Sequence to Logic with History Gates

Javid Dadashkarimi, Alex Fabbri and Dragomir Radev
Contact Information
Email javid.dadashkarimi@yale.edu
Phone +1 (475) 731 4330

Problem

1. Generating logic form of human language is a fresh way with neural architecture.
2. The logic form of the sequence usually preserves information in the human side with similar tokens.
3. recently a copying mechanism is proposed where increases the probability of the source tokens through decoding [1].

Introduction

1. Recently [1] proposed an augmented pointer network for sequence to logic.
2. This is a very straightforward mechanism to pay attention to the source sequence.
   - x: what rivers flow through colorado?
   - y: answer (A, river(A), traverse {A, B}, const {B, stateid (colorado)})
3. In this paper we step forward and introduce history gates to take whole the source vocabulary into account for decoding (e.g., 'flow through' and 'traverse').

Results

<table>
<thead>
<tr>
<th>ID</th>
<th>calendar</th>
<th>housing</th>
<th>publications</th>
<th>recipes</th>
</tr>
</thead>
<tbody>
<tr>
<td>attn</td>
<td>0.565</td>
<td>0.883</td>
<td>0.476</td>
<td>0.837</td>
</tr>
<tr>
<td>j1(sj, cj)</td>
<td>0.589</td>
<td>0.885</td>
<td>0.508</td>
<td>0.832</td>
</tr>
<tr>
<td>j2(sj, cj, σ)</td>
<td>0.613</td>
<td>0.890</td>
<td>0.439</td>
<td>0.821</td>
</tr>
<tr>
<td>j3(sj, cj, σ)</td>
<td>0.548</td>
<td>0.883</td>
<td>0.493</td>
<td>0.821</td>
</tr>
<tr>
<td>j4(sj, cj, tanh)</td>
<td>0.554</td>
<td>0.873</td>
<td>0.471</td>
<td>0.826</td>
</tr>
<tr>
<td>j5(sj, cj, σ)</td>
<td>0.554</td>
<td>0.873</td>
<td>0.407</td>
<td>0.806</td>
</tr>
<tr>
<td>j6(sj, cj, σ)</td>
<td>0.524</td>
<td>0.867</td>
<td>0.407</td>
<td>0.806</td>
</tr>
</tbody>
</table>

Discussion and Conclusion

1. The performance of the method is not independent of the domain.
2. History gates are general form of the traditional copying mechanism.
3. The proposed method generally achieves consistent improvements in token level accuracy.
4. For example in the publication data: authors of articles published in 2004 both the data and 2004 were selected from history not the source sequence. 2004 is already in the source sequence. The reason is that dates and numbers are very common in this domain.

In this paper we introduced the history gates for sequence to sequence encoder/decoder system. The previous copying mechanism could boost performance by giving chance to the source tokens to appear in target. History gates gives this possibility for all the source vocabulary that might have semantic relationship with current context.

References


Future Research

1. transfer learning: for using hidden information from one domain in another.
2. machine translation: for out out vocabulary words (e.g., generating Barack Obama in French by an English sentence).
3. recommender system: recomending movies for a user based on her/his most recent interests and also a long history.