Machine translation report

Uwe Muegge
Overview

The Machine Translation Project has reached its first milestone: We successfully completed post-editing (in a customized Trados environment) a typical guide section originally translated in a standard machine-translation system. As this process was carefully monitored, we were able to measure how the translation productivity possible in this type of integrated machine translation environment compares to the translation productivity typical for our existing translation process. The results of this initial project are very encouraging: Our tests indicate a productivity gain of approx. 40%.

Contents

PROCESS DESCRIPTION ................................................................. 2
CHALLENGES AND SOLUTIONS .................................................. 2
SEQUENCE OF PROCESS STEPS ................................................ 2
EXPECTATIONS ................................................................................ 3
IMPROVE PERFORMANCE OF HUMAN LINGUISTS ..................... 3
FINDINGS ...................................................................................... 4
WORD COUNT COMPARISON ...................................................... 4
BREAKDOWN OF TIME ............................................................... 4
BREAKDOWN OF POST-EDITING EFFORT .................................... 5
EVALUATION .................................................................................. 6
MEETING EXPECTATIONS .......................................................... 6
RECOMMENDATIONS ................................................................. 8
EXPAND SCALE OF PROJECT .................................................... 8
USE CONTROLLED INPUT .......................................................... 8
Process Description

Challenges and solutions

Traditional rule-based machine-translation environments are typically lacking in two important areas:

- time and cost penalties resulting from manual dictionary creation,
- time and cost penalties resulting from the lack of adequate post-editing tools.

I have developed and implemented solutions to overcome both of these challenges:

- a high-performance automatic terminology extraction process that dramatically speeds up the dictionary-creation process
- a selective approach to dictionary coding (building an MT dictionary traditionally is a very time-consuming process)
- a unique process for integrating machine-translation into our Trados translation-memory environment.

Sequence of process steps

Using standard tools (linguatec PersonalTranslator, Trados Workbench) as well as the above-mentioned customized processes I implemented a unique integrated machine-translation environment.

Here is an overview of the most important steps in this translation process:

<table>
<thead>
<tr>
<th>Step-Table</th>
<th>Integrated Machine-Translation Process</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td>Automatic extraction of new terminology</td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td>Extension of master terminology database (TDB/MultiTerm)</td>
</tr>
<tr>
<td><strong>Step 3</strong></td>
<td>Export of master terminology database into MT dictionary</td>
</tr>
<tr>
<td><strong>Step 4</strong></td>
<td>Machine translation</td>
</tr>
<tr>
<td><strong>Step 5</strong></td>
<td>Integration of machine-translated text into Trados Workbench</td>
</tr>
<tr>
<td><strong>Step 6</strong></td>
<td>Post-editing of machine-translated output by a human linguist</td>
</tr>
</tbody>
</table>
Expectations

Improve performance of human linguists

The stated purpose of this project was not to replace human linguists but to make human linguists more productive. For this reason, the results the machine translation system produced did not have to be perfect. The project would be considered a success if the overall time required to produce a translation was shorter than the time required to produce a human translation.

- The target was a 30% increase in translator productivity
  (the target output for translators is currently 2000 words per day, i.e. we were expecting to see the post-editor process 2600 words per day)

In previous translation projects of technical documentation I have seen the following distribution of the post-editing effort required to produce publishable quality:

- 20% practically no post-editing effort (machine translation perfect or can be perfected with a single keystroke)
- 50% minor post-editing effort (30% of a given segment needs to be changed or, in the case of terminology, terms need to be retrieved from database)
- 30% major post-editing effort (more than 30% of a given segment needs to be changed)

Note:
Despite measurable gains, translators/post-editors in a test environment more often than not have the feeling that the post-editing process takes as long as if not longer than translating from scratch. This is why metrics are an important component of our testing environment.
Findings

Word count comparison

<table>
<thead>
<tr>
<th>Daily Productivity</th>
<th>3002 words per day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated Machine-Translation Environment</td>
<td></td>
</tr>
<tr>
<td>Traditional Translation-Memory Environment</td>
<td>2000 words per day</td>
</tr>
</tbody>
</table>

Looking at the translation output in isolation, this test indicates the process and the tools employed in this integrated machine translation environment boost translator productivity by 50%.

Breakdown of time

<table>
<thead>
<tr>
<th>Time Breakdown per Task</th>
<th>Required Time</th>
<th>MT-specific?</th>
<th>Background Task?</th>
</tr>
</thead>
<tbody>
<tr>
<td>MultiTerm export</td>
<td>15 min</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Automatic term.-extraction</td>
<td>4 min.</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Manual editing of the above</td>
<td>22 min</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Researching 180 new terms</td>
<td>6 hrs.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exporting MultiTerm into MT</td>
<td>32 min.</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Machine-translation</td>
<td>44 min</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Human post-editing</td>
<td>8 hrs.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>File-handling tasks</td>
<td>10 min.</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

As terminology work performed by terminologists is excluded from the translation metrics in our standard environment, I am also excluding all non-MT-specific terminology work from the efficiency calculation.

Result:

- **8 hrs.** for human post-editing effort
- **25 min.** for other MT-specific interactive tasks
- **Total productivity increase >40%**
**Findings (Continuation)**

**Breakdown of post-editing effort**

<table>
<thead>
<tr>
<th>Breakdown of post-editing effort</th>
<th>Correct - S.</th>
<th>Correct - L.</th>
<th>Minor Edit</th>
<th>Major Edit</th>
<th>Untransl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segments</td>
<td>74</td>
<td>3</td>
<td>79</td>
<td>100</td>
<td>5</td>
</tr>
<tr>
<td>Percentage</td>
<td>28%</td>
<td>1%</td>
<td>30%</td>
<td>39%</td>
<td>2%</td>
</tr>
</tbody>
</table>

The numbers in the above table and graphic represent the values for the absolute and relative post-editing effort in various categories:

- 59% of segments either needed only **minor or practically no post-editing**;
- the **majority** of segments produced by the machine-translation system were useful for the **human post-editor**.

**Note:**

Many of the **short sentences correctly translated** by the machine-translation system were

- **bulleted sentence fragments**
- **headings**
- **index entries** (index entries with subentries were not translated correctly; see **below**)
Evaluation

Meeting expectations

With a daily output of approx. 3000 words, we clearly met and exceeded the stated goal for this test, i.e. translating 2600 words per day. Looking at the raw numbers, translating in the unique translation environment created for this project resulted in a productivity increase of 50%. Adjusting this number for MT-specific user intervention the productivity gain is still approx. 40%, which is a very impressive result.

Note:
In a conversation with Daniel Grasmick (head of the Multilingual Technology Group, SAP) at the Translating and the Computer 24 conference, this expert confirmed that SAP is also achieving a 40% productivity increase by using machine translation for certain document types.

It is interesting to note that the distribution of the post-editing effort differed markedly from what I observed in previous machine-translation projects.

<table>
<thead>
<tr>
<th>Distribution of post-editing effort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical result</td>
</tr>
<tr>
<td>20% no post-editing effort</td>
</tr>
<tr>
<td>50% minor post-editing effort</td>
</tr>
<tr>
<td>30% major post-editing effort</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Explanation:

- The high percentage of perfect or near-perfect translations is due to the high frequency of short and very short segments such as bulleted text, headings, index entries, etc.
- The low percentage of minor edits and the high percentage of major edits is due to the fact that the authors of our guides generate free-flowing text generally in the form of long and complex sentences.
- The relatively high percentage of untranslated segments is due in part to the long and complex sentences used in the source document and in part to an error in a filter I wrote.
Potential for improvement

Even though many machine-translated segments were beyond repair, i.e. had to be re-translated from scratch, the majority of machine translated segments was useful for the post-editor. The most useful segments, of course, were

- short segments (15 words or less)
- simple sentences (subject-predicate-object structure).

Index entries meet both of the above criteria and should be considered ideal candidates for machine translation. However, the authors of our guides use subentries, which results in a syntax where two words are being joined by a separator, which in turn renders the expression untranslatable for our machine-translation system.

Example:

```
{xe "Forecasting\[characteristics"]}
```

This issue can be fixed with a simple macro that can easily be integrated into our translation environment.

Decimal groups are an important feature of almost any financial document we translate.

Example:

```
1,329.52 (US standard) 1\,329.52 (SI standard)
```

In a translation-memory environment, translators need to push hotkeys for each decimal-group conversion – provided the system recognizes a decimal group as such (an unresolved issue in tables, for instance), otherwise this conversion must be performed manually. A test I carried out outside of this project indicates that our machine translation system automatically converts all decimal groups correctly. Automatic conversion of decimal groups improves productivity not only during the translation stage but also during the editing stage, as this feature basically eliminates the need for manually checking for correct decimal-group conversion.
Recommendations

Expand scale

In my opinion, the highly positive results of this test warrant the expansion of the Machine Translation Project to include languages other than German and text types other than user guides. I would recommend asking volunteers in the other language teams to perform short (or longer, time permitting) test translations following the established process. Like in our pilot, future expanded tests would involve

- preparing a comprehensive dictionary before translation (terminology extraction and terminology research to extend MultiTerm functionality)
- carefully monitoring and measuring post-editing results.

Use controlled input

It is a well-known fact, once again confirmed by this test, that long and complex sentences are a serious challenge for machine-translation systems. When using a machine-translation system to translate into multiple languages, as recommended above, it might prove to be efficient to pre-edit the source text for machine translation. Using one of the currently available commercial tools, pre-editing would involve

- dividing long sentences into shorter sentences,
- simplifying the sentence structure of complex sentences,
- clarifying ambiguities,
- substituting synonyms,
- correcting spelling and grammar errors overlooked at the authoring stage.