Using WordNet for linking UWs to the UNL UW System

Luis Iraola

Facultad de Informática, Universidad Politécnica de Madrid
Campus de Montegancedo, 28660 Madrid, Spain
luis@opera.dia.fi.upm.es

Abstract. This paper presents the work done with the Spanish-UNL dictionary compiled at the Spanish Language Centre in order to enrich the universal words it contained with the supplementary semantic information required to produce a master entries dictionary. Focusing on a subset of the Spanish-UNL dictionary, namely on the substantives it contains, the work has consisted in automatically enrich the universal word associated with each substantive with the semantic information required to link the universal word to the Universal Word System. For this process, WordNet has been employed as an external source of semantic information and used in addition to semantic features already present in the dictionary. The results achieved are not final and further work is required for a fully automatic, high quality semantic enrichment of the current entries. However, the work done shows the fruitfulness of the approach and its outcome has contributed to the creation of a master entries dictionary.

1   Introduction

A UNL dictionary in which language entries are associated with universal words (UW for short) can be viewed as a repository of UWs and as such does not organise its contents in any way. It links a set of UWs with lexical items of a specific language, each entry having no relation with any other. The necessity of establishing certain relations between UWs arises when considering several desirable features of the UNL system:

– Setting the combinatory possibilities of each UW with respect to any other UW regarding the conceptual relations that may link them and the attributes they may accept.
– Enabling a "fall-back" generation mechanism for those UWs that are not linked with head words in a given language at a given time. Those UWs would be replaced with semantically close but linked UWs so allowing generation to continue.

In order to support these features, a network with the set of UWs as nodes and semantic relations as arcs has been proposed. Such network is called the UNL UW System [1, 2]. Therefore, and in order to build the UW System, UNL Language Centres have to modify their respective UNL language dictionaries for including such new information. Once modified, the new master entries dictionaries will be the repository from which current language dictionaries will be produced as well as the UNL Knowledge Base will be created.