Multi-target Machine Translation with Multi-synchronous Context Free Grammars

Graham Neubig, Philip Arthur, Kevin Duh
Nara Institute of Science and Technology (NAIST), Japan

Motivation

- When translating into language T1, equivalent translations into a second language T2 can help

Proposed Framework

- Build on the well-known synchronous context-free grammars (SCFG)

Learning MSCFGs

- Learn from tri-lingual parallel data

1) Alignment

<table>
<thead>
<tr>
<th>Target 1 (T1)</th>
<th>Target 2 (T2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>京都条约</td>
<td>Kyoto Treaty</td>
</tr>
<tr>
<td>京都制裁</td>
<td>Kyoto Sanction</td>
</tr>
<tr>
<td>京都议定书</td>
<td>Kyoto Protocol</td>
</tr>
</tbody>
</table>

2) Phrase Extraction

OUN → 联合国
a → 批准
ratifié → 批准了
a ratified → ratified

Decoding w/ MSCFGs

- Two ways to handle increased search space due to two language models

Experiments

MultiUN Corpus:
- Parallel, T1 LM data: 100,000 Sentences
- T2 LM data: 4,000,000 Sentences

<table>
<thead>
<tr>
<th>S:</th>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>en</td>
<td>ar, es, fr, ru, zh (all combinations)</td>
<td></td>
</tr>
</tbody>
</table>

Baseline: SCFG-based 1-target Hiero Grammar
Proposed: MSCFG-based 2-target Hiero Grammar

Result 1: Does second target Help?

- e.g. T2=es (best results)

Answer:
- Yes! In most cases accuracy improves
- Particularly effective in similar languages

Result 2: Influence of T1 LM strength?

- e.g. T1=fr, T2=es

Answer:
- As expected, works best when T1 LM is less strong:
  Covers weakness of T1

Result 3: Influence of T2 language?

- e.g. T2=zh (worst results)

Answer:
- When T2 is very different results less good, due to rule extraction constraints

Results on decoding, model size, etc. in paper

Try out the code/scripts!
http://phontron.com/project/naacl2015