Off-line Answer Extraction for Dutch Question Answering

CLIN
December 17th 2004

Jori Mur
Humanities Computing
University of Groningen
Answer extraction for question answering

- Task of question answering (QA): Find an answer in a large text collection to a question posed in natural language.

- Answer extraction: Extract answers by looking at frequently occurring patterns.
Example of frequent answer structure

- What is the capital of [country]?
  Wat is de hoofdstad van Bulgarije?

  [answer], the capital of [country]
  Sofia, de hoofdstad van Bulgarije

- The [country-adj] capital [answer]
  De Bulgaarse hoofdstad Sofia
Off-line answer extraction

- Extract answers from a text before the QA-process and store them

- Two ways to do that (Mur et al., 2004)
  - Regular Expressions
  - Syntactic Patterns
Overview

- Off-line answer extraction
- Our QA-system
- Patterns
- Table Creation
- Table lookup
- Evaluation and Results
- Conclusion and Future work
Off-line answer extraction for Dutch QA

Our QA-system: Joost

Off-line extraction
Overview
Our QA-system
Patterns
Table Creation
Table lookup
Evaluation & Results
Conclusion

Our QA-system:

1. Question
2. Parsed corpus
3. Question Analysis
4. Keywords & Question type
5. Table lookup
6. Tables
7. Online Information Retrieval
   NLP Documents reranking
8. Answers

University of Groningen
Off-line answer extraction for Dutch QA

Parsing with Alpino

Off-line extraction
Overview
Our QA-system
Patterns
Table Creation
Table lookup
Evaluation & Results
Conclusion
Categories

Included:
Abbreviations, Capitals, Currency, Functions, Inhabitants, Locations, Manner of death

Excluded:
Definition, Date, Measure and 'Which'
Off-line answer extraction for Dutch QA

Patterns

hoofdstad (noun)  mod  Country (adj)

hoofdstad (noun)  app  Capital (name)
Table Creation

... capital('Brazilië',5,'Brasilia','AD19950207-0050-277-1.xml').
capital('Brazilië',5,'Brasilia','NH19941222-0114-911-1.xml').
capital('Brazilië',5,'Brasilia','NH19950127-0035-244-8.xml').
capital('Brazilië',5,'Brasilia','NH19950217-0007-45-1.xml').
capital('Brazilië',5,'Brasilia','NH19950224-0043-306-4.xml').
capital('Bretagne',1,'Rennes','NH19940204-0076-551-1.xml').
capital('Bulgarije',5,'Sofia','AD19940422-0032-220-2.xml').
capital('Bulgarije',5,'Sofia','NH19940103-0037-264-2.xml').
capital('Bulgarije',5,'Sofia','NH19950421-0087-773-5.xml').
capital('Bulgarije',5,'Sofia','NH19951019-0133-1065-2.xml').
capital('Bulgarije',5,'Sofia','NH19951026-0144-993-2.xml').
capital('Burundi',1,'Bujumbura','NH19950413-0027-189-2.xml').
...
Table Lookup

Q: *Wat is de hoofdstad van Bulgarije?*

Question analysis ->
question type: capital, keyword(s): *Bulgarije*


A: *Sofia*
Evaluation

- **Data**

- **Training CLEF 2003**
  AD and NRC 1994/1995
  450 questions, 370 with answer

- **Testing CLEF 2004**
  AD and NRC 1994/1995
  200 questions, 185 with answer

- **MRR**

  The score for a question is the reciprocal of the rank at which the first correct answer occurs or 0 if no correct answer is found. The score for the whole system is then the mean over the set of questions.
Results (1)

MRR for questions answered by Qatar

<table>
<thead>
<tr>
<th>Corpus</th>
<th># Questions</th>
<th>MRR Qatar</th>
<th>MRR Online</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLEF 2003</td>
<td>65/370</td>
<td>0.73</td>
<td>0.40</td>
</tr>
<tr>
<td>CLEF 2004</td>
<td>24/185</td>
<td>0.71</td>
<td>0.46</td>
</tr>
</tbody>
</table>
### Results (2)

**CLEF 2003**

370 questions

<table>
<thead>
<tr>
<th></th>
<th>Without tables</th>
<th>With tables</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRR</td>
<td>0.49</td>
<td>0.54</td>
</tr>
<tr>
<td># correct on 1st</td>
<td>145 (39%)</td>
<td>174 (47%)</td>
</tr>
</tbody>
</table>

- 65 (18%) questions answered by Qatar
Results (3)

CLEF 2004
185 questions

<table>
<thead>
<tr>
<th>Without tables</th>
<th>With tables</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRR</td>
<td>0.51</td>
</tr>
<tr>
<td># correct on 1\textsuperscript{st}</td>
<td>77 (42%)</td>
</tr>
<tr>
<td></td>
<td>81 (44%)</td>
</tr>
</tbody>
</table>

24 (13%) questions answered by Qatar
Conclusion

- Precision is high for Off-line module
- Performance increases
- Too few questions are answered
- Unexpected question types will never be answered
Future work

- Is it possible to answer more questions with Qatar?
  - Using ontology to classify 'which' questions
  - Discovery of patterns by machine learning techniques
  - Coreference resolution (maybe for Definition questions)
Questions?
Off-line answer extraction for Dutch QA
## Evaluation & Results (4)

### Table lookup per Category (CLEF 2004)

<table>
<thead>
<tr>
<th>Categories</th>
<th>MRR</th>
<th># questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital</td>
<td>1.00</td>
<td>3 of 3</td>
</tr>
<tr>
<td>Inhabitants</td>
<td>1.00</td>
<td>1 of 1</td>
</tr>
<tr>
<td>Location</td>
<td>1.00</td>
<td>3 of 20</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>1.00</td>
<td>1 of 4</td>
</tr>
<tr>
<td>Manner of death</td>
<td>0.667</td>
<td>3 of 5</td>
</tr>
<tr>
<td>Function</td>
<td>0.545</td>
<td>11 of 24</td>
</tr>
<tr>
<td>Currency</td>
<td>0.500</td>
<td>2 of 5</td>
</tr>
</tbody>
</table>