Motivation: First-order graph-based parsers cannot capture sufficient subtree information. Most of transition-based parsers cannot perform global optimization.

Observation: In a projective parse tree, the whole subtree rooted at a headword forms a contiguous sequence (i.e., span) in the surface order. We call such a span-headword pair as headed span. A projective parse tree can be viewed as a collection of headed spans, similar to the span-based view of constituency parsing.

Method: We decompose the score of a dependency tree into scores of headed spans into a single child span.

Conc: the rule used to obtain a headless child span from a headed span.

Our algorithm ≈ hook trick + head-splitting trick

The hook trick: reduce subtree into headless spans because the linking of heads and the concatenation of subtrees can be separated.

The head-splitting trick: split each subtree into a left and a right fragment. We can transform the above rules to R-CONC, L-CONC, and CONC using the head-splitting trick.

Parsing as deduction

The corresponding recursive formulas:

\[ \beta_{i,j+1} = s_{i,j+1}^{\text{span}} \]

\[ \alpha_{i,j} = 0 \]

\[ \beta_{i,k} = \alpha_{i,k-1} + \alpha_{k,j} + s_{i,j}^{\text{span}} \]

\[ \alpha_{i,j} = \max_{i<k<j}(\alpha_{i,k} + \alpha_{k,j}), \max_{i<k<j}(\beta_{i,j+k}) \]

- \( \alpha_{i,j} \): the accumulated score of span \((i,j)\) serving as a left or right child span
- \( \beta_{i,j+k} \): the accumulated score of the headed span \((i,j,k)\).
- S-CONC: the rule used to concatenate two consecutive child spans into a single child span.
- C-CONC: the rule used to concatenate left and right child span \((i,k-1)\) and \((k,j)\) along with the root word-span \((k-1,k)\) to form a headed span \((i,j,k)\).
- HEADLESS: the rule used to obtain a headless child span from a headed span.

Table 1. Labeled Attachment Score (LAS) on twelve languages in UD 2.2.

<table>
<thead>
<tr>
<th></th>
<th>PTB</th>
<th>CTB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UAS</td>
<td>LAS</td>
</tr>
<tr>
<td>MFVI2O</td>
<td>94.78</td>
<td>91.04</td>
</tr>
<tr>
<td>TreeCRF2O</td>
<td>94.72</td>
<td>91.65</td>
</tr>
<tr>
<td>HierPtr</td>
<td>97.22</td>
<td>95.71</td>
</tr>
<tr>
<td>Biaffine+MM†</td>
<td>97.24</td>
<td>95.73</td>
</tr>
</tbody>
</table>

Table 2. Results for different model on PTB and CTB.