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

Rui Wang  
Saarland University  
66123 Saarbrücken, Germany  
rwang@coli.uni-sb.de

Günter Neumann  
LT-lab, DFKI  
66123 Saarbrücken, Germany  
neumann@dfki.de

Abstract

Based on our promising results in the RTE-3 challenge (66.9% of accuracy) using a precision-oriented dependency-based approach, we explored further extensions of this perspective. This year, we achieve 70.6% for the two-way task and 61.4% for the three-way task.

Highlights

Time Anchoring Component for RTE (Wang and Zhang, 2008)  
The Event Structure of Named-Entities  
The Tree Skeleton Structure (Wang and Neumann, 2007)  
Two Backup Modules (Wang and Neumann, 2007)

The System

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 bit Holyfield’s ear.

H: In 1996 Mike Tyson bit Holyfield’s ear.

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.

T: For their discovery of ulcer-causing bacteria, Australian doctors Robin Warren and Barry Marshall have received the 2005 Nobel Prize in Physiology or Medicine.

H: Robin Warren was awarded a Nobel Prize.

Experimental Results

<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>72.0%</td>
<td>74.0%</td>
<td>71.5%</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>69.5%</td>
<td>76.5%</td>
<td>73.2%</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>

Answers Run1(2) Run2(2) Run3(2) Answers Run1(3) Run2(3) Run3(3)
---------------------------------------------------------------
Yes(500) 66.6% 61.4% 74.8% Yes(500) 68.2% 66.6% 72.8%  
No(500) 67.8% 58.4% 66.4% No(150) 38.7% 41.3% 33.3%  
Unknown(350) 61.4% 47.1% 54.9%  
All(1000) 67.2% 69.9% 70.6% All(1000) 61.4% 56.0% 60.6%  

References


Acknowledgements

The work was partially supported by a research grant from BMFT 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.

HyLaP: http://hylap.dfki.de/  
QALL-ME: http://qallme.itc.it/