Project no. 027391

Project acronym: LT4eL
Project title: Language Technology for eLearning

Instrument Specific Targeted Research Project

Thematic Priority Information Society Technology

D3.1b Ontology validation on the basis of (multilingual) search – first cycle

Due date of deliverable: 30-11-2007
Actual submission date: 21-12-2007

Start date of project: 1-12-2005 Duration: 30 Months

Organisation name of lead contractor for this deliverable: Bulgarian Academy of Sciences (IPP-BAS)

Revision [1]

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	x
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Ontology Validation Deliverable

Contents

1 Title

D3.1b Ontology validation on the basis of (multilingual) search – first cycle

2 Summary

During the first year of the project we developed a methodology for ontology creation which reflected both - the most common ontology engineering approaches and the goal of the project , namely - studying the role of ontologies in eLearning. We constructed an ontology following this methodology.

We created an English lexicon aligned to the ontology. For each domain class in the ontology we presented at least one term in the lexicon. This lexicon then was used for the annotation of the English learning objects.

In parallel to this task we performed a pilot study on creation of lexicons and their alignment to the ontology in several languages - Bulgarian, Dutch, German and Romanian. We created 200 entries for each language.

On the basis of the work done during the first year we proceeded with the work during the second year accordingly. The following tasks were performed:

- Creation of Lexicons, which covered the whole ontology and all the languages;
- Creation of Concept Annotation Grammars for all the languages;
- Semantic annotation of Learning objects with concepts;
- Formal evaluation of ontology-based semantic search vs. textual search;
- Crosslingual search and integration within ILIAS.
- Ontology revision;

In this deliverable we describe the work done with respect to the tasks, described above. The deliverable is divided into two parts: (1) the first part deals with the ontology validation and evolution; (2) the second part is devoted to the lexicon creation and the annotation of learning objects.

Here we first repeat the ontology creation methodology, which was developed in the first year. Then we present some approaches to ontology evaluation and how we applied it to our ontology. On the base of the evaluation results we revised our ontology.

3 Ontology Creation Methodology

In this section we present the steps of the ontology construction, which were adopted in the project:

- Processing of the Keywords

We started with the list of keywords, annotated by all partners in their corpora of learning materials. The reason to do this is that the ontology would be used within the project for the annotation of the learning objects with ontological information (e. g. concepts, instances).

All the keywords were translated into English, because English is the common language for the partners in the project.

- Initial classification of keywords

The keywords were sorted by domains. From them we selected only the keywords that fell into the domain of Computer Science.

- Definition collection

In order to register all (or more important) the meanings of the keywords that are ambiguous in the domain, we collected definitions of them from different sources. We added more than one definition for a meaning if the different sources highlighted different aspects of the meaning. For example, if the meaning is about a device, one definition might highlight the function, another - the material, from which the device was constructed, etc.

- Definition selection and Sense differentiation

From the definitions collected on Internet we selected one most appropriate definition or created a new definition, based on several definitions from different sources. When a keyword had more than one meaning we produced several definitions.

- Formalisation of the meanings

The informal definitions from the previous step were formalized as concepts and represented as OWL-DL statements.

- Link to an upper ontology

We linked the domain ontology to an upper ontology (DOLCE) in order to get benefits from the methodology of its construction (OntoClean methodology), and to allow the inheritance of some general relations from it.

As a tool for mapping the domain concepts defined in step 2 to the concepts in DOLCE we used OntoWordNet. OntoWordNet is based on WordNet 1.6 and it is aligned to DOLCE by the developers of DOLCE.

In the mapping from the domain concepts to DOLCE we also kept the intermediate concepts from OntoWordNet which facilitated the mapping.

- Addition of new Concepts

Addition of new concepts was done in the following cases: (1) several compound keywords were selected, but the head noun was not selected as a separate keyword, but this head noun denoted a more general concept; (2) some of the meanings in OntoWordNet were marked as belonging to the domain of computer science, but were not in the list of the keywords; (3) some concepts were necessary for the definitions of the concepts

constructed on the basis of keywords.

- Addition of Relations

Addition of relations defines in better way the concepts. The basic relations were inherited from the upper ontology.

- Documentation

We have been keeping track on the source of information. At the end of the development of the ontology each concept will have an explanation of its content in English.

- Lexicons

The lexicons in all languages of the project were constructed on the basis of the formal definitions in the ontology. The lexicons were used to facilitate the annotation of the learning objects with ontological information in each language. Also, they help the users of different languages to 'read' the ontology in their own language within ILIAS.

Summary:

At the end of the first year the ontology contained: 1054 concepts and 6 instances. The concepts were divided into: domain ontology concepts - 707, DOLCE concepts - 41, OntoWordNet (concepts not in the domain ontology, but necessary for the alignment between DOLCE and the domain ontology - 306).

This ontology was used in the annotation of the learning objects in the languages of the project and in the experiments for searching of learning materials in different scenarios.

4 Ontology evaluation

In our work we accept the ideas of [3] that the evaluation of an ontology is part of the life cycle of its development, and therefore it takes place in parallel to this development. We also support the opinion in [5] according to which 'Good ontologies are the ones that serve their purpose'. In order to do so, ontologies have to convey the appropriate piece of knowledge in the appropriate way and at the appropriate time.

The main usage of the ontology in LT4eL is for the annotation of the learning materials with ontological information and for facilitating the (multilingual) search for appropriate parts of the learning material.

Evaluation of ontologies is an active research area. Although there already exist many proposals for evaluation of different aspects of the ontology development, usage, evolving, there is no a widely accepted evaluation methodology. Here we present a short overview of the existing methodologies. At the end we list the methods on which we relied to evaluate the coverage of the ontology and its structure.

The overview presented here is based on [1], [2], [3] and [4]. The first classification (see [1]) of the methodologies for evaluation of ontologies is the "black box" evaluation methodologies, which are oriented towards testing the performance of tasks of the whole system, versus "glass box" evaluation methodologies, which are oriented towards testing different components of a system. The black box evaluation depends on the participation of a human user who performs the corresponding task as an end user of the system. Thus, the evaluation becomes very much application dependant and hence needs specific evaluation criteria. The glass box evaluation is additionally divided into three stages: (i) evaluation of an ontology in its pre-modelling stage, (ii) evaluation of an ontology in its modelling stage, and (iii) evaluation of an ontology after its release. The evaluation at stage 1 is in fact an evaluation of materials available to the ontology engineer when he or she starts the creation of the ontology. Such materials include other ontologies to be

re-used in the creation of the new one, database schemata, domain oriented text corpora, etc. Before the actual usage of these sources, the ontology engineer needs to evaluate their quality and appropriateness with respect to the creation of the new ontology. The evaluation at stage 2 comprises regular checks on the consistency of the ontology, logical errors or methodological errors with respect to the ontology construction methodology. At this stage also the ontology content is evaluated - whether the domain model represents the necessary concepts. The evaluation at stage 3 can involve comparison of the ontology to other knowledge sources (other ontologies or other conceptualizations of the same domain). Usually this evaluation is performed by people other than the ontology developers.

According to [4] there are three dimensions to which ontologies can be evaluated:

- the structural dimension which concentrates on syntax and

the formal semantics of the ontology (representation as a graph). Here the topological, logical and meta-logical properties of an ontology can be measured;

- the functional dimension reflects the intended usage of

the ontology and its components;

- the usability-profiling dimension focuses on the

communication context of an ontology. The structural measurements include the following, among others: depth of the graph, breadth of the graph, tangledness, consistency, etc. The functional measurements refer to the correspondence between the ontology as a model, and some conceptualisation of the domain. The metrics, which are usually used in the evaluation of systems in Information Retrieval or NLP, such as precision, recall and accuracy are not straightforward applicable to ontology evaluation (see [5]), because it is not clear what exactly the ontology is compared to. However, [4] defines the following sources of information that can be used: (1) experts' judgment and (2) a data set assumed as a qualified expression of experts' judgment.

The last point we would like to mention here is the approach taken by [3] which interconnects the evolution of an ontology and its evaluation. The authors define the quality of an ontology as a function of the context of the ontology and then the evolution is considered as an optimisation procedure that maximizes the function. Although in our work we do not define such a function we use their idea of interconnecting the development of the ontology and the regular evaluation of the intermediate states of the ontology. With respect to the above criteria for the quality of an ontology we tried to reflect all of them in the process of development of our ontology.

Applicability to LT4eL ontology

The specificity of LT4eL project and the usage of the ontology in it allowed us to include elements of the evaluation of the ontology in the first year of the project as much as the selected methodology of the construction of the ontology lead us to an ontology which is of good quality with respect to its usage in the project. Thus, the evaluation of the ontology was done mainly with respect to the following three criteria:

- *Coverage of the domain*

As a first requirement we selected the coverage of the ontology concepts, which was necessary for the annotation of the learning objects in the corpora of all partners. The second requirement was the ontology to cover the domain uniformly.

The first requirement was ensured by the following steps in the creation of the ontology and its usage in the project: (1) the initial source of terms to start the creation of the ontology were the keywords manually annotated in the learning object of all partners.

Thus, we covered the concepts highlighted by these keywords in the learning materials of the partners; (2) the annotation of the learning object with ontological information was a test of the coverage. During the annotation of learning objects with ontological information there was feedback on the missing concepts necessary for the annotation of the learning objects. After the completion of the annotation, the reported cases were analysed and appropriate extension of the ontology was done. In addition to this we did checks on the automatically selected keywords in several languages.

The second requirement was ensured by the following steps in the creation of the ontology: (1) the collection of definitions from Internet helped us to identify all (or most important) meanings of the selected keywords, not just these used in the learning objects of the partners; (2) we added some concepts that were registered in other sources. Here mainly we added the concepts from OntoWordNet that were marked to be in the domain of computer science.

- *Ontology structure*

Here we mainly checked the structure of the ontology with respect to OntoClean methodology. We achieved this by aligning the ontology to DOLCE which was already constructed with respect to OntoClean methodology. Additionally, we checked the concepts with many subconcepts in order to discover missing generalizations. Such new concepts were added to the ontology.

- *Ontology usage*

The ontology will be evaluated with respect to its real usage. This task will be mainly done during the rest of the project although it has already started.

Summary:

During the first year our main concern was to ensure appropriate coverage of the ontology and it conformity to the OntoClean methodology. The coverage was ensured by the available resources: the keywords annotated by all partners.

During the second year the coverage was additionally checked by feedback from the annotation of the learning objects of all partners. The conformity to OntoClean was ensured by aligning to DOLCE upper ontology.

Improvements on the ontology were also implemented on the basis of the evaluation of the semantic search as it was incorporated within ILIAS. The main suggestions were with respect to the following aspects:

- The navigation over the ontology needs to be done with

respect not only to the domain ontology, but also with respect to the upper part of the ontology;

- Ontology has to cover not only the learning objects, but

also the expectations of the user even in the cases when a given concept is not mentioned in the text of the learning objects;

- Reports on non-understandable concepts, i.e. concepts

which do not reflect the knowledge and intuition of the users.

The evaluation of the usage of the ontology will be done during the rest of the project. This process has already started.

5 Ontology revision

As a consequence of the evaluation of the ontology - the coverage check with respect to the LOs annotation, the feedback from the partners and the alignment to the upper ontology, we revised the ontology by performing the following actions:

- Extending the domain part of the ontology;
- Restructuring the ontology hierarchy;
- Simplification of the upper part of the ontology.

We extended the domain part of the ontology by adding 204 new domain concepts. There were two sources of information here. First, we received feedback from the partners on missing concepts or too general concepts based on the their work on learning object annotation. The second source of new concepts were the places in the ontology were the concepts were presented unevenly. All of these cases could be considered as problems in the initial structure of the domain ontology. Starting with the keywords annotated in the learning objects, the initial set of concepts reflected the following factors: (1) the intuition of the annotators about what a keyword is; (2) the coverage of the learning objects; (3) some of the middle level concepts were rarely mentioned in the domain texts. As a result of this, some concepts are missing from the ontology because they were not mentioned in the learning object at all, or because they were mentioned in such places were they were not considered keywords. Some concepts are too general because in some cases only sub-parts of noun phrases were annotated as keywords, but the maximal noun phrases in fact correspond to concepts appropriate to be in the ontology. The third factor plays role for the deeper structure and better informativeness of the ontology.

The concepts mentioned in the learning objects, but not annotated as keywords were collected by the partners and reported to us. The uneven structure of the ontology was repaired on the basis of information from other sources such as Wikipedia and some terminological lexicons. The same sources were used for addition of middle level concepts. In this way we restructured also the hierarchy of the ontology towards a more intuitive way.

In order to be able to provide the upper part of the ontology to the users of LMS, we decided to substitute DOLCE upper ontology with its simplified version DOLCE Ultralight (http://wiki.loa-cnr.it/index.php/LoaWiki:Ontologies). This version of the ontology uses friendly names and comments for classes and properties, has simple restrictions for classes. In addition to substitution of the DOLCE with DOLCE Ultralight we reduced the number of concepts that originated from OntoWordNet. This reduction was possible because some of the concepts in OntoWordNet reflected the hierarchy of much bigger ontology than LT4eL domain ontology. Thus, there were concepts necessary in OntoWordNet for defining different branches, which however in LT4eL became just one such branch. Consequently, some of the concepts could be deleted.

As a result of this revision of the ontology the current version contains 1185 concepts. Their distribution is as follows: 910 concepts in the domain ontology, 106 concepts in DOLCE Utralight (we kept the whole upper ontology because some of the concepts define the range of some important relations), 169 concepts in OntoWordNet.

The revision of the ontology will be done again on the basis of the feedback from the next evaluations of the semantic search in the project.

At the moment the parters are constructing lexicons to be aligned to the upper part of the ontology.

6 Formal evaluation of ontology semantic search vs. textual search

We consider the search a basic functionality of a LMS that ontology has to improve. The actual improvement will be evaluated within real usage of ILIAS, but we decided to perform some formal comparison between ontology semantic search and textual search.

In a small experiment, which is part of the evaluation of the search function, simple text search and semantic search have been compared. The task which was the basis of this evaluation was as follows: two terms which were also lexical entries in the lexicon of the language under investigation have been chosen as parts of a query (the equivalents of the terms program and slides in each of the languages). The query has been run on the document set with and without the ontology. This resulted in two sets of paragraphs as the respective retrieval results. The conceptual annotation has been used to identify the paragraphs in these documents. The conjunction of these two result sets has been investigated by a researcher and each paragraph rated as either relevant or irrelevant to the search. The retrieval results of both methods (in the following: text search and semantic search) have been weighted against the set of relevant paragraphs with the well-know measures of recall and precision. Both values have been combined in an F-measure. The F-measure is used to compare the results.

The experiment has been run for six languages: Bulgarian, Dutch, English, German, Polish and Portuguese. The F-measures for both text search and semantic search are presented in Table 1. The gain is due to improvements in both recall and precision. It is significant for all languages. The gain is the lowest for Portuguese, because there were only a small number of returned documents. Also, there is visible variation between the languages.

Language	Text Search	Semantic Search
Bulgarian	56,25	91,30
Dutch	47,50	94,12
English	27,96	79,42
German	36,00	59.26
Polish	12,50	50,00
Portuguese	28,67	33,33

Table 1: F-measures for full text search and semantic search in six languages.

Another factor that played a role in the results was the context: the narrower the context (e.g. sentences), the better the results, and vice versa. As has been said before, the conceptual search produces results only in those cases where the search words are in the lexicon and thus matched to concepts in the ontology. This has been the case in the evaluation example. In the case where the search word does not match a lexical item, the text search as well as the keyword-based search is used as a fallback strategy.

7 References

[1] Jens Hartmann, Peter Spyns, Alain Giboin, Diana Maynard, Roberta Cuel, Mari Carmen Suárez-Figueroa and York Sure. 2005. D1.2.3 Methods for ontology evaluation. Deliverable D1.2.3 EU-IST Network of Exccellence IST-2004-507482 KEWB.

[2] Janez Brank, Marko Grobelnik, Dunja Mladeni?. 2005. D1.6.1 Ontology Evaluation. Deliverable D1.6.1 EU-IST Project IST-2003-506826 SEKT.

[3] Peter Haase, York Sure. 2005. D3.1.2 Incremental Ontology Evolution-Evaluation. Deliverable D3.1.2 EU-IST Project IST-2003-506826 SEKT.

[4] Aldo Gangemi, Carola Catenacci, Massimiliano Ciaramita, and Jos Lehmann. 2006.

Modelling Ontology Evaluation and Validation. To appear in Proceedings of ESWC2006, Springer.

[5] Christopher Brewster, Harith Alani, Srinandan Dasmahapatra, and Yorick Wilks. 2004. Data Driven Ontology Evaluation. International Conference on Language Resources and Evaluation, Lisbon, Portugal.

8 Scientific papers on the ontology, lexicons and annotation of learning objects

The following papers have been written, presented on conferences and published or they will be published. The full papers are in the appendix. In the following, the bibliographical data are listed.

- Kiril Simov, Petya Osenova: Applying Ontology-Based Lexicons to the Semantic Annotation of Learning Objects. Proceedings of the RANLP 2007 workshop: Natural Language Processing and Knowledge Representation for eLearning Environments.
- Lothar Lemnitzer and Kiril Simov and Petya Osenova and Eelco Mossel and Paola Monachesi. Using a domain-ontology and semantic search in an eLearning environment. International Conference on Engineering Education, Instructional Technology, Assessment, and E-learning (CISSE-EIAE 07).
- Cristina Vertan, Paola Monachesi, Kiril Simov, Petya Osenova, Lothar Lemnitzer, Alex Killing and Diane Evans. Crosslingual retrieval in an eLearning environment. Proceedings of The 10th Congress of the Italian Association for Artificial Intelligence (AIIA 2007).
- Lothar Lemnitzer, Cristina Vertan, Alex Killing, Kiril Simov, Diane Evans, Dan Cristea, Paola Monachesi: Improving the search for learning objects with keywords and ontologies. Appeared in: Duval, Erik; Klamma, Ralf; Wolpers, Martin (Eds.) Creating New Learning Experiences on a Global Scale. Second European Conference on Technology Enhanced Learning, Lecture Notes in Computer Science , Vol. 4753, pp. 202-216.
- Eelco Mossel: Crosslingual Ontology-Based Document Retrieval. Proceedings of the RANLP 2007 workshop: Natural Language Processing and Knowledge Representation for eLearning Environments.
- Paola Monachesi, Lothar Lemnitzer and Kiril Simov *Language Technology for eLearning*. Poster presentation at KP7-congres. 13 February 2007, Den Haag. The Netherlands.

Accepted for presentation:

- Paola Monachesi, Kiril Simov, Eelco Mossel, Petya Osenova and Lothar Lemnitzer. What ontologies can do for eLearning. Will be presented at: IMCL 2008 (http://www.imcl-conference.org/).