Broad-Coverage Semantic Parsing: A Transition-Based Approach

Daniel Hershcovitch, Omri Abend and Ari Rappoport
Hebrew University of Jerusalem
Outline

- Background
- Conversion-Based Parsing
- Broad-Coverage Semantic Parsing
Outline

● Background

● Conversion-Based Parsing

● Broad-Coverage Semantic Parsing
Given a sequence of tokens $w = w_1, \ldots, w_n$,

A (labeled) directed graph $(V, E)$ where $\{w_i\} \subseteq V$ is a grounded representation of $w$.

Examples:

Dependency/constituency trees, UCCA...
Syntactic representation is sensitive to formal variations.

Dependency:

Constituency:
Universal Conceptual Cognitive Annotation (UCCA)

Abend and Rappoport (2013)
Structural Properties

Properties required for full semantic coverage in grounded representations:

1. **Multiple parents (DAG).**

2. **Non-projectivity (discontinuity).**

3. **Non-terminal nodes.**
Broad-Coverage Parsing

- Non-projective dependency parsing
- Discontinuous constituency parsing
- Semantic dependency parsing (SDP)
Some frequent constructions do not have one clear head, e.g. coordination, some multi-word expressions, compounds.

Coordination in SDP:

Oepen et al. (2015)
Structural Properties in UCCA

H  parallel scene  D  adverbial  N  connector
A  participant  E  elaborator  R  relator
P  process  S  state  C  center
G  ground  L  linker  F  function

After graduation John moved to Paris

John went home

John and Mary

John gave everything up
Coordination represented by one parent node

Non-Terminal Units
Remote edges denote implicit relations
Multi-word expression annotated as one unit
Corpus Statistics

160K tokens from English Wikipedia
+25K tokens from *Twenty Thousand Leagues Under the Sea*

<table>
<thead>
<tr>
<th></th>
<th>Train</th>
<th>Dev</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td># passages</td>
<td>281</td>
<td>35</td>
<td>43</td>
</tr>
<tr>
<td># sentences</td>
<td>4021</td>
<td>537</td>
<td>608</td>
</tr>
<tr>
<td># nodes</td>
<td>277,587</td>
<td>40,700</td>
<td>45,047</td>
</tr>
<tr>
<td>% terminal</td>
<td>42.41</td>
<td>42.8</td>
<td>42.66</td>
</tr>
<tr>
<td>% non-term.</td>
<td>57.59</td>
<td>57.20</td>
<td>57.34</td>
</tr>
<tr>
<td>% discont.</td>
<td>0.52</td>
<td>0.55</td>
<td>0.47</td>
</tr>
<tr>
<td>% &gt;1 parent</td>
<td>2.29</td>
<td>1.89</td>
<td>2.21</td>
</tr>
</tbody>
</table>
Outline

- Background
- Conversion-Based Parsing
- Broad-Coverage Semantic Parsing
After graduation, John moved to Paris.

**B:** Buffer of nodes to process, initialized to the list of tokens.

**S:** Stack of partially processed nodes, initially just the root.

**G:** Graph of constructed edges.
After graduation, John moved to Paris.
Transitions to create new nodes

After graduation, John moved to Paris.
1. Convert UCCA to dependency and constituency trees.
2. Apply existing transition-based parsers.
3. Convert back to UCCA.

Dependency parsers:
MaltParser (Nivre 2003), Stack–LSTM Parser (Dyer et al. 2015)

Constituency parser:
UPARSE (Maier 2015): discontinuous constituency parser
Scores on the *Wiki* test set:

<table>
<thead>
<tr>
<th></th>
<th>Primary</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LP</td>
<td>LR</td>
<td>LF</td>
<td></td>
</tr>
<tr>
<td>Constituency Tree Conversion</td>
<td>UPARSE</td>
<td>64</td>
<td>67.3</td>
<td>65.4</td>
</tr>
<tr>
<td></td>
<td>Upper Bound</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Remote</td>
<td>LP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Primary</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LP</td>
<td>LR</td>
<td>LF</td>
<td></td>
</tr>
<tr>
<td>Dependency Tree Conversion</td>
<td>Malt_{arc-standard}</td>
<td>63.4</td>
<td>57.3</td>
<td>60.1</td>
</tr>
<tr>
<td></td>
<td>Malt_{arc-eager}</td>
<td>63.9</td>
<td>57.9</td>
<td>60.5</td>
</tr>
<tr>
<td></td>
<td>LSTM</td>
<td>73.2</td>
<td>66.2</td>
<td>69.2</td>
</tr>
<tr>
<td></td>
<td>Upper Bound</td>
<td>93.8</td>
<td>83.7</td>
<td>88.4</td>
</tr>
</tbody>
</table>

Upper bound is due to lossy conversion algorithms.
• MaltParser: perceptron/SVM

• UPARSE: perceptron

• Stack-LSTM parser: recurrent neural network + continuous features
Outline

- Background
- Conversion-Based Parsing
- Broad-Coverage Semantic Parsing
Discontinuous DAG parser

After graduation, John moved to Paris.
Scores on the *Wiki* test set and the *20K leagues* set:

<table>
<thead>
<tr>
<th></th>
<th>Primary</th>
<th></th>
<th>Remote</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LP</td>
<td>LR</td>
<td>LF</td>
<td>LP</td>
</tr>
<tr>
<td>BSP</td>
<td>62.4</td>
<td>56</td>
<td>59</td>
<td>15.3</td>
</tr>
<tr>
<td>BSP&lt;sub&gt;Tree&lt;/sub&gt;</td>
<td>63.8</td>
<td>56.5</td>
<td>59.9</td>
<td>11.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out-of-domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BSP</td>
<td>60.6</td>
<td>53.9</td>
<td>57.1</td>
<td>20.2</td>
</tr>
<tr>
<td>BSP&lt;sub&gt;Tree&lt;/sub&gt;</td>
<td>60.2</td>
<td>52.8</td>
<td>56.2</td>
<td>10.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*(BSP<sub>Tree</sub> trained on converted trees without remote edges)*
The structural desiderata of grounded semantic parsing is not supported by today’s parsers.

We present a transition-based system that does.

Encouraging results with UCCA suggest that NN-based classification may be helpful for better performance.
Future Work

- Neural network for BSP classifier
- Improved conversions
- Beam search
- More languages, e.g. German
Thank you
References