An Accuracy-Oriented Divide-and-Conquer Strategy for Recognizing Textual Entailment

Rui Wang
Saarland University
rwang@coli.uni-sb.de

Guenter Neumann
DFKI GmbH
neumann@dfki.de
Outline

• The Architecture(s)

• Precision-Oriented Modules
  • The TACTE module
  • The NE-Oriented module
  • The Tree Skeleton module

• Results & Conclusion
The Architecture(s)
A Common Architecture

- Preprocessing

  - T-H pair

  - NE Recognition

  - Parser

  - WSD

  - Anaphora Resolution

  - …

- Post-Processing

- Entail?
An Alternative Architecture

T-H pair

Splitter

Split 1
Entail?

Split 2
Entail?

Split 3
Entail?

Split 4
Entail?

…
Entail?

Entail?

Post-Processing
Requirements

- Divide
  - A good split

- Conquer
  - Precision-oriented / Highly confident
The Workflow

The RTE Corpus → Preprocessing Module → PO Module 1, PO Module 2, ..., Backup Module 1, Backup Module 2, ..., Voting Module → YES/NO/Unknown
The Precision-Oriented Modules
The TACTE Module

- Input: <T> & <H>
- Result: Yes or No
- NER: SProUT
- Parser: Stanford Parser
- Lexical Resources:
  - WordNet
  - VerbOcean
Temporal Expression Anchoring (TAC)

- Temporal Expression Extraction
  - SProUT

- Temporal Expression Anchoring
  - Default reference date for both $<T>$ and $<H>$
  - Explicit vs. relative temporal expressions
    - e.g. July 5th vs. last Friday
  - Granularity
    - [second < minute < hour < pofd < dofw < day < weeknumber < pofm < month < pofy < year]

(Reference date: Friday, Oct 24th, 1997)

1. The defense secretary William Cohen announced plans on last Thursday. → Thursday, Oct 16th, 1997

2. The earthquake shook the province of Mindanao at 3:08 p.m this afternoon. → 15:08, Friday, Oct 24th, 1997
Event Extraction - Preprocessing

- Preprocessing: dependency parsing (Stanford parser)
An Example

(Entailment = No)

- **<T>** Released in 1995, Tyson returned to boxing, winning the World Boxing Council title in 1996. The same year, however, he lost to Evander Holyfield, and in a 1997 rematch he bit Holyfield’s ear, for which he was temporarily banned from boxing.

- **<H>** In 1996 Mike Tyson bit Holyfield’s ear.

**<T>** 1995: released (verb)
1996: winning (verb)
1997: rematch (noun), bit (verb)

**<H>** 1996: bit (verb)
Another Example

(Entailment = Yes)

• <T> Lima, Jan. 10, '90, the national police reported that over 15,000 people have been arrested in Lima in a dragnet aimed at uncovering the assassins of former Defense Minister Enrique Lopez Albujar Trint, who was murdered in a terrorist attack, yesterday.

• <H> Enrique Lopez Albujar Trint was killed on Jan. 9 '90.

<T> 10-01-1990 (Jan. 10, '90): …
09-01-1990 (yesterday): murdered
<H> 09-01-1990 (Jan. 9 '90): killed
The NE-Oriented Module
The Event Structure

- \(<\text{Event, Time, Location, List<Participants>>}>\)
  - Time: the TACTE system (Wang and Zhang, 2008)
  - Location: the GeoCLEF system (Wang and Neumann, 2008)
  - Participants: the Stanford NER system (Finkel et al., 2005)
An Example

- Pair: YES
  - T: A controversial part of the agreement is the release of Lebanese prisoner Samir Kantar, a militant serving a 542-year sentence for killing two men and a four-year-old girl in a 1979 raid on northern Israel. The brutality of that attack horrified Israelis.
  - H: In 1979 Israel was attacked.

- Events
  - T: [Event:raid, Time:[1979], Location:[Israel]]
  - H: [Event:attacked, Time:[1979], Location:[Israel]]
An Example with Multiple Events

- Pair: YES
  - T: *Spain appeared hardest hit by the protests today*. An estimated 100,000 farmers drove tractors through Madrid and dozens of other Spanish cities, warning of more aggressive action if there is no agreement to *compensate* them for higher fuel costs by October.
  - H: *Spain stages fuel protests*.

- Events
  - T1: [Event:[hit], Time:[today]]
  - T2: [Event:[appeared], Location:[Spain]]
  - T3: [Event:[compensate], Time:[October]]
  - H: [Event:[stages], Location:[Spain]]
The Tree Skeleton Module

- Pair: id="61" entailment="YES" task="IE" source="RTE"
  - Text:
    Although they were born on different planets, Oscar-winning actor Nicolas Cage's new son and Superman have something in common, both were named Kal-el.

- Hypothesis:
  Nicolas Cage's son is called Kal-el.
Tree Skeleton

Dependency Tree of H of pair (id=61):

Text: Nicolas Cage's son is called Kal-el.
Tree Skeleton

Dependency Tree of H of pair (id=61):

- Text: *Nicolas Cage's son is called Kal-el.*
Dependency Tree of T
of pair (id=61):

Tree Skeleton (cont.)
Results & Conclusion
Settings of the Whole System

- Main modules
  - The TACTE system (TAC-M)
  - The Event system (NE-M)
  - The Tree Skeleton system (TS-M) (Wang and Neumann, 2007)

- Backup modules (Wang and Neumann, 2007)
  - The triple similarity (Tri-BM)
  - The bag-of-words similarity (BoW-BM)

- Two issues
  - When to apply the module (Coverage)
  - How good is the module (Precision)
Results (2-way)

<table>
<thead>
<tr>
<th>Tasks</th>
<th>TAC-M</th>
<th>TS-M</th>
<th>NE-M</th>
<th>BoW-BM</th>
<th>Tri-BM</th>
<th>Run1</th>
<th>Run2</th>
<th>Run3</th>
</tr>
</thead>
<tbody>
<tr>
<td>IR(300)</td>
<td>75.0%/4</td>
<td>76.5%/85</td>
<td>61.0%/164</td>
<td>63.3%</td>
<td>54.3%</td>
<td>66.0%</td>
<td>72.3%</td>
<td>71.7%</td>
</tr>
<tr>
<td>QA(200)</td>
<td>90.0%/10</td>
<td>73.2%/82</td>
<td>54.8%/93</td>
<td>49.0%</td>
<td>53.5%</td>
<td>73.0%</td>
<td>72.0%</td>
<td>74.0%</td>
</tr>
<tr>
<td>SUM(200)</td>
<td>83.3%/6</td>
<td>74.5%/51</td>
<td>55.2%/67</td>
<td>63.5%</td>
<td>54.0%</td>
<td>64.0%</td>
<td>69.5%</td>
<td>71.5%</td>
</tr>
<tr>
<td>IE(300)</td>
<td>72.7%/11</td>
<td>74.2%/128</td>
<td>46.7%/152</td>
<td>50.0%</td>
<td>50.0%</td>
<td>66.7%</td>
<td>66.3%</td>
<td>66.7%</td>
</tr>
<tr>
<td>All(1000)</td>
<td>80.6%/31</td>
<td>74.6%/346</td>
<td>54.3%/477</td>
<td>56.5%</td>
<td>52.8%</td>
<td>67.2%</td>
<td>69.9%</td>
<td>70.6%</td>
</tr>
</tbody>
</table>

- **Run1**: TAC-M, TS-M, and Tri-BM
- **Run2**: TAC-M, TS-M, and BoW-BM
- **Run3**: TAC-M, TS-M, NE-M, and Tri-BM, BoW-BM
# Results (3-way)

<table>
<thead>
<tr>
<th>Answers</th>
<th>Run1(2)</th>
<th>Run2(2)</th>
<th>Run3(2)</th>
<th>Answers</th>
<th>Run1(3)</th>
<th>Run2(3)</th>
<th>Run3(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes(500)</td>
<td>66.6%</td>
<td>81.4%</td>
<td>74.8%</td>
<td>Yes(500)</td>
<td>68.2%</td>
<td>66.6%</td>
<td>72.8%</td>
</tr>
<tr>
<td>No(500)</td>
<td>67.8%</td>
<td>58.4%</td>
<td>66.4%</td>
<td>No(150)</td>
<td>38.7%</td>
<td>41.3%</td>
<td>33.3%</td>
</tr>
<tr>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>Unknown(350)</td>
<td>61.4%</td>
<td>47.1%</td>
<td>54.9%</td>
</tr>
<tr>
<td>All(1000)</td>
<td>67.2%</td>
<td>69.9%</td>
<td>70.6%</td>
<td>All(1000)</td>
<td>61.4%</td>
<td>56.0%</td>
<td>60.6%</td>
</tr>
</tbody>
</table>

- **Run1**: TAC-M, TS-M, and Tri-BM, BoW-BM
- **Run2**: TAC-M, TS-M, NE-M (partial), and Tri-BM, BoW-BM
- **Run3**: TAC-M, TS-M, NE-M, and Tri-BM, BoW-BM

- If BoW-BM = YES & Tri-BM = NO then **CONTRACTION**
- If BoW-BM = YES & Tri-BM = YES then **ENTAILMENT**
- Others **UNKNOWN**

An Example

• Pair: YES
  • T: A French court on Wednesday sentenced serial killer Michel Fourniret and his wife to life in prison for the murder of seven girls and young women.
  • H: Michel Fourniret was sentenced to life imprisonment.

• Events
  • T: [Event:[sentenced], Time:[on Wednesday], Roles:[Michel Fourniret]]
  • H: [Event:[sentenced], Roles:[Michel Fourniret]]
An Error

• Pair: YES
  • T: Two Britons have died in a light aircraft plane crash in north west Italy, the Foreign Office has said.
  • H: A plane crashes in Italy.

• Events
  • T1: [Event:died, Location:Italy]
  • T2: [Event:crash, Location:Italy]
  • H: [Event:crashes, Location:Italy]

• How to know the corresponding events
  • Similarity vs. Relatedness
Others’ Work

• NE features

• Precision-based RTE

• Natural Logic
Conclusion & Future Work

• Divide
  • Basic linguistic processing
    → Simple cases of entailment

• Conquer
  • Precision-oriented modules
    → More accurate and more modules

• Integration
  • The voting model
    → A uniform representation/theory
Acknowledgements

- The work was partially supported by a research grant from BMBF to the DFKI project HyLaP (FKZ: 01 IW F02) and the EC-funded project QALL-ME.

- Special thanks to Yajing Zhang for the TAC system.

- Thank you!
Publications

