Interactive machine translation

Vicent Alabau
valabau@iti.upv.es

May 31, 2013

1 Index

1 Machine translation and computer-assisted translation ▷ 2
2 Interactive machine translation ▷ 7
3 Human evaluation of IMT ▷ 40
4 Adaptation in interactive machine translation ▷ 50
5 IMT API in CasMaCat ▷ 61
6 Bibliography ▷ 68
3 Machine translation (MT)

- Existing MT technologies are currently seen as promising approaches to help produce high-quality translations (HQT) cost-effectively.

- However, the current state of the art in machine translation is still very far from allowing fully automatic HQT.

- (Pre-) post-editing

- While the number of errors and bad constructions is high, “post-editing” can make the result useful.

- Many problems could have been avoided by making the source text “simpler”.

```
  source  pre-editing  translation  post-editing  target
```
**HQT: Computer-assisted translation (CAT)**

- CAT is a form of translation wherein a human translator translates texts using computer software designed to support and facilitate the translation process. (From Wikipedia)

- Other names: computer-aided translation, computer–assisted translation, machine–aided translation, machine-assisted translation.

- Historically, CAT and MT have been considered as different but close technologies [Kay, MT 1997].

**HQT: Human-machine interaction (HMI)**

- In classical MT and CAT, the interaction between human translators and machines is very limited.

- At the origin: HMI focused on disambiguation of the source text or for updating user dictionaries or for searching through dictionaries [Slocum, CL 1985][Whitelock et al., COLING 1986].

- **Novel idea**: A complete MT system is used to produce target sentence hypotheses, which can be accepted or amended by a human translator. Each correct text segment is then used by the MT system as additional information to achieve further, hopefully improved suggestions. [Foster et al., MT 1997] [Barrachina et al., CL 2008][Casacuberta et al., CACM 2008].

- A slightly different approach for interaction with MT systems was to present selectable phrase translation options [Koehn, 2009].
The Multimodal Interactive Machine Translation Paradigm

Challenges of multimodal interaction

- Feedback
- Adaptation
- Multimodality

Index

1 Machine translation and computer-assisted translation ▶ 2
2 Interactive machine translation ▶ 7

3 Human evaluation of IMT ▶ 40
4 Adaptation in interactive machine translation ▶ 50
5 IMT API in CasMaCat ▶ 61
6 Bibliography ▶ 68
Diagram of a SMT system

Post-editing: example

Translating the source sentence “Click OK to close the print dialog” into Spanish
(the reference is “Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión”):
Post-editing: example

Translating the source sentence “Click OK to close the print dialog” into Spanish (the reference is “Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión”):

System: Haga clic para cerrar el diálogo de impresión

User: Haga clic en para cerrar el diálogo de impresión
Post-editing: example

Translating the source sentence “Click OK to close the print dialog” into Spanish
(the reference is “Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión”):

System: Haga clic para cerrar el diálogo de impresión
User:   Haga clic en para cerrar el diálogo de impresión
User:   Haga clic en ACEPTAR para cerrar el diálogo de impresión

System: Haga clic para cerrar el diálogo de impresión
User:   Haga clic en para cerrar el diálogo de impresión
User:   Haga clic en ACEPTAR para cerrar el cuadro diálogo de impresión
Post-editing: example

Translating the source sentence “Click OK to close the print dialog” into Spanish (the reference is “Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión”):

System: Haga clic para cerrar el diálogo de impresión
User:   Haga clic en para cerrar el diálogo de impresión
User:   Haga clic en ACEPTAR para cerrar el diálogo de impresión
User:   Haga clic en ACEPTAR para cerrar el cuadro diálogo de impresión
User:   Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión

CASMACAT Interactive machine translation May 31, 2013
Post-editing: example

Translating the source sentence “Click OK to close the print dialog” into Spanish (the reference is “Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión”):

System: Haga clic para cerrar el diálogo de impresión
User: Haga clic en para cerrar el diálogo de impresión
User: Haga clic en ACEPTAR para cerrar el diálogo de impresión
User: Haga clic en ACEPTAR para cerrar el cuadro diálogo de impresión
User: Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión
User: Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión

Result: Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión

TOTAL: Four word-strokes

Introduction to interactive machine translation (IMT)

- IMT (or interactive CAT) aims to increase the overall (MT + human) productivity by incorporating human correction activities within the translation process itself

**Main idea in IMT:**

- Use a MT system to produce target text segments that can be accepted or amended by a human translator; these correct(ed) segments are then used by the MT system as additional information to achieve further, hopefully improved suggestions
- IMT can be seen as an auto-completion system that is sensitive to the source text
- auto-completion can predict full sentences of parts of it
Diagram of an interactive MT system

Human-machine (keyboard) interactive process: example

Translating the source sentence (x) “Click OK to close the print dialog” into Spanish:

System (∩y_s): Haga clic para cerrar el diálogo de impresión
Human-machine (keyboard) interactive process: example

Translating the source sentence (x) “Click OK to close the print dialog” into Spanish:

System ($\hat{y}_s$): Haga clic para cerrar el diálogo de impresión
User ($y_p$): Haga clic en

System ($\hat{y}_s$): ACEPTAR para cerrar el diálogo de impresión
User ($y_p$): Haga clic en
System ($\hat{y}_s$): ACEPTAR para cerrar el diálogo de impresión
Human-machine (keyboard) interactive process: example

Translating the source sentence (x) “Click OK to close the print dialog” into Spanish:

System (ŷ_s): Haga clic para cerrar el diálogo de impresión
User (y_p): Haga clic en
System (ŷ_s): ACEPTAR para cerrar el diálogo de impresión
User (y_p): Haga clic en ACEPTAR para cerrar el cuadro

System (ŷ_s): de diálogo de impresión
Human-machine (keyboard) interactive process: example
Translating the source sentence (x) “Click OK to close the print dialog” into Spanish:

System ($y_s$): Haga clic para cerrar el diálogo de impresión
User ($y_p$): Haga clic en
System ($y_s$): ACEPTAR para cerrar el diálogo de impresión
User ($y_p$): Haga clic en ACEPTAR para cerrar el cuadro
System ($y_s$): de diálogo de impresión
User ($y_p$): Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión
Result ($\hat{y}$): Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión

TOTAL: Two word-strokes
Human-machine (keyboard) interactive process

- In each iteration, a correct prefix ($y_p$) of the target sentence is available and the IMT system computes its best (or $N$-best) translation suffix hypothesis ($\hat{y}_s$) to complete this prefix.

- Given $y_p\hat{y}_s$, the IMT cycle proceeds by letting the user establish a new, longer acceptable prefix.

This prefix is typically formed by $y_p$, followed by an initial part of $\hat{y}_s$ accepted by the user, followed by text obtained by means of additional user keystrokes generally aimed to amend remaining incorrect parts of $\hat{y}_s$.

This prefix becomes a new $y_p$, thereby starting a new IMT prediction cycle.

- Ergonomics and user preferences dictate exactly when the system can start its new cycle, but typically, it is started after each user-entered word or even after each new user keystroke.

* The interaction is performed from left-to-right. The user model does not foresee users going back to fix errors to the left of the cursor.

Interactive machine translation: the original idea

- These ideas were studied in the TransType (TT) project [Foster et al. EMNLP 2002] and have been thoroughly explored in the TransType-2 (TT2) project [Barrachina et al. CL 2008]

- In TT, the IMT system suggests the best target word that follows the given prefix, however in TT2, the IMT system suggests the best complete suffix.
Text prediction for interactive machine translation

- Given a source text \( x \) and a “correct” prefix \( y_p \) of the target text, search for a suffix \( \hat{y}_s \), that maximizes the posterior probability over all possible suffixes:

\[
\hat{y}_s = \arg\max_{y_s} \Pr(y_s | x, y_p)
\]

- Taking into account that \( \Pr(y_p | x) \) does not depend on \( y_s \), we can write:

\[
\hat{y}_s = \arg\max_{t_s} \frac{\Pr(y_p y_s | x)}{\Pr(y_p | x)} = \arg\max_{t_s} \Pr(y_p y_s | x)
\]

- \( \Pr(y_p y_s | x) \): Log-linear models (or SFST if \( \Pr(y_p y_s, x) \) is used)

- MT is a particular case, where \( y_p = \lambda \)

- Main difference of IMT vs. MT: search over the set of suffixes

IMT search: \( n \)-best lists

\textbf{N-best based approach:}

- For each source sentence, a \( n \)-best list is generated which represents the top \( n \) translations according to the translation model.

- In each IMT iteration, the \( n \)-best list is searched for a best hypothesis compatible with the prefix given in this iteration.

- Error-correcting smoothing (edit distance) is used to allow for user-given prefixes that may not exist in the \( n \)-best list.

- Computation is carried out in an incremental manner: in each iteration the results from the previous iteration are updated.
**IMT search: \( n \)-best lists**

**source:** Para ver la lista de recursos

**reference:** To view a listing of resources

- **8.13** To view the resources list
- **8.41** To view a resource list
- **9.32** To view a list of resources
- **9.85** To view a listing resources
- **10.45** To see the list of resources
- **10.72** To view a listing of resources
IMT search: \textit{n}-best lists

\textbf{source:} Para ver la lista de recursos

\textbf{reference:} To view a listing of resources

-8.13 To view the resources list

-8.41 To view a resource list

-9.32 To view a list of resources

-9.85 To view a listing resources

-10.45 To see the list of resources

-10.72 To view a listing of resources
**IMT search: \( n \)-best lists**

- Error-correcting smoothing (edit distance) is used to allow for user-given prefixes that may not exist in the \( n \)-best list.

**source:** Para ver la lista de recursos

**reference:** To see a list of resources

-8.13 To view the resources list
-8.41 To view a resource list
-9.32 To view a list of resources
-9.85 To view a listing resources
-10.45 To see the list of resources
-10.72 To view a listing of resources
IMT search: word graphs

High speed is needed because typically a new system hypothesis must be produced in real time after each user keystroke [Och et al. EACL 2003] [Barrachina et al. CL 2008].

**WORD-GRAF BASED APPROACH:**

- Word-graphs are compact structures for n-best list. There exists efficient algorithms to operate them.

- For each source sentence, a word graph representing all its possible translations according to the translation model is generated.

- In each IMT iteration, the word graph is searched for a best path compatible with the prefix given in this iteration.

- Error-correcting smoothing (edit distance) is computed using word graph algorithms. Computation is also carried out in an incremental manner.

Translation word graph corresponding to the source “seleccionar el siguiente”
IMT search

Translation word graph corresponding to the source “seleccionar el siguiente”
Example of human-machine (keyboard) interaction

Source sentence (x): Load your originals into the Document Feeder

System (ŷ_s): Cargue los originales en la
Example of human-machine (keyboard) interaction

Source sentence (x): Load your originals into the Document Feeder

System (ŷ_s): Cargue los originales en la

User (ŷ_p): Cargue los originales en e

System (ŷ_s): l alimentador de originales
Example of human-machine (keyboard) interaction

Source sentence (x): Load your originals into the Document Feeder

System (\(\hat{y}_s\)): Cargue los originales en la

User (\(y_p\)): Cargue los originales en e

System (\(\hat{y}_s\)): al alimentador de originales

User (\(y_p\)): Cargue los originales en el alimentador de originales

Output (\(\hat{y}\)): Cargue los originales en el alimentador de originales
Evaluating MT and IMT systems

FOUR MEASURES

- **Translation Word Error Rate (TWER):**
  Minimum number of word insertions, deletions and substitutions needed to edit the system output into a (single) target reference.

- **Word-Stroke Ratio (WSR):**
  Number of user interactions that are necessary to achieve the reference target divided by the number of running words. In each interaction only one wrong word is changed.

- **Translation Character Error Rate (CER):**
  Minimum number of character insertions, deletions and substitutions needed to edit the system output into a (single) target reference.

- **Key-Stroke Ratio (KSR):**
  Number of key-strokes that are necessary to achieve a (single) target reference divided by the number of running characters. In each interaction only one character is changed.

Benchmark Xerox printer manuals corpus

<table>
<thead>
<tr>
<th>Data</th>
<th>English</th>
<th>Spanish</th>
<th>English</th>
<th>German</th>
<th>English</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train Sent. pairs</td>
<td>56K</td>
<td>53K</td>
<td>49K</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Run. words</td>
<td>572K</td>
<td>657K</td>
<td>583K</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>8K</td>
<td>12K</td>
<td>19K</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Sentences</td>
<td>1 125</td>
<td>984</td>
<td>996</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Run. words</td>
<td>7.6K</td>
<td>9.4K</td>
<td>10.0K</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out of Voc.</td>
<td>341</td>
<td>362</td>
<td>552</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perplexity</td>
<td>107</td>
<td>60</td>
<td>169</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary of IMT results with the Xerox corpus

<table>
<thead>
<tr>
<th>DATA: SFST (1-best)</th>
<th>XRCE2</th>
<th>WSR</th>
<th>TWER</th>
</tr>
</thead>
<tbody>
<tr>
<td>English-Spanish</td>
<td>27.4</td>
<td>43.1</td>
<td></td>
</tr>
<tr>
<td>Spanish-English</td>
<td>31.7</td>
<td>51.4</td>
<td></td>
</tr>
<tr>
<td>English-French</td>
<td>65.1</td>
<td>73.8</td>
<td></td>
</tr>
<tr>
<td>French-English</td>
<td>58.5</td>
<td>71.9</td>
<td></td>
</tr>
<tr>
<td>English-German</td>
<td>55.4</td>
<td>81.3</td>
<td></td>
</tr>
<tr>
<td>German-English</td>
<td>55.0</td>
<td>78.5</td>
<td></td>
</tr>
</tbody>
</table>

Similar results were achieved with phrase-based models

[Barrachina et al. CL 2008]

---

Benchmark EU bulletin corpus

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Train</td>
<td>214K</td>
<td>5.9M</td>
<td>84K</td>
<td></td>
<td></td>
<td></td>
<td>800</td>
<td></td>
<td></td>
<td></td>
<td>96</td>
<td>108</td>
</tr>
<tr>
<td>Run. words</td>
<td></td>
<td>6.6M</td>
<td>97K</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20K</td>
<td>140</td>
</tr>
<tr>
<td>Vocabulary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>107</td>
<td>113</td>
</tr>
<tr>
<td>Test</td>
<td>223K</td>
<td>6.5M</td>
<td>87K</td>
<td></td>
<td></td>
<td></td>
<td>800</td>
<td></td>
<td></td>
<td></td>
<td>22K</td>
<td>113</td>
</tr>
<tr>
<td>Run. words</td>
<td></td>
<td>6.1M</td>
<td>152K</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>21K</td>
<td>119</td>
</tr>
<tr>
<td>Out of Voc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22K</td>
<td>113</td>
</tr>
<tr>
<td>Perplexity</td>
<td>215K</td>
<td>6.0M</td>
<td>85K</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>24K</td>
<td>119</td>
</tr>
</tbody>
</table>
## Summary of IMT results with the EU corpus

<table>
<thead>
<tr>
<th>DATA: SFST (1-best)</th>
<th>EU WSR</th>
<th>TWER</th>
</tr>
</thead>
<tbody>
<tr>
<td>English-Spanish</td>
<td>52.1</td>
<td>55.8</td>
</tr>
<tr>
<td>Spanish-English</td>
<td>48.5</td>
<td>52.5</td>
</tr>
<tr>
<td>English-French</td>
<td>62.2</td>
<td>53.9</td>
</tr>
<tr>
<td>French-English</td>
<td>60.5</td>
<td>49.2</td>
</tr>
<tr>
<td>English-German</td>
<td>49.6</td>
<td>65.5</td>
</tr>
<tr>
<td>German-English</td>
<td>44.0</td>
<td>66.6</td>
</tr>
</tbody>
</table>

Similar results were achieved with PB models

[Barrachina et al. CL 2008]

---

## Summary of IMT results with KSR

### Xerox corpus: printer manuals

<table>
<thead>
<tr>
<th>Corpus</th>
<th>CER</th>
<th>PKSR</th>
<th>KSR</th>
</tr>
</thead>
<tbody>
<tr>
<td>English–Spanish</td>
<td>21.6</td>
<td>16.8</td>
<td>11.8</td>
</tr>
<tr>
<td>English–French</td>
<td>47.9</td>
<td>35.6</td>
<td>24.5</td>
</tr>
<tr>
<td>English–German</td>
<td>55.2</td>
<td>39.1</td>
<td>28.3</td>
</tr>
</tbody>
</table>

### EU-TT2 corpus: legal documents

<table>
<thead>
<tr>
<th>Corpus</th>
<th>CER</th>
<th>PKSR</th>
<th>KSR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish–English</td>
<td>37.4</td>
<td>26.7</td>
<td>16.7</td>
</tr>
<tr>
<td>French–English</td>
<td>32.8</td>
<td>23.4</td>
<td>14.3</td>
</tr>
<tr>
<td>German–English</td>
<td>43.5</td>
<td>30.7</td>
<td>20.1</td>
</tr>
</tbody>
</table>

**Post-editing. PKSR.** post-editing keystroke ratio. CER with auto-completion
Using the mouse as an additional information source

- Up to now, the only interface between the underlying SMT engine and the user was the **keyboard**, and the utility of the **mouse** was to position the cursor in the appropriate place before typing in a word.

- However, when the user wants to correct the translation hypothesis, he needs to click previously on the position he intends to correct, therefore:
  - he is validating a prefix up to the position where he has set the cursor
  - he is indicating that whatever comes after that cursor position is incorrect and he wants it to be replaced.

- Moreover, successive click actions can be used to ask for different alternatives to the current incorrect word [Sanchis et al. MLMI 2008].

---

Using the mouse as an additional information source

Translating the source sentence “Click OK to close the print dialog” into Spanish:

(The symbol “↑” is used to indicate the position where a mouse action is performed)

(The reference sentence is “Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión”)

---
Using the mouse as an additional information source

Translating the source sentence “Click OK to close the print dialog” into Spanish:
(the symbol “↑” is used to indicate the position where a mouse action is performed)
(The reference sentence is “Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión”)
System: Haga clic para cerrar el diálogo de impresión

User: Haga clic ↑ para cerrar el diálogo de impresión
Using the mouse as an additional information source

Translating the source sentence “Click OK to close the print dialog” into Spanish:
(the symbol “↑” is used to indicate the position where a mouse action is performed)
(The reference sentence is “Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión”)

System: Haga clic para cerrar el diálogo de impresión

User:  Haga clic↑ para cerrar el diálogo de impresión

System: en ACEPTAR para cerrar el diálogo de impresión
Using the mouse as an additional information source

Translating the source sentence “Click OK to close the print dialog” into Spanish:
(the symbol “↑” is used to indicate the position where a mouse action is performed)
(The reference sentence is “Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión”)
System: Haga clic para cerrar el diálogo de impresión
User: Haga clic para cerrar el diálogo de impresión
System: en ACEPTAR para cerrar el diálogo de impresión
User: Haga clic en ACEPTAR para cerrar el diálogo de impresión
System: marco de diálogo de impresión
User: Haga clic en ACEPTAR para cerrar el cuadro
Using the mouse as an additional information source

Translating the source sentence “Click OK to close the print dialog” into Spanish:
(the symbol ↑ is used to indicate the position where a mouse action is performed)
(The reference sentence is “Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión”)
System: Haga clic para cerrar el diálogo de impresión
User: Haga clic ↑ para cerrar el diálogo de impresión
System: en ACEPTAR para cerrar el diálogo de impresión
User: Haga clic en ACEPTAR para cerrar el ↑ diálogo de impresión
System: Haga clic en ACEPTAR para cerrar el marco de diálogo de impresión
User: Haga clic en ACEPTAR para cerrar el cuadro
de diálogo de impresión
User: Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión
Output: Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión
Using the mouse as an additional information source

Translating the source sentence “Click OK to close the print dialog” into Spanish:
(The symbol “↑” is used to indicate the position where a mouse action is performed)
(The reference sentence is “Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión”)
System: Haga clic para cerrar el diálogo de impresión
User: Haga clic ↑ para cerrar el diálogo de impresión
System: en ACEPTAR para cerrar el diálogo de impresión
User: Haga clic en ACEPTAR para cerrar el ↑ diálogo de impresión
System: Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión
User: Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión
Output: Haga clic en ACEPTAR para cerrar el cuadro de diálogo de impresión
TOTAL: Two mouse actions + one word stroke

CASMACAT Interactive machine translation May 31, 2013

Using the mouse as an additional information source

As in pure text IMT the system suggests an optimal suffix:

\[
\hat{y}_s = \text{argmax}_{t_s} \Pr(y_s \mid x, y_p)
\]

Mouse actions can be considered as a constrained search problem

\[
\hat{y}_s = \text{argmax}_{t_s: t_s \neq \hat{t}_l} \Pr(y_s \mid x, y_p, \hat{y}_l)
\]

where \(\hat{y}_l\) is the suffix generated in the previous iteration, already discarded by the user, and \(\hat{t}_l\) is the first word in \(\hat{t}_l\).

Alternatively, using the property of joint distributions:

\[
\hat{y}_s = \text{argmax}_{t_s: t_s \neq \hat{t}_l} \Pr(x, y_p, y_s \mid \hat{y}_l)
\]
Evaluating the mouse as an additional information source

Two Measures

- **Word-Stroke Ratio (WSR):**
  Number of user interactions that are necessary to achieve the reference target divided by the number of running words. In each interaction only one wrong word is changed.

- **Mouse Action Ratio (MAR):**
  Ratio between the number of mouse actions required by the user to achieve the final reference sentence and its total number of words.

Results with the Xerox corpus

<table>
<thead>
<tr>
<th>DATA:</th>
<th>Baseline</th>
<th>With mouse actions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAR</td>
<td>WSR</td>
</tr>
<tr>
<td>XRCE2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>En-Es</td>
<td>10.0</td>
<td>27.4</td>
</tr>
<tr>
<td>Es-En</td>
<td>13.5</td>
<td>31.7</td>
</tr>
<tr>
<td>En-Fr</td>
<td>12.6</td>
<td>65.1</td>
</tr>
<tr>
<td>Fr-En</td>
<td>13.5</td>
<td>58.5</td>
</tr>
<tr>
<td>En-De</td>
<td>13.6</td>
<td>55.4</td>
</tr>
<tr>
<td>De-En</td>
<td>15.7</td>
<td>55.0</td>
</tr>
</tbody>
</table>

[Sanchis et al. MLMI 2008]
Results with the EU corpus

<table>
<thead>
<tr>
<th>DATA:</th>
<th>Baseline</th>
<th></th>
<th>With mouse actions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>EU</td>
<td>MAR</td>
<td>WSR</td>
<td>MAR</td>
<td>WSR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>En-Es</td>
<td>15.9</td>
<td>52.1</td>
<td>52.3</td>
<td>44.9</td>
</tr>
<tr>
<td>Es-En</td>
<td>13.8</td>
<td>48.5</td>
<td>48.2</td>
<td>40.8</td>
</tr>
<tr>
<td>En-Fr</td>
<td>23.5</td>
<td>62.2</td>
<td>62.3</td>
<td>56.4</td>
</tr>
<tr>
<td>Fr-En</td>
<td>14.4</td>
<td>60.5</td>
<td>60.5</td>
<td>53.4</td>
</tr>
<tr>
<td>En-De</td>
<td>15.6</td>
<td>49.6</td>
<td>49.8</td>
<td>43.1</td>
</tr>
<tr>
<td>De-En</td>
<td>14.4</td>
<td>44.0</td>
<td>43.8</td>
<td>36.3</td>
</tr>
</tbody>
</table>

[Sanchis et al. MLMI 2008]

Index

1 Machine translation and computer-assisted translation ▷ 2
2 Interactive machine translation ▷ 7
3 Human evaluation of IMT ▷ 40
4 Adaptation in interactive machine translation ▷ 50
5 IMT API in CasMaCat ▷ 61
6 Bibliography ▷ 68
Human evaluation: A typical session in TT2

Human evaluation in TT2 project

- Assessment: to measure the overall time required to translate a test corpus using the IMT system and without any system.

- Six professional translators, recruited from the two translation agencies. Five rounds during 18 months using Xerox corpus.

- Human evaluation: Overall, an IMT system can allow translators to increase their productivity while maintaining high-quality (about a save of 10-15% of human effort) depending on the task [Casacuberta et al. CACM 2008].

- Drop-down menus were perceived as not useful. N-best options were too similar and there were lots of text to read on each interaction.
MIPRCV prototype

- black → word being edited
- yellow → words visible in working area

- keyboard shortcuts:
  - validate current word: ENTER
  - insert after current word: CTRL+SPACE
  - insert before current word: CTRL+SHIFT+SPACE
  - delete word after current word: CTRL+SUPR
  - reject current word: click o CTRL+UP
  - go back to previously rejected word: CTRL+DOWN

- glossary (5-best IBM1 translations)
  
  http://cat.iti.upv.es/imt/

- IMT system disabled

- word autocompletion (wordlist)

- glossary (5-best IBM1 translations)
MIPRCV prototype

- 10 non-professional translators, 40 sentences per user
- PE was a better system than the IMT system
- IMT achieved more consistent translation
- If the first translation is wrong PE is harder
- Many suffix changes demand cognitive effort
- Complaints regarding UI interaction
- Bugs in the prediction engine
- Bugs in UI make good features bad

MIPRCV prototype

you will have to file a tax return in the May following the end of the tax year.

déberá presentar declaración en la mayo de tras el final del ejercicio fiscal.

1. The system’s suggestions were helpful
2. The system was cumbersome to use

Quick overview
- The translation system is a normal text area where you can interact with it as usual with the following enhancements:
  - Each time you press a key, the system will try to suggest a suitable continuation.
  - To accept the suggestions, press ENTER. Otherwise, just keep typing.
  - If you want to toggle the suggestions, press ESC.
- Your feedback is greatly appreciated, please submit your comments at any time.

System status
Suggestions are enabled.

Dictionary
Click on a word in the source sentence to look for it in an English-Spanish dictionary.
MIPRCV prototype

- IMT required less time and interactions than PE
- No significant correlation was found between KSMR and duration
- Small significant correlation between duration and difference of consecutive suffixes
- Small significant correlation between user KSMR and satisfaction
- Summary of IMT comments:
  - When you change the prefix, corrected words in suffix change
  - Suggestions at the end of the sentence are specially annoying
  - ‘I had to disable suggestion because they were annoying’
  + Occasionally, IMT is very useful and the predictions are correct

Conclusions of MIPRCV prototype

- UI design is crucial
- Larger scale human evaluation is needed
- Improve search algorithms to:
  - Keep user corrections when the user modifies the prefix
  - Suggestions should only change if they are for the better
  - Limit the number of changes of suggestion w.r.t. current suffix

[Alabau et al. EAMT 2012]
Index

1 Machine translation and computer-assisted translation ▷ 2

2 Interactive machine translation ▷ 7

3 Human evaluation of IMT ▷ 40

4 Adaptation in interactive machine translation ▷ 50

5 IMT API in CasMaCat ▷ 61

6 Bibliography ▷ 68
Diagram of an interactive MT system: Adaptation

- Human interaction offers another unique opportunity to improve IMT system’s behavior by tuning the translation models.

- In each iteration, the text obtained by means of additional user keystrokes to correct the suggestion produced by the IMT systems together with the corresponding aligned source segments can generally be converted into new, fresh training data, useful for adapting the system to changing environment.

- Adaptation is nowadays an open problem in IMT and only heuristic ideas are proposed to take advantage of the new, fresh training data.
Adaptation techniques

- In the TT framework, a cache technique both for language models (unigrams, bigrams and trigrams) and translation models (model IBM1 and IBM2) is used [Nepveu et al. EMNLP 2004].

- The output of post-editing has been used for adaptation [Callison-Burch et al. EAMT 2004].

- Another technique is based on translation memories, in order to store the data corrected by the human [Biçini & Dymetman CICling 08].

- Techniques for topic-adaptation in statistical machine translation

SMT and Online Learning

- SMT allows us to translate a given source text without human intervention

- Output of SMT systems can be supervised to obtain high-quality translations

- User feedback can be used to extend the statistical models of the SMT system

- Online learning fits naturally in two well-known SMT applications:
  - Post-editing (PE)
  - ITP

- Online learning can be used to estimate:
  - Parameters of the translation and language models
  - Scaling factors of the log-linear combination
Online Learning of translation and language models

- HMM models are used here for:
  - smoothing
  - generating word alignment matrices

- update counts and normalize the language model
- incremental EM algorithm to train the HMM models
- sufficient statistics are the expected counts of dictionary and alignment probabilities
- word alignment matrices are used to extract phrase counts
- phrase model probabilities are estimated from these updated phrase counts:

\[
p(\tilde{x}|\tilde{y}) = \frac{c(\tilde{x}, \tilde{y})}{\sum_{\tilde{x}'} c(\tilde{x}', \tilde{y})}
\]

- Learning from previously estimated models (English-French Xerox)

<table>
<thead>
<tr>
<th></th>
<th>ITP system</th>
<th>BLEU</th>
<th>KSMR</th>
<th>LT (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eng-Fre</td>
<td>batch</td>
<td>33.7±2.0</td>
<td>33.9±1.3</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>online</td>
<td>42.2±2.2</td>
<td>27.9±1.3</td>
<td>0.09</td>
</tr>
</tbody>
</table>

- Learning from scratch:
  - 10,000 sentences randomly extracted from the English-Spanish Xerox corpus
Online Learning of SMT scaling factors

- Discriminative Ridge Regression to update weights given by MERT
- Good hypotheses within a $n$-best list score higher, bad hypotheses lower
- Originally designed for conventional SMT (PE)
- Establish correlation between difference in translation quality and difference in score
- Find $\tilde{\lambda}_t$ such that $R_X \cdot \tilde{\lambda}_t \propto l_X$, with
  - $R_X$ difference of values in $\tilde{h}$ between every $y \in n$-best and best hypothesis $y^*$
  - $l_X$ difference in quality between every $y \in n$-best and best hypothesis $y^*$
- Extension to IMT: modified mDDR extracts features from word-graphs to measure translation quality

Corpora used:
- Spanish–English Europarl (training and development)
- News-Commentary (test)

Two different scenarios:
- PE: quality measured with TER
- ITP: quality measured with WSR

Baseline with Moses decoder
Online Learning of SMT scaling factors

- DDR works well to improve TER
- mDRR shows not significant improvements in WSR

Online Learning conclusions

- Online learning techniques for SMT have been proposed
- Such techniques allow to incrementally update the parameters of the generative models and their scaling factors
- Estimation of scaling factors in ITP is being studied
- Empirical results clearly show the utility of online learning in SMT and ITP
Index

1 Machine translation and computer-assisted translation > 2
2 Interactive machine translation > 7
3 Human evaluation of IMT > 40
4 Adaptation in interactive machine translation > 50
5 IMT API in CasMaCat > 61
6 Bibliography > 68

IMT API in CasMaCat

```javascript
/**
* Initialization method
* @param url {String} Server URL to connect to
*/
self.connect = function(url) {
  ...
  self.server = new io.connect(url, ioOptions);
  ...
};

/**
* Start predictive session.
* @param {Object} @setup obj
* @trigger startSessionResult
* @return {Object}
* errors {Array} List of error messages
* data {Object}
* @setup data
* elapsedTime {Number} ms
*/
startSession: function(obj) {
  this.checkConnection();
  this.server.emit('startSession', {data: obj});
},
```
IMT API in CasMaCat

/** Send decoding results for the current segment.
* @param {Object}
* @setup obj
*   target {String} Segment text
*   caretPos {Number} Index position of caret cursor
*   [numResults] {Number} How many results should be retrieved (default: 1)
* @trigger setPrefixResult
* @return {Object}
*   errors {Array} List of error messages
*   data {Object}
*     @setup data
*       source {String} Verified source
*       sourceSegmentation {Array} Verified source segmentation
*       elapsedTime {Number} ms
*       nbest {Array} List of objects
*     @setup nbest
*       target {String} Result
*       targetSegmentation {Array} Segmentation of result
*       elapsedTime {Number} ms Time to process each result
*       [author] {String} Technique or person that generated the target result
*       [alignments] {Array} Dimensions: source * target
*       [confidences] {Array} List of floats for each token
*       [quality] {Number} Quality measure of overall hypothesis
*       [priorities] {Array} List of integers, where each integer indicates a group of tokens
* */
setPrefix: function(obj) {
  this.checkConnection();
  this.server.emit('setPrefix', {data: obj});
},

CASMACAT Interactive machine translation May 31, 2013

IMT API in CasMaCat

/** Reject received suffix for the current segment.
* @param {Object}
* @setup obj
*   target {String} Segment text
*   caretPos {Number} Index position of caret cursor
*   [numResults] {Number} How many results should be retrieved (default: 1)
* @trigger rejectSuffixResult
* @return {Object}
*   errors {Array} List of error messages
*   data {Object}
*     @setup data
*       source {String} Verified source
*       sourceSegmentation {Array} Verified source segmentation
*       caretPos {Number} Verified index position of caret cursor
*       elapsedTime {Number} ms
*       nbest {Array} List of objects
*     @setup nbest
*       target {String} Result
*       targetSegmentation {Array} Segmentation of result
*       elapsedTime {Number} ms Time to process each result
*       [author] {String} Technique or person that generated the target result
*       [alignments] {Array} Dimensions: source * target
*       [confidences] {Array} List of floats for each token
*       [quality] {Number} Quality measure of overall hypothesis
*       [priorities] {Array} List of integers, where each integer indicates a group of tokens
* */
rejectSuffix: function(obj) {
  this.checkConnection();
  this.server.emit('rejectSuffix', {data: obj});
},

CASMACAT Interactive machine translation May 31, 2013
IMT API in CasMaCat

/**
 * End predictive session for the current segment.
 * @trigger endSessionResult
 * @return {Object}
 * errors {Array} List of error messages
 * data {Object}
 * @setup data
 * elapsedTime {Number} ms
 */
endSession: function () {
    this.checkConnection();
    this.server.emit('endSession');
},
/**
 * Event handler
 * @param {Mixed} String or Array of strings name of trigger
 * @param {Function} Callback
 */
self.on = function (ev, fn) {
    ...
};
/**
 * Trigger event
 */
self.trigger = function () {
    console.log("trigger", arguments);
    self.server.$emit.apply(self.server, arguments);
};
/**
 * Retrieves decoding results for the current segment.
 * @param {Object}
 * @setup obj
 * ... elapsedTime {Number} ms
 */
self.ping = function (obj) {...};
/**
 * Configures server as specified by the client.
 * @param {Object} Server-specific configuration
 * @trigger configureResult
 * @return {Object}
 * errors {Array} List of error messages
 * data {Object}
 * @setup data
 * ms {Number} Original ms
 * elapsedTime {Number} ms 0 by definition
 */
self.configure = function (obj) {...};
IMT API in CasMaCat

```javascript
/**
 * Validates source-target pair.
 * @param {Object}
 * @setup obj
 * source {String}
 * target {String}
 * @trigger validateResult
 * @return {Object}
 * errors {Array} List of error messages
 * data {Object}
 * @setup data
 * elapsedTime {Number} ms
 */
self.validate = function(obj) {...};

/**
 * Resets server.
 * @trigger resetResult
 * @return {Object}
 * errors {Array} List of error messages
 * data {Object} Response data
 * @setup data
 * elapsedTime {Number} ms
 */
self.reset = function() {...};
```

index

1 Machine translation and computer-assisted translation ▶ 2
2 Interactive machine translation ▶ 7
3 Human evaluation of IMT ▶ 40
4 Adaptation in interactive machine translation ▶ 50
5 IMT API in CasMaCat ▶ 61
6 Bibliography ▶ 68
Bibliography


Bibliography


