Levenshtein-distance-based post-processing

shared task spotlight

Antal van den Bosch

ILK / CI and AI, Tilburg University

Ninth Conference on Computational Natural Language Learning (CoNLL-2005)
Ann Arbor, MI, June 29-30, 2005

Global models $\approx$ Post-processing

- Many systems:
  - local model vs. global model
- Goal of global model:
  - correcting mislabelings caused by “blind” decisions of a more local model
- Methods:
  - Probabilistic language models over argument labeling language
  - Distance-based error correction over argument labeling language

Hand-crafted post-processing

- “eraser” script by Erik Tjong Kim Sang (2004):
  - For each verb, if there is any double A0 - A5 in the sentence, delete the one furthest from the verb.
  - Can only improve precision, not recall.
  - Will typically lower recall, because it will delete the incorrect one of a double occasionally.

Data-driven post-processing

- Idea: argument labeling correction as spelling correction.
  - Classical solution: Levenshtein or string-edit distance (Levenshtein, 1965). Sum over:
    - Deletion: distance++
    - Insertion: distance++
    - Substitution: distance++
  - Closest found argument patterns contains corrections that need to be applied.
    - Predicted: emphasize $A_0V A_1 A_0$
    - Nearest in training data and PropBank at distance 1: emphasize $A_0V A_1$
    - Correction: delete final $A_0$ in predicted string

Data-driven post-processing (2)

- Implements deletions and replacements
- Does not perform insertions
  - Does not know where to “insert”
- Levenshtein-based correction
  - Should be able to improve precision, like “eraser”
  - Might improve recall, due to correct replacements
- Can be applied to all systems!
  - $>$500 deletions, $>$200 replacements for some systems on WSJ dev & test
  - $<$100 deletions, $<$75 replacements for systems that already have a global model

Levenshtein post-processing:

Example 1

- System marquez, Brown: 40 deletions, 30 replacements.
  - bend $A_1 V A_1$
  - bend $A_0 V A_1$
  - love $A_1 A_0 V A_1$
  - love $A_0 V A_1$
  - unite $A_1 V A_2$
  - unite $A_1 V$
Levenshtein postprocessing:

Example 2

- System punyakanok, Brown: 20 deletions, 17 replacements. E.g.,
  - search V A0
  - search V A1
  because in data:
  - search A0 V
  - search A0 V A1

Levenshtein postprocessing:

Example 3

- System pradhan, WSJ dev: 111 deletions, 72 replacements. E.g.,

```
[ [ That dividend ] A1 is almost double the 35% currently taken out of Farmers by B.A.T ] A1, the spokesman added .
```

Levenshtein post-processing: effect

<table>
<thead>
<tr>
<th>system</th>
<th>post</th>
<th>glob</th>
<th>WSJ test</th>
<th>WSJ test LPP</th>
<th>Brown test</th>
<th>Brown test LPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>punyakanok</td>
<td>✓</td>
<td>☐</td>
<td>78.54</td>
<td>78.18</td>
<td>87.75</td>
<td>87.56</td>
</tr>
<tr>
<td>highlight</td>
<td>☐</td>
<td>✓</td>
<td>78.24</td>
<td>78.15</td>
<td>87.73</td>
<td>87.56</td>
</tr>
<tr>
<td>NIST</td>
<td>☐</td>
<td>✓</td>
<td>77.34</td>
<td>77.18</td>
<td>87.07</td>
<td>86.93</td>
</tr>
<tr>
<td>pradhan</td>
<td>✓</td>
<td>☐</td>
<td>78.53</td>
<td>78.19</td>
<td>87.70</td>
<td>87.56</td>
</tr>
<tr>
<td>el</td>
<td>☐</td>
<td>☐</td>
<td>78.54</td>
<td>78.17</td>
<td>87.69</td>
<td>87.55</td>
</tr>
<tr>
<td>sutton</td>
<td>☐</td>
<td>✓</td>
<td>76.55</td>
<td>76.81</td>
<td>86.80</td>
<td>86.66</td>
</tr>
<tr>
<td>cohn</td>
<td>☐</td>
<td>✓</td>
<td>77.52</td>
<td>77.18</td>
<td>87.12</td>
<td>86.96</td>
</tr>
<tr>
<td>johansson</td>
<td>☐</td>
<td>✓</td>
<td>74.64</td>
<td>75.21</td>
<td>86.05</td>
<td>85.99</td>
</tr>
<tr>
<td>moschitti</td>
<td>☐</td>
<td>✓</td>
<td>74.35</td>
<td>75.86</td>
<td>86.79</td>
<td>86.63</td>
</tr>
<tr>
<td>che</td>
<td>☐</td>
<td>✓</td>
<td>73.31</td>
<td>73.90</td>
<td>85.49</td>
<td>85.34</td>
</tr>
<tr>
<td>surdeanu</td>
<td>✓</td>
<td>☐</td>
<td>72.48</td>
<td>72.48</td>
<td>85.26</td>
<td>85.21</td>
</tr>
<tr>
<td>pradhan</td>
<td>☐</td>
<td>✓</td>
<td>66.73</td>
<td>83.83</td>
<td>85.16</td>
<td>89.85</td>
</tr>
</tbody>
</table>

Blue means red shows bold number and cell color: effect > 0.2