Using OntoNotes PropBank and Syntax

Dilek Hakkani-Tür

dilek@icsi.berkeley.edu

Collaborators: Benoit Favre, ICSI
Sibel Yaman, ICSI
Gokhan Tur, SRI
Outline

• Syntax for Punctuation Insertion
  ♦ Sentence Boundaries
  ♦ Comma Boundaries

• Syntax and Semantics for Extraction of 5Ws

• Feedback/Conclusions
The summit, the first in seven years, had long been held up by an impasse over the attendance of Mugabe. Portugal, which holds the EU presidency, gave in to African leaders and invited Mugabe, making the summit possible.

The summit, the first in seven years, had long been held up by an impasse over the attendance of Mugabe. Portugal, which holds the EU presidency, gave in to African leaders and invited Mugabe, making the summit possible.

Further Processing (IE, MT, etc.)
Sentence Segmentation

- Word boundary classification with local features.
  - Features:
    - Lexical: words
    - Structural: speaker changes
    - Prosodic: Pitch, energy, and durations
    - Local and sequence classifiers.

- Using syntax: use parser as a language model?
Forming the Sentence
Hypothesis Lattice

- Huge search space
- Slow parsers
Forming the Sentence Hypothesis Lattice

- Non-sentence boundary with a high score according to the local model
- Sentence boundary with a high score according to the local model
- Candidate boundary

P(S) estimated by Berkeley Parser
Experiments

- F-measure on TDT-4 test set
- Local Classifier: Boosting

<table>
<thead>
<tr>
<th>Method</th>
<th>Mandarin (F-measure)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boosting</td>
<td>68.8</td>
</tr>
<tr>
<td>Boosting + syntax (BP, WSJ)</td>
<td>68.7</td>
</tr>
<tr>
<td>Boosting + syntax (BP, TDT-4 ASR, self-training)</td>
<td>70.1</td>
</tr>
<tr>
<td>Boosting + syntax (BP, OntoNotes)</td>
<td>71.1</td>
</tr>
</tbody>
</table>

- Training parser for ASR and training parser for speech both help.
- Future challenges:
  - supervised training using ASR data with reference parse trees.
  - Parsing for segmenting conversational speech
Conclusions and Feedback

- Training parsers for genre is beneficial.
  - More speech data (both conversations and news).
- Richer annotation for speech data:
  - Disfluencies including types (repair, repeat, filled pause, etc.).
  - Dialog act tags (e.g. backchannels and interruptions for conversations)
Outline

• Syntax for Punctuation Insertion
  ✷ Sentence Boundaries
  ✷ Comma Boundaries

• Syntax and Semantics for Extraction of 5Ws

• Feedback/Conclusions
Answer Extraction for 5W Questions

• Given 5W questions, find the exact answer.
• Find the 5Ws of sentences.

• Syntactic parses with function tags
  ✷ UMD parser (M. Harper)
  ✷ SRI parser (W. Wang)
  ✷ Berkeley parser

• Preprocessing:
  ✷ Detect quotes
  ✷ Find passives
  ✷ Find top level predicates

• Rules

The United States has had its share of warm weather this year.

WHO: The United States
WHAT: has had its share of warm weather
WHERE: null
WHEN: this year
WHY: null
Answer Extraction for 5W Questions

• Semantic role labeler
  ✷ Based on Pradhan et al.’s ASSERT.
  ✷ Trained using OntoNotes PropBank annotations.
  ✷ Uses Berkeley Parser with function tags, also trained using OntoNotes.

TARGET: **had**
ARG0: The United States
ARG1: its share of warm weather
ARGM-TMP: this year

WHO: The United States
WHAT: has had its share of warm weather
WHERE: null
WHEN: this year
WHY: null
Experiments for 5W Extraction

- Re-scoring/Re-ranking to select one output.
  - 2-state and 6-state HMM.
- Experiments with OntoNotes reference transcriptions, split into train/dev/test:

<table>
<thead>
<tr>
<th>WHO</th>
<th>Parser Correct</th>
<th>Parser Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRL Correct</td>
<td>90.8%</td>
<td>2.9%</td>
</tr>
<tr>
<td>SRL Incorrect</td>
<td>4.9%</td>
<td>1.4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WHAT</th>
<th>Parser Correct</th>
<th>Parser Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRL Correct</td>
<td>61.4%</td>
<td>11.9%</td>
</tr>
<tr>
<td>SRL Incorrect</td>
<td>11.3%</td>
<td>15.4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WHERE</th>
<th>Parser Correct</th>
<th>Parser Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRL Correct</td>
<td>39.7%</td>
<td>12.5%</td>
</tr>
<tr>
<td>SRL Incorrect</td>
<td>20.1%</td>
<td>27.7%</td>
</tr>
</tbody>
</table>

Best (Parser): 95.7%
Oracle: 98.6%
Merged: 94.0%

Best (SRL): 73.3%
Oracle: 84.6%
Merged: 74.8%

Best (Parser): 59.8%
Oracle: 72.3%
Merged: 67%
Conclusions and Feedback

• Training parsers for genre is beneficial.
  - More speech and web data.
• Function tags help 5W answer extraction, even though OBJ is not labeled, and hence wasn’t generated by the parser.
• Data from older dates: is ASR using them already?
• Active Learning:
  - Select data for parsing, for SRL, for ASR?