

Hermeto: A NL-UNL Enconverting Environment

Ronaldo Martins, Ricardo Hasegawa and M. Graças V. Nunes

Núcleo Interinstitucional de Lingüística Computacional (NILC)
Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação
Av. do Trabalhador São-Carlense, 400. CEP 13560-970. São Carlos – SP – Brasil
ronaldo@nilc.icmc.usp.br; gracac@icmc.usp.br; rh@nilcicmc.usp.br
<http://www.nilc.icmc.usp.br>

Abstract. This paper aims at presenting and describing HERMETO, a computational environment for fully-automatic, both syntactic and semantic, natural language analysis. HERMETO converts a list structure into a network structure, and can be used to enconvert from any natural language into the Universal Networking Language (UNL). As a language-independent platform, HERMETO should be parameterized for each language, in a way very close to the one required by the UNL Center's EnConverter. However, HERMETO brings together three special distinctive features: 1) it takes rather high-level syntactic and semantic grammars; 2) its dictionaries support attribute-value pair assignments; and 3) its user-friendly interface comprises debug, compiling and editing facilities. In this sense, HERMETO is said to provide a better environment for the automatic production of UNL expressions.

1 Introduction

In the UNL System [1], natural language (automatic) analysis has been carried out either by the EnConverter (EnCo) [2] or, more recently, by the Universal Parser (UP) [3], both provided by the UNL Center. In the first case, enconverting from natural language (NL) to Universal Networking Language (UNL) is supposed to be conducted in a fully-automatic way, whereas in the second case a full-fledged human tagging of the input text should be carried out before NL analysis is triggered. In both cases, results have not been adequate. EnCo's grammar formalism, as well as UP's tagging needs, are rather low-level, and requires a human expertise seldom available. In what follows, we present an alternative analysis system, HERMETO, developed at the Interinstitutional Center for Computational Linguistics (NILC), in Sao Carlos, Brazil, which has been used for automatic enconverting from English and Brazilian Portuguese into UNL. Due to its interface debugging and editing facilities, along with its high-level syntactic and semantic grammar and its dictionary structure, it is claimed that HERMETO may provide a more user-friendly environment for the production of UNL expressions than EnCo and UP.

The structure of this paper is as follows. The second section, on motivation, addresses the context in which the HERMETO initiative was conceived and the goals ascribed to the system. The third section presents HERMETO's architecture. HERMETO's functioning is briefly detailed in section four (on resources) and five