et. al 05,08), completely different learning approach.

<table>
<thead>
<tr>
<th>Remember</th>
<th>V: remember</th>
</tr>
</thead>
<tbody>
<tr>
<td>what</td>
<td>A1: patient</td>
</tr>
<tr>
<td>Daddy</td>
<td>A0: agent</td>
</tr>
<tr>
<td>said</td>
<td>V: say</td>
</tr>
</tbody>
</table>
The nature of the semantic roles

- A key assumption of the SRL is that the semantic roles are abstract:
  - This is a result of the PropBank annotation scheme: verb specific roles are grouped into macro-roles (e.g., Dowty, 1991)

<table>
<thead>
<tr>
<th>Like:</th>
<th>Give:</th>
<th>Have:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg0: Liker</td>
<td>Arg0: Giver</td>
<td>Arg0: Owner</td>
</tr>
<tr>
<td>Arg1: Object of affection</td>
<td>Arg1: Thing given</td>
<td>Arg1: Possession</td>
</tr>
<tr>
<td></td>
<td>Arg2: Recipient</td>
<td></td>
</tr>
</tbody>
</table>
3 corpora of child-directed speech

- **Annotated** parental utterances using PropBank annotation scheme
- Speech to Adam, Eve, Sarah (Brown, 1973)
- Adam 01-23 (2;3 - 3;2)
  - Train on 01-20: 3951 propositions, 8107 arguments
- Eve 01-20 (1;6 - 2;3)
  - Train on 01-18: 4029 propositions, 8499 arguments
- Sarah 01-90 (2;3 - 4;1)
  - Train on 01-83: 8570 propositions, 15599 arguments
Testing the Baby SRL

- Constructed test sentences like those used in experiments with children
  - Unknown verbs & two animate nouns force the SRL to rely on syntactic knowledge
- "Adam krads Daddy!"

\[
\begin{align*}
\text{Adam} & \quad \text{Mommy} \\
\text{Daddy} & \quad \text{Ursula} \\
\ldots & \\
\end{align*}
\quad \text{krad} \quad \begin{align*}
\text{Adam} & \quad \text{Mommy} \\
\text{Daddy} & \quad \text{Ursula} \\
\ldots & \\
\end{align*}
\]
BabySRL: Key Components

- **Representation:**
  - Theoretically motivated representation of the input
  - Shallow, *abstract*, sentence representation consisting of
    - # of nouns in the sentence
    - Noun Patterns (1\textsuperscript{st} of two nouns)
    - Relative position of nouns and predicates

- **Learning:**
  - Guided by knowledge kids have
    - Classify words by *part-of-speech* (distinct states)
    - Identify *arguments and predicates*
    - Determine the role arguments take
BabySRL: Early Results

- Fine grained experiments with how language is represented
- Test different **levels of representation**

Primary Focus:
- Hypothesis: number and order of nouns important
  - Once we know some nouns, can use them to represent structure
  - NPattern gives count and placement: first of two, second of three, etc.

Alternative:
- Verb Position
  - Target argument is before or after verb

Key Finding:
- **NPattern reproduces errors in children**
  - Promotes A0-A1 interpretation in transitive, but also intransitive sentences
  - Verb position **does not make this error**
    - Incorporating it recovers correct interpretation
    - **But:** requires the ability to recognize the predicate, a harder (and later) task

- Target argument is before or after verb
BabySRL: Key Components

- **Representation:**
  - Theoretically motivated representation of the input
  - Shallow, abstract, sentence representation consisting of
    - # of nouns in the sentence
    - Noun Patterns (1st of two nouns)
    - Relative position of nouns and predicates

- **Learning:**
  - Guided only by knowledge kids have
    - Classify words by part-of-speech
    - Identify arguments and predicates
    - Determine the role arguments take
Minimally Supervised BabySRL

- Protocol: Provide little prior knowledge & only high level semantic feedback
  - Defensible from psycholinguistic evidence

- Unsupervised “parsing”
  - Identifying part-of-speech states

- Argument Identification
  - Identify Argument States
  - Identify Predicate States

- Argument Role Classification
  - Labeled Training using predicted arguments

- Learning is done from CHILDES corpora
- Indirect supervision driven from very week scene feedback

Learning with Indirect Supervision

Input + Distributional learning

Structured Intermediate Representation (no supervision)

Binary weak supervision for the final decision
Various levels of feedback, from roles, to superset roles

Classify role of arguments using features of the intermediate representation

Identify verbs as non-nouns with consistent argument patterns

Identify arguments with seed nouns, and augment with those in same states.

Latent Representation: Replace pipeline for arguments and predicates by treating structure of sentence (arg/predicates) as hidden structure

Unsupervised HMM: initialized with knowledge of function words.

Hand in “Hand me the spoon” and “Put the spoon in my hand” are in distinct states.

Various levels of feedback, from roles, to superset roles
Joint Semantics and Structure

Input Sentence (text) → Hidden Structure → HMM Identified Arguments

Hidden Structure

Set of Possible Interpretations → Single Possible Interpretation

Meaning from World
Online Latent Structure Training

Input: Sentences with *labels
Output: Two linear separators:
\[ u_t \] for hidden structure, \( w_t \) for role classification

For each sentence:

Find best joint hidden structure and labeling
\[ (h^*_i, y^*_i) \gets \arg \max_{h \in H, y \in Y} \left[ u_t \cdot \Phi_w(x_i, h, y) + w_t \cdot \Phi_u(x_i, h) \right] \]

Update \( u \) to predict \( h^* \)
\[ h' \leftarrow \arg \max_h u_t \cdot \Phi_u(x_i, h) + C \cdot 1[h \neq h^*_i] \]
\[ u_{t+1} \leftarrow u_t + \alpha_u(\Phi_u(x_i, h^*_i) - \Phi_u(x_i, h')) \]

Update \( w \) based on \( h^* \) to predict \( y^* \)
\[ y' \leftarrow \arg \max_y w_t \cdot \Phi_w(x_i, h^*_i, y) + C \cdot 1[y \neq y^*_i] \]
\[ w_{t+1} \leftarrow w_t + \alpha_w(\Phi_w(x_i, h^*_i, y^*_i) - \Phi_w(x_i, h^*_i, y)) \]
\[ t \leftarrow t + 1 \]

Use HMM identified nouns to constrain possible structures during inference and hidden structure (H)

Instead of true labels, use more ambiguous constraints

Unordered set of true labels

Unordered superset of true labels
Latent BabySRL Conclusions

- Semantics can provide cue for identifying structure, but ambiguous semantics needs some help
  - Small set of seed nouns provide a plausible structure that can constrain and begin to learn with plausible semantic ambiguity

- Developed online latent structure classifier that integrates constraints from both semantic feedback and syntactic knowledge
  - Demonstrate alternative means of supervision, incorporate partial knowledge from multiple available sources as feedback signal
  - Comparable to early stage in child language acquisition
Summary

- BabySRL: a realistic computational model for verb meaning acquisition via the structure-mapping theory.
  - Representational Issues
  - A machine learning framework that learns through combining multiple psycholinguistically plausible partial sources of information

- It is possible to begin learning sentence level semantics – Verb Meaning – once the learner is able to identify some nouns
  - And expects abstract semantics from the sentence
  - Simple, early representations are robust to noisy input and feedback

- Next Steps:
  - Bootstrap Language, handle growing complexity (argument structure)
  - Handle multiple predicates (verbs, prepositions,...)
  - Discourse: modeling missing argument

Thank You

1. Can the minimal feedback we use be “computed” with a vision system?
2. Perhaps we *must* start from scratch...