Automatic Detection and Classification of Social Events

Apoorv Agarwal
Owen Rambow
Columbia University

October 10, 2010
Overall Goal

Extract a social network from text where nodes are people and links are *social events*
Overall Goal

Extract a social network from text where nodes are people and links are *social events*
Social Event:(Agarwal et. al. 2010) An event between two people or group of people where at least one party is aware of the other party and aware of the event. Types:

- Interaction event (INR): both parties mutually aware
- Cognition event (COG): only one party aware of the other

Figure: Interaction (INR) and Cognition (COG) social events respectively
We annotated social events for part of Automatic Content Extraction (ACE) data

ACE already has annotations for entities, relations and events but:

- Our definition of social event is conceptually different from ACE since we require reasoning about cognitive states of people
- [Toujan Faisal], 54, {said} [she] was {informed} of the refusal by an [Interior Ministry committee] overseeing election preparations
- Sequence and Tree kernels on dependency and phrase structure trees (Alessandro Moschitti and his research group)
- SqGRW: *Toujan_Faisal nsubj T1-Individual said ccomp informed prep by T2-Group pobj committee*
Experiments and Results

Experimental Set-up:
- 138 ACE documents: 172 INR, 174 COG, 1291 No relation classes
- SVM with kernels: 5-fold cross-validation

<table>
<thead>
<tr>
<th>Kernel</th>
<th>Event Detection (% F1)</th>
<th>Classification % Acc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>Under</td>
<td>Over</td>
</tr>
<tr>
<td>PET_GR</td>
<td>38.9</td>
<td>48.5</td>
</tr>
<tr>
<td>PET_GR_SqGRW</td>
<td>38.0</td>
<td>48.5</td>
</tr>
<tr>
<td>GR_SqGRW</td>
<td>36.2</td>
<td>47.3</td>
</tr>
<tr>
<td>GRW_SqGRW</td>
<td>25.0</td>
<td>47.1</td>
</tr>
<tr>
<td>GR_GRW_SqGRW</td>
<td>32.6</td>
<td>46.8</td>
</tr>
</tbody>
</table>

- Event detection is harder than event classification
- Combination of structures derived from phrase structure trees and dependency trees work best
- SqGRW plays a role in both best performing system