Annotating Spatial Containment Relations Between Events

Kirk Roberts, Travis Goodwin, and Sanda Harabagiu
• Problem
• Previous Work
• Schema
• Corpus
• Future Work
Problem

- Natural language documents contain a large number of event mentions.
- Many resources tie events to their location and participants—e.g., FrameNet, PropBank, NomBank, ISO-Space.
- However, very few events have explicit spatial information—Only TODO% of events have a location in our data.
Malaysia's Anwar loses appeal in sodomy case ATTENTION – ADDS Anwar quote, background

KUALA LUMPUR, April 18 (AFP)

Malaysia's Appeal Court Friday refused to overturn the conviction and nine-year jail sentence imposed on ex-deputy prime minister Anwar Ibrahim for sodomy.

Anwar, 56, who this week completed four years in prison on a corruption charge, now faces an earliest possible release date of April 14, 2009 if he is given one third remission of his sentence for good behaviour.

The former heir-apparent to Prime Minister Mahathir Mohamad, who says he was framed for political reasons, told reporters after his appeal was rejected: "You must remember its pre-selected judges, most junior of the court of appeal, and its all scripted.

"What do you expect?"

Mahathir's sacking of the charismatic and popular Anwar in September 1998 rocked Malaysian politics, with thousands of people taking to the streets to demonstrate their support for him.

Within weeks he was arrested and charged with sodomising an official driver several years previously and with abusing his powers to cover up the offence.

The conduct of the trials was widely criticised internationally and the United States still lists Anwar as a political prisoner.

Anwar was told Monday that he had been granted a standard one-third remission of a six-year corruption sentence for good behaviour, and immediately began to serve the nine-year sentence for sodomy.
Problem

• **Heuristic:** choose closest *location* (token-wise)
  – TODO: cite
• Completely unprincipled
• Ignores the complex relationship that *locations* have with *events*
  – Events may occur *within* a location, *include* a location, *occur near* a location, etc.
Rafiq Hariri submitted his resignation during a 10-minute meeting with the head of state at the Baabda presidential palace.
As Egyptian columns retreated, Israel’s aircraft attacked them, using napalm bombs. The attacks destroyed hundreds of vehicles and caused heavy casualties in Sinai.
• Problem
• **Previous Work**
• Schema
• Corpus
• Future Work
SpatialML

- Identifies
  - locations (PLACE)
  - trajectories (PATH)
  - containment relations (LINK)
- Disambiguates toponyms (e.g., “Paris”)
ISO-Space

- Identifies
  - Events (EVENT and MOTION)
  - Location containment relations (EVENT_LOCATION and EVENT_PATH)
  - Relations between locations (QSLINK and S_FUNCTION)

John drove to Boston.
Missing Piece

• Detailed representations of event locations, but very few events have explicit locations
• Spatial information is often stuck in silos
• Solution: event-event relations that describe spatial relationships in much the same way as
  – SpatialML describes location-location relations
  – ISO-Space describes event-location relations
  – E.g., a spatial analog of TimeML that, but designed with implicit relations in mind
• Problem
• Previous Work
• **Schema**
• Corpus
• Future Work
Rafiq Hariri submitted his resignation during a 10-minute meeting with the head of state at the Baabda presidential palace.
Schema

- Containment relations derived from RCC-8 while accounting for common ambiguities in natural language
Schema

- **Event-Event** (containment) – 5 sub-types
- **Event-Participant** – 3 sub-types
- **Event-Location** – 3 sub-types
As Egyptian columns retreated, Israel’s aircraft attacked them, using napalm bombs. The attacks destroyed hundreds of vehicles and caused heavy casualties in Sinai.
Example

The soldier sent his ballot for the Maine election.
Example
Example
Example
Example
• Problem
• Previous Work
• Schema
• **Corpus**
• Future Work
Corpus

- Subset of **SpatialML** corpus
  - 160 newswire documents
- Gold-standard **locations** pre-annotated
- Automatic tagging for **events** and candidate **participants** (person/org entities)
- Difficult annotation task, fairly low agreement
  - Several rounds of annotation/inspection
Corpus

- Containment relation distribution:
Simple Baseline

• To approximate the difficulty of the task, we evaluated with an **SVM** classifier and a simple feature set:
  
  – **EW**: Both event words (**submitted**:::meeting)
  
  – **EL**: Both event lemmas (**submit**:::meeting)
  
  – **WB**: Words between events (**his, resignation, during, a, 10-minute**)
  
  – **HN**: Hypernyms of events (**refer**:::gathering, **send**:::gathering...)
## Simple Baseline

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>P</th>
<th>R</th>
<th>F₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>EW</td>
<td>13.7</td>
<td>31.0</td>
<td>19.0</td>
</tr>
<tr>
<td>EL</td>
<td>14.7</td>
<td>33.4</td>
<td>20.4</td>
</tr>
<tr>
<td>WB</td>
<td>14.3</td>
<td>57.9</td>
<td>22.9</td>
</tr>
<tr>
<td>HN</td>
<td>17.9</td>
<td>32.6</td>
<td>23.1</td>
</tr>
<tr>
<td>EW + HN</td>
<td>19.7</td>
<td>31.4</td>
<td>24.2</td>
</tr>
<tr>
<td>EL + HN</td>
<td>19.4</td>
<td>31.5</td>
<td>24.0</td>
</tr>
<tr>
<td>WB + HN</td>
<td>25.8</td>
<td>39.9</td>
<td>31.3</td>
</tr>
<tr>
<td>EW + WB + HN</td>
<td>27.8</td>
<td>37.5</td>
<td>31.9</td>
</tr>
<tr>
<td>EL + WB + HN</td>
<td>27.6</td>
<td>38.0</td>
<td>32.0</td>
</tr>
<tr>
<td>EW + EL + WB + HN</td>
<td>29.1</td>
<td>35.5</td>
<td>32.0</td>
</tr>
</tbody>
</table>
## Simple Baseline

<table>
<thead>
<tr>
<th>Relation Type (5-way)</th>
<th>Feature Set</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EW</td>
<td>58.3</td>
</tr>
<tr>
<td></td>
<td>EL</td>
<td>57.7</td>
</tr>
<tr>
<td></td>
<td>WB</td>
<td>52.1</td>
</tr>
<tr>
<td></td>
<td>HN</td>
<td>54.9</td>
</tr>
<tr>
<td></td>
<td>EL + EW</td>
<td>57.9</td>
</tr>
<tr>
<td></td>
<td>WB + EW</td>
<td>53.1</td>
</tr>
<tr>
<td></td>
<td>HN + EW</td>
<td>54.8</td>
</tr>
</tbody>
</table>
• Problem
• Previous Work
• Schema
• Corpus
• Future Work
Future Work

• For annotation:
  – Finish annotation effort (abt. 2/3 complete)
  – Integration with other spatial information standards (e.g., ISO-Space)
Future Work

• For classification, it is clear that basic NLP features are insufficient

• Further integration of event structures to represent complex event semantics
  – Events narratives (Chambers & Jurafsky, 2008), (Bejan, 2008)
  – Event coreference
  – Discourse relations
  – Relative spatial bounds of events (e.g., a football game is larger than a pass or kick)