**DiscoTK: Using Discourse Structure for Machine Translation Evaluation**

Shafiq Joty, Francisco Guzmán, Lluís Màrquez and Preslav Nakov

Qatar Computing Research Institute

**Discourse for MT Evaluation**

- Discourse structure helps MT evaluation [Guzmán et al., 2014]
- We present two metrics that consider discourse information
  - DiscoTK\_light only uses discourse
  - DiscoTK\_party also uses metrics from ASIYa
- DiscoTK\_party is the best performing metric at WMT14

**Method**

Compute discourse similarity between Hyp and Ref
- RST-parse Hyp and Ref (Joty et al., 2012)
- RST trees are transformed to five different representations
- We use syntactic tree kernel [Collins & Duffy, 2002] to measure the similarity between two discourse trees
  - Use this similarity as a segment-level score
  - For system-level, average segment level scores

Combine discourse similarity with existing metrics (ASIYa)
- Uniform linear interpolation
- Tuned (MaxEnt pairwise learning)

**Results**

<table>
<thead>
<tr>
<th>Metric</th>
<th>Tuning</th>
<th>Segment Level</th>
<th>System Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>WMT12</td>
<td>WMT13</td>
</tr>
<tr>
<td>DiscoTK_light</td>
<td>na</td>
<td>0.254</td>
<td>0.264</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.884</td>
<td>0.922</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.860</td>
<td>0.872</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.865</td>
<td>0.871</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.865</td>
<td>0.871</td>
</tr>
<tr>
<td>ASIYa</td>
<td></td>
<td>0.257</td>
<td>0.231</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.941</td>
<td>0.928</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.936</td>
<td>0.943</td>
</tr>
</tbody>
</table>

**Summary**

- DiscoTK\_light competitive at system-level
- Tuned DiscoTK\_party improves over ASIYa both at segment- and system-level
- Tuning helps consistently
- We improve over the best WMT12, WMT13 results

**Future Work**

- Learn with preference kernels from a syntactic-semantic-discourse tree representation
- Go beyond the sentence-level