The Patent Machine Translation Task
— Summary of NTCIR-9 and Plans for NTCIR-10 —

Isao Goto (NICT)
Bin Lu (City Univ. of Hong Kong / Hong Kong Institute of Education)
Ka Po Chow (Hong Kong Institute of Education)
Eiichiro Sumita (NICT)
Benjamin K. Tsou (Hong Kong Institute of Education / City Univ. of Hong Kong)
Summary of NTCIR-9
Motivation

- There is a significant **practical need** for patent translation.
  - to understand patent information written in foreign languages.
  - to apply for patents in foreign countries.
- Patents constitute one of the **challenging domains**.
  - Patent sentences can be quite **long** and contain **complex structures**.
  - Translation between **languages with largely different word order** is difficult for **long** sentences.
Goals of PatentMT

- To develop **challenging** and **significant practical** research into patent machine translation.

- To **investigate** the **performance** of state-of-the-art machine translation systems in terms of patent translations involving Japanese, English, and Chinese.

- To **compare** the effects of **different methods** of patent translation by applying them to the same test data.

- To **create** publicly-available **parallel corpora of patent documents** and human evaluations of MT results for patent information processing research.

- To **drive machine translation research**, which is an important technology for cross-lingual access of information written in unknown languages.

- The ultimate goal is **fostering scientific cooperation**.
# Findings of Previous Patent Translation Tasks

<table>
<thead>
<tr>
<th>NTCIR-7</th>
<th><strong>Human evaluation</strong></th>
<th><strong>RBMT</strong> was better than SMT for JE and EJ.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CLIR evaluation</strong></td>
<td>SMT was better than RBMT for EJ.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ The translations were used as bag-of-words.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>▪ This means that <strong>word selection</strong> by SMT was better than that by RBMT.</td>
<td></td>
</tr>
</tbody>
</table>

| NTCIR-8 | **Automatic evaluation** | A hybrid system (RBMT with statistical post edit) achieved the best score for JE. |
## Comparison of NTCIR-7, 8, and 9

<table>
<thead>
<tr>
<th></th>
<th>NTCIR-7</th>
<th>NTCIR-8</th>
<th>NTCIR-9</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language</strong></td>
<td>Japanese to English</td>
<td>Japanese to English</td>
<td>Chinese to English</td>
</tr>
<tr>
<td></td>
<td>English to Japanese</td>
<td>English to Japanese</td>
<td>Japanese to English</td>
</tr>
<tr>
<td><strong>Human evaluation</strong></td>
<td>Adequacy Fluency</td>
<td>No human evaluation</td>
<td>Adequacy Acceptability</td>
</tr>
<tr>
<td><strong>Extrinsic evaluation</strong></td>
<td>CLIR</td>
<td>CLIR</td>
<td>No extrinsic evaluation</td>
</tr>
<tr>
<td><strong>Number of participants</strong></td>
<td>15</td>
<td>8</td>
<td>21</td>
</tr>
</tbody>
</table>

At NTCIR-9, participants can choose subtasks from three language directions, including **Chinese to English**.
Notable Findings at NTCIR-9

- **SMT** was the **best** system for **Chinese to English** and **English to Japanese** patent translation.
  - This is the **first time** for SMT to be **demonstrated equal or better** quality than that of the top-level RBMT for **English to Japanese** patent translation.
  - The **pre-ordering** method of NTT-UT for SMT is very effective for **English to Japanese** patent translation.

- **80%** of patent sentences could be understood in the **best system** for **Chinese to English** patent translation.

- **RBMT** was the **best system** for **Japanese to English** patent translation.
Remaining Issues of NTCIR-9

- Practical evaluation
  - The quality of translated sentences was evaluated at NTCIR-9.
  - More practical evaluations are also expected.
Plans for NTCIR-10
Outline of the Plans for NTCIR-10

- Three subtasks:

<table>
<thead>
<tr>
<th>Subtasks</th>
<th>Training data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese to English</td>
<td><strong>1 million</strong> sentence pairs</td>
</tr>
<tr>
<td>Japanese to English</td>
<td>Approximately <strong>3.2 million</strong> sentence pairs</td>
</tr>
<tr>
<td>English to Japanese</td>
<td></td>
</tr>
</tbody>
</table>

(Subtasks and training data are the same as at NTCIR-9)

- Participants select subtasks in which they wish to participate.
- Large scale **parallel corpora** and **new test sets** will be provided.
- **Practical evaluation** will be added (under consideration).
- **Human evaluation** will be carried out.
## Differences from NTCIR-9

<table>
<thead>
<tr>
<th>Practical Evaluation (under consideration)</th>
<th><strong>New</strong>: To explore <strong>practical</strong> MT performance in appropriate fields for patent machine translation.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intrinsic Evaluation</strong></td>
<td><strong>Similar to the NTCIR-9 evaluation.</strong> Quality of translated sentences will be evaluated.</td>
</tr>
<tr>
<td></td>
<td><strong>Additions:</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Chronological evaluation</strong> Comparison between NTCIR-10 and NTCIR-9 to <strong>measure progress</strong>.</td>
</tr>
<tr>
<td></td>
<td><strong>Multilingual evaluation</strong> Comparison of CE and JE translations using the <strong>same English reference</strong> will be added.</td>
</tr>
</tbody>
</table>
Possible Approach to Practical Evaluation

Can the machine translated texts show that patent document A should be rejected?

English

Rejected patent document A

English

Human with patent examiner experience

Text that shows patent document A should be rejected.

Manual translation

Preparation

Evaluation

Japanese/Chinese

English

MT

(The feasibility of this is under investigation. We are working hard to make necessary arrangements.)
Why is it so exciting to participate in?

- **Patents** are one of the *challenging domains* for MT.
  - Patent sentences could be quite *long* and contain *complex structures*.
  - Translation between languages with largely different word order is difficult for *long* sentences.

- Participants will receive *evaluation results* for their MT quality.

- Participants can use *large-scale patent parallel* and *monolingual corpora*.

- Participants can choose subtasks from three language directions, including the language direction of **Chinese to English**.

- We look forward to many groups participating in PatentMT at NTCIR-10!