

ACCEPT

SEVENTH FRAMEWORK PROGRAMME

THEME ICT-2011.4.2(a)

Language Technologies

ACCEPT

Automated Community Content Editing PorTal

www.accept-project.eu

Starting date of the project: 1 January 2012

Overall duration of the project: 36 months

Report on assistance: Report on novel types of assistance to monolingual post-editors

Workpackage n° 7

Name: Monolingual Postediting

Deliverable n° 7.2

Name: Report on assistance: Report on novel types of assistance to monolingual post-editors

Due date: 31 December 2013

Submission date: 19 December 2013

Dissemination level: PU

Organisation name of lead contractor for this deliverable: University of Edinburgh

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 288769.



Contents

Objectives of the Deliverable 3

Display of Alternative Word Translations 3

 Motivation 3

 Implementation 4

 Conclusions 4

Confidence Measures 4

 Motivation 4

 Conclusions 5

Request of Alternative Translations of Phrases 5

 Motivation 5

 Implementation 5

 Conclusions 6

References 6

Report on assistance: Report on novel types of assistance to monolingual post-editors

Objectives of the Deliverable

The objective of WP7 is to collect edits from users who perform monolingual post-editing in a volunteer or community collaboration context. The main objective of this deliverable is to give an overview of the types of assistance that are being explored in the ACCEPT project in this type of use case.

The project collaborates closely with the EU project CASMACAT which runs contemporaneously and shares one partner (University of Edinburgh), the leader for this work package. One of the main objectives of the CASMACAT project (besides cognitive studies of translator behaviour, adapting machine translation systems to their usage in CAT tools, and the development of an open source workbench) is the development of novel types of assistance for translators.

Given the overlap of objectives of the CASMACAT project, the ACCEPT project is mainly focused on the exploration of use of novel types of assistance in the community translation setting, rather than developing novel types of assistance in parallel. This is reflected in the low number of person months allocated to this WP (8 PM in total over 3 years) and its focus on user studies, which are reported in Deliverables 7.1.1, 7.1.2, and 7.1.3.

This deliverable on assistance for community translators provides an overview of types of assistance that have been considered, are currently explored, and are planned for use in the ACCEPT project for the use by monolingual post-editors. Overall, the focus of this assistance is to indicate the reliability of the machine translation output (the starting point for the monolingual post-editor) and the provision of alternatives.

Specifically, we describe here the following types of assistance:

- Display of alternative translations for words
- Confidence measures
- On-demand generation of alternative translations for arbitrary phrases.

Use of this type of assistance requires significant implementation effort, part of which is described in the deliverables for WP5 *Portal Integration*.

Display of Alternative Word Translations

Motivation

Since the monolingual post-editor has either no access to the source document or cannot glean any insights from it, they are limited to the words that are present in the machine translation output. This

is especially a disadvantage if some of the word choices of the machine translation system are wrong.

The ambiguity of source language words is often unexpected to the monolingual translator. A machine translation system may produce the word *bench* when it should have produced the word *bank* because these translated words have the same source translation in German (*die Bank*). The translator hopefully will be able to spot that a word is out of place (*I invested £20,000 in a bench*) but may not be able to do much with this information.

Thus, it would be useful to the translator to see alternative translations that the machine translation system considered but discarded.

Implementation

While the number of possible translations for input words and phrases is very large, it would be very confusing to display them all to the user at the same time. Therefore, we decided to implement a version that displays alternative word translations only on request by the user.

For more details on the implementation, please refer to Deliverable 7.1.1 (Year 1).

Conclusions

We explored the use of this type of assistance in a user study in Year 1. When offering the translation options to monolingual translators working on Symantec forum data, we can measure faster editing both on German and English. However, some caution is necessary since different texts were post-edited under the different conditions.

Having said that, we have every reason to believe that this is useful functionality for monolingual translators.

Confidence Measures

Motivation

An important concern is that monolingual post-editing will accept wrongly translated output, because it sounds plausible. The role of a strong language model may lead to conflicting outcomes here: while it helps the machine translation system to produce more fluent output, it also leads to more well-formed target language that gives readers a wrong sense of confidence in the translation.

There has been increased interest in the research community in work on confidence measures (now also commonly referred to as “quality estimation”), such as the shared task at the 2013 Workshop on Statistical Machine Translation, co-organized by the related projects MATECAT and CASMACAT. Confidence measures indicate how likely a translation is actually to be correct. Just as human translators (or humans working on any task) often have a sense of how well they performed on a task (*“I would say ‘A’, but I really do not know”*), these automatic methods would do the same for a machine translation system.

Such confidence measures would inform the monolingual post-editor about the reliability of translations. A translation that looks fluent and plausible, but has a low confidence score would

prompt the user to take a closer look, maybe question its veracity given the overall document context.

Confidence measures have been used in user studies in the related EU projects, but without much success. In an in-house field trial of the CASMACAT project, the bilingual translators stated that they would like such a feature, but the current quality of the confidence measures is not high enough to be useful¹. The results of the WMT shared task, especially on word-level confidence measures prompted similarly skeptical assessments (Bojar et al., 2013).

Conclusions

We are currently not pursuing the use of confidence measures in monolingual post-editing, due to concerns about the lack of reliability of these measures and limited resources to implement and test such measures.

Request of Alternative Translations of Phrases

Motivation

Given the positive results of offering alternative word translations to monolingual post-editors, a promising extension is to offer the re-translation of larger parts of a translated sentence.

Such a functionality may offer several purposes, the main two being (1) displaying alternative lexical choices and formulations may inspire the post-editor to find more idiomatic language, and (2) making the ambiguity of the source visible by indicating each meaning with a different translation.

Implementation

While the display of alternative word translations can be obtained from the translation table of the machine translation model, the selection of a larger part of the sentence requires more complex processing:

- The monolingual post-editor marks up a sequence of words in the current translation of the sentence, for which she wants to see alternative translations.
- The target language word sequence has to be mapped to the source, thus identifying a sequence of source words.
- We then consult the phrase translation table for these source words to generate alternative translation in a simplified version of the full machine translation decoding process.
- Given a list of possible translations, we apply a number of filters and scorers to select a list of 5 alternative translations shown to the user.

There are several ways this process could be refined and extended. For instance, given the availability of monolingual paraphrase corpora (Ganitkevitch et al., 2013) we could also extend the list of possible translations by looking for synonym formulations.

¹ See CASMACAT Deliverable 1.2.

Both automatic evaluation and manual evaluation of the approach has been quite promising, with about 50% accuracy of one of the displayed 5 translations according to manual bilingual judges with access to the source.²

The paraphrasing service is currently being implemented as a server process that responds to API calls. Note that this requires significant re-engineering, since the ACCEPT portal is less strongly tied to the machine translation system as in the current CAT applications.

Conclusions

Given the encouraging results from intrinsic evaluation, we are currently planning user studies of this type of assistance early in year 3 of the project.

References

- Ondrej Bojar, Christian Buck, Chris Callison-Burch, Christian Federmann, Barry Haddow, Philipp Koehn, Christof Monz, Matt Post, Radu Soricut, and Lucia Specia :
Findings of the 2013 Workshop on Statistical Machine Translation
In *Proceedings of the Eighth Workshop on Statistical Machine Translation, 2013*.
- Juri Ganitkevitch, Benjamin Van Durme, Chris Callison-Burch:
PPDB: The Paraphrase Database
In *Proceedings of NAACL-HLT, 2013*.

² See CASMACAT Deliverable 3.2.