**Universal Dependency Parsing with a General Transition-Based DAG Parser**

Daniel Hershcovich\(^1,2\) & Omri Abend\(^2\) & Ari Rappoport\(^2\)

\(^1\)The Edmond and Lily Safra Center for Brain Sciences
\(^2\)School of Computer Science and Engineering
The Hebrew University of Jerusalem

{danielh,oaabend,arir}@cs.huji.ac.il

**Learning to parse enhanced dependencies jointly with basic Universal Dependency Parsing.**

gslab.com/CoNLL-2018/HUI

We extend **TUPA** [2, 3], a general DAG parser originally designed for **UCCA**: transition-based parser supporting reentrancy (DAG), discontinuity (non-projectivity) and non-terminal nodes.

**Transitions:**

- **SHIFT:**
- **REDUCE:**
- **NODE:**
- **RIGHT-EDGE:**
- **LEFT-EDGE:**
- **SWAP:**
- **FINISH:**

**Parser state:**

- Source
- Target
- Graph
- Node
- Enhanced

**Transition classifier:**

<table>
<thead>
<tr>
<th>Transition</th>
<th>MLP</th>
<th>BILSTM</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEAD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSUBJ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Root</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Example:**

- He went straight to work and finished the job efficiently and promptly.
- We were made to feel very welcome.

**Unified DAG Format**

We convert UD into a **UCCA**-like format supported by **TUPA**, by inserting non-terminal nodes.

**UCCA** (Universal Conceptual Cognitive Annotation): cross-lingual semantic representation [1]. Nodes are scenes/concepts. Primary edges form a tree. Remote edges (dashed) allow reentrancy.

**Enhanced Dependencies**

Some UD treebanks contain enhanced graphs with additional or augmented edges [5, 4].

- Conjoined predicates and arguments:
  - head
  - nsubj
  - aux
  - nsubj:pass
  - advmod
  - advcl
  - xcomp

- Null nodes due to elided predicates, case information:
  - he
  - was
  - all
  - happy
  - holidays
  - and
  - moved
  - E9
  - back
  - on
  - earth

- **Raising:**
  - We were made to feel very welcome.
  - He had a robe that was made back in the 90s.

- **Relative clause:**
  - He had a robe that was made back in the 90s.

**Results**

<table>
<thead>
<tr>
<th>Dataset</th>
<th>TUPA (official)</th>
<th>TUPA (unofficial)</th>
<th>UDPipe + CoreNLP 72.10</th>
<th>UDPipe + CoreNLP 72.08</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDPIpe + CoreNLP</td>
<td>77.66</td>
<td>21.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remote</td>
<td>72.10</td>
<td>57.13</td>
<td>72.33</td>
<td>49.12</td>
</tr>
<tr>
<td>Embed.</td>
<td>71.82</td>
<td>54.54</td>
<td>72.33</td>
<td>54.54</td>
</tr>
<tr>
<td>NER</td>
<td>69.23</td>
<td>42.29</td>
<td>75.49</td>
<td>21.68</td>
</tr>
<tr>
<td>PE3</td>
<td>62.07</td>
<td>56.82</td>
<td>59.56</td>
<td>21.68</td>
</tr>
<tr>
<td>Big treebanks</td>
<td>56.35</td>
<td>66.63</td>
<td>70.81</td>
<td>21.68</td>
</tr>
<tr>
<td>Small treebanks</td>
<td>36.74</td>
<td>55.01</td>
<td>74.38</td>
<td>21.68</td>
</tr>
<tr>
<td>All treebanks</td>
<td>53.69</td>
<td>65.80</td>
<td>75.49</td>
<td>21.68</td>
</tr>
</tbody>
</table>

**References**

1. [Enhanced LAS-F1 on test treebanks.](https://tinyurl.com/semeval-ucca)
2. [UDPipe + CoreNLP for enhanced dependencies.](https://tinyurl.com/semeval-ucca)