論文の内容の要旨

Semi-Automatic Conversion of Natural Language Text into Concept Description Language (CDL)
（自然言語テキストの Concept Description Language (CDL) への半自動変換）

氏名　タコア　レネヴェイ　フランシスコ　アントニオ
Tacoa　Renevey　Francisco　Antonio

Semantic Computing is a field that combines elements of semantic analysis, natural language processing and data mining. Semantic Computing aims to solve three fundamental problems: to understand and express the intentions of users in a machine-understandable format; to understand and express the meaning of computational contents in a machine-understandable format; and mapping the semantics of user with that of content for the purpose of content retrieval, management, creation, etc. Machine-understandable format is very important for Semantic Computing since it is the way for computers to understand and communicate with users and vice-versa. At the present time, the Concept Description Language (CDL) is an artificial language that aims to become the standard to represent the meaning expressed in natural language texts in the form of a semantic graph, where nodes are unambiguous concepts and edges are relations between concepts. Unfortunately, there are not many tools to demonstrate its functionality and the few tools that are available are based on a fully automatic process (which increases the risk of incorrect conversion) or manual process (which takes long time for big amounts of text). In this work I propose a conversion mechanism to address the problem of converting natural languages into Concept Description Language (CDL) in a semi-automatic way. Since natural languages are ambiguous, I have included a Word Sense Disambiguation (WSD) method that selects sense candidates for words only when they have a high probability of being correct. Otherwise, it asks the user to select the right senses. The process begins with syntactic analysis, followed by WSD, conversion into semantic information and ends with the output in the available CDL formats: textual (understandable by computers) and graphical (understandable by humans). An optional step is to store the conversion into an XML file. One contribution of this work is that the methodology requires little human effort and improves the correctness of the conversion into CDL. Therefore, it will allow researchers to construct and evaluate CDL-annotated datasets efficiently.