Frank Palma Gomez¹, Subhadarshi Panda², Michael Flor³, Alla Rozovskaya¹
¹ CS Department, Queens College, ² CS Department, CUNY Graduate Center, ³ ETS

Task and Motivation

Given a sentence and target word, generate challenging distractors.
A challenging distractor is a word that is semantically similar to the target but does not make the sentence valid.
Distractors are created manually by language experts, which is time consuming and expensive.

Approach

We use round-trip neural machine translation (NMT) to generate a diverse set of distractors.
Intuition: Lexical confusions, common among English learners, will manifest in round-trip NMT as back-translated words that are semantically close to the target.
We use 5 language pairs to generate translations: Czech, German, French, Italian and Russian.

Contributions

We propose a novel round-trip NMT method to generate challenging distractors.
Using diverse language pairs yield diverse/challenging distractors.
Comparison with state-of-the-art methods (Word2Vec, BERT) shows our approach is superior.

Evaluation

Left: Average number of automatic distractors generated per sentence.
Right: Union of all languages retrieves 31.9% of all gold distractors.
Union of all languages produces more distractors than individual language systems, thereby demonstrating that each language system contributes unique distractors.

State-of-the-Art Comparison

Round-trip NMT generates more difficult and semantically close candidate distractors compared to state-of-the-art NLP approaches.

Sentence: Police told everyone to stay at home until the ____ of thieves had been captured.
Target Word: band
Top-5 Word2Vec Distractors: keyboardist, vocalist, drummer, quintet, guitarist
Top-5 BERT Distractors: recommendations, assertions, observations, results, conclusions
Round-trip NMT Distractors: crew, group, orchestra, gang, squad

Manual Evaluation: Our approach produces significantly more valid distractors compared to state-of-the-art methods, Word2vec and BERT.

<table>
<thead>
<tr>
<th>Method</th>
<th>% of valid distractors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rater 1</td>
</tr>
<tr>
<td>MT-no-ranking</td>
<td>67.9</td>
</tr>
<tr>
<td>Word2vec</td>
<td>57.2</td>
</tr>
<tr>
<td>BERT</td>
<td>22.7</td>
</tr>
<tr>
<td>MT with re-ranking (Word2vec)</td>
<td>50.4</td>
</tr>
<tr>
<td>MT with re-ranking (BERT)</td>
<td>27.7</td>
</tr>
</tbody>
</table>