### Results Overview

<table>
<thead>
<tr>
<th>Subtask</th>
<th>System</th>
<th>BLEU</th>
<th>(\Delta\text{baseline})</th>
<th>(\Delta\text{competitor})</th>
<th>RIBES</th>
<th>(\Delta\text{baseline})</th>
<th>(\Delta\text{competitor})</th>
<th>Adequacy</th>
<th>(\Delta\text{baseline})</th>
<th>(\Delta\text{competitor})</th>
<th>Accept</th>
<th>(\Delta\text{baseline})</th>
<th>(\Delta\text{competitor})</th>
</tr>
</thead>
<tbody>
<tr>
<td>EnJa</td>
<td>Primary</td>
<td>39.48</td>
<td></td>
<td>+7.58(5BM)</td>
<td>78.13</td>
<td>+6.13(5BM)</td>
<td>+9.08(5BM)</td>
<td>3.67</td>
<td>+1.07(5BM)</td>
<td>+0.16(5BM)</td>
<td>0.69</td>
<td>+0.22(5BM)</td>
<td>+0.03(5BM)</td>
</tr>
<tr>
<td>EnJa</td>
<td>Best Single</td>
<td>38.81</td>
<td>+6.91(5BM)</td>
<td>+4.78(5BM)</td>
<td>77.82</td>
<td>+5.82(5BM)</td>
<td>+8.77(5BM)</td>
<td>2.75</td>
<td>+0.13(5BM)</td>
<td>-0.92(5BM)</td>
<td>0.49</td>
<td>+0.02(5BM)</td>
<td>-0.22(5BM)</td>
</tr>
<tr>
<td>JaEn</td>
<td>Primary</td>
<td>28.35</td>
<td>-0.60(5BM)</td>
<td>-3.34(5BM)</td>
<td>71.95</td>
<td>+1.31(5BM)</td>
<td>-2.05(5BM)</td>
<td>2.58</td>
<td>+0.13(5BM)</td>
<td>-0.92(5BM)</td>
<td>0.49</td>
<td>+0.02(5BM)</td>
<td>-0.22(5BM)</td>
</tr>
<tr>
<td>JaEn</td>
<td>Best Single</td>
<td>27.78</td>
<td>-1.17(5BM)</td>
<td>-3.91(5BM)</td>
<td>67.32</td>
<td>-3.32(5BM)</td>
<td>-6.72(5BM)</td>
<td>2.02</td>
<td>+0.13(5BM)</td>
<td>-0.92(5BM)</td>
<td>0.49</td>
<td>+0.02(5BM)</td>
<td>-0.22(5BM)</td>
</tr>
<tr>
<td>ZhEn</td>
<td>Primary</td>
<td>30.26</td>
<td>-0.46(5BM)</td>
<td>-9.28(G1-1)</td>
<td>76.47</td>
<td>-0.74(5BM)</td>
<td>-6.80(G1-1)</td>
<td>3.23</td>
<td>-0.06(5BM)</td>
<td>-0.80(G1-1)</td>
<td>0.49</td>
<td>+0.02(5BM)</td>
<td>-0.22(G1-1)</td>
</tr>
<tr>
<td>ZhEn</td>
<td>Best Single</td>
<td>30.74</td>
<td>+0.02(5BM)</td>
<td>-8.70(G1-1)</td>
<td>76.62</td>
<td>-0.59(5BM)</td>
<td>-6.65(G1-1)</td>
<td>3.23</td>
<td>-0.06(5BM)</td>
<td>-0.80(G1-1)</td>
<td>0.49</td>
<td>+0.02(5BM)</td>
<td>-0.22(G1-1)</td>
</tr>
</tbody>
</table>

### Syntactic Pre-ordering (EnJa, JaEn)

**English-to-Japanese: Head Finalization** [Isozaki+ WMT 2010] (~5% BLEU gain)
- Japanese is a head-final language
- Articles (a, an, the) and noun plural form are not needed
- Verb arguments are specified by particles and case markers

*Bridge such gaps with simple rules on HPSG parse tree*

**Japanese-to-English: Rule-based Chunk Reordering** (~1% BLEU gain)
- Series of hand-crafted reordering rules (see the paper for details)
- Japanese chunk-based dependency parser “Cabocha”

**Fast monotone decoding by WFSTS** (EnJa, JaEn)
- Decoding becomes (almost) monotone by pre-ordering
- Efficiently solved by WFSTS (~3x faster than Moses!)

**Domain Adaptation of Word Alignment** (ZhEn)
- Use external bilingual corpora to improve word alignment
- NIST OpenMT 2008 data (107M words after cleaning)

### Acknowledgments:
We’d like to thank the organizers for their great effort on this PatentMT task. We also thank Dr. Takaaki Hori & Dr. Shinji Watanabe for their help on the use of WFSTS, and Prof. Hideo Isozaki for valuable discussion & suggestion about EJ patent translation.

---

**References**:
H. Isozaki+, “Head Finalization: A Simple Reordering rule for SOV languages”, in WMT, 2010

---

**GMBR-based System Combination** *(All subtasks)*

**Theory** [Duh+ UCLNP 2011]

**MBR decision rule**:

\[ \text{argmin}_f \sum_{e \in M(f)} L(e | f) \]

- hypothesis space for input \(f\)
- \(L(e | f)\) is decomposed loss function
- Any sentence-wise metrics can be used: sentBLEU, RIBES, RIBES+sentBLEU, ...

**GMBR decision rule**:

\[ \text{argmin}_{e \in N(f)} \sum_k \theta_k L_k(e | e') \]

- N-best hypotheses for input \(f\)
- \(L_k(e | e')\) is decomposed loss function (e.g., n-gram precision, brevity penalty, rank correlation coeff.)

This is a *Learning to Rank* problem:

*If true loss* on dev. set reference: \(L(e_1 | f) < L(e_2 | f)\) (i.e., \(e_1\) is better than \(e_2\))

*We choose* \(\theta\) such that:

\[ \sum_{e \in N(f)} \sum_k \theta_k L_k(e | e') < \sum_{e' \in M(f)} \sum_k \theta_k L_k(e' | e) \]

---

**Implementation --- Ranking SVM**

**Decoding**

- Decode sentence \(f\)
- Decode \(e\) for \(f\)

**Calculate Loss**

- \(L(e | f)\) is decomposed loss function

**Pairwise comparison**

- \(\Delta\text{baseline}\) = \(L(e_1 | f) - L(e_2 | f)\)
- \(\Delta\text{competitor}\) = \(L(e_1 | f) - L(e_2 | f)\)

**Evaluation**

**Subtask**

*One of the combined systems is from Univ. of Tokyo, others from NTT. See UT team presentation for details.*