Sentence Generation using Fan Theories

TANVI SAHAY
University of Massachusetts Amherst

Goals

• Application of natural language processing techniques for sentence generation
• Performance assessment based on Domain Knowledge

Database

Fan theories!
Based on the previous events, what do fans of the show predict will happen next

- Theories are non-repetitive
- Data base has high noise
- Need of Domain Knowledge

Cleganebowl!
R + L = J

Tools Analyzed

Stanford Log-linear Part-Of-Speech Tagger
Stanford Named Entity Recognizer (NER)
Stanford Deterministic Coreference Resolution System
Stanford Open Information Exracttion (OpenIE)

Models

Tokenized text + bigram LM


OpenIE relation tuples + bigram LM

| 1.0 | Jon Snow | be | second |
| 0.97 | Jon Snow | coming of | Azor Ahai |

OpenIE Named Entity Information + bigram LM

Arya’s friend is Gendry
His was temporarily freed

OpenIE relation tuples + HMM

Character-level LSTM

Results

10 sentences generated for most famous characters and rated based on both grammatical correctness and domain relevance

Top 5 most talked about characters (NER):

Jon Arya Cersei Dany Ned

Example Sentences:

- Arya get married
- Arya’s joins up with band
- Arya arrived back trying
- Arya winterfell
- (arya stark’s kill list. in season 3) Arya (the the

Average Scoring of models

Average score

<table>
<thead>
<tr>
<th>Model Number</th>
<th>Grammar Rating</th>
<th>Domain Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.00</td>
<td>6.02</td>
</tr>
<tr>
<td>2</td>
<td>6.92</td>
<td>6.58</td>
</tr>
<tr>
<td>3</td>
<td>7.20</td>
<td>6.00</td>
</tr>
<tr>
<td>4</td>
<td>6.46</td>
<td>6.46</td>
</tr>
</tbody>
</table>

Sentence scoring for model 1

<table>
<thead>
<tr>
<th>Score</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Conclusions

• Model 3 performs best, owing to the training data containing minimum noise and only relevant sentences.
• The Stanford POS Tagger, NER and Coreference resolution systems perform well despite the noisy data.
• In general, most models are able to map grammatical correctness better than domain relevance.
• Noise removal improves the quality of sentences being generated but to a limited extent only.
• Despite being powerful generative models character LSTMs fail on small and noisy databases.

Scope of Improvement

Further noise removal using the resolved co-references

Word-level LSTM