UAIC Participation at RTE5

Adrian Iftene
Mihai Alex Moruz
{adiftene,mmoruz}@info.uaic.ro
Overview

- 3-way RTE5 System
  - Newly added components
  - Positive and Negative rules
  - Results

- Pilot task
  - Application of QA techniques
  - Results

- Conclusions

- Further work
RTE Competition

The **two-way RTE task** (2005-2009) is to decide whether:
- **T entails H - ENTAILMENT**
- **T does not entail H - NO ENTAILMENT**

- The **three-way RTE task** (2007-2009) is to decide whether:
  - **T entails H - ENTAILMENT**
  - **T contradicts H - CONTRADICTION**
  - **The truth of H cannot be determined on the basis of T - UNKNOWN**
System Presentation

600 pairs (T, H)

Preparation

Minipar module

Dependency trees for (T, H) pairs

TreeTagger

GATE

Numbers & Dates

LingPipe module

Named entities for (T, H) pairs

DIRT

NE rule

Fitness rule

Negation rule

Contradiction rule

Wikipedia

Background knowledge

Wordnet

VerbOcean

Google API

•Pre-processing

•Resources

•Main module

Final result

Iftene, Moruz – TAC, 2009
Preparation and New Modules

- We replace “hasn’t” with “has not”, “isn’t” with “is not”, “couldn’t” with “could not” and pad with some punctuation.

- In the case of Named Entities of type JOB and LANGUAGE, we additionally used GATE, which contains finer-grained classes of entities.

- In order to cope with misspelled words (particularly Named Entities) we used the Google API.
Example

Text

WordNet

verb\textsubscript{T} (V)

noun (N)

adv (A)

loc

mod

subj

verb\textsubscript{H} (V)

noun (N)

NE (N)

Hypothesis

DIRT, VerbOcean

DirT: solve=resolve 0.31453
DirT: convict=arrest 0.28895
DirT: convict=acquit 0.302455, OPPOSITE
VerbOcean: increase<>decrease,
VerbOcean: leave<>stay

WordNet: trouble=problem
WordNet: talk=discussion

Acronyms, BK

Acronym: EU=European Union
BK: 16 [is] sixteen

Iftene, Moruz – TAC, 2009
Rules

- For every type of possible answer we will present the rules that lead to it.

- Possible cases are:
  - Entailment cases
  - No entailment cases
    - Contradiction cases
    - Unknown cases
Entailment Cases

- Every type of mapping: direct (lexical) or indirect (using knowledge bases)

- Verb similarity is computed using DIRT
  - passed away ≈ has died

- For named entity we use an acronym database and background knowledge
  - United States ≈ US
  - Basel in Switzerland ≈ European City
Entailment Cases (cont.)

- For nouns and adjectives we use WordNet and some of the relations from eXtended WordNet to look up synonyms.
- For every transformation with DIRT or WordNet, we will consider local fitness to be the similarity value indicated by these resources.
- Stop words from the hypothesis artificially increase the value of global fitness and are ignored.

Iftene, Moruz – TAC, 2009
Entailment Cases for Numbers

- When numbers from T or H are separated by “and” or “,”, add them
  - T: 10 people were killed and more than 30 died \( \approx \) H: killing more than 40 people

- Positive rules for Numbers (context rules): quantification words: at least, more than, less than, over, under, etc.
  - T: at least 80 percent \( \approx \) H: more than 70 percent
Contradiction Cases

For every verb subtree, we check for words such as “not”, “never”, “may”, “might”, “cannot”, etc. and modify the negation value of the verb

- T: New Line Cinema has announced that movie director Peter Jackson will never be allowed to work ...
- H: New Line wants to work with Peter Jackson.

Verbs in the long infinitive receive special treatment
Contradiction Cases (cont.)

- When before the long infinitive we have “refuse”, “deny”, “ignore”, “plan”, “intend”, “proposal”, “able”, etc.
- We consider a combination of synonyms from WordNet and antonyms from WordNet or opposites from VerbOcean.
Unknown Cases

- When verbs are modified by words such as “may”, “can”, “should”, “could”, “must”, “might”, “infrequent”, “rather”, “probably”, etc.
  - T: ...*could eventually be taken over* ... and H: “... *is taken over*...”

- Related to verbs in the infinitive, we will consider as Unknown those cases which are not included in the contradiction cases.

- If we cannot map a NE from H, either directly or by using the acronym database and background knowledge, the result for the current pair is *Unknown*.
Unknown Cases (cont.)

- We make an exception from the named entity rule when the type of named entity is *first name*
  - T: *The man accused of killing Ms. Zapata, ...*
  - H: *Angie Zapata has been killed with a fire extinguisher.*

- In this case we only insert a penalty in the global fitness.
Unknown Cases (cont…)  

- If any of the numbers in the text or the hypothesis has an attached unit of measure, it is always kept
  
  - T: At *least 14 people have been killed in a suicide bomb attack in southern Sri Lanka, police say*. The telecoms minister was among about *35 people* injured in the blast at the town of Akuressa…
  
  - H: *35 government officials* were injured by a suicide bomber in Akuressa
## Results in RTE5

<table>
<thead>
<tr>
<th>Answer Type</th>
<th>In Gold</th>
<th>Correct offered by our system</th>
<th>Total offered by our system</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entailment</td>
<td>300</td>
<td>260</td>
<td>379</td>
<td>68.60%</td>
<td>86.67%</td>
<td>76.58%</td>
</tr>
<tr>
<td>Contradiction</td>
<td>90</td>
<td>22</td>
<td>44</td>
<td>50.00%</td>
<td>24.44%</td>
<td>32.84%</td>
</tr>
<tr>
<td>Unknown</td>
<td>210</td>
<td>128</td>
<td>177</td>
<td>72.32%</td>
<td>60.95%</td>
<td>66.15%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>600</strong></td>
<td><strong>410</strong></td>
<td><strong>600</strong></td>
<td></td>
<td></td>
<td><strong>68.33%</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Answer Type</th>
<th>In Gold</th>
<th>Correct offered by our system</th>
<th>Total offered by our system</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>300</td>
<td>260</td>
<td>379</td>
<td>68.60%</td>
<td>86.67%</td>
<td>76.58%</td>
</tr>
<tr>
<td>No</td>
<td>300</td>
<td>181</td>
<td>221</td>
<td>81.90%</td>
<td>60.33%</td>
<td>69.48%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>600</strong></td>
<td><strong>441</strong></td>
<td><strong>600</strong></td>
<td></td>
<td></td>
<td><strong>73.50%</strong></td>
</tr>
</tbody>
</table>
# Ablation Tests

<table>
<thead>
<tr>
<th>System Description</th>
<th>2-way (73.5 %)</th>
<th>3-way (68.33 %)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P (%)</td>
<td>C (%)</td>
</tr>
<tr>
<td>Without DIRT</td>
<td>73.33</td>
<td>0.17</td>
</tr>
<tr>
<td>Without WordNet</td>
<td>72.50</td>
<td>1.00</td>
</tr>
<tr>
<td>Without Acronyms</td>
<td>73.33</td>
<td>0.17</td>
</tr>
<tr>
<td>Without BK</td>
<td>72.33</td>
<td>1.17</td>
</tr>
<tr>
<td>Without NE rule</td>
<td>67.33</td>
<td><strong>6.17</strong></td>
</tr>
<tr>
<td>Without the Negation rule</td>
<td>73.50</td>
<td>0.00</td>
</tr>
<tr>
<td>Without the Contradiction rule</td>
<td>71.50</td>
<td><strong>2.00</strong></td>
</tr>
<tr>
<td>Without additional processing steps</td>
<td>69.33</td>
<td>4.17</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>14.85</strong></td>
<td><strong>12.49</strong></td>
</tr>
</tbody>
</table>
## Ablation Tests Overview

<table>
<thead>
<tr>
<th>System description</th>
<th>RTE-3 (69.13 %)</th>
<th>RTE-4 (72.1 %)</th>
<th>RTE-5 (73.5 %)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>C</td>
<td>WR</td>
</tr>
<tr>
<td>Without DIRT</td>
<td>68.76</td>
<td>0.37</td>
<td>0.54</td>
</tr>
<tr>
<td>Without WordNet</td>
<td>68.00</td>
<td>1.13</td>
<td>1.63</td>
</tr>
<tr>
<td>Without Acronyms</td>
<td>68.38</td>
<td><strong>0.75</strong></td>
<td><strong>1.08</strong></td>
</tr>
<tr>
<td>Without BK</td>
<td>67.75</td>
<td>1.38</td>
<td>2.00</td>
</tr>
<tr>
<td>Without the NE rule</td>
<td>57.58</td>
<td><strong>11.55</strong></td>
<td><strong>16.71</strong></td>
</tr>
<tr>
<td>Without the Negation rule</td>
<td>67.63</td>
<td>1.50</td>
<td>2.17</td>
</tr>
<tr>
<td>Without the Contradiction rule</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>No additional processing steps</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>16.68</strong></td>
<td><strong>24.13</strong></td>
<td><strong>18.3</strong></td>
</tr>
</tbody>
</table>
Pilot Task

- Extraction of text from a series of newspaper articles that yielded positive entailment for a given set of hypotheses
  - the texts are not modified in any way as compared to the original source
  - a large numbers of candidate pairs, as for every one of the nine topics there are about ten hypotheses
Pilot Task (cont.)

- In order to reduce the search space, we have made use of a technique used for our question answering systems.
- Using Lucene, we have indexed the articles from each topic at the sentence level.
- We have built queries for all the hypotheses.
- The snippets with the highest chance of yielding positive entailment are clustered around the top scoring snippets.
In order to determine the entailment value of the candidate pairs (approx. 1700 in all), we have applied a lightweight version of our entailment system.

<table>
<thead>
<tr>
<th>Result</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Micro-average</td>
<td>51.12%</td>
<td>22.88%</td>
<td>31.61%</td>
</tr>
<tr>
<td>Macro-average Topic</td>
<td>53.03%</td>
<td>24.08%</td>
<td>33.12%</td>
</tr>
<tr>
<td>Macro-average Hypothesis</td>
<td>46.55%</td>
<td>26.42%</td>
<td>33.71%</td>
</tr>
</tbody>
</table>
Conclusions

- Main idea of our TE system is to map every node from hypothesis to a node from text, either *lexically* or *semantically*.
- The rules regarding Named Entity processing were more elaborate.
- Preprocessing for our RTE-5 is more elaborate.
- RTE-5 also introduced a pilot task, to which we applied QA techniques to reduce the solution space.
Further Work

- Using semantic roles
  - “LOSAIL, Qatar (AFP) - Torrential rain caused the season-opening Qatar MotoGP to be cancelled on Sunday, …”
  - “Valentino Rossi won the season-opening Qatar MotoGP.”
  - “the season-opening Qatar MotoGP did not finish” ≈ “the season-opening Qatar MotoGP was cancelled”
- In order to win a race, the race must finish, the winner must finish that race and the winner needs to be first when he finishes
Further Work (cont.)

- Enhancing entities with ontological knowledge
  - “He has long been linked to some of the world's most notorious conflicts, allegedly supplying arms to former Liberian dictator Charles Taylor and Libyan leader Colonel Gaddafi.”
  - “Gaddafi is the Liberian dictator.”

- In ontological knowledge, we find that a person can only have one occupation
- We attempt to unify the property sets in the text with those in the hypothesis
Acknowledgments

- **Students:**
  - Alexandra Balahur-Dobrescu, Daniel Matei

- **NLP group of Iasi:**
  - Supervisor: Prof. Dan Cristea
  - Maria Husarciuc, Ionut Pistol, Marius Raschip, Diana Trandabat

- **Support from SIR-RESDEC project**
THANK YOU!