Summary

- Tracking entity states while summarizing table information improves adequacy while preserving fluency.
- Incorporating the writer information is also helpful for selecting appropriate data and writing better summary.

Data-to-Text Generation

Task: Given a set of records, the system should produce a fluent and an adequate summary.

Problem:
- Template system could generate the adequate text but not fluent.
- NN based system could generate the fluent text but not adequate.

<table>
<thead>
<tr>
<th>Adequacy</th>
<th>Fluency</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

For document-level generation, the system has an influence of the writer's bias.

Entity Tracking

Idea: Equipping an NN-based generation system with the entity-tracking module enables capturing of the saliency and the coherence while preserving the fluency.

Method: Inspired by EntityNLM [1], our model dynamically updates each entity state and the tracking module.

Experimental results

Baseline:
- Template: game results, 6 prominent players, next game information.
- Wiseman et al. [2]: Encoder-Decoder w/ Attention
- Puduppully et al. [3]: Two stage model

1. Predicting the sequence of data records
2. Generating a summary conditioned on the predicted sequence

Evaluation Metrics:
- RG: the ratio of the correct relations of all of the extracted relations.
- CS: the F1 score of the relations extracted from the generated summary against those from the reference summary.
- CO: the normalized Damerau-Levenshtein Distance (DLD) between the sequences of relations extracted from the generated and reference summary.
- BLEU

Writer Bias Evaluation:
- For Puduppully et al., concatenate the writer embedding w for decoder's input.
- For our model, concatenate w for hidden state of LM and Tracking module states for composing h_w:

Annotation: To train the model in a fully supervised way, we obtain the following annotations using the information extraction system developed by Wiseman et al. [2].

References

1. Ji et al. - Dynamic Entity Representations in Neural Language Models, In EMNLP 2017
2. Wiseman et al. - Challenges in Data-to-Document Generation, In EMNLP 2017
3. Puduppully et al. - Data-to-Text Generation with Content Selection and Planning, In AAAI 2019