

Question Answering (continued)

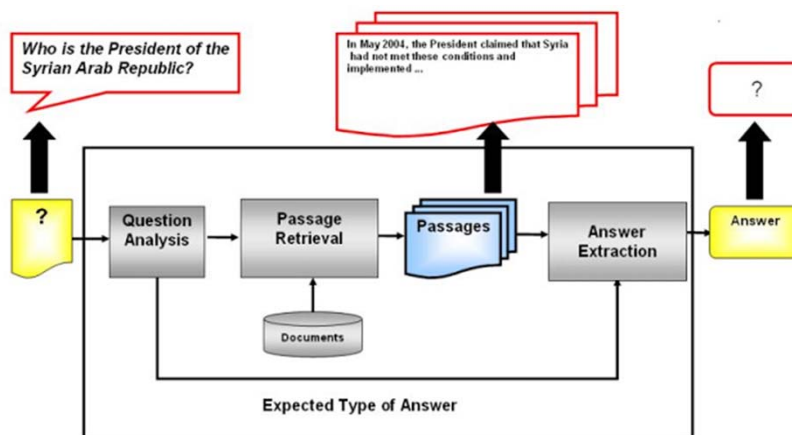
Lecture 22: November 22, 2013

CS886-2 Natural Language Understanding
University of Waterloo

CS886 Lecture Slides (c) 2013 P. Poupart

1

Common Architecture



CS886 Lecture Slides (c) 2013 P. Poupart

2

Passage Retrieval

- Task: given some relevant documents, retrieve candidate sentences, paragraphs or sections
- Rule-based systems or machine learning techniques
- Rules or features based on the presence/absence of
 - The desired answer type
 - Question keywords or n-grams
 - Question entities and relations

CS886 Lecture Slides (c) 2013 P. Poupart

3

Answer Type Detection

- Factoid Questions
 - Location, person, organization, date/time, etc.
 - Answer often consists of one or a few words
 - Answer types often correspond to entity types
- Non-Factoid Questions
 - Definitions, lists, descriptions, reasons, etc.
 - Answer often consists of a phrase (multiple words)
 - Answer types may be detected by patterns, e.g., regular expressions

CS886 Lecture Slides (c) 2013 P. Poupart

4

Non-Factoid Questions

- Examples
 - What is autism?
 - “... development disorders such as autism...”
 - What is caldera?
 - “... the Long Valley caldera, a volcanic crater 19 km long ...”
- Pattern-based detection
 - <AP> such as <QP>
 - <QP>, a <AP>

AP means answer phrase and QP means question phrase

CS886 Lecture Slides (c) 2013 P. Poupart

5

Pattern Discovery

1. Manually build a list of mention pairs
E.g. {<autism, development disorders>, <caldera, volcanic crater>}
2. Query the web with instances of those mention pairs
3. Extract sentences that contain both mentions in a pair
4. Extract regular expressions that match words and punctuation between mention pairs
5. Keep all patterns that have sufficiently high precision

CS886 Lecture Slides (c) 2013 P. Poupart

6

Precision

- Measure precision of a pattern by querying the web with the question phrase and the pattern.
- Precision can be approximated by the fraction of mentions that match the answer phrase
- Example
 - Web query: <such as autism>?
 - Answers retrieved: development disorders(20), disease (10), syndrome (10)
 - Precision: $20/40 = 50\%$

CS886 Lecture Slides (c) 2013 P. Poupart

7

Answer Extraction

- Task: extract one or several answers consisting of sequences of words ordered by relevance
 - Keep surrounding text to help users verify/understand the answer
- Example: How tall is mount Everest?
 1. The official height of mount Everest is 29035 feet
 2. Mount Everest is the highest peak in the Himalayas at more than 8km high
 3. Mount Everest is taller than 25000 feet

CS886 Lecture Slides (c) 2013 P. Poupart

8

Answer Extraction

- **Rule-based Systems**
 - Rules based on regular expressions and other patterns
- **Machine Learning Techniques**
 - Learn mapping techniques from features to word sequences
- **Logical Inference Techniques**
 - Convert text to first-order logic and reason about the answer
 - Does not take into account ambiguities
 - Not used anymore

CS886 Lecture Slides (c) 2013 P. Poupart

9

Rules/Features for Answer Extraction

- Answer type
- Pattern match
 - Templates, regular expressions, n-grams
- Keyword distance
- Punctuation location
- Entity/relation match

CS886 Lecture Slides (c) 2013 P. Poupart

10

Evaluation of Factoid Answers

- Output: ranked list of answers
- **Mean reciprocal rank (MRR)**
 - Average of the inverse rank of the first correct answer for each question

$$MRR = \frac{\sum_{i=1}^N \frac{1}{rank_i}}{N}$$

CS886 Lecture Slides (c) 2013 P. Poupart

11

State of the Art

- Conferences for the evaluation of QA systems
 - **TREC** (Text REtrieval Conference) (USA)
 - Question Answering Track
 - **NTCIR** (National institute of informatics Test Collection for Information Retrieval) (Japan)
 - Multi-lingual question answering track
 - E.g., Questions asked in English, but answers returned in Japanese
 - **CLEF** (Conferences & Labs of the Evaluation Forum) (EU)
 - QA4MRE track (Question Answering for Machine Reading Evaluation)
 - Extract or verify answer in a single document

CS886 Lecture Slides (c) 2013 P. Poupart

12

Interactive QA

- Traditional QA: single question and answer
- Interactive QA: sequence of questions and answers
 - Context
 - Refine/disambiguate questions
 - Refine/disambiguate answers
 - Feedback
- Dialogue systems
 - Conversation
 - Chatbots
 - Spoken dialog systems