Overview of the NTCIR-10 Cross-lingual Link Discovery Task

Ling-Xiang Tang†, In-Su Kang‡, Fuminori Kimura*, Yi-Hsun Lee*, Andrew Trottman*, Shlomo Geva*, Yue Xu†
†Queensland University of Technology ‡Kyungsun University  *Ritsumeikan University  †Academia Sinica  ‡University of Otago

What’s New

New Tasks

New Participants

New Challenges

GROUP  AFFILIATION
DCU  Dublin City University
III  Institute for Information Industry
KECIR  Shenyang Aerospace University
KMI  The Open University
KSLP  Kyungsung University
NTHU  National Tsing Hua University
OKSAT  Osaka Kyioiku University
QUT  Queensland University of Technology
RDLL  Ritsumeikan University
UKP  TU Darmstadt

New Collections

<table>
<thead>
<tr>
<th>LANG</th>
<th>#DOC</th>
<th>SIZE</th>
<th>DATE OF DUMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td>404,620</td>
<td>3.6GB</td>
<td>11/01/2012</td>
</tr>
<tr>
<td>English</td>
<td>3,381,772</td>
<td>33.0GB</td>
<td>04/01/2012</td>
</tr>
<tr>
<td>Japanese</td>
<td>858,610</td>
<td>9.8GB</td>
<td>04/01/2012</td>
</tr>
<tr>
<td>Korean</td>
<td>297,913</td>
<td>2.2GB</td>
<td>22/01/2012</td>
</tr>
</tbody>
</table>

Evaluation

Precision$_{1F} = \frac{\text{number of correct links}}{\text{number of identified links}}$

Recall$_{1F} = \frac{\text{number of correct links}}{\text{number of links in query}}$

A2F Evaluation

Precision$_1$ and Recall$_1$

Precision$_K$ and Recall$_K$

System Evaluation Metrics

LMAP = $\frac{\sum_{t=1}^{T} \sum_{i=1}^{M} \phi(t \in i \cap K)}{\sum_{t=1}^{T} \sum_{i=1}^{M} \phi(t \in i)}$

where $n$ is the number of identified items (articles for F2F or anchors in A2F), and $m$ is the number of topics (source articles used in evaluation); $K$ is the number of topics for item $t$. $R - Prec = \frac{\sum_{t=1}^{T} \sum_{i=1}^{M} \phi(t \in i \cap R)}{\sum_{t=1}^{T} \sum_{i=1}^{M} \phi(t \in i)}$

where $n$ is the number of topics, and $P_i \cap R$ is the precision at $R$ where $R$ is the number of unique items in the $q$ of topic $t$. Precision-at-$N$ is computed using the average precision for all topics (source articles) at a pre-defined position $N$ in the results list. Values of $N$ were chosen as: 5, 10, 20, 30, 50, and 250.

Evaluation Framework

Conclusions

Many good approaches were seen in the CJK to English cross-lingual link discovery tasks.

The evaluation methods distinguish the effective and less effective CLLD algorithms.

Where is still work to be done: the PR curves are still below 0.35

Segmentation seems to help, so team KECIR and team KSLP achieved good A2F evaluation scores with manual assessment in the CJE and KJE tasks separately.

The top performing teams include KMI, OKSAT who employed a unified cross-lingual linking method achieved very good results in different language subtasks even different link directions.

What’s Next?

PubMeb Patent

CLLD Personalisation

What’s Next?