Gradable quality translations through mutualization of human translation and revision, and UNL-based MT and coedition

Christian Boitet

GETA, laboratoire CLIPS,
385 rue de la bibliothèque - BP 53, 38041 Grenoble Cedex 9 - France
Christian.Boitet@imag.fr

Abstract. Translation of specialized information for end users into many languages is necessary, whether it concerns agriculture, health, etc. The quality of translations must be gradable, from poor for non-essential parts to very good for crucial parts, and translated segments should be accompanied with a measured and certified "quality level". We sketch an organization where this can be obtained through a combination of "mutualized" human work and automatic NLP techniques, using the UNL language of "anglosemantic" graphs as a "pivot". Building the necessary multilingual lexical data base can be done in a mutualized way, and all these functions should be integrated in a "Montaigne" environment allowing users to access information through a browser and to switch to translating or postediting and back.

Keywords: MT, HAMT, UNL, interactive disambiguation, pivot editing, revision sharing, interlingual representation, text-UNL coedition, multilingual communication, mutualization.

1 Introduction

Translation of specialized information into many languages is necessary, notably in agriculture, but also for health and other domains, because it is often crucial for final users, who don't master the source language. Quality should be very high, at least for the crucial parts. In many cases, also, it is urgent to use the information, and only automated translation could offer a solution. At the same time, resources are scarce, especially to produce high quality translations. Does that mean that nothing can be done ? No, of course.

The first idea which comes to mind is to "mutualize" the translation effort. That becomes possible thanks to the wide availability of Internet. There is always a minority of targeted readers who understand the source language, and could produce good translations. Also, they would translate only a fraction of their time, so that, even with machine helps which may be developed by and by, it is reasonable to assume that not every part of every document could be translated in this way. Why not, then, use "rough" machine translation (MT), or even "active reading helps" (annotations of