ASQA2 –
Academia Sinica Question Answering System on C-C and E-C Subtasks


Academia Sinica, Taiwan
aska@iis.sinica.edu.tw
Outline

- Overview
- Major Extensions
  - English Question Classification
  - Answer Filtering with Answer Template
  - Answer Ranking with SCO-QAT Feature
- Error Analysis
- Conclusion
Overview

- ASQA1 System (NTCIR R5)
- ASQA2 System (NTCIR R6) participated in C-C and E-C subtasks
- ASQA2 focuses on
  - Cross-lingual QA (EC)
  - Syntactic information
  - Global information
- Post-hoc evaluation of ASQA2 on NTCIR R5 test set
  - C-C RU-Accuracy: 0.445 → 0.555
  - C-C R-Accuracy: 0.375 → 0.395
ASQA2 System (NTCIR 5)

Chinese Question

Passage Retrieval

Answer Extraction

Answer Filter

Answer Ranking

Chinese Answer

English Question

Google Translate

Lucene

NER

EAT

Answer Template

Others

SCO-QAT

KE

CQC

NER

EQC

NTCIR-6 CLQA CC subtask:
• R-Accuracy: 0.52
• RU-Accuracy: 0.553

NTCIR-6 CLQA EC subtask:
• R-Accuracy: 0.253
• RU-Accuracy: 0.34
Outline

- Overview
- Major Extensions
  - English Question Classification
  - Answer Filtering with Answer Template
  - Answer Ranking with SCO-QAT Feature
- Error Analysis
- Conclusion
English Question Classification

- SVM
- Features for SVM EQC model
  - word bi-gram
  - first word
  - first two words
  - question wh-word
  - question informer
  - question informer bi-gram
Question Informer for English Questions

- Answer type informer span (Krishnan et al. 2005)
  - a short (typically 1–3 word) subsequence of question tokens that are adequate clues for question classification
  - How much does an adult elephant weigh?

- Predicted by a Conditional Random Field (CRF) model

- Training data set (5,500 questions)
  - UIUC QC dataset (Li and Roth, 2002)
  - Question informer dataset (Krishnan et al., 2005)

- Features: Word, POS, heuristic informer, Parser Information, Question wh-word, length, position.

- 0.939 F-score
Accuracy of English Question Classification by SVM

English Question Classification

<table>
<thead>
<tr>
<th>Method</th>
<th>Top 1 Accuracy (Fine)</th>
<th>Top 1 Accuracy (Coarse)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WB</td>
<td>86.00%</td>
<td>86.00%</td>
</tr>
<tr>
<td>WB+WH</td>
<td>89.33%</td>
<td>89.33%</td>
</tr>
<tr>
<td>WB+WH+QIF</td>
<td>92.00%</td>
<td>92.00%</td>
</tr>
<tr>
<td>WB+WH+QIFB</td>
<td>95.33%</td>
<td>95.33%</td>
</tr>
<tr>
<td>WB+WH+F1+F2</td>
<td>90.67%</td>
<td>90.67%</td>
</tr>
<tr>
<td>WB+WH+QIF</td>
<td>94.00%</td>
<td>94.00%</td>
</tr>
<tr>
<td>WB+WH+QIFB</td>
<td>95.33%</td>
<td>95.33%</td>
</tr>
</tbody>
</table>

NCTIC-6, May 15-18, 2007, National Center of Sciences, Tokyo, Japan

Outline

- Overview
- Major Extensions
  - English Question Classification
  - Answer Filtering with Answer Template
  - Answer Ranking with SCO-QAT Feature
- Error Analysis
- Conclusion
Answer Filters

- Goal
  - Reducing the number of answers without damaging the upper bound of answer accuracy
  - Improving the performance of answer ranking since unrelated answers are removed

- EAT (Expected Answer Type) Filter
- AT-based Filter
Answer Templates

- Syntactic patterns for capturing relations between question terms and answers

- Similar to Surface Patterns used in some QA researches
  - Trained from Question-Answer pairs
  - Gather passages by sending question keywords and the answer

- But different in some ways:
  - Generated by local sequence alignment
  - Not targeting to a specific question type
  - No bootstrapping
Generate and Apply 
Answer Templates

846 QA pairs

Template Generation by Sequence Alignment

Template Selection

Answer templates

AT-based Filter

Template Matching and Relation Construction

Answers
Templates generated by local alignment

- 因/Cbb/O 台中縣/Nc/LOC 議長/Na/OCC 顏清標/Nb/PER 涉嫌/VK/O..
  - 清朝/Nd/O 台灣/Nc/LOC 巡撫/Na/OCC 劉銘傳/Nb/PER 所/D/O..
  - LOC OCC PER (contains only Semantic-tag)

- 被/P/O 大陸/Nc/LOC 國家/Na/O 主席/Na/OCC 江民/Nb/O 形容為/VG/O..
  /COMMA/O 香港/Nc/LOC 行政/Na/O 長官/Na/OCC 董建華/Nb/PER 近日..
  - 俄羅斯/Nc/LOC 男子/Na/O 選手/Na/OCC 史莫契柯夫/Nb/O 在/P/O..
  - LOC Na OCC Nb (template contains POS-tag)

- 由/P/O 建業/Nc/O 所長/Na/OCC 張龍憲/Nb/PER 擔任/VG/O
  由/P/O 安侯/Nb/O 所長/Na/OCC 魏忠華/Nb/PER 擔任/VG/O
  - 由 N 所長 PER 擔任 (template contains paritial POS-tag, word)

- 在/P/O 卡達首都/Nc/LOC 多哈/D/PER,LOC 舉行/VC/O
  於/P/O 國父紀念館/Nc/ORG - 舉行/VC/O
  在/P/O 國父紀念館/Nc/ORG 廣場/Nc/O 舉行/VC/O
  - PNc - 舉行 (template with don't care ‘-’)

Priority of template tag types
  Word > Semantic-tag > POS-tag
Template Selection

- Apply the generated templates to the retrieved passages of training questions
- If there is a passage of which the matched parts contains the answer and some question key terms (with semantic-tag, Nb, or verb), the template will be retained.
- 126 answer templates are selected
### Use Answer Templates to Filter Answers

**Question:** 女演員/OCC 蜜拉索維諾/PER 獲得/VJ 奧斯卡/Nb/ORG 最佳/A 女配角/OCC 獎/Na 是/SHI 因/Cbb 哪/Nep 部/Nf 電影/Na

<table>
<thead>
<tr>
<th>Passage₁</th>
<th>...... 而/Cbb 奪得/VC 一九九五/Neu 奧斯卡/Nb 最佳/A 女配角/OCC 的/DE 殊榮/Na …</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Template₁:</strong></td>
<td>VC Neu Nb A OCC - Na</td>
</tr>
<tr>
<td><strong>Relation₁:</strong></td>
<td>{奪得/VC，奧斯卡/Nb，女配角/OCC}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Passage₂</th>
<th>… 蜜拉索維諾/PER 在/O/P/O 「/O/PAR 非強力春藥/ART 」/PAR 中/Ncd …… 獲/VJ 奧斯卡/Nb 獎/Na …</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Template₂:</strong></td>
<td>PER P PAR ART PAR – DE Na X VJ Nb</td>
</tr>
<tr>
<td><strong>Relation₂:</strong></td>
<td>{蜜拉索維諾/PER，非強力春藥/ART，獲/VJ，奧斯卡/Nb}</td>
</tr>
</tbody>
</table>

| Relation₃: | {奪得/VC，奧斯卡/Nb，女配角/OCC，蜜拉索維諾/PER，非強力春藥/ART，獲/VJ} |

- Only answers found in final relations are retained
- If there is no answer found in the relations, retain all the answers
Answer Template Performance

- **NTCI R6 CLQA C-C**
  - Question Coverage: 37.3%
  - RU-Accuracy: 0.911

- **NTCI R6 CLQA E-C**
  - Question Coverage: 25.3%
  - RU-Accuracy: 0.632
Outline

- Overview
- Major Extensions
  - English Question Classification
  - Answer Filtering with Answer Template
  - Answer Ranking with SCO-QAT Feature
- Error Analysis
- Conclusion
Answer Ranking

- Answer score
  
  \[ AnswerScore(ans) = \sum_{i} w_i \cdot feature_i \]

- Local features: only use the information in the containing passage

- Global features: use information from all the returned passages
SCO-QAT (Sum of Co-occurrence of Question and Answer Terms)

\[ SCOQAT(\text{ans}) = \sum_{c \in QTermCombination} \frac{\text{Freq}(c, \text{ans})}{\text{Freq}(c)} \]
Input Question $Q$: 誰是泰國總理 (Who is the prime minister of Thailand?)

Question Terms: 泰國 總理

2 Candidate Answers: 乃川, 史柏柴

4 Retrieved Passages =

泰國副總理史柏柴也指責避險基金及投機客加重亞洲金融危機。
泰國總理乃川計畫在未來十年中將軍方將領減少四分之三。
泰國總理乃川今天諭令泰國駐印尼大使館，若印尼局勢惡化，應準備撤回泰僑事宜。
泰國總理乃川已指示駐印尼大使館預先準備撤僑事宜，並為泰國僑民設立廿四小時不....

$\Rightarrow$

SCO-QAT(乃川) $= \frac{F(乃川, 泰國)}{F(泰國)} + \frac{F(乃川, 總理)}{F(總理)} + \frac{F(乃川, 泰國, 總理)}{F(總理)} = \frac{3}{4} + \frac{3}{4} + \frac{1}{4} = 2.25$

SCO-QAT(史柏柴) $= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 0.75$
SCO-QAT RU-Accuracy on NTCIR-5 test data

NTCIR-5 C-C

NTCIR-5 E-C

SCO-QAT
Outline

- Overview
- Major Extensions
  - English Question Classification
  - Answer Filtering with Answer Template
  - Answer Ranking with SCO-QAT Feature
- Error Analysis
- Conclusion
Error Analysis of ASQA2 at NTCIR R-6

- **C-C**
  - Answer Ranking (28.6%)
  - Question Classification (19.0%)
  - Time Constraint (15.9%)
  - Others

- **E-C**
  - Wrong Translation (37.9%)
  - Answer Ranking (20.4%)
  - Synonym (10.7%)
  - Others
Conclusion

- We have successfully built an EC QA system by enhancing the CC version with *Google translate* and EQC.

- In both C-C and E-C subtasks:
  - Syntactic information is helpful (Answer Template)
  - Global information is helpful (SCO-QAT)
誰是泰國總理？

ASQA Results:
Cost (sec): 6.719

兀川
——泰國總理兀川八十六歲的父親談著子女之選
udn_auc_19980208_0021

兀川
泰國總理兀川八十六歲的父親談著子女之選
udn_auc_19980317_0192

兀川
泰國總理兀川今天致信泰國駐印度大使館，並向印度總理表示喜愛與支持，並準備撤回泰國的行為，兀川
udn_auc_19980514_0298

兀川
泰國總理兀川已指示泰國駐印度大使館預先準備撤回事務，並於泰國優良地位的行動
udn_auc_19980515_0290

兀川
泰國總理兀川昨天表示，泰國經濟景氣不佳，但泰國已著手科學創新
udn_auc_19981112_0079
Thank You

Intelligent Agent Systems Lab (IASL),
Academia Sinica, Taiwan