Using Explicit Semantic Models to Track Situations across News Articles

Earl Wagner, Jiahui Liu, Larry Birnbaum, Ken Forbus, James Baker
Northwestern University
Intelligent Information Laboratory
& Qualitative Reasoning Group

Using Online News to Track Situations

• Read an article about an event
  – What happened before?
  – What might happen next?
  – Who are the people involved in this?
• Interested in many situations evolving over time
  – Want alerts when something new happens
• Explore use of structured representation
Technical Challenges

• Scale: Perform better with more data
• Aggregate information from multiple accounts
  – Complementary, redundant, or contradictory
  – Reduce noise due to inaccurate extraction
• Integrate IR, IE, & semantics
  – Use semantic models as structured contexts to drive IR & IE
  – Contexts provide structure for aggregation

Brussell Overview

• Generates structured summary of situation
  – Extracts info about events and entities
  – Links to textual sources of extracted information
  – User can track over time with STT/Brussell + Cyc
• Represents situations as scripts
  – Events are scenes, entities are roles / actors
• Uses scripts to retrieve/extract info and organize for presentation to user
• Currently supports terrorism script types:
  – kidnappings, sieges
Brussell Demo I

Brussell Demo II
Brussell Demo III

Brussell Demo IV
Implementation: Database

• Retrieves news articles from database
• ~400,000 articles in a MySQL database
  – Acquired from April 2004 to present
  – About 500/day retrieved via RSS feeds
  – Sources: NYTimes, Washington Post, Yahoo! News (AP, Reuters, UPI), etc.
  – Standard MySQL text indexing and querying

Implementation: Retrieval

• Perform query combining script type keywords and user supplied parameters, e.g.,
  – “+<name> +(kidnap* abduct* capt* seiz*)”
  – Here, kidnap victim’s name is “seed” with specificity
• For each article that comes back
  – Split article into text fragments, for now sentences
• For each fragment with seed
  – Look for scene terms in fragment
Implementation: Extraction

- Fastus-style Cascaded Finite State Transducer
  - But top-down rather than bottom-up
  - Don’t extract something unless you know what you’re going to do with it
- Each level above triggers patterns below
  - Abduction scene: “X ‘kidnapped’ Y.”
  - Search for descriptions of kidnapper in region X and kidnap victim in region Y
- Generate scene instance and incorporate into current script

Implementation: Aggregation and Voting

- Aggregate references within script’s hierarchy and count as votes, e.g.
  - # of times scene referenced -> votes for its occurrence
  - # of times person name appears with nationality X vs. nationality Y
- Semantic model structures and constrains aggregation
  - Scene choice
What’s New?

• Like Message Understanding Conference (MUC) systems, read article, fill out templates with event info
  – But, templates are part of larger structures
    • Share parameters
    • Represent relations
• Like old-style NL systems
  – But deeper understanding than from one article
  – How to aggregate info from the 100 or 10000 results from Google News?

Evaluation

• AP list of 36 kidnappings of foreigners in Iraq in October 2004
• Hand assembled reference data of 34 kidnapping script instances
• For each script instance, compute precision and recall based on
  – Scene choice
  – Scene data
    • Date and time, and location
  – Role data
    • Biographical
• At the time of evaluation, the corpus contained approximately 250,000 articles retrieved from April 2004 to February 2006
Brussell’s Performance

- Precision:
  - Mean 73%
  - Standard Deviation 19%
- Recall:
  - Mean 59%
  - Standard Deviation 20%
- Performance comparable to MUC-7 systems
  - Simple extraction techniques + voting + hierarchical models

Relationship between Performance and Corpus Size

- Evaluation was done on subsets of the corpus of different sizes
- Expectation: With larger corpus, both precision and recall on slots would increase using voting
Relationship between Performance and Corpus Size I

What’s Wrong?

- This analysis aggregates over all extracted slots
  - Previously filled slots and newly filled slots are lumped together
- Maybe improvement in previously filled slots is obscured by errors in newly filled slots
Improving Syntactic Competence

- False positive due to complex grammatical structure
- Syntactic preprocessing: simplifying the sentence syntactically

- “A South Korean hostage threatened with execution in Iraq has been killed, officials in Seoul have confirmed,”

  ==>

- “A South Korean hostage has been killed.”

  (S (NP (NP A South Korean hostage)) (VP threatened (PP with (NP execution)) (PP in (NP Iraq)))) (VP has VP been (VP killed))) , (NP (NP officials) (PP in (NP Seoul))) (VP have (VP confirmed .)))
Overall Precision

Distribution of Overall Precision of 34 Test Cases

<table>
<thead>
<tr>
<th></th>
<th>Average</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without SA</td>
<td>73%</td>
<td>19%</td>
</tr>
<tr>
<td>With SA</td>
<td>81%</td>
<td>23%</td>
</tr>
</tbody>
</table>

Overall Recall

Distribution of Overall Recall of 34 Test Cases

<table>
<thead>
<tr>
<th></th>
<th>Average</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without SA</td>
<td>59%</td>
<td>20%</td>
</tr>
<tr>
<td>With SA</td>
<td>44%</td>
<td>22%</td>
</tr>
</tbody>
</table>
Digging Deeper: Performance on Specific Cases

- Good
  - Big: Margaret Hassan
  - Small: Micah Garen
- Bad: Eugene Armstrong
  - Problems
    - Info for group members conflated
    - Name Eugene is also treated as location
  - These highlight need for semantic constraints applied at multiple levels

Future

- GATE for IE, Lucene for IR
- Smarter voting - not just at leaves
  - Kidnapper: “John or Abu”, “Smith or Zarqawi”
  - Scene occurs in a single location at a single date and time - otherwise create multiple scenes
- Rely on user to merge/split problem cases
- Deeper integration of IR and IE
  - Crystallize around certainty
  - Perform queries to fill in unfilled/uncertain slots
Summary

• Semantic models provide
  – Driver for IR, IE
  – Structure for aggregating then voting
  – Organization for presentation to user