The Informative Role of WordNet in Open-Domain Question Answering

Marius Pa_ca and Sanda M. Harabagiu
(NAACL 2001)

Presented by Shauna Eggers
CS 620  February 17, 2004
Introduction

- **Information Extraction**: not just for keywords anymore!
  - Massive document collections (databases, webpages) require more sophisticated search techniques than keyword matching
  - Need way to focus and narrow search → improve precision

- **One solution**: Open-Domain Q/A
  - Find answers to natural language questions from large document collections
  - Examples:
    - “What *city* is the capital of the United Kingdom?”
    - “Who is the first *private citizen* to fly in space?”
  - Text Retrieval Conferences (TREC) evaluate entered systems; show that this sort of task can be performed with “satisfactory accuracy” (Voorhees, 2000)
Q/A: Previous Approach

- Captures the *semantics* of the question by recognizing:
  - expected answer type (i.e., its semantic category)
  - relationship between the answer type and the question concepts/keywords

- The Q/A process:
  - Question processing – Extract concepts/keywords from question
  - Passage retrieval – Identify passages of text relevant to query
  - Answer extraction – Extract answer words from passage

- Relies on standard IR and IE Techniques
  - Proximity-based features
    - Answer often occurs in text near to question keywords
  - Named-entity Recognizers
    - Categorize proper names into semantic types (persons, locations, organizations, etc)
    - Map semantic types to question types ("How long", "Who", "What company")
Problems

- NE assumes all answers are named entities
  - Oversimplifies the generative power of language!
  - What about: “What kind of flowers did Van Gogh paint?”

- Does not account well for morphological, lexical, and semantic alternations
  - Question terms may not exactly match answer terms; connections between alternations of Q and A terms often not documented in flat dictionary
  - Example: “When was Berlin’s Brandenburger Tor erected?” → no guarantee to match built
  - Recall suffers
WordNet to the rescue!

1. **WordNet can be used to inform all three steps of the Q/A process**
   1. Answer-type recognition (*Answer Type Taxonomy*)
   2. Passage Retrieval (“specificity” constraints)
   3. Answer extraction (recognition of keyword alternations)

2. **Using WN’s lexico-semantic info: Examples**
   - “What kind of flowers did Van Gogh paint?”
     - Answer-type recognition: need to know (a) answer is a kind of flower, and (b) sense of the word flower
     - WordNet encodes 470 hyponyms of flower sense #1, flowers as plants
     - Nouns from retrieved passages can be searched against these hyponyms
   - “When was Berlin’s Brandenburger Tor erected?”
     - Semantic alternation: erect is a hyponym of sense #1 of build
Interactions between WN and Q/A

**Question Processing**
- Expected Answer Type
- Keyword Alternations

**Document Processing**
- Index
- Passage Retrieval

**Answer Processing**
- Answer Extraction

WordNet

Question → 
Documents → 
Answer(s) → 

WN in Answer-type Recognition

- **Answer Type Taxonomy**
  - a taxonomy of answer types that incorporates WN information
  - Acts as an “ontological resource” that can be searched to identify a semantic category (representing answer type)
  - Used to associate found semantic categories with a named entity extractor
  - So, still using an NE, but not bound to proper nouns; have found a way to map NEs to more general semantic categories

- Developed on principles conceived for Q/A environment (rather than as general onto principles)
  - Principle 1: Different parts of speech specialize the same answer type
  - Principle 2: Selected word senses are considered
  - Principle 3: Completeness of the top hierarchy
  - Principle 4: Conceptual average of answer types
  - Principle 5: Correlating the Answer Type Taxonomy with NEs
  - Principle 6: Mining WordNet for additional knowledge
Answer Type Taxonomy (example)
WN in Passage Retrieval

- Identify relevant passages from text
  - Extract keywords from the question, and
  - Pass them to the retrieval module

- “Specificity” – filtering question concepts/keywords
  - Focuses search, improves performance and precision
  - Question keywords can be omitted from the search if they are too general
  - Specificity calculated by counting the hyponyms of a given keyword in WordNet
    - Count ignores proper names and same-headed concepts
    - Keyword is thrown out if count is above a given threshold (currently 10)
WN in Answer Extraction

- If keywords alone cannot find an acceptable answer, look for alternations in WordNet!

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q196: Who wrote “Hamlet”??</td>
<td>before the young playwright has written Hamlet – and Shakespeare seizes the opportunity</td>
</tr>
<tr>
<td><strong>Morphological Alternation:</strong></td>
<td><strong>wrote → written</strong></td>
</tr>
<tr>
<td>Q136: Who is the queen of Holland?</td>
<td>Princess Margrit, sister of Queen Beatrix of the Netherlands, was also present</td>
</tr>
<tr>
<td><strong>Lexical Alternation:</strong></td>
<td><strong>Holland → Netherlands</strong></td>
</tr>
<tr>
<td>Q196: What is the highest mountain in the world?</td>
<td>first African country to send an expedition to Mount Everest, the world’s highest peak</td>
</tr>
<tr>
<td><strong>Semantic Alternation:</strong></td>
<td><strong>mountain → peak</strong></td>
</tr>
</tbody>
</table>
Evaluation

- Pa_ca/Harabagiu approach measured against TREC-8 and TREC-9 test collections
- WN contributions to Answer Type Recognition
  - Count number of questions for which acceptable answers were found; 3GB text collection, 893 questions

<table>
<thead>
<tr>
<th>Method</th>
<th># questions with correct answer type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat dictionary (baseline)</td>
<td>All 227 (32%)</td>
</tr>
<tr>
<td>A-type taxonomy (static)</td>
<td>All 445 (64%)</td>
</tr>
<tr>
<td>A-type taxonomy (dynamic)</td>
<td>All 463 (67%)</td>
</tr>
<tr>
<td>A-type taxonomy (dynamic + answer patterns)</td>
<td>All 533 (76%)</td>
</tr>
</tbody>
</table>
### Evaluation (2)

#### WN contributions to Passage Retrieval

**Impact of keyword alternations**

<table>
<thead>
<tr>
<th>Alternation Type</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>No alternations enabled</td>
<td>55.3%</td>
</tr>
<tr>
<td>Lexical alternations enabled</td>
<td>67.6%</td>
</tr>
<tr>
<td>Lexical + semantic alternations enabled</td>
<td>73.7%</td>
</tr>
<tr>
<td>Morphological expansions enabled</td>
<td>76.5%</td>
</tr>
</tbody>
</table>

**Impact of specificity knowledge**

<table>
<thead>
<tr>
<th>Specificity knowledge</th>
<th>TREC-8</th>
<th>TREC-9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not included</td>
<td>133 (65%)</td>
<td>463 (67%)</td>
</tr>
<tr>
<td>Included</td>
<td>151 (76%)</td>
<td>515 (74%)</td>
</tr>
</tbody>
</table>
Conclusions

- Massive lexico-semantic information must be incorporated into the Q/A process
  - Using such information encoded in WN improved system precision by 147% (qualitative analysis)

- Visions for future:
  - Extend WN so that online resources like encyclopedias can link to WN concepts
    - Answer questions like: “Which classic rock group first performed live in Alburquerque?”
  - Further improve Q/A precision with WN extension projects
    - Eg, “finding keyword morphological alternations could benefit from derivational morphology, a project extension of WordNet” (Harabagiu et al., 1999)