Information Extraction based Approach for the NTCIR-9 1CLICK Task
Makoto P. Kato  Meng Zhao  Kosetsu Tsukuda  Yoshiyuki Shoji
Takehiro Yamamoto  Hiroaki Ohshima  Katsumi Tanaka
Graduate School of Informatics, Kyoto University

Methods

**Query Classification**

Multi class SVM

<table>
<thead>
<tr>
<th>Feature</th>
<th># of features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has Wikipedia article</td>
<td>1</td>
</tr>
<tr>
<td>Frequency of parts-of-speech</td>
<td>44</td>
</tr>
<tr>
<td>Query unigram</td>
<td>85</td>
</tr>
<tr>
<td>Sentence pattern</td>
<td>2</td>
</tr>
<tr>
<td>Number of documents containing expanded query</td>
<td>15</td>
</tr>
<tr>
<td>Has travel services</td>
<td>1</td>
</tr>
<tr>
<td>Number of search results</td>
<td>1</td>
</tr>
<tr>
<td>Terms in search results</td>
<td>39</td>
</tr>
<tr>
<td>Total</td>
<td>185</td>
</tr>
</tbody>
</table>

**CELEBRITY**

Attribute-based

- **Query**: Kyoto Royal Hotel
  - www.royalhotel.com
  - address OR access OR ...

Sentence-based

- Hayao Miyazaki
  - ContainsQuery

<table>
<thead>
<tr>
<th>Attribute name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>USA</td>
</tr>
<tr>
<td>Residence</td>
<td>Las Vegas, Nevada, U.S.</td>
</tr>
</tbody>
</table>

**LOCATION**

Check in: 15:00
Check out: 10:00

**DEFINITION**

**QA**

Yahoo! Chiebukuro

S = \{s_1, s_2, s_3, ..., s_n\}

s_i: hospitality towa ...

LexRank

S_n

\[ BA_1 \]

\[ BA_2 \]

\[ BA_3 \]

Q_1

Q_2

Q_3

Information Summarizer

Summarized Text

Results

- The object identification problem should be tackled (e.g. Kanazawa University)
- QA queries achieved high score only when exactly the same questions are available

Where is Yumura-onsen?

Which is taller, Tsutenkaku or Utsunomiya tower?

Kanazawa University