Knowledge Engineering Suite: A Tool to Create Ontologies for an Automatic Knowledge Representation in Intelligent Systems

Tânia C. D'Agostini Bueno, Hugo cesar Hoeschl[2], André Bortolon[1], Eduardo Mattos[1], Cristina Souza Santos[1]

[1] Instituto de Governo Eletrônico, Inteligência Jurídica e Sistemas

[2] Universidade Livre de Florianópolis

Abstract. The present work is focused on a computational structure called Knowledge Engineering Suite, an ontological engineering tool to support the construction of ontologies to assign an automatic text indexation of documents. This tool is a collaborative environment and was based on observations made on the Semantic Web, UNL (Universal Networking Language) and WordNet. We use both a knowledge representation technique called DCKR and psychoanalytic studies, focused mainly on Lacan and his language theory to organize ontologies.

1 Introduction

Most recently, the notion of ontology is being so popular in fields such as intelligent information integration, information retrieval on the Internet, and knowledge management. The reason is partly due to what they promise: a shared and common understanding of some domain that can be communicated through people and computers [1]. Different developments with a worldwide range have a reference in cooperative work such as a UNL (Universal Networking Language) [2], WordNet [3] and Semantic Web [4] through the construction of ontologies using collaborative tools. In the present development, we create a tool to support the Knowledge Engineering process by assisting developers in the design and implementation of ontologies in a specific domain. This tool, called Knowledge Engineering Suite, allows the organization of a knowledge base established on the relationship between relevant expressions from a context. In earlier works, we used a methodology called DCKR (Dynamically Contextualized Knowledge Representation) [5]. DCKR allows the construction of a knowledge base, improving the construction of the domain ontology and the automatic representation of cases in knowledge-based systems, either in the juridical area [6], or in a knowledge management domain [7]. The main intention of this process is to allow an automatic process of text indexing, on the basis of a controlled vocabulary (ontologies). DCKR is a methodology of knowledge representation whose approach is centered in a dynamic process acquisition of the

C. D'Agostini Bueno T., cesar Hoeschl H., Bortolon A., Mattos E. and Souza Santos C. (2005).
Knowledge Engineering Suite: A Tool to Create Ontologies for an Automatic Knowledge Representation in Intelligent Systems.
In *Proceedings of the 2nd International Workshop on Natural Language Understanding and Cognitive Science*, pages 168-171
DOI: 10.5220/0002572401680171

knowledge of texts, defined through the elaboration of a controlled vocabulary and a dictionary of terms, associated to an analysis of frequency of the words and indicative expressions of the context.

2 The Knowledge Engineering Suite

The Knowledge Engineering Suite is an Ontological Engineering Tool for collaborative-networked works on the Web, built to facilitate knowledge sharing between the Knowledge Engineering team and the Specialist team. The Suite allows the building of relationships between complex terms, considering its concepts in the specific domain of the application.

The Suite is an editor of ontologies structured in a way to allow an automatic text indexation in Knowledge Based Systems.

This computational environment of shared access has two main objectives: organization and representation of knowledge, and updating of the Knowledge Base.

It is basically composed by four modules, which are:

1. Register. It allows entering with new *indicative expressions*. The user defines the topic and sub-topic in which s/he will insert a new indicative expression. A domain can be categorized in innumerable topics and sub-topics;

2. Search. It informs about other *indicative expressions* already registered on the base, which have some phonetic similarity with the term typed. This tool allows the verification of possible typing errors, besides preventing the registration of the same term more than once. It is a search system based on similarity. It supplies the user with a list of similar indicative expressions present in the knowledge base in alphabetical order after consultation made by the user. It is used in the registers, in the edition and the administration module. The *indicative expressions* can be registered in multiple topics, with different relations;

3. Relationship Editor. Ontology construction (insertion and consistency checking). It allows the building of the relationship tree, always considering the similarity between all the terms registered and the ones already existing on the base. These relationships allow Knowledge Based Systems to expand the search context. The fields with all the relationships available to be formed are presented. They are the following: -synonyms; -Related terms; -This is type of; - It is a type of this; - This is part of; - It is part of this. Each relationship has a weight related to the defined indicative expression in the search by the user (synonyms –0,99; related terms – 0,75; homonyms, hyponyms, hypernyms and meronyms – 0,4). Therefore, the organization of the tree allows the dynamic definition of the weights of the *indicative expressions* according to the entrance of the user, the same *indicative expressions* can be different types of relations, according to domain allowed.

4. Administration Environment. The knowledge integration and the validation between words are made in accordance with the context of topics and sub topics This topic is organized into three levels: - High Level - which allows us to insert topics and sub topics, to validate exclusions, to include and exclude users, to verify productivity of each user and to verify descriptions of the dictionaries, topics and sub topics and indicative expressions; - Medium level- which allows to verify productivity and historical data; and, Low level- which allows to verify descriptions.

The definition of the related concepts implies a wide research or specialists experienced in the matter.

An identifiable limit doesn't exist for this attribute. Then it is important to observe the application of the terms in concrete cases. The specialists are doing this task by a technological structure and by a methodology called Mind Engineering [8].

All the concepts, linked to each other, generate a semantic-like network. This network improves the Knowledge based capacity of the systems to recognize concepts even if is not in the text. Levels, indicating the "distance" between two concepts, organize the network. These levels are used later in the similarity measure.

It is important to highlight that this structure of contextualized ontologies allows automatic information indexed by the system and a knowledge acquisition that gives more qualitative answers in the retrieval process.

3 Elaborating synchronicity in a collaborative networked organization

There are many different techniques of Knowledge Acquisition. We created Mind engineering to help developing the following process (DCKR methodology): 1. Inventory of the entire domain (classification of all sources of digital information that will be in the system database). 2. Application of the word frequency extractor based on the database inventoried; 3. Comparison between extractor results with the specialist's needs. 4. Construction of a representative vocabulary of the domain, by the specialist and knowledge engineers. 5. Application of the semantic extractor tool on the database; using the representative vocabulary (indicative expressions). 6. Definition of a list of words based on the evaluation of the results of the frequency of the indicative expressions found in the inventory. 7. Construction of the ontologies in the Knowledge Engineering Suite based on this controlled vocabulary. 8. Definition of synonyms, related terms, homonyms, hyponyms, hypernyms and meronyms.

That is, it did not have any synchronization problems, therefore the deep knowledge of the area specialists of the AI technique that was being applied in the system modeling (e.g., Case-Based Reasoning) [9] allowed a transference of knowledge for the computational language in a very positive way for the final target of the systems.

Basically, Mind Engineering is a process that involves the study of people, processes and technologies, through three premises: 1. knowledge sharing; 2. visualization and 3. relevance definition. It is the synchronization of these factors with an only objective: to allow knowledge or expertise on a certain domain to be totally understood through a computational system, more specifically an ontological engineering tool acting as a mechanism of knowledge acquisition.

The continuous sharing of the established visions makes the specialists and engineers work better in cooperation in the construction of the ontologies of the domain. This productive process is continuous and can establish changes in elapsing the implantation of the system.

4 Conclusions

The Knowledge Engineering Suite enables a cooperative work among people in different places, structuring a continuous knowledge base and easy visualization (knowledge tree) through relationship nets and supplies an exceptional coherence among the semantic relations of those that are called 'indicative expressions', mainly by the support of all this computational structure during the process. This allowed the knowledge engineer and the specialist to develop much more than the knowledge of the domain, but abilities such as conscience itself, disciplines, persistence, and empathy.

References

1. Duineveld, A. J. et al, 1999. WonderTools? A comparative study of ontological engineering tools. Twelfth Workshop on Knowledge Acquisition, Modeling and Management.Voyager Inn, Banff, Alberta, Canada.
2. UNL. Universal Networking Language. Available in: http://www.unl.ias.unu.edu/unlsys/index.html. Access in: 19 jan. 2004.
3. Wordnet. Available in: http://www.cogsci.princeton.edu/~wn/. Access in: 19 jan. 2004.
4. Semantic Web. Available in: http://www.w3.org/2001/sw/. Access in: 19 jan. 2004.
5. Hoeschl, Hugo. C. et al, 2003. Structured Contextual Search For The Un Security Council. Proceedings of the fifth International Conference On Enterprise Information Systems. Anger, France, v.2. p.100 – 107.
6. Bueno, Tânia Cristina D'agostini; Hoeschl, Hugo Cesar; Bortolon, André; Mattos, Eduardo da Silva; Ribeiro, Marcelo Stopanowski. Analyzing the use of dynamic weights in legal case based system. In: Ninth International Conference On Artificial Intelligence And Law, 2003, Edimburgo. Proceedings of the Conference. New York: ACM, 2003. v. 1, p. 136-141.
7. RIBEIRO, Marcelo Stopanovski; MATTOS, Eduardo da Silva; BUENO, Tânia Cristina D' Agostini; HOESCHL, Hugo Cesar. KMAI- Knowledge Management With Artificial Intelligence. In: The Symposium on Professional Practice In AI, In The First IFIP International Conference on Artificial Intelligence Application and Innovations, 2004, Toulouse. 2004.
8. Bueno, Tania C. D., Hoeschl, Hugo, Bortolon, Andre, Barcia, Ricardo. Engineering of Minds: The Synchronicity Between Artificial Intelligence and the Management of Knowledge in Collaborative Networked Organizations. Proceedings IADIS International Conference www/internet 2004. Madrid, Espanha.
9. Kolodner, J. Case-Based Reasoning. Morgan Kaufmann, Los High, CA. 1993.