IMTKU Question Answering System for World History Exams at NTCIR-12 QA Lab2

Department of Information Management
Tamkang University, Taiwan

myday@mail.tku.edu.tw
IMTKU
Question Answering System for World History Exams at NTCIR-12 QA Lab2
IMTKU Textual Entailment System for Recognizing Inference in Text at NTCIR-9 RITE

Department of Information Management
Tamkang University, Taiwan

Min-Yuh Day
myday@mail.tku.edu.tw

Chun Tu

NTCIR-9 Workshop, December 6-9, 2011, Tokyo, Japan
IMTKU Textual Entailment System for Recognizing Inference in Text at NTCIR-10 RITE-2

Department of Information Management
Tamkang University, Taiwan

Min-Yuh Day
Chun Tu
Hou-Cheng Vong
Shih-Wei Wu
Shih-Jhen Huang

myday@mail.tku.edu.tw

NTCIR-10 Conference, June 18-21, 2013, Tokyo, Japan
IMTKU Textual Entailment System for Recognizing Inference in Text at NTCIR-11 RITE-VAL

Tamkang University

2014

Yu-Hsuan Tai
Yu-An Lin
Shang-Yu Wu
Yu-Hsuan Tai
Cheng-Chia Tsai

Huai-Wen Hsu
Che-Wei Hsu
Ya-Jung Wang
En-Chun Tu
2016
IMTKU Question Answering System for
World History Exams at NTCIR-12 QA Lab2

Department of Information Management
Tamkang University, Taiwan

Min-Yuh Day  Cheng-Chia Tsai  Wei-Chun Chung  Hsiu-Yuan Chang  Tzu-Jui Sun  Yuan-Jie Tsai  Jin-Kun Lin  Cheng-Hung Lee
Yu-Ming Guo  Yue-Da Lin  Wei-Ming Chen  Yun-Da Tsai  Cheng-Jhih Han  Yi-Jing Lin  Yi-Heng Chiang  Ching-Yuan Chien

myday@mail.tku.edu.tw

NTCIR-12 Conference, June 7-10, 2016, Tokyo, Japan
Outline

• IMTKU Question Answering System Architecture
• IMTKU System Description
• Performance
• Discussions
• Conclusions
Who was the U.S. president in 2012?

Barack Obama
Highlights

• IMTKU (Information Management at TamKang University) Question Answering System for World History Exams at NTCIR-12 QA Lab2

• Phase-1
  — 6 End-to-End QA runs results for English subtask for National Center Test for University Admissions and Secondary exams subtask.

• Phase-3
  — 7 End-to-End QA run results for English and Japanese subtask for Nation Center Exams and Secondary exams subtask.
IMTKU
System Description
Question Analysis

1

Question (XML)

JA & EN Translator

NER & POS Tagger

Question Type Identification

Keyword Extraction

Question Analysis Result (XML)

Complex Essay
Simple Essay
True-or-False
Factoid
Slot-Filling
Unique

JA&EN Translator
Stanford CoreNLP

NTCIR-12 Conference, June 7-10, 2016, Tokyo, Japan
Japanese: 欧陽脩や蘇軾は、唐代を代表する文筆家である。

English (provided by organizer): Ouyang Xiu and Su Shi are writers representative of the Tang period.

English (JA & EN Translator by Google Translate): Ouyang Xiu and Su Shi is a writer representative of the Tang Dynasty.
Raw Data:
Wang Anshi, who lived during the Song period, carried out reforms called the New Policies (xin fa).

POS tagger and NER:
Question Analysis Result

<KEY_TERM_SET LANGUAGE="EN">
  <KEY_TERM RANK="1" SCORE="1">time</KEY_TERM>
  <KEY_TERM RANK="2" SCORE="0.95">Tsuda</KEY_TERM>
  <KEY_TERM RANK="3" SCORE="0.9">Umeko</KEY_TERM>
  <KEY_TERM RANK="4" SCORE="0.85">journey</KEY_TERM>
  <KEY_TERM RANK="5" SCORE="0.8">USA</KEY_TERM>
  <KEY_TERM RANK="6" SCORE="0.75">China</KEY_TERM>
  <KEY_TERM RANK="7" SCORE="0.7">language</KEY_TERM>
  <KEY_TERM RANK="8" SCORE="0.65">schools</KEY_TERM>
  <KEY_TERM RANK="9" SCORE="0.6">students</KEY_TERM>
  <KEY_TERM RANK="10" SCORE="0.55">modernization</KEY_TERM>
  <KEY_TERM RANK="11" SCORE="0.5">policy</KEY_TERM>
  <KEY_TERM RANK="12" SCORE="0.45">foreign</KEY_TERM>
  <KEY_TERM RANK="13" SCORE="0.4">government-funded</KEY_TERM>
  <KEY_TERM RANK="14" SCORE="0.35">Tsuda</KEY_TERM>
  <KEY_TERM RANK="15" SCORE="0.3">Umeko</KEY_TERM>
</KEY_TERM_SET>
Document Retrieval

Question Analysis Result (XML)

Keyword Index Matching

Wikipedia Search

Wikipedia

Preprocessing

Lucene

Wikipedia Knowledge Base

Lucene Search

IR run Result (XML)
Zen in the USA

Spiral Jetty, Utah, USA

Sheng Yen also founded dharma centers in the USA

Numbers of first-language speakers

The following schools are located within the Gloucester Public Schools District:
Answer Extraction

- Question Analysis Result (XML)
- Document Retrieval Result (XML)

1. True-or-False
2. Factoid
3. Slot-Filling
4. Unique

- Sentence Scoring
- Answer Scoring

Answer Extraction Result (XML)
During this time, the Song court retreated south of the Yangtze River and established their capital at Lin'an. Although the Song Dynasty had lost control of the traditional birthplace of Chinese civilization along the Yellow River, the Song economy was not in ruins, as the Southern Song Empire contained 60 percent of China's population and a majority of the most productive agricultural land. Southern Tang was conquered in 976 by the Northern Song Dynasty. Ouyang Xiu and Su Shi are writers representative of the Tang period. Yan Zhenqing is a calligrapher representative of the Song period. Wang Anshi, who lived during the Song period, carried out reforms called the New Policies (xin fa). Qin Hui came into conflict with the party in favor of war, concerning the relationship with the Yuan.
Answer Generation

- Question Analysis Result (XML)
- Document Retrieval Result (XML)
- Answer Extraction Result (XML)

Combination and Matching Strategy

- Complex Essay
- Simple Essay

Answer (XML)
<answer match_type="exact" order="-1" choices="" format_string=""
length_limit="-1">
    <expression_set>
        <expression>
            Jessica Motaung
        </expression>
    </expression_set>
</answer>
Answer Validation

Answer Extraction Result (XML)

Cosine Similarity

Machine Learning (SVM)

Answer (XML)
<combine>
  <sentence>China is the largest trading nation in the world and plays a vital role in international trade, and has increasingly engaged in trade organizations and treaties in recent years</sentence>
  <answer>Song is to manage maritime trade, established a Maritime Affairs.</answer>
  <option>1</option>
  <dict>0.177122977108019</dict>
</combine>

<combine>
  <sentence>China is the largest trading nation in the world and plays a vital role in international trade, and has increasingly engaged in trade organizations and treaties in recent years</sentence>
  <answer>In the Middle Ages of the Champagne region, large-scale regular city is opened.</answer>
  <option>2</option>
  <dict>0.180775381515547</dict>
</combine>

<combine>
# IMTKU Phase-1 Official Runs

<table>
<thead>
<tr>
<th>IMTKU Official Runs</th>
<th>Question Analysis results(QA)</th>
<th>IR run results(RS)</th>
<th>End-to-End QA run results(FA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Center Test</td>
<td>Center-1999--Main-WorldHistoryB_IMTKU_EN_QA_01</td>
<td>Center-1999-Main-WorldHistoryB_IMTKU_EN_RS_01</td>
<td>Center-1999-Main-WorldHistoryB_IMTKU_EN_FA_01</td>
</tr>
<tr>
<td></td>
<td>Center-1999--Main-WorldHistoryB_IMTKU_EN_QA_02</td>
<td>Center-1999-Main-WorldHistoryB_IMTKU_EN_RS_02</td>
<td>Center-1999-Main-WorldHistoryB_IMTKU_EN_FA_02</td>
</tr>
<tr>
<td></td>
<td>Center-1999--Main-WorldHistoryB_IMTKU_EN_QA_03</td>
<td>Center-1999-Main-WorldHistoryB_IMTKU_EN_RS_03</td>
<td>Center-1999-Main-WorldHistoryB_IMTKU_EN_FA_03</td>
</tr>
<tr>
<td></td>
<td>question_format_type_set_def</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>answer_type_set_def</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second Stage Examination</td>
<td>D792W10_IMTKU_EN_QA_01</td>
<td>D792W10_IMTKU_EN_RS_01</td>
<td>D792W10_IMTKU_EN_FA_01</td>
</tr>
<tr>
<td></td>
<td>D792W10_IMTKU_EN_QA_02</td>
<td>D792W10_IMTKU_EN_RS_02</td>
<td>D792W10_IMTKU_EN_FA_02</td>
</tr>
<tr>
<td></td>
<td>D792W10_IMTKU_EN_QA_03</td>
<td>D792W10_IMTKU_EN_RS_03</td>
<td>D792W10_IMTKU_EN_FA_03</td>
</tr>
</tbody>
</table>
## IMTKU at NTCIR-12 QA-Lab2
### Phase-1 Performance

<table>
<thead>
<tr>
<th>RUN</th>
<th>Correct rate</th>
<th>Total score</th>
<th>Average score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.293</td>
<td>31</td>
<td>0.756</td>
</tr>
<tr>
<td>2</td>
<td>0.244</td>
<td>27</td>
<td>0.659</td>
</tr>
<tr>
<td>3</td>
<td>0.000</td>
<td>0</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**IMTKU at result of multiple choice questions**

<table>
<thead>
<tr>
<th>RUN</th>
<th>Correct rate</th>
<th>ROUGE-1 score</th>
<th>ROUGE-2 score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0.032</td>
<td>0.005</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0.008</td>
<td>0.000</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0.032</td>
<td>0.005</td>
</tr>
</tbody>
</table>

**IMTKU at result of free description questions**

NTCIR-12 Conference, June 7-10, 2016, Tokyo, Japan
# IMTKU at NTCIR-12 QA-Lab2
## Phase-1 Performance

<table>
<thead>
<tr>
<th>RUN</th>
<th>CORRECT</th>
<th>INCORRECT</th>
<th>UNANSWERED</th>
<th>Total</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>19</td>
<td>10</td>
<td>12/41</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>21</td>
<td>10</td>
<td>10/41</td>
<td>27</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>41</td>
<td>0/41</td>
<td>0</td>
</tr>
</tbody>
</table>

Correct, Incorrect, Unanswered and score of IMTKU National Center Test 3 runs

<table>
<thead>
<tr>
<th>RUN</th>
<th>CORRECT</th>
<th>INCORRECT</th>
<th>UNANSWERED</th>
<th>WITHHOLDING</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>8</td>
<td>22</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>8</td>
<td>22</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>8</td>
<td>22</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Correct, Incorrect, Unanswered and score of IMTKU Second-stage Examination 3 runs
# IMTKU Phase-3 Official Runs

<table>
<thead>
<tr>
<th>IMTKU Official Runs</th>
<th>Question Analysis results(QA)</th>
<th>IR run results(RS)</th>
<th>End-to-End QA run results(FA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Center Test</td>
<td>Center-2011--Main-WorldHistoryB_IMTKU_EN_QA_01</td>
<td>Center-2011-Main-WorldHistoryB_IMTKU_EN_RS_01</td>
<td>Center-2011-Main-WorldHistoryB_IMTKU_EN_FA_01</td>
</tr>
<tr>
<td></td>
<td>Center-2011--Main-WorldHistoryB_IMTKU_EN_QA_02</td>
<td>Center-2011-Main-WorldHistoryB_IMTKU_EN_RS_02</td>
<td>Center-2011-Main-WorldHistoryB_IMTKU_EN_FA_02</td>
</tr>
<tr>
<td></td>
<td>Center-2011--Main-WorldHistoryB_IMTKU_EN_QA_03</td>
<td>Center-2011-Main-WorldHistoryB_IMTKU_EN_RS_03</td>
<td>Center-2011-Main-WorldHistoryB_IMTKU_EN_FA_03</td>
</tr>
<tr>
<td></td>
<td>Center-2011--Main-WorldHistoryB_IMTKU_JA_QA_01</td>
<td>Center-2011-Main-WorldHistoryB_IMTKU_JA_RS_01</td>
<td>Center-2011-Main-WorldHistoryB_IMTKU_JP_FA_01</td>
</tr>
<tr>
<td></td>
<td>Center-2011--Main-WorldHistoryB_IMTKU_JA_QA_02</td>
<td>Center-2011-Main-WorldHistoryB_IMTKU_JA_RS_02</td>
<td>Center-2011-Main-WorldHistoryB_IMTKU_JP_FA_02</td>
</tr>
<tr>
<td></td>
<td>Center-2011--Main-WorldHistoryB_IMTKU_JA_QA_03</td>
<td>Center-2011-Main-WorldHistoryB_IMTKU_JA_RS_03</td>
<td>Center-2011-Main-WorldHistoryB_IMTKU_JP_FA_03</td>
</tr>
</tbody>
</table>
# IMTKU Phase-3 Official Runs

<table>
<thead>
<tr>
<th>IMTKU Official Runs</th>
<th>IR run results(RS)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Second Stage Examination</strong></td>
<td></td>
</tr>
<tr>
<td>M792W10_IMTKU_EN_RS_01</td>
<td>M792W10_IMTKU_JP_RS_01</td>
</tr>
<tr>
<td>M792W10_IMTKU_EN_RS_02</td>
<td>M792W10_IMTKU_JP_RS_02</td>
</tr>
<tr>
<td>M792W10_IMTKU_EN_RS_03</td>
<td>M792W10_IMTKU_JP_RS_03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IMTKU Official Runs</th>
<th>Combination Run</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>National Center Test</strong></td>
<td>Center-2011-Main-WorldHistoryB_KitAi_IMTKU_EN_FA_01</td>
</tr>
</tbody>
</table>
# IMTKU at NTCIR-12 QA-Lab2
## Phase-3 Performance

<table>
<thead>
<tr>
<th>Run</th>
<th>Lang.</th>
<th>Correct rate</th>
<th>Total score</th>
<th>Average score</th>
</tr>
</thead>
<tbody>
<tr>
<td>RUN01</td>
<td>EN</td>
<td>0.194</td>
<td>20</td>
<td>0.556</td>
</tr>
<tr>
<td>RUN02</td>
<td>EN</td>
<td>0.194</td>
<td>20</td>
<td>0.556</td>
</tr>
<tr>
<td>RUN03</td>
<td>EN</td>
<td>0.139</td>
<td>14</td>
<td>0.389</td>
</tr>
<tr>
<td>RUN01</td>
<td>JA</td>
<td>0.222</td>
<td>24</td>
<td>0.667</td>
</tr>
<tr>
<td>RUN02</td>
<td>JA</td>
<td>0.083</td>
<td>8</td>
<td>0.222</td>
</tr>
<tr>
<td>RUN03</td>
<td>JA</td>
<td>0.222</td>
<td>24</td>
<td>0.667</td>
</tr>
</tbody>
</table>

IMTKU at result of multiple choice questions in Phase-3
# IMTKU at NTCIR-12 QA-Lab2

## Phase-3 Performance

<table>
<thead>
<tr>
<th>EN</th>
<th>CORRECT</th>
<th>INCORRECT</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RUN01</td>
<td>7</td>
<td>29</td>
<td>20</td>
</tr>
<tr>
<td>RUN02</td>
<td>7</td>
<td>29</td>
<td>20</td>
</tr>
<tr>
<td>RUN03</td>
<td>5</td>
<td>31</td>
<td>14</td>
</tr>
</tbody>
</table>

Correct, Incorrect and score of IMTKU National Center Test English 3 run

<table>
<thead>
<tr>
<th>JA</th>
<th>CORRECT</th>
<th>INCORRECT</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RUN01</td>
<td>4</td>
<td>32</td>
<td>12</td>
</tr>
<tr>
<td>RUN02</td>
<td>3</td>
<td>33</td>
<td>8</td>
</tr>
<tr>
<td>RUN03</td>
<td>8</td>
<td>28</td>
<td>24</td>
</tr>
</tbody>
</table>

Correct, Incorrect and score of IMTKU National Center Test Japanese 3 runs
IMTKU at NTCIR-12 QA-Lab2
Phase-3 Performance

<table>
<thead>
<tr>
<th>TeamID</th>
<th>Comb</th>
<th>Correct rate</th>
<th>Total score</th>
<th>Average score</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMTKU</td>
<td>KitAi</td>
<td>0.333</td>
<td>34</td>
<td>0.944</td>
</tr>
</tbody>
</table>

IMTKU at result of multiple choice questions for combination run in Phase-3

<table>
<thead>
<tr>
<th>Combination</th>
<th>CORRECT</th>
<th>INCORRECT</th>
<th>Total</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RUN with KitAi</td>
<td>12</td>
<td>24</td>
<td>12/36</td>
<td>34</td>
</tr>
</tbody>
</table>

Correct, Incorrect and score of IMTKU with KitAi National Center Test English Combination run.
Discussions

• Systematic Error of XML output (IMTKU FA run03 on Phase-1)

• Limitations of the IMTKU System

o(symbol-symbol-symbol):
(1) a→b→c (2) a→c→b (3) b→a→c (4) b→c→a

symbol-symbol:
(1)(A) → a (2)(A) → b (3)(B) → a (4)(B) → b

(symbol-TF)*2
(1)a- Correct b- Correct (2)a- Correct b- Incorrect
(3)a- Incorrect b- Correct (4)a- Incorrect b- Incorrect

(symbol-term_other)*2:
(1)(A)- Huguang(B)- Gongsuo (guild)
(2)(A)- Huguang(B)- Zujie (concession)
(3)(A)- Suhu (Jiangzhe)(B)- Gongsuo (guild)
(4)(A)- Suhu (Jiangzhe)(B)- Zujie (concession)
Conclusions

- IMTKU Question Answering System for World History Exams at NTCIR-12 QA Lab2
  - Phase-1
    - 6 End-to-End QA runs results for English subtask for National Center Test for University Admissions and Secondary exams subtask.
  - Phase-3
    - 7 End-to-End QA run results for English and Japanese subtask for Nation Center Exams and Secondary exams subtask.
- IMTKU with KitAi for National Center Test English Combination run
IMTKU Question Answering System for World History Exams at NTCIR-12 QA Lab2

Department of Information Management
Tamkang University, Taiwan

Min-Yuh Day
Cheng-Chia Tsai
Wei-Chun Chung
Hsiu-Yuan Chang
Tzu-Jui Sun
Yuan-Jie Tsai
Jin-Kun Lin
Cheng-Hung Lee

Yu-Ming Guo
Yue-Da Lin
Wei-Ming Chen
Yun-Da Tsai
Cheng-Jhih Han
Yi-Jing Lin
Yi-Heng Chiang
Ching-Yuan Chien

myday@mail.tku.edu.tw

NTCIR-12 Conference, June 7-10, 2016, Tokyo, Japan