Synergies in learning words and their referents

Mark Johnson¹, Katherine Demuth¹, Michael Frank² and Bevan Jones¹

1: Macquarie University, 2: Stanford University, 3: University of Edinburgh

Abstract

We show how to construct adaptor grammars that perform word segmentation and map the words they learn to objects.

The non-linguistic context permits our “one topic per collocation” model to use more information about each collocation, at the cost of increased complexity.

The non-linguistic context permits our “one topic per collocation” model to use more information about each collocation, at the cost of increased complexity.

Prior work: mapping words to referents

- Input to learner:
  - word sequence: Is that the pig?
  - objects in nonlinguistic context: DOG, PIG

- Learning objectives:
  - identify utterance topic: PIG
  - identify word-topic mapping: PIG — PIG

Frank et al (2009) “topic models” as PCFGs

- Prefix each sentence with possible topic marker, e.g., PIG/DOG
- PCFG rules designed to choose a topic from possible topic marker and propagate it through sentence
- Each word is either generated from sentence topic or null topic
- Simple grammar modification requires at most one topical word per sentence
- Bayesian inference for PCFG rules and trees corresponds to Bayesian inference for word and sentence topics using topic model (Johnson, 2010)

Prior work: segmenting words in speech

- Running speech does not contain “pauses” between words
- Elman (1990) and Brent et al (1996) studied segmentation using an artificial corpus
- Child directed utterance: Is that the pig?

Synergies: an advantage of interactive learning

An interactive learner can take advantage of synergies in acquisition of linguistic components

- Semantic bootstrapping: semantics is learnt first, and used to predict syntax (Pinker, 1984)
- Syntactic bootstrapping: syntax is learnt first, and used to predict semantics (Gleitman, 1991)
- Conventional view of lexical acquisition, e.g., Kuhl (2004)
- Child first learns the phoneme inventory, which it then uses to learn phonological forms of words in the lexicon, ...

2. Interactive acquisition of all linguistic components together

- corresponds to joint inference for all components of language
- stages in language acquisition might be due to:
  - child’s input may contain more information about some components
  - some components of language may be learnable with less data

Two hypotheses about language acquisition

1. Pre-programmed staged acquisition of linguistic components
   - Semantic bootstrapping: semantics is learnt first, and used to predict syntax (Pinker, 1984)
   - Syntactic bootstrapping: syntax is learnt first, and used to predict semantics (Gleitman, 1991)

2. Interactive acquisition of all linguistic components together
   - corresponds to joint inference for all components of language
   - stages in language acquisition might be due to:
     - child’s input may contain more information about some components
     - some components of language may be learnable with less data

Does non-linguistic context help segmentation?

<table>
<thead>
<tr>
<th>Model</th>
<th>segmentation</th>
<th>topics</th>
<th>sentence-referent accuracy</th>
<th>f-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>unigram</td>
<td>not used</td>
<td>0.543</td>
<td>0.702</td>
<td>0.325</td>
</tr>
<tr>
<td>unigram</td>
<td>any number</td>
<td>0.537</td>
<td>0.719</td>
<td>0.347</td>
</tr>
</tbody>
</table>

Does better segmentation help topic identification?

<table>
<thead>
<tr>
<th>Task</th>
<th>identity object (if any) this sentence is about</th>
<th>Model</th>
<th>segmentation</th>
<th>topics</th>
<th>sentence-referent accuracy</th>
<th>f-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>unigram</td>
<td>not used</td>
<td>0.799</td>
<td>0.503</td>
<td>0.495</td>
<td></td>
<td></td>
</tr>
<tr>
<td>unigram</td>
<td>any number</td>
<td>0.728</td>
<td>0.440</td>
<td>0.493</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Does better segmentation help topic identification?

<table>
<thead>
<tr>
<th>Task</th>
<th>identity head nouns of NPs referring to topical objects (e.g., PIG in input PIG)</th>
<th>Model</th>
<th>segmentation</th>
<th>topics</th>
<th>sentence-referent accuracy</th>
<th>f-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>unigram</td>
<td>not used</td>
<td>0.149</td>
<td>0.321</td>
<td>0.636</td>
<td></td>
<td></td>
</tr>
<tr>
<td>unigram</td>
<td>any number</td>
<td>0.220</td>
<td>0.321</td>
<td>0.636</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Conclusions and future work

- Adaptor Grammars can express a variety of useful HDP models
  - generic AG inference code makes it easy to explore models
- There seem to be synergies a learner could exploit when learning word segmentation and word-object mappings
  - incorporating word-topic mapping improves segmentation accuracy
  - improving segmentation accuracy improves topic detection and acquisition of topical words

Caveat: results seem to depend on details of model

Future work:
- extend expressive power of AGs (e.g., phonology, syntax)
- richer data (e.g., more non-linguistic context)
- more realistic data (e.g., phonological variation)