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
Outline

• IMTKU RITE System Architecture
• Highlight of Features
  – Syntactic and Semantic Features
• Performance and Discussion
• Demo of RITE.IM.TKU
• Call for participating in IEEE EM-RITE 2013, 2014, 2015, ...
Overview

• IMTKU (Information Management at TamKang University) textual entailment system for recognizing inference in text at NTCIR-10 RITE-2 (Recognizing Inference in Text)

• Hybrid approach
  – integrate semantic features and machine learning techniques for recognizing inference in text at NTCIR-10 RITE-2 task

• We submitted 3 official runs for BC, MC and RITE4QA subtask

• IMTKU is ranked #1 in the CS-RITE4QA subtask of NTCIR-10 RITE-2 task
IMTKU System Architecture for NTCIR-10 RITE-2

XML Train Dataset of RITE Corpus (T1, T2 Pairs)

- Preprocessing
- Feature Generation
- Feature Selection
- Training Model (SVM Model)
- Evaluation of Model (k-fold CV)

XML Test Dataset of RITE Corpus (T1, T2 Pairs)

- Preprocessing
- Feature Generation
- Feature Selection
- Use model for Prediction
- Predict Result (Open Test)

CKIP AutoTag (POS Tagger)
HIT TongYiCiLing
WordNet
Dependency Parser
Negation Antonym

NTCIR-10 Conference, June 18-21, 2013, Tokyo, Japan
IMTKU System Architecture for NTCIR-9 RITE

Preprocessing

RITE Corpus (T1, T2 Pairs)

CKIP AutoTag (POS Tagger)

HIT Dependency Parser

SINICA BOW

HIT TongYiCiLing

Chinese Antonym

Feature Extraction

Machine Learning Module

Knowledge-Based Module

Voting Strategy Module

Predict Result (BC)/(MC)

Similarity Evaluation
Semantic and Syntactic features

1. String Length/Length Difference/Ratio
2. Longest Common Substring
3. Char-based Edit Distance
4. Word Length/Difference/Ratio
5. Word-based Edit Distance
6. Noun/Verb Number
7. Word Semantic (Synonym) Similarity
8. WordNet Similarity
9. Negation
10. Antonym
11. Dependency Parser
Recognizing Inference in Text (RITE)

T1: 香港的主權和領土是在1997年由英國歸還給中國的
(T1: Hong Kong's sovereignty and territories were returned to China by the United Kingdom in 1997.)

T2: 1997年香港回歸中國
(T2: Hong Kong was returned to China in 1997.)

BC  YES
Recognizing Inference in Text (RITE)

T1: 車諾比病毒在1999年4月總共造成超過200萬台電腦無法開機
(T1: CIH caused severe boot problems in more than 200 million computers in April, 1999)

T2: 1999年4月車諾比病毒總共造成逾200萬台電腦無法開機
(T2: CIH caused severe boot problems in over 200 million computers in April, 1999)
Recognizing Inference in Text (RITE)

T1: 車諾比病毒在1999年4月總共造成超過200萬台電腦無法開機

(T1: CIH caused severe boot problems in more than 200 million computers in April, 1999)

T2: 1999年4月車諾比病毒總共造成逾200萬台電腦無法開機

(T2: CIH caused severe boot problems in over 200 million computers in April, 1999)
Word Semantic (Synonym) Similarity

HIT TYCCL

• 19 synonyms of “World” (世界)
  – Di01A01=世界, 世, 世上, 大地, 天下, 天底下, 全世界, 環球, 全球, 舉世, 中外, 寰宇, 五洲, 海內, 海內外, 五湖四海, 大千世界, 大世界, 普天之下
  – TYCCL Scoring Function:
    • $((t-r) + 1) / t$
    • $((19-1)+1)/19 = 19/19 = 1$
  – World 世界
    • Di01A01=世界:1.0000,
      Di14C04=世風:0.5000,
      Dd05B03=領域:0.3333
Negation

• **52 Chinese negation words list**
• Examples:
  - 沒
  - 不
  - 否
  - 無
  - 非
  - 未
  - 免
  - 別
  - 莫
  - ...
  - 沒有
  - 無法
  - 尚未
  - 未可
  - 未得
  - 未必
  - 未聞
  - 未有
  - 未定
  - ...

NTCIR-10 Conference, June 18-21, 2013, Tokyo, Japan
Antonym

• 568 antonym pair list
• Examples:
  - 開心 苦悶
  - 開心 傷心
  - 開心 難過
  - 快樂 難過
  - 快樂 傷心
  - 高興 傷心
  - 高興 難過
  - 高興 痛苦
  - 幸福 痛苦
  - 高興 憤怒
  - 高興 掃興
  - ...

NTCIR-10 Conference, June 18-21, 2013, Tokyo, Japan
# IMTKU BC Subtask
## Official Runs

<table>
<thead>
<tr>
<th>IMTKU Subtask Official Runs</th>
<th>Resources</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>RITE-2-IMTKU-CT-BC-01</td>
<td>Bilingual Wordnet, HIT TongYiCiLing, Stanford Parser</td>
<td>Antonym, Negation, Word Based Similarity, Token Based Similarity, Lexical overlap, Text Pair Length, Token Length, WorkNet Similarity, Tree Edit Distance</td>
</tr>
<tr>
<td>RITE-2-IMTKU-CS-BC-01</td>
<td>Bilingual Wordnet, HIT TongYiCiLing</td>
<td>Antonym, Negation, Word Based Similarity, Token Based Similarity, Lexical overlap, Text Pair Length, Token Length, WorkNet Similarity</td>
</tr>
<tr>
<td>RITE-2-IMTKU-CT-BC-02</td>
<td>Bilingual Wordnet, HIT TongYiCiLing</td>
<td>Antonym, Negation, Word Based Similarity, Token Based Similarity, Lexical overlap, Text Pair Length, Token Length, WorkNet Similarity</td>
</tr>
<tr>
<td>RITE-2-IMTKU-CS-BC-02</td>
<td>Stanford Parser</td>
<td>All syntactic and semantic features (except Stanford Parser)</td>
</tr>
</tbody>
</table>
IMTKU at NTCIR-10 RITE-2 Task Performance

CT-BC Subtask Accuracy

RITE-2-IMTKU-CT-BC-01

Bilingual Wordnet, HIT TongYiCiLing, Stanford Parser

Antonym, Negation, Word Based Similarity, Token Based Similarity, Lexical overlap, Text Pair Length, Token Length, WorkNet Similarity, Tree Edit Distance

Accuracy: 0.663
# IMTKU MC Subtask
## Official Runs

<table>
<thead>
<tr>
<th>IMTKU Subtask Official Runs</th>
<th>Resources</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>RITE-2-IMTKU-CT-MC-01</td>
<td>Stanford Parser</td>
<td>Longest Common Substring, Word Length Ratio, Text Length, Similarity between t1 and t2, Tree Edit Distance</td>
</tr>
<tr>
<td>RITE-2-IMTKU-CS-MC-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RITE-2-IMTKU-CT-MC-02</td>
<td>Bilingual Wordnet, HIT TongYiCiLing, Stanford Parser</td>
<td>Integrated Semantic features and Machine Learning Approach</td>
</tr>
<tr>
<td>RITE-2-IMTKU-CS-MC-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RITE-2-IMTKU-CT-MC-03</td>
<td>Bilingual Wordnet, HIT TongYiCiLing</td>
<td>Longest Common Substring, Word Length Ratio, Text Length, Similarity between t1 and t2, Tree Edit Distance</td>
</tr>
<tr>
<td>RITE-2-IMTKU-CS-MC-03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
IMTKU at NTCIR-10 RITE-2 Task Performance

CT-MC Subtask Accuracy

RITE-2-IMTKU-CT-MC-01

Stanford Parser

Longest Common Substring, Word Length Ratio, Text Length, Similarity between t1 and t2, Tree Edit Distance
## IMTKU RITE4QA Subtask Official Runs

<table>
<thead>
<tr>
<th>IMTKU Subtask Official Runs</th>
<th>Resources</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>RITE-2-IMTKU-CT-RITE4QA-01</td>
<td>Stanford Parser</td>
<td>Antonym, Negation, Word Based Similarity, Token Based Similarity, Lexical overlap, Text Pair Length, Token Length, WorkNet Similarity</td>
</tr>
<tr>
<td>RITE-2-IMTKU-CS-RITE4QA-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RITE-2-IMTKU-CT-RITE4QA-02</td>
<td>Bilingual Wordnet, HIT TongYiCiLing</td>
<td>Antonym, Negation, Word Based Similarity, Token Based Similarity, Lexical overlap, Text Pair Length, Token Length</td>
</tr>
<tr>
<td>RITE-2-IMTKU-CS-RITE4QA-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RITE-2-IMTKU-CT-RITE4QA-03</td>
<td>HIT TongYiCiLing</td>
<td>Longest Common Substring, Text Length, Text Length Ratio, Antonym, Negation</td>
</tr>
<tr>
<td>RITE-2-IMTKU-CS-RITE4QA-03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
IMTKU at NTCIR-10 RITE-2 Task Performance

CS-RITE4QA Subtask
MRR

RITE-2-IMTKU-CS-RITE4QA-03

HIT TongYiCiLing

Longest Common Substring,
Text Length, Text Length Ratio,
Antonym, Negation
NTCIR-10 RITE-2 CT-RITE4QA

NTCIR-10 Conference, June 18-21, 2013, Tokyo, Japan
IMTKU at NTCIR-10 RITE-2 Task Performance

CT-RITE4QA Subtask
MRR

RITE-2-IMTKU-CT-RITE4QA-03

HIT TongYiCiLing

Longest Common Substring, Text Length, Text Length Ratio, Antonym, Negation
Discussions

• Issues of Definition in RITE MC between NTCIR-9 and NTCIR-10:
  – Definition of NTCIR-9 MC subtask :
    • “A 5-way labeling subtask to detect (forward / reverse / bidirection) entailment or no entailment (contradiction / independence) in a text pair.”
  – Definition of NTCIR-10 MC subtask :
    • “A 4-way labeling subtask to detect (forward / bidirection) entailment or no entailment (contradiction / independence) in a text pair.”
  – Misused NTCIR-9 MC labels on NTCIR-10 MC test datasets where “Reverse” label should be excluded.
## IMTKU Experiments for NTCIR-10 RITE-2 Datasets

<table>
<thead>
<tr>
<th>Datasets</th>
<th>10 Fold CV Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>RITE2_CT_dev_test_bc_g.txt</td>
<td>68.85%</td>
</tr>
<tr>
<td>(RITE2 BC Dev + Test Dataset: 1321 + 881 = 2202 pairs)</td>
<td></td>
</tr>
<tr>
<td>RITE1_CT_r1000_dev_test_bc_g.txt</td>
<td>73.83%</td>
</tr>
<tr>
<td>(Random select 1000 pairs from RITE1 BC Dev+ Test Dataset)</td>
<td></td>
</tr>
<tr>
<td>RITE1_CT_dev_test_bc_g.txt</td>
<td>72.29%</td>
</tr>
<tr>
<td>(RITE1 BC Dev + Test Dataset: 421 + 900 = 1321 pairs)</td>
<td></td>
</tr>
<tr>
<td>RITE1_CT_dev_bc_g.txt (gold standard)</td>
<td>72.21%</td>
</tr>
<tr>
<td>(RITE1 BC Development Dataset: 421 pairs)</td>
<td></td>
</tr>
</tbody>
</table>
**IMTKU Experiments for NTCIR-9 RITE Datasets**

<table>
<thead>
<tr>
<th>Datasets</th>
<th>10 Fold CV Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>RITE1_CT_dev_bc_g.txt (gold standard) (BC Development Dataset: 421 pairs)</td>
<td>76.48%</td>
</tr>
<tr>
<td>RITE1_CT_test_bc_g.txt (BC Test Dataset: 900 pairs)</td>
<td>66.33%</td>
</tr>
<tr>
<td>RITE1_CT_dev_test_bc_g.txt (BC Dev+Test Dataset: 421+900 = 1321 pairs)</td>
<td>67.67%</td>
</tr>
</tbody>
</table>
IMTKU Textual Entailment System for Recognizing Inference in Text at NTCIR-10 RITE-2

Demo

http://rite.im.tku.edu.tw

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

myday@mail.tku.edu.tw

2013/06/19
NTCIR-10 Conference, June 18-21, 2013, Tokyo, Japan
IMTKU Textual Entailment System
Department of Information Management, Tamkang University

Demo

- Example:  ● Sample 1  ● Sample 2  ● Sample 3

- Textual (T1):  一九九七年香港回歸中國

- Hypothesis (T2):  香港的主權和領土是在一九九七由英國歸還給中國的。

Result:  No 0.653509

Detail:

<table>
<thead>
<tr>
<th>Word Similarity</th>
<th>0.8747</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Net Similarity</td>
<td>18.55</td>
</tr>
<tr>
<td>Word Net Similarity Ratio</td>
<td>23.08333333333333</td>
</tr>
<tr>
<td>Word Net Similarity Short</td>
<td>30.33333333333333</td>
</tr>
<tr>
<td>Negation Number Difference</td>
<td>0</td>
</tr>
<tr>
<td>Antonym Number Difference</td>
<td>0</td>
</tr>
<tr>
<td>T1: 一九九七年香港回歸中國</td>
<td></td>
</tr>
<tr>
<td>T2: 香港的主權和領土是在一九九七由英國歸還給中國的。</td>
<td></td>
</tr>
<tr>
<td>T1 CKIP:  ?(QUESTIONCATEGORY) 一九九七年 (N) 香港 (N) 回歸 (Vr) 中國 (N)</td>
<td></td>
</tr>
</tbody>
</table>

http://rite.im.tku.edu.tw

In conjunction with IEEE IRI 2013

San Francisco, USA
August 14, 2013

https://sites.google.com/site/emrite2013/
Textual Entailment and Paraphrase are inference tasks of natural language processing (NLP) for automatically detecting entailment, paraphrase, and contradiction in texts. The aim of this workshop is to provide a forum for original high-quality research contributions on empirical methods for recognizing inference in text as well as multidisciplinary research opportunities.

Topics of interest include but are not limited to practical areas that span a variety of aspects of empirical methods for recognizing inference in text including:

- Guidelines, standards, best practices and models for the construction and annotation of Textual Entailment datasets
- Evaluation of Knowledge Resources for Textual Entailment
- Recognizing Inference in Text
- Recognizing Textual Entailment
Conclusions

• Issues of definitions and datasets at NTCIR9-RITE and NTCIR-10 RITE-2

• Online demo system RITE.IM.TKU
  – http://rite.im.tku.edu.tw

• Welcome to join IEEE EM-RITE 2013, 2014, 2015, ...
Q & A

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