Using Summary techniques to Answer Definition Questions

CS5244 Course Project

JIA Yanjun
Motivation: Improve information seeking accuracy in Digital Library

IR (keyword->Documents) vs. QA (Question->Answer)

Definition questions: For example, “Who is Fred Durst?”
The response is a list of sentences.

Goal: The answer should be exhaustive while maintaining the cohesion and eliminating redundancy.

Summary: a set of sentences extracted verbatim from a text which cover major substance of that text

Goal: include as much relevant information as possible for which the user is looking and excludes extraneous and redundant information.
Introduction (cont.)

- In the limit question answering and summarization may merge as research areas.
- Concern: Redundancy and Diversity
- Redundancy -> MMR
- Diversity -> Clustering
System Architecture

Questions

Document Retrieval

Sentence Retrieval

Sentence Clustering

Sentence Ranking

Sentence Selection

Pattern Matching

Density-based weight

Definition
Approach

- K-means Clustering Algorithm
- Pattern Matching: Hybrid approach
- Density-based Ranking: SiteQ+
- MMR Summary Generation
Evaluation

- Recall = number of correct concepts retrieved / number of concepts in the ground truth list.
- Precision = 1 - (Length-Allowance)/Length
- An allowance of 200 non-white-space characters is given for each correct concept retrieved.

\[ F = \frac{(\beta^2 + 1)RP}{\beta^2 P + R} \]

- \( \beta \) is set to 5 to reflect the emphasis gave to recall and adjust for the crudeness of the length approximation to true precision.
Evaluation (cont.)

- **SETUP**
  - A corpus consisting of 160 documents taken from TREC corpus
  - 16 questions: selected from those in TREC QA Track 2004+Manually Construct
- **RESULT**

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Cluster</th>
<th>MMR</th>
<th>Cluster+MMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>0.313</td>
<td>0.314</td>
<td>0.360</td>
<td>0.379</td>
</tr>
<tr>
<td>R</td>
<td>0.304</td>
<td>0.325</td>
<td>0.358</td>
<td>0.361</td>
</tr>
<tr>
<td>F5</td>
<td>0.305</td>
<td>0.327(+7.2%)</td>
<td>0.367(+20.3%)</td>
<td>0.388 (+27.2%)</td>
</tr>
</tbody>
</table>
Conclusion

• By measuring the answer extraction performance in isolation, the preliminary experiment shows that summary techniques are very useful in definitional question answering system.

• More extensive experiments and detailed error analysis is needed.

• Explore more summary techniques in definitional QA systems to produce a coherent natural language definition.