Towards Zero-resource Cross-lingual Entity Linking
Shuyan Zhou, Shruti Rijhwani, Graham Neubig
School of Computer Science, Carnegie Mellon University

I. Background

Cross-lingual Entity Linking (XEL)

Associate the detected name entities in a source language to an English KB (e.g., English Wikipedia)

Challenges

Resources are scarce for low-resource languages

- English Wikipedia
- Source language Wikipedia
- Bilingual entity maps
- Multilingual embeddings

Our Contributions

Towards true zero-resource XEL

- Extensive evaluation on the effect of resource restrictions on existing XEL systems in true low-resource settings, instead of simulated ones.
- Three improvements towards ameliorating the large degradation in low-resource settings

II. Baseline XEL System: an approximation to existing works

Candidate Generation Models

- WikiMention: heavy resource reliance

Eertiraa (om, mention) → Eertiraan (om, entity)

Featuarization (Base)

- $f_1^2$: Entity co-occurrence frequency
- $f_2^1$: Entity prior given the mention

Non-iterative Linear Inference Model (Greedy)

$$s(e_i|D) = s_y(e_i|D) + s_t(e_i|D)$$

- Linear feature combination
- Non-iterative score combination

III. Baseline Results in True Low-resource Settings

Performance (%)

<table>
<thead>
<tr>
<th>Languages (Wikipedia Size)</th>
<th>gold candidate recall</th>
<th>Base+Greedy 9th accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- The Wikipedia resources strongly influence the gold candidate recall
- Simple context-insensitive method approaches the upper bound (i.e., gold candidate recall) in low-resource languages
- More sophisticated model brings little or no improvement to low-resource languages

IV. Our Improved XEL System

👍: Calibrated Candidate List Combination

Improving gold candidate recall

- Add light-resource candidate generation model Pivoting

Design calibrated entity scores

$$p_{l,j} = \frac{\exp(\gamma s_{l,j})}{\sum_{k} \exp(\gamma s_{k,j})}$$

👍: More Intelligent Feature Set (Feat)

Exploiting available resources

- $f_2^0$: Smoothed entity co-occurrence frequency
- $f_3^0$: Entity embedding similarity
- ...

👍: Belief Update Neural Network (Burn)

Combining features efficiently

- Non-linear feature combination using multi-layer perceptron with activation functions
- Iterative score combination

V. Final Results in True Low-resource Settings

- The calibrated combination consistently improves the gold candidate recall compared to other two models
- Burn model outperforms Greedy model in all four languages
- The advantage of Feat varies among languages
- Our improved XEL system results in gains of 6-23% among all four languages compared to the baseline system