Integration of human and machine translation

Marcello Federico
Fondazione Bruno Kessler
MT Marathon, Prague, Sept 2013
Motivation

- **Human translation (HT)**
  - worldwide demand for translation services has accelerated, due to globalization and growth of the Information Society

- **Gap between MT and HT**
  - MT has improved significantly but independently from HT
  - MT research has not directly addressed how to improve HT
  - Today professional translators barely use MT

- **The unavoidable adoption of MT**
  - Post-editing experiments have shown great promise
  - Integration of HT and MT is still an open problem!
Questions

- How do human translators work?
- What tools do they use?
- How is productivity measured?
- How can MT help human translators?
- What are important problems to solve?
- Why should MT researchers care?
- …
Outline

- Typical translation-industry workflow
- Computer assisted translation tools
- Simple MT-CAT integration
- Two undergoing research projects
  - their research challenges
  - their experimental platforms
Translation Project

All our translators got a CAT tool!

Language Service Provider

Scenario
Scenario

I’m the project manager
Computer Assisted Translation (CAT) is the dominant technology in the translation industry

CAT tools: special text editors supporting many document formats and integrating information from different sources.
CAT Tools

Source/target text is split into segments
Translation progresses segment by segment
Provides helps from different sources:
  • spell checkers
  • dictionaries
  • terminology managers
  • concordancers
  • translation memory (TM)
  • and recently machine translation (MT)
CAT Tool
But I must explain to you how all this mistaken idea of denouncing pleasure and praising pain was born and I will give you a complete account of the system, and expound the actual teachings of the great explorer of the truth, the master-builder of human happiness.

Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut edit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt.

No one rejects, dislikes, or avoids pleasure itself, because it is pleasure...

Neque porro quisquam est, qui dolorem ipsum quia voluptas sit amet, consectetur, adipisci velit...
Terminology

- **Terms**: words and compound words that in specific contexts have specific meanings.

- **Termbase**: database consisting of terms and related information, usually in multilingual format.
<table>
<thead>
<tr>
<th>Term</th>
<th>concorrenza sleale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>IT-Italian</td>
</tr>
<tr>
<td>Target</td>
<td>EN-English</td>
</tr>
<tr>
<td>Domain</td>
<td>LAW</td>
</tr>
<tr>
<td>Reliability</td>
<td>3 (reliable)</td>
</tr>
<tr>
<td>Term reference</td>
<td>Enc Giuridica,Treccani,Roma,vol.VII,1988,s.v.concorrenza II;Codice Civile art.2598</td>
</tr>
<tr>
<td>Date</td>
<td>29/09/2009</td>
</tr>
</tbody>
</table>

**Definition**

an attempt to do better than another company by using techniques which are not fair, such as importing foreign goods at very low prices or by wrongly criticising a competitor's products

**Definition reference**

3 (Dict of Accounting, Collin-Joliffe, 1992)

**Term**

unfair competition

**Date**

29/09/2009
Concordance

- **Concordance**: occurrence of a word in a texts together with its context.
- Bilingual concordances show use of words in parallel texts.
She felt very sleepy, when suddenly a White rabbit with pink eyes ran close by her.

nor did Alice think it so unusual to hear the rabbit say to itself "Oh dear! Oh dear! I shall be too late!"

But when the rabbit actually took a watch out of its waistcoat-pocket, and looked at it, and then hurried on, Alice started to her feet, for she remembered that she had never before seen a rabbit with either a waistcoat-pocket or a watch to take out of it, and she ran across the field after it, and was just in time to see it pop down a large rabbit-hole under the hedge.

The rabbit-hole went straight on like a tunnel for some way, and then dipped suddenly down, so suddenly that Alice had no time to think about stopping herself before she found herself falling down what seemed to be a very deep well.
Translation Memory

- Incrementally stores translated segments. Given a new source segment it looks for **perfect or fuzzy matches**

- Matches are ranked (100%-matches on top) and presented to the user as translation **candidates** for post-editing

- A TM can be shared among and simultaneously updated by several translators working on the same project

- TMs model the style and terminology of the customers
Translation Memory

When does it help?
- on highly repetitive, such as technical manuals
- on new versions of previously translated manuals
- when several translators are working on the same project

How does it help?
- speeds up translation process
- ensures consistency across different translators

Limitations
- number of useful matches found is generally small (5-10%)
Machine Translation

Machine translation in general decomposes the language translation process into a sequence of rule applications.

**In statistical MT:**
- the translation process is expressed as a search problem that computes an **optimal** sequence of rules to apply
- translation rules are automatically extracted from a **large parallel corpus** and a **stochastic model** is defined over the translation rules, that is optimised to best fit the data
- according to the employed stochastic model, the sequence of rules may generate **linear or hierarchical** structures.
Machine Translation

When does it help?
- language pairs supported by large parallel data
- translation directions between close languages
- training data represent well task data

How does it help?
- provides good draft translation to start with
- avoid translating easy/repetitive fragments

Limitations
- translations may lack of global coherence
- bad translations cause waste of time, loss of trust
## TM versus MT

<table>
<thead>
<tr>
<th>Capabilities</th>
<th>TM</th>
<th>MT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can it start from scratch?</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Does it improve during usage?</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Can it instantly learn a new translation?</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Does it consider context of the segment?</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Can it retrieve 100% matches?</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Can it create <strong>new</strong> 100% matches?</td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>

TM and MT are rather complementary!
Simple MT Integration

TM backed up by MT

How to evaluate the impact of MT?
Translation cost

Translation projects are quoted on word basis

Price per-word depends on:

- domain
- languages
- urgency
- quality
- TM matches

From a research perspective we are interested in the impact of MT on user experience and productivity
Human productivity

Daily productivity of translators is highly variable … and also translations vary significantly among translations.

To evaluated the impact of MT technology we have to consider both subjective and objective criteria:

- usability, user preferences, …
- variations in productivity
  - **effort**: e.g. human TER
  - **speed**: e.g. word/hour, sec/word
Simple MT Integration

Integrated logging functions in plug-in and run experiments with 2 domains x 2 directions x 4 versions = 16 translators
Simple MT Integration

MT helps, what next?

- **TM**: Blue bars
- **TM+MT**: Light blue bars
- **Time Gain (%)**: Black line

### Language-Domain
- **EN>DE IT**
- **EN>DE LE**
- **EN>IT IT**
- **EN>IT LE**

### Comparison
- **Words/Hour**
  - EN>DE IT: 550
  - EN>DE LE: 650
  - EN>IT IT: 400
  - EN>IT LE: 550

### Time Gain
- **50.00%**
Seamless integration of MT into the CAT workflow

Research:

- **self-tuning** machine translation
- **user-adaptive** machine translation
- **informative** machine translation

Enterprise level open-source CAT tool

Extensive field testing with professional translators
MateCat Tool

Simple. Web based.
**MateCat Tool**

Data collection and logging for in-depth analysis

---

### Job 984 - Editing Log

**Summary**

<table>
<thead>
<tr>
<th>Words</th>
<th>Avg Secs per Word</th>
<th>% of MT</th>
<th>% of TM</th>
<th>% of Time</th>
<th>Total Time to edit</th>
<th>Avg Ed Effort</th>
<th>% of words in too SLOW edits</th>
<th>% of words in too FAST edits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1041</td>
<td>5.4</td>
<td>90%</td>
<td>4%</td>
<td>1%</td>
<td>100h 25m 20s</td>
<td>10%</td>
<td>1%</td>
<td>1%</td>
</tr>
</tbody>
</table>

**Editing Details**

<table>
<thead>
<tr>
<th>Secs/Word</th>
<th>Job ID</th>
<th>Segment ID</th>
<th>Words</th>
<th>Suggestion source</th>
<th>Match percentage</th>
<th>Time-to-edit</th>
<th>Post-editing effort</th>
</tr>
</thead>
<tbody>
<tr>
<td>105.3</td>
<td>984</td>
<td>1044</td>
<td>26</td>
<td>Machine Translation</td>
<td>80%</td>
<td>10m 55s</td>
<td>20%</td>
</tr>
</tbody>
</table>

---

Segment:

You can move a volume to a new I/O group to balance the workload across the system without stopping host activity to the volumes.

Suggestion:

È possibile spostare un volume di I/E in un nuovo gruppo di bilanciare il carico di lavoro tra il sistema host senza dover arrestare l'attività per i volumi.

Translation:

È possibile spostare un volume in un nuovo gruppo I/O per bilanciare il carico di lavoro nel sistema senza arrestare l'attività dell'host sul volume.

DIFF View:

È possibile spostare un volume in un nuovo gruppo I/O per bilanciare il carico di lavoro nell'host senza dover arrestare l'attività per l'host sul volume.

---

...
Cognitive studies of translator behaviour
  - based on key logging and eye tracking
MT research:
  - interactive translation prediction
  - interactive editing
  - adaptive translation models
Open source workbench
Field test by translation agency and volunteers
Interactive MT Prediction

In May, she speaks about her return to town: "I will go there this week, but I will only stay until the 19th, because on the 20th I must present myself at work."

En mayo, habla de su regreso a la ciudad: "Voy allí esta semana, pero me limitaré a quedarse hasta el siglo XIX, porque en la 20ª debo presentarme en el trabajo".

Left-to-right post-editing

MT prediction
Interactive Editing

Multimodal Interaction
1. MT stands for machine translation.
2. Our research aims to make it more useful to translators.
3. All software will be seamlessly integrated into a Web-based CAT tool.
4. Tutto il software verrà integrato in un CAT tool basato su Web.
1. MT stands for machine translation.
2. Our research aims to make it more useful to translators.
3. Our MT technology will be seamlessly integrated into a Web-based CAT tool.
4. All software will be

1. MT è l'acronimo di traduzione automatica.
2.
User-adaptive MT

• Discriminative re-ranking methods

  K. Wäschle et. al, Generative and discriminative method for online adaptation in SMT, MT Summit 2013.

• Cache based language and translation models

  N. Bertoldi et. al, Cache-based online adaptation for machine translation enhanced computer assisted translation, MT Summit 2013.

• Discriminative online adaptation of dense/sparse features

  P. Mathur et al., Online learning approaches in computer assisted translation, WMT 2013.
The goal of this work package is to develop methods and system architectures to adapt statistical machine translation (SMT) systems.

L'objectif de cet atelier de travail est de développer des méthodes et des architectures de système d'adapter les systèmes de traduction automatique statistique (TAS).

We distinguish several types of adaptation:
Informative MT
Informative MT

• Score MT vs TM suggestions
  – Show most useful suggestion in first position

• Filter out bad MT suggestions
  – Avoid translator wasting time, loosing trust

• Provide reliable partial information
  – Show suggestions of important/difficult words
1. MT stands for machine translation.
2. Our research aims to make it more useful to translators.

### Translation matches

<table>
<thead>
<tr>
<th>Source</th>
<th>Target</th>
<th>MT</th>
<th>TM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our research aims to make it more useful to translators.</td>
<td>La nostra ricerca mira a renderla più utili per i traduttori</td>
<td>90%</td>
<td>75%</td>
</tr>
<tr>
<td>Our goal is to make it more useful to interpreters.</td>
<td>Il nostro obiettivo è di renderlo più utili per gli interpreti.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**MateCat Tool**

---

**Informative MT**

**MT and TM suggestions**

**Filtering and ranking**

**MT Server**

- MT decoder
- QE engine

**TM Server**
Open experimental platforms

Easy to install and run post-editing experiments
Google-compliant MT API, MyMemory TM API
comparing different MT engines, and much more
Suited for remote access (e.g. crowdsourcing)
Test Protocol

Day 1
Translation of 50% of doc with MT1 (domain adapted)

Day 2
translation of rest of doc with MT2 (project adapted)
Self-tuning MT

Average gain: 22.25%  
Average gain: 10.71%

This protocol introduces secondary effects: learning curve of users about system and document.
Better Test Protocol

MT systems randomly switched inside the same document: user does not know where the suggestions come from.
Conclusions

- Integration of HT and MT is still an open issue
  - very challenging research problems
- Open experimental infrastructure
  - permits to evaluate how useful MT is
- Interested to try our CAT tools?
  - [www.matecat.com](http://www.matecat.com) simple UI, industry ready
  - [www.casmacat.eu](http://www.casmacat.eu) enhanced UI, research oriented
Thank you!