A Specialized RTE System for the PETE Task

Rui Wang and Yi Zhang
Saarland University and DFKI GmbH, Germany
Outline

- Motivations
- System Description
- Experiment Results
- Discussion
The Task

- Parser Evaluation using Textual Entailment (PETE)
  - Textual Entailment $\rightarrow$ Parser Evaluation
  - Parser Output $\rightarrow$ Recognizing Textual Entailment (RTE)

- An Example
  - T: *The man with the hat was tired.*
  - H-1: *The man was tired.* (YES)
  - H-2: *The hat was tired.* (NO)
Motivation I: Specialized RTE

- Specialized RTE Modules (Wang and Neumann, 2009)
  - Temporal expressions
  - Named-Entities
  - Inference/Paraphrase Rules
  - Etc.

- Textual Entailment Specialized Datasets (Bentivogli et al., 2010)
Motivation II: Syntax vs. Semantics

- Parsing
  - Deep grammars (Syn → Sem)
  - Shallow semantic parsers (i.e. semantic role labeler) (Sem → Syn)

- RTE
  - Syntactic phenomena only (this task)

- No clear-cut boundary
System Architecture

- Preprocessing
- Dependency Path Extraction
- Feature-based Classification
HPSG Parsing

- English Resource Grammar (Flickinger, 2000)
  - Precision-Oriented broad coverage HPSG grammar
  - Semantic representation based on MRS
  - Strict grammaticality checking
  - [http://www.delph-in.net/erg/](http://www.delph-in.net/erg/)

- PET Parser (Callmeier, 2001)
  - Efficient parsing with unification-based grammars
  - [http://wiki.delph-in.net/moin/PetTop](http://wiki.delph-in.net/moin/PetTop)
HPSG Parsing (cont.)

+ Ungrammatical Hs
  + T: They could be done or left undone and nobody really gave a damn.
  + H: Somebody could be done the undone.

  + T: He went upstairs to phone Crumb.
  + H: Somebody went Crumb.

  + T: Sequa makes and repairs jet engines.
  + H: Somebody jets the engines.

  + *H: Indicator is the monthly tally.
Preprocessing

+ Syntactic Dependency Tree
Preprocessing (cont.)

+ Semantic Dependency Graph

```
-Root- Devotees of the market question the value of the work national service would perform
  1 devotee NNS of IN the DT market NN question NN the DT value NN of IN the DT work NN national JJ service NN would MD perform VB

-Root- Value is questioned
  1 value NNP be VBZ question VBN
```
Dependency Path Extraction

+ Syntactic Dependency Tree
  + Direct path finding

+ Semantic Dependency Graph
  + Dijkstra’s algorithm (Dijkstra, 1959)

+ Joint Dependency Graph
  + Different weights for syntactic and semantic dependencies
  + Dijkstra’s algorithm
Feature Model

+ An example
  + T: explain – A1/SemEdge → may – TMP/SynEdge → then – AMOD/SynEdge → again
  + H: explain – AM-TMP/SemEdge → again

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Joint</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>No Sem</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>No Syn</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
Experiment Settings

- Dependency Parsers
  - MSTParser (McDonald et al., 2005)
  - MaltParser (Nivre et al., 2007)

- Semantic Role Labeler (Zhang et al., 2008)

- HPSG Parsing
  - English Resource Grammar (Flickinger, 2000)
  - PET Parser (Callmeier, 2001)

- SVM-based Classifier: UniverSVM (Collobert et al., 2006)
## Experiment Results

<table>
<thead>
<tr>
<th></th>
<th>MSTParser+SRL</th>
<th>MaltParser+SRL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Joint</td>
<td>No Sem</td>
</tr>
<tr>
<td>+GC</td>
<td>0.5249</td>
<td>0.5116</td>
</tr>
<tr>
<td></td>
<td>(-1.3%)</td>
<td>(-2.0%)</td>
</tr>
<tr>
<td>-GC</td>
<td>0.5216</td>
<td>0.5050</td>
</tr>
</tbody>
</table>

- HPSG-based grammaticality checking helps
- Semantic dependency helps
- Parser variation
Examples

+ Gold: YES
  + T: Devotees of the market question the value of the work national service would perform.
  + H: Value is questioned.

+ Syntactic dependency
  + T: value ←SBJ– is –VC → question
  + H: value ←OBJ– question

+ Semantic dependency
  + T: value ←A₁– question
  + H: value ←A₁– question
Examples (cont.)

+ **Gold: NO**
  + T: All she had to do was **put $15,000 in a certificate of deposit, or qualify for a $10,000 personal line of credit.**
  + H: She had to **qualify in a certificate of deposit.**

+ **Syntax**
  + A: put – **TMP → in**
  + B: put – **PUT → in**

+ **Semantics**
  + A: **qualify** – **AM-LOC/SemEdge → in**
  + B: **qualify** – **A1/SemEdge → all ← A0/SemEdge – put – A2/SemEdge → in**
Examples (cont.)

+ Gold: NO
  + T: And many in the young cast bear striking resemblances to American TV and movie personalities known for light roles.
  + H: Many bear something to American TV.

+ Syntax
  + A: bear –OBJ → resemblances –NMOD → to
  + B: bear –ADV → to

+ Semantic
  + A: no links between “bear” and “to”
  + B: bear –A2 → to
Summary

- Deeper parsing is useful (vs. shallow ones)
- Different parsers have different behavior
- Application-based evaluation
  - One scenario
  - Indirect evaluation
  - Difficult to do error analysis
Future Work

- Parser evaluation using other applications
- Make full use of HPSG parsing
  - Currently only grammaticality checking
- HPSG-MRS semantic information
Thank YOU!

Questions?