**SUMMARY**

Recognizing Textual Entailment is an important and hard basic research that can be applicable to many research fields, e.g. Question Answering, Text Summarization, Information Retrieval and Information Extraction. We, LTI team, developed an adaptable feature-based approach that combines multiple complementing features motivated by our analysis on data as well as linguistic insights. We evaluated our system in the Japanese tracks of the BC, Entrance Exam and RTE4QA subtasks at NTCIR-9 RITE, which marked relatively good scores among participants.

**Task Definition (binary-class)**

Input: sentence $t_1$ and $t_2$

Output: binary label indicating whether the premise $t_1$ entails the hypothesis $t_2$ (meaning that if a human reading $t_1$ would infer that $t_2$ is most likely true).

**APPRAOCH**

**Overview**

Given the analysis result, we decided to take a supervised machine learning approach with a careful design on features motivated by linguistics. The classification models we chose are SVM with the linear kernel (in the BC subtasks) and MaxEnt (in the Entrance Exam and RTE4QA subtasks).

**Resources**

- **WordNet Synonymy**
- **WordNet Semantic Relatedness** – We developed and released a tool that can calculate semantic similarity of two words, using NICT’s Japanese WordNet.
- **Wikipedia Hyponymy** – hypernym-hyponym resource automatically extracted from Wikipedia using NICT’s Hyponymy extraction tool.
- **Wikipedia Redirect** – We developed and released a tool which can extract and utilizes Wikipedia’s redirect information to be used for solving alternative forms of the same concept.

**Features**

**Morpheme overlap** (through continuous-to-categorical conversion); **Basic Element** (Hovy et al., 2006; Fukumoto, 2007) soft-structural overlap; **polarity** (fires when a mismatch of sentiment polarity is captured between $t_1$ and $t_2$); **Morpeme Diff** (takes a diff between $t_1$ and $t_2$), and makes an entailment recognition decision on the different morphemes using character-level heuristic soft-matching; **Quote** – this feature is an N-Label indicator, which fires when a quoted content in $t_1$ occurs in $t_2$. Intuition: what’s written and what’s said (or reported in quotation) have different likelihood of being true.

**Quantification** – fires when there is a mismatch in quantification expression which also indicates an N-label. The quantifier cues: 限ってonly; 常にalways etc.

**Which Features are Useful?**

All-but-one feature ablation study

<table>
<thead>
<tr>
<th>Feature</th>
<th>BC</th>
<th>EXAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>All features</td>
<td>62.1%</td>
<td>N/A</td>
</tr>
<tr>
<td>- Morpheme Overlap</td>
<td>61.0%</td>
<td>59.1%</td>
</tr>
<tr>
<td>- BE Overlap</td>
<td>54.2%</td>
<td>68.8%</td>
</tr>
<tr>
<td>- Quote</td>
<td>61.3%</td>
<td>67.7%</td>
</tr>
<tr>
<td>- Polarity</td>
<td>55.9%</td>
<td>-0.2%</td>
</tr>
<tr>
<td>- Quantification</td>
<td>62.5%</td>
<td>68.9%</td>
</tr>
<tr>
<td>- Morpheme Diff</td>
<td>57.4%</td>
<td>-0.2%</td>
</tr>
</tbody>
</table>

**RESULTS**

Run 01: BE only baseline.
Run 02: BE + char-overlap baseline.
Run 03: Proposed approach.

* 5-fold cross-validation

**CONCLUSION**

We presented the LTI’s system participated in NTCIR-9 RITE. Through analysis, we assumed that multiple linguistic phenomena must be captured, and there is a need of adaptability in a Textual Entailment recognition system. We experimentally showed that they are reasonable assumptions to make. Tools we built, i.e. WS4J1 and Wikipedia Redirect2, are released as open source software.

1 http://code.google.com/p/w4j/
2 http://code.google.com/p/wikipedia-redirec/

Future work: elaborate more on capturing linguistic modalities. Especially, recognizing epistemic modality, or committed and non-committed belief could be a sophisticated extension of the Quote feature we used.