Chunking-based Question Type Identification for Multi-sentence Queries

Mineki Takechi (Fujitsu Ltd.)
Takenobu Tokunaga (TITECH)
Yuji Matsumoto (NAIST)
SIGIR 2007 Workshop Focused Retrieval

Open Domain QA and Web Q&A Services

- Open domain question-answering: e.g., TREC, NTCIR
  - Query type in TREC and NTCIR
    - Mainly factoid and single sentence questions
    - e.g., “When did the Jurassic Period end?”
    - Non-factoid and multi-sentence queries (MSQs) studied not enough
- The Web question-answering service: e.g., FMWORLD, MSN, Oshiete! goo
  - Queries in Web question-answering services
    - Multiple sentences and multiple questions
    - Factoid and Non-factoid questions mixed in a query

Example Query in Web Q&A services

- S1: Even when I sleep enough every night, I’m very tired all day...
- Question A
  - S2: My friends tell me that these symptoms resemble depression, but what is the definition of depression?
- S3: In my office, I have no time to relax because of my post.
- Question B
  - S4: My wife is concerned about my recent condition and recommends that I see the doctor.
  - S5: I’m not sure whether I should see the doctor.
  - S5*: Should I see the doctor?
  - S6: Please tell me your advice if you have same experience.

Contents

- Query processing and annotation
- Proposed technique
- Evaluations
- Summary
Overview of Query Processing

Question Types in MSQs

- Identifying all questions in MSQs
  - Extract questions described in natural languages and identify their (semantic) question types.
- Various question types proposed
  - Mainly address single sentence queries
  - Depending on tasks and based on answer types
  - TREC/NTCIR: Name, Location, Organization...
  - QA services [Kurohashi2000][Matsui2003]: How-to, Request, ...
- Not many studies of question types in the Japanese Web Q&A services

Question Type Annotation to Real Queries

- “Oshiete go!”: Web Q&A service, 2234 queries (in 2001, 2006)
- Assign question types to multi-sentence queries; four factoid and five non-factoid types

<table>
<thead>
<tr>
<th>Type</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes-No</td>
<td>Do your browser show “save” tab?</td>
</tr>
<tr>
<td>How-to</td>
<td>Ask methods to solve problems</td>
</tr>
<tr>
<td>Reason</td>
<td>Ask a reason</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Ask sentiments or opinions How do you feel new digital camera released by A Limited?</td>
</tr>
</tbody>
</table>

- Check inter-annotator agreement of question type annotation

Annotation results

- Queries including multiple questions: 56% (1252/2234 articles)
  - Dominantly exist in real queries

<table>
<thead>
<tr>
<th>Question Type</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes-No</td>
<td>43%</td>
</tr>
<tr>
<td>Non-factoid</td>
<td>19%</td>
</tr>
<tr>
<td>Mix of factoid and</td>
<td>19%</td>
</tr>
<tr>
<td>non-factoid</td>
<td></td>
</tr>
<tr>
<td>factoid</td>
<td></td>
</tr>
<tr>
<td>Name, Location, Time</td>
<td>19%</td>
</tr>
</tbody>
</table>

- Inter-annotator agreement on the F-measure (MUC98) without sentences including multiple question items
  - High agreement: Yes-No, Location, How-to : 0.8
  - Low agreement: Description : 0.5
Observations of Multi-sentence Queries

- Many non-question sentences: greetings and apologies
- Ellipses and reference expressions often exist in question

  “After catching a cold, many small, white dots appeared on the downside of my tongue. What do that seem to be?”

  – To know the antecedent of “that” and semantic class of “what”, the first sentence is needed.

- One question described in multiple sentences often appear

  “Plants that can grow in indoor or small beranda. Additionally, I wants them to be easy to care for beginner like me, so do anyone know such plants?”

- The sentence including multiple questions is existed but few and mostly stereotyped
  – Excluding these cases in the current stage of study

Question Type Identification for MSQs

- Required new unit of question type identification
  – Questions in MSQs often described by multiple sentences

  Question segment
  – A minimal set of sentences to identify a question type

  Question type identification for MSQs
  – Extract all question segments in a query and identify their question types

Previous Work of QTI for MSQs

- Mainly pattern matching approaches for single sentence queries but some machine learning based methods [Li02, Zhang03, Suzuki05, Tamura05]

  – Two pass processing: question extraction and the type identification
  – Accuracy of core sentence extraction: F-measure 0.897. But not clear in type identification for MSQs

- High computational cost using Support Vector Machines [Vapnik95]
  – Required reduce the computational cost

Proposed Technique
Overview of Proposed Techniques

- **Chunking-based question type identification**
  - Chunking: identification of semantic “chunks” e.g., noun phrase, named entity, ...
  - A chunk in our task equals to a question segment
  - Sentence chunking: assign a chunk tag to each sentence
  - Chunk tags: represented by a pair of the portion in the question segment and the question type

  ex. 1 B: the first sentence of the question segment and D: Description type => B-D
  ex. 2 I: the second or more sentence and How-to => I-W

Chunking-based Question Type Identification

<table>
<thead>
<tr>
<th>Labels</th>
<th>Query</th>
<th>Question Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-D</td>
<td>Even when I sleep enough every night, I’m very tired all day.</td>
<td>Description</td>
</tr>
<tr>
<td>0</td>
<td>My friends tell me that these symptoms resemble depression, but what is the definition of depression?</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>In my office, I have no time to relax because of my post.</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>My wife is concerned about my recent symptoms and recommends that I see the doctor.</td>
<td></td>
</tr>
<tr>
<td>B-W</td>
<td>How do other directors like me manage their work stress?</td>
<td>How-to</td>
</tr>
<tr>
<td>I-W</td>
<td>Please let me know if you have good advice.</td>
<td></td>
</tr>
</tbody>
</table>

Machine Learning Approach with the CRFs

- Perform question segmentation and the type identification concurrently; more efficient than the method in previous work

- Conditional Random Fields(CRFs): single exponential model for the joint probability of the entire sequence of labels given the observation sequence
  - Advantageous in computational cost and the flexibility to design the parameters of the model

- Using CRF++: the CRF implementation by [Kudo06]

Evaluations
Experimental Settings

- Features: the uni-gram and bi-gram of words
  - Comparing the performance when using the feature of functional words alone with one when using all words
- Baseline: n words from the head and tail of sentence
- Traditional representation of chunk
  - IO/IOB1/IOB2/IOE1/IOE2/IOBES [Kudo2000]
- Varying the number of sentence exploited in the features of CRF

Feature Matrix for Chunker

- Rows of binary feature vector based on words occur in each sentence

<table>
<thead>
<tr>
<th></th>
<th>Group A: m frequent POS</th>
<th>Group B: n POS at end of sentence</th>
<th>chunk tags</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>feature1</td>
<td>feature2 ...</td>
<td></td>
</tr>
<tr>
<td>s1</td>
<td>Even when ...</td>
<td>W1 W2 ... Wm</td>
<td></td>
</tr>
<tr>
<td>s2</td>
<td>My friend ...</td>
<td>W1 W2 ... Wm</td>
<td>B-D</td>
</tr>
<tr>
<td>s3</td>
<td>In my office ...</td>
<td>nil nil ... nil</td>
<td></td>
</tr>
<tr>
<td>s4</td>
<td>My wife ...</td>
<td>nil nil ... nil</td>
<td></td>
</tr>
<tr>
<td>s5</td>
<td>-</td>
<td>W1 nil ... Wm</td>
<td></td>
</tr>
<tr>
<td>s6</td>
<td>Please let me ...</td>
<td>W1 nil ... Wm</td>
<td>B-W</td>
</tr>
<tr>
<td>s7</td>
<td>I need some ...</td>
<td>W1 W2 ... Wm</td>
<td></td>
</tr>
</tbody>
</table>

Experimental Results and Discussion

- Segmentation: F-measure 0.60 (Bi-gram, window size = five)
  - When adding the words and the contexts, improved the accuracy
- Type identification: F-measures (window size = one)
  - When adding the contexts of words, incline the accuracy

<table>
<thead>
<tr>
<th>How-to</th>
<th>0.50</th>
<th>Uni-gram</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>0.34</td>
<td>Uni-gram</td>
</tr>
<tr>
<td>Reason</td>
<td>0.35</td>
<td>Uni+bi-gram</td>
</tr>
<tr>
<td>Description</td>
<td>0.34</td>
<td>Uni+bi-gram</td>
</tr>
</tbody>
</table>

- Different tendency with the contexts of words by question types
- Error Analysis: errors often occur when anaphora resolution and word sense disambiguation are required
  - e.g., “Could you give me any advice”; no antecedent in adjacent sentence

Summary

- Question type identification for multi-sentence queries characterized low computational cost
  - Handling the query including multiple questions of single sentence query in a same flame work
  - Disappointed accuracies in some experiments yet
- Required anaphora resolutions and word sense disambiguation for queries required
- Consider different tendency between question segmentation and type identification for same features
Thank you