**Introduction**

- We propose to model definition extraction problem using Conditional Random Fields.
- Previous works focus on glossary sentence identification (at the sentence level). We tackle the problem of obtaining the exact bounds of the term and its definition.

**Problem Formalization**

- Assign each input word $w_i$ an annotation $a_i = \text{(T)erm, (D)efinition, (O)ther}$.
- Recover definitional sentences that contains both a term and its definition.

**Best Solution Explored:**

2-Step Serial Word-Level CRF Model

- Base classifier exploits lexical, orthography, dictionary and corpus features.
- Augment with dependency parsing and shallow parsing features.
- Utilize results from term classification and incorporate into definition classification.

**Macroscopic**

- Definition location distribution over 10 quantiles.
- Definition sentences tend to occur towards the beginning of documents.

- 45% of the detected terms are multi-word terms.
- Definition length is more varied. 75% are between 5-16 words.

**Main Experiment Results**

<table>
<thead>
<tr>
<th>System / Feature Class</th>
<th>Term</th>
<th>Definition</th>
<th>Overall*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
<td>F₁</td>
</tr>
<tr>
<td>1: Baseline (Lexical + Orthography + Dictionary + Corpus)</td>
<td>0.50</td>
<td>0.35</td>
<td>0.41</td>
</tr>
<tr>
<td>2: (1) + shallow parsing</td>
<td>0.50</td>
<td>0.40</td>
<td>0.45</td>
</tr>
<tr>
<td>3: (2) + dependency parsing</td>
<td>0.50</td>
<td>0.41</td>
<td>0.45</td>
</tr>
<tr>
<td>4: (3) + 2-stage [DefMiner]</td>
<td>0.50</td>
<td>0.41</td>
<td>0.45</td>
</tr>
<tr>
<td>5: (3) + Reverse 2-stage</td>
<td>0.50</td>
<td>0.40</td>
<td>0.44</td>
</tr>
<tr>
<td>6: (3) + Term Oracle</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

*The result is reported for experiments on our manually annotated W00 corpus. Evaluation on token (word) level.

**Conclusion**

- We introduced DefMiner, a sequence labeling system that identifies scientific terms and their definitions.
- Improved system accuracy by exploiting a small set of shallow and dependency parsing features.
- Serial classification (term $\Rightarrow$ definition) boosts the performance significantly.
- Applied to a large corpus of scientific publications, highlighting trends and applications.

**Term-Level Microscopic**

- Frequentely defined terms in the ACL Corpus.
- Extracted terms often be fit into 3 categories: resources, methodologies, and evaluation metrics.

**Applying DefMiner to the ACL Anthology Reference Corpus**

- Across publication categories.
- Term length distribution.
- Density of definitions increases in workshop / conference papers over time.
- Around 40% of the definitions introduced in 2004 have been seen before.

**WCL Dataset Results**

Comparing to
(Navigli and Velardi, 2010)

<table>
<thead>
<tr>
<th>System</th>
<th>Term (Word Level)</th>
<th>Definition (Word Level)</th>
<th>Sentence Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P/R/F₁</td>
<td>P/R/F₁</td>
<td>P/R/F₁</td>
</tr>
<tr>
<td>DefMiner</td>
<td>.82/.78/.80</td>
<td>.82/.78/.81</td>
<td>.92/.79/.85</td>
</tr>
<tr>
<td>N&amp;V’10</td>
<td>-/-/-</td>
<td>-/-/-</td>
<td>-.96/.61/.77</td>
</tr>
</tbody>
</table>