**Abstract**

We introduce a system description of Toshiba Machine Translation System for WAT2014. We participated in two tasks, namely Japanese-English (JE) translation and Japanese-Chinese (JC) translation. In each task, we submitted two results; one is a result of a rule-based translation system, and the other is a result which is an output of statistical post editing trained with the ASPEC training corpora. In both tasks, output by statistical post editing shows improvement in machine evaluation, but we obtained different results from human evaluation.

**Toshiba Machine Translation System**

We have been developing a Rule-Based Machine Translation (RBMT) system. The core functions can realize both high performance and flexibility of customization by using a large volume of dictionaries (rules) including translation knowledge. Although Statistical-based Machine Translation (SMT) has practical translation performance in the target domain, it is extremely high cost to develop parallel corpora in wide-domain for commercial use.

**Purpose.** We applied statistical approach to RBMT system in order to improve its performance and analyze its availability.

**System:** A combination of RBMT and Statistical Post Editing (SPE) with selected technical term dictionaries.

**Selecting Technical Term Dictionaries**

- **Available Technical Term Dictionaries**
  - JE: 32 dictionaries, 2M words
  - JC: 14 dictionaries, 0.46M words

- **Selected Technical Term Dictionaries**
  - JE: 4 dictionaries, 0.5M words (Chemistry, Medicine 1&2, Science)
  - JC: 2 categories, 48K words (Chemistry, Biology)

**Context-aware Machine Translation**

Our RBMT system has following functions for context-aware translation. But we cannot confirm effectiveness of these function in WAT2014 task.

**Results**

<table>
<thead>
<tr>
<th>System</th>
<th>BLUE</th>
<th>RIES</th>
<th>HUMAN</th>
<th>RANK</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBMT</td>
<td>15.69</td>
<td>0.69</td>
<td>20.2%</td>
<td>1.38</td>
</tr>
<tr>
<td>+SPE</td>
<td>20.61</td>
<td>0.71</td>
<td>23.2%</td>
<td>1.45</td>
</tr>
<tr>
<td>SMT</td>
<td>18.45</td>
<td>0.65</td>
<td>-</td>
<td>2.24</td>
</tr>
</tbody>
</table>

SPE achieved improvements of 31.4% for Japanese-English and 42.2% for Japanese-Chinese in BLUE. On the other hand, in RANK of Japanese-English, RBMT showed better translation than SPE. We found improvements and worse related to the difference as follows:

- **Improvements**
  - Some phrases including noun, verb and adjective are post-edited to better phrases. Because a number of vocabulary in Chinese dictionaries are significantly smaller than that of English dictionaries, the effect of SPE for Chinese may be large.

- **Mistreatment**
  - However, deletion, that translated phrases by RBMT get worse by post-editing (e.g., "interface mold" -> "mold"), have high proportion of failures of SPE. Furthermore, tense disagreement (e.g., RBMT generated past form but SPE modified present form) and number disagreement occurred.

**Alternative approach using a monolingual corpus**

We have been used statistical approach in our commercial systems. It is effective to apply monolingual corpus to word selection in the specific domain. However, we did not use these approaches in the WAT2014 because ASPEC includes multi-domain contents.