iGLU

Written by Ronaldo Martins

Thursday, 18 February 2010 10:07 - Last Updated Monday, 13 September 2010 14:53

The UNDL Foundation has launched the project **i**(**GL**>**U**) - from glosses into UNL - which aims at UNL-izing the glosses of 27,255 entries extracted from an abridged version of the WordNet3.0. Results are expected to be incorporated into the UNL Knowledge Base (UNL KB), which codifies the most systematic part of the meaning conveyed by natural language words.

The WordNet is a large lexical database of English, where nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), which are interlinked by means of conceptual-semantic and lexical relations. The WordNet was developed at the Cognitive Science Laboratory of Princeton University, and has been widely used for computational linguistics and natural language processing. The version 3.0 was released in December 2006 and contains 117,659 synsets, each expressing a distinct concept.

The project iGLU intends to map WordNet glosses from English into UNL. The project is divided into two main phases: the first one (iGLU#1) addresses a subset of 27,255 synsets and is supposed to be carried out in a predominantly human basis; the second one (iGLU#2) focuses on the remaining 90,404 synsets and it is expected to be mainly automatic. In iGLU#1, linguists are supposed to UNL-ize WordNet definitions through the UNL Editor, a graph-based UNL authoring tool available at the UNL

dev

. Decisions are stored in a UNL-ization memory, which comprises mappings between lexical items of English and Universal Words. Information on attributes and relations are also encoded. These data will be used in the second phase, when the UNL-ization process is expected to be performed by IAN - the UNDL Foundation Interactive ANalyzer -, under development. IAN requires much less human intervention than the UNL Editor, and it is a first step towards a fully-automatic natural language analysis system.

Results of the project iGLU are expected to be used not only in compiling the UNL-ization memory, but also in populating the UNL Knowledge Base, which is an essential part of the architecture of the UNL system. It will improve the quality of word sense disambiguation and enhance the capability of information retrieval and extraction through UNL.

The iGL>U is an open and public project, and its results will be available for all. Public discussions will be available at the
[UNLization forum](#)
, and final results will be available through the
[UNL](#)
arium
. In order to join the project, you must hold a CUP - Certificate of Proficiency in UNL, which is

iGLU

Written by Ronaldo Martins
Thursday, 18 February 2010 10:07 - Last Updated Monday, 13 September 2010 14:53

electronically issued by
 [VALERIE](#)

.