Towards a systematic process in the use of UNL to support multilingual services

Jesús Cardeñosa, Carolina Gallardo, Edmundo Tovar

Departamento de Inteligencia Artificial- Facultad de Informática
Universidad Politécnica de Madrid
Campo de Montegancedo, s/n
28660 Madrid- Spain
{carde, carolina, etovar}@opera.dia.fi.upm.es

Abstract. The UNL Programme of the United Nations University (UNU) was launched in 1996 aiming at the elimination of linguistic barriers in Internet. Now, eight years later, UNL is not ready to support real applications due to several circumstances. This eight-year period can be divided in two: a first four-year period devoted to the formal definition of UNL as a formal language (under the sponsorship of the Institute of Advanced Studies –IAS– of the UNU) and the remaining four years devoted to the technical experimentation of UNL. A new period is starting right now, which could be the period of maturity at all levels, especially at technical and business levels. In this paper, the authors summarize the more significant experiences until now, their conclusions and the set of procedures to produce marketable multilingual services. This kind of work will be the work of the UNL consortium during the next two years before launching UNL to the market.

1   Introduction

The natural evolution of UNL as a project and as a Programme is the support of useful applications for a multilingual society. Apart from other uses of UNL, like cross-lingual information retrieval or support for ontologies, the more understable use and possibly the easiest application, is the support of multilingual services, that is, to represent contents written in any language and to generate any other language [1].

UNL is not conceived to become a (fully automatic) machine translation system (MT hereafter). Up to date, MT systems based on the transfer architecture have achieved reasonable results, always involving pairs of languages. These systems are somehow handicapped by their *language coverage*. In other words, a transfer based system involving N languages requires the development of N*(N-1) systems, that ends up with the consequent combinatorial explosion of the number of systems to be developed as the number of languages grows.

On the other hand, interlingua-based MT systems show, in principle, a highly attractive advantage over transfer systems: interlingua-based systems do not grow exponentially as the number of language increases since for a system to support N languages, only 2*N systems have to be developed. The ATLAS system [2] and the PIVOT system [3] in open domains, and Mikrokosmos [4] and Kant [5] in restricted