A Practical Evaluation of an Integrated Translation Tool during a Large Scale Localisation Project

Reinhard Schäler

Department of Computer Science, University College Dublin,
Belfield, Dublin 4, Ireland
reinhard@nova.ucd.ie

Abstract

This paper reports on the assessment of computer assisted translation tools for a large localisation company and the practical evaluation of one such tool.

Introduction

Over the past two years, a number of integrated translation tools which include a translation editor, an on-line terminology database and a translation memory system became commercially available. Most of these tools were targeted at the software localisation industry, where the ability to produce consistent translations of highly repetitive texts within tight deadlines determines valuable market shares in an increasingly competitive market.

The level of sophistication these tools have reached is still quite embryonic. Potential users have, in most cases, no practical experience with the issues involved in the application of such tools. They seek advice from experts who they feel can evaluate the systems in question for them, but even the experts have not yet agreed on a standard evaluation approach.

This paper reports on the assessment of computer assisted translation tools for a large localisation company and the practical evaluation of one such tool during a large scale localisation project which involved the translation of around 1 million words into 9 languages at several different locations.

The localisation of software and software documentation, for a variety of reasons, is an ideal application area for the new generation of electronic translation tools (Translators' Workbenches), which integrate a translation editor, an electronic dictionary look-up system and a translation memory feature (cf. Ishida 1994):

Text in software and the accompanying documentation is highly repetitive (up to 30%).

Translation often starts at a β-stage. When the translation has to be updated to the final production version, Translators' Workbenches can detect modifications made to the original source file and automatically insert text that has previously been translated.

New releases of the same product have to be localised at ever decreasing intervals (every 9 months approximately).

In this paper, we report that even one of the most sophisticated tools currently available still contains a considerable number of undesirable features many of which cannot, realistically, be identified during an initial test period. We also note that there are major critical issues not connected to any particular tool which have up to now been overlooked by both developers and potential users.

Project Background

In 1992, we were approached by one of the major localisation companies in Ireland to assist them with the assessment of two Translators' Workbenches. Our initial brief was to recommend one of the two translation tools for use in the company's operation.

There are no standard evaluation procedures for potential buyers of translation memory systems and none of the evaluation methods proposed for machine translation systems (cf. Machine Translation 1993) is suitable because they are either too costly (translation memory systems cost a fraction of MT systems) or they are simply inappropriate (because of the conceptual difference between MT and CAT tools).

Our assessment, therefore, had to rely on our knowledge of the localisation industry and its requirements, and was limited by the resources at our disposal. Our recommendation of the TRADOS Translator's Workbench II (TWII) was based largely on practical issues of cost and portability.

In autumn 1993, the company decided to use the tool for a large scale project (10 languages, roughly one million words/language, 60 translators 10 countries). As far as we are aware, this was the first time that TWII (one of the market leaders in the area of CAT) was put to a major practical test.

In order to help with evaluation of the tool, the translators were asked to fill in a one page question-

naire. Unfortunately, the response rate was very low even after we had sent out a reminder (2 out of 10)[1]. Unsurprisingly, by far the best source of information about the performance of the tools under industrial conditions proved to be the queries and the Performance Reports forwarded to our support engineers. It is these reports, in addition to an, at times, overwhelming number of queries communicated by telephone and fax, on which the following practical evaluation of the TRADOS Translators' Workbench II is based.

A Practical Evaluation

All modules, the translation editor, the MultiTerm terminology database, the translation memory module and the utility programs, revealed undesirable features over the course of the project. Of these by far the most serious flaws were discovered in the Translation Memory (TM) module which had been expected to be the most useful module. Because of the impact this had on the overall performance of the tool, we will concentrate on some of the unexpected features of the TM module.

Segmentation of the source file: A number of problems were caused by the way in which TWII segments a file. What TWII identifies as a segment is rigidly defined by the program (TRADOS 1992). Apart from adding entries to a user definable abbreviation list, there is no way to modify this definition to cater for project specific circumstances. Should a text not contain any formatting tags, TWII will simply fail to segment it properly. This is unacceptable in view of the fact that translators often have to deal with incorrectly or partially formatted text, and TWII should be capable of correctly segmenting plain ASCII files.

Definition of a 100% match: The automatic search for and the "translation" of previously translated sentences is generally seen as probably the most useful feature of this type of program (S. Bell 1993). Our experience showed that the definition of a 100% match is not quite as clear as the developers and, admittedly, we ourselves had believed. Indeed, the definition of a 100% match is one of the most problematic features.

Storing Target Language Sentence Pairs: No safety mechanisms prevented translators and editors form corrupting TM by storing target language sentence pairs in TM.

The most frequent interference with the integrity of translation memories occured when a translated segment was edited outside TM mode. This problem had been explained to translators but under the

This was in line with the experience made during other MT evaluation projects (Vasconcellos 1994) and confirmed the view that questionnaires are not an efficient way to gather information from users.

pressure of deadlines and the need to produce translations quickly, the warning was often overlooked. Once a segment had been edited outside TM mode, the automatic "translation" features of TWII became useless.

Translation memory attributes: TWII can only store a single translation for any particular source segment into TM, unless translators use different TM attributes for their translations. Sometimes, however, it proved necessary to translate the same source segment occurring twice into two different target segments. This again made the automatic "translation" of source text impossible as TWII is not able to switch between different attributes automatically.

Conclusions

The most interesting feedback we received was the comment that translation memory based systems are probably very useful for inexperienced translators who could draw on the knowledge and the examples supplied by the translation memory. Professionals, however, who translate up to 7,000 words a day using a dictaphone and the services of two highly skilled typists would almost certainly be slowed down by the "cumbersome" translation process enforced by the program.

Without any doubt, the new generation of integrated translation tools are of great value to the localisation industry. They can speed up the translation process, improve the quality of the translation by insuring a higher degree of consistency and bring down translation costs significantly.

However, our experience with the TRADOS Translators' Workbench II, probably the best integrated tool currently available, has shown that, while translators will have to acquire the necessary technical skills to work effectively with these tools, developers will have to work more closely with their customers in order to produce tools which meet their requirements.

REFERENCES

S. Bell (1993) "Translators' Workbench II", *Language International* 5(3), 5-7

R. Ishida (1994) "Future Translation Workbenches: Some Essential Requirements", *The Localisation Industry Standards Association (LISA) Forum Newsletter* (III) February, 3-12

Machine Translation (1993), Special Issue on the Evaluation of MT Systems, (8)1-2

TRADOS (1992) *Translators' Workbench II, User's Guide*, Stuttgart.

M. Vasconcellos (1994) "The Current State of MT Usage", *The Localisation Industry Standards Association (LISA) Forum Newsletter*, (III) February, 21-29