Background & Motivation

User-adaptation in post-editing is CRUCIAL:
1. To overcome domain shifts between training data and translated materials
2. To prevent frustrations related to post-editing
3. To boost efficiency of translators and (possibly) quality

BUT:
Most adaptation approaches rely on imprecise automatic alignment methods
→ We present an interface to collect user-generated phrase-alignments, which are then used in an adaptive SMT engine
→ Our approach is evaluated in a user study

Repetitiveness in Patent Translation
WO 2007000372 A1: Sheathed element glow plug
– A sheathed element glow plug (1) is to be placed inside a chamber (3) of an internal combustion engine.
– The sheathed element glow plug (1) comprises heating body (2) that has a glow tube (6) connected to a housing (4).

WO 2007031371 A1: Sheathed element glow plug
– A sheathed element glow plug (1) serves for arrangement in a chamber of an internal combustion engine.
– The sheathed element glow plug comprises a heating body...

Example
Learn:
sheathed element glow plug, Glühkerze

Example
Learn:
a0 → eine | is2 to be placed, X1 → wird X1 eingebaut | a chamber, of6 a7 → eines | combustion engine, Verbrennungsmotor

Weight Adaptation
- Pairwise ranking updates to weigh many sparse features
- Per coordinate learning rates used to prevent too harsh changes

User Study
Subjects
19 students, 13 prospective translators, 6 CS students, 4 different mother tongues

Data
Titles and abstracts of patent documents, filtered by length, clustered by similarity

Environment
Controlled environment in a computer pool, 90 minute sessions

Machine translation
Hierarchical phrase-based system built from title/abstract training data, good baseline translation results

Task
Post-edit about 500 words from English into German, each task is shared by two subjects

Results
<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Estimated Δ</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBLEU,</td>
<td>+6.8 ± 2.1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>HTER</td>
<td>-5.3 ± 1.9</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>normalized time</td>
<td>-118 ms</td>
<td></td>
</tr>
</tbody>
</table>

LMMEM analysis: Estimated differences in the response variables contrasting non-adaptive to adaptive systems along with associated p-values, if p ≤ 0.05

Conclusions
1. Novel graphical interface with phrase-alignments for a new form of interactive post-editing
2. Alignment can be used for adaptation of the translation model
3. User study shows significant reductions in manual effort and slight speed improvement