Topic Maps e-Learning Portal Development

Kamila Olsevicova
University of Hradec Kralove, Czech Republic
Kamila.Olsevicova@uhk.cz

Abstract: Topic Maps, ISO/IEC 13250 standard, are designed to facilitate the organization and navigation of large collections of information objects by creating meta-level perspectives of their underlying concepts and relationships. The underlying structure of concepts and relations is expressed by domain ontologies. The Topics Maps technology can become the core of an e-learning portal that will integrate different kinds of information and knowledge resources, available in the educational institution – this idea was explored in the Ph.D. dissertation of the author. The offered portal solution promises to bring advantages both for content consumers (students) and content providers (teachers, administrative staff), but numerous problems hinder the practical implementation of this portal and therefore it requires certain changes in the functioning of the educational institution and asks teachers, teaching assistants and e-courses designers to change their routines and to develop new skills. In the paper we offer a new methodology for development and maintenance of the Topic Maps e-learning portal and we briefly present a pilot application.

Keywords: e-Learning portal, Ontology engineering, Knowledge methodology, Topic Maps, Omnigator

1. Introduction

In the last three decades numerous approaches have appeared which have tried to adopt information and communication technologies for the purpose of learning and education. The term "e-learning" was accepted for expressing the effort to transform educational processes through application of different up-to-date electronic media and to customize learning to student's needs in terms of study style, time and space. The dramatic growth of local and wide area computer networks accelerated the evolution of the phenomenon of online education (also Internet-based education, web-based education, education via computer-mediated communication, virtual education).

Online education is realized in virtual study environments (also online education systems), e.g. WebCT, Blackboard or LearningSpace. Virtual study environments can be used within the context of the traditional education system, in part-time education, or to facilitate distance-education and therefore they are interesting for different kinds of educational institutions. Through virtual study environments, e-courses are managed and provided to students. The general shift of emphasis towards the Semantic web idea has to be reflected by the developers of particular web-based applications. In the educational context, it means thinking about the next generation of virtual study environments that would overcome insufficiencies of current systems such as information overflow, lack of customization, necessity of manual management and updating of stored instructional e-content, high set-up costs etc.

The term e-learning implies a shift in emphasis in education: we talk about "computer assisted learning", not about "computer assisted teaching".

This implies an effort to stimulate the student's own activity, and de-emphasises the role of the teacher. The objectivist approach where the educational process was conducted by teachers, and students demonstrated their knowledge level in tests, where they reproduced facts and information, is replaced by a constructivist theory, based on the opinion that students have to develop their own approaches to understand and organize knowledge. Accentuation of knowledge structures, their internalization by students and the environments that must enable this internalization – it all makes us think about adaptation of knowledge management (KM) principles into the educational area and to understand e-learning solutions as pure KM solutions.

The paper is focused on the need for an innovative virtual study environment, developed in accordance with the Semantic web idea and reflecting the theories of student's internal processes of knowledge organization. We describe a novel solution, an ontology-driven Topic Maps portal that mainly addresses tasks of content management of e-courses and which integrates different information and knowledge resources of the educational institution.

2. Topic maps technology

Topic maps (TM), the ISO/IEC 13250 standard, are designed to facilitate the organization and navigation of large collections of information objects by creating meta-level perspectives of their underlying concepts and relationships. The underlying structure of concepts and relations are

Reference this paper as:
Olsevicova K (2006) "Topic Maps e-Learning Portal Development" *The Electronic Journal of e-Learning* Volume 4 Issue 1, pp 59-66, available online at www.ejel.org

expressed by domain ontologies, or by other modeling formalisms, e.g. subject categorizations, classifications or schemas, relational or object-oriented schemas, indices and thesauruses.

Topic maps promise to solve the following tasks:

- The problem of metadata – the TM application operates with metadata records on available information resources. These records are not a part of resources, but are stored and managed independently.
- The network structure of links – interconnection of topics (and so, also resources) minimizes the risk of being lost in the information space, although the search possibilities are maximized.
- The structure of knowledge in the domain – the TM standard was developed for explicit modeling of knowledge and simple navigation in knowledge resources, therefore it visualizes the terminology of the domain and allows each user to view the terminology in scopes which reflect the needs and abilities of the user.

The Topic Maps model defines three basic building blocks: topic, association and occurrence which together form "TAO of Topic Maps", as (Pepper, 2000) humorously says. Other concepts which extend the expressive power of TM are those of scope, theme and published subject:

- Topics are computer representations, either of particular subjects of our world, or abstract categories that exist in this world. Through the topics, subjects transform to computerized objects, about which any statements can be linked to the TM. A particular structure of topics depends on the application domain, authors and users of the TM application. Each topic has three types of characteristics: names, occurrences and roles in associations.
- Occurrences of topics are relevant information sources about topics. The occurrences are characteristics of topics and are true in the context of the particular scope. The occurrence is a string value - either a statement about the topic, or a link to the resource (monograph, article, image, sound file etc.). The occurrences are inserted to the TM using suitable identifiers (URI, HyTime addressing etc.). Through the occurrences, information resources are included in the TM application, although they are not changed inside, i.e. no additional internal tags are necessary to be inserted into the resources to describe their content.
- Associations are essential for establishing the network structure of topics. The associations simulate human associative reasoning. Two or more topics are in an association, if there is some relationship between them. Each topic plays a role in a particular association.
- Scopes express the range of validity of topics' characteristics, especially they establish context for the validity of topics' names and their roles in associations. Through scopes, TM applications can be divided into semantic slices based on different criteria (langauge, prerequisite knowledge, access rights etc.).
- Themes are defined as sets of topics used to specify the scope.
- Published subjects are entities, placed and maintained in an advertised place, to standardize and reuse non-addressable subjects in various TM applications and to facilitate merging different TM applications. Agreement is up to the humans defining and using published subjects.

3. e-Learning portal

Our basic idea is to use a TM application (more precisely, a web portal based on the TM standard) as a gate to all information, and if possible, to all knowledge resources of the educational institution. This idea has two main advantages.

- Integration of currently separated resources of information and knowledge at the university, such as scheduling applications, students' agenda information systems, digital and classic libraries, shared directories etc.
- More user friendly, more intuitive navigation of the info-space, enabled by ontologies which are the core of TM applications and whose application is innovative in contrast with websites organized using predefined categories and sections.

In virtual study environments, TM technology is applicable in the following way. Each e-course is focused on certain discipline which has its own terminology. This terminology is conceptualized by the discipline (domain) ontology. The TM application of study resources can be built above this ontology. Such TM application visualizes the discipline terminology, which helps students to understand the structures of studied disciplines. Together with the discipline ontologies, used for subject categorization of resources, it is possible to apply a kind of course ontology for arranging units and elements that together form the course content. Therefore through the TM-based portal, teachers can define the recommended order of resources (presentations, documents, exercises etc.) to be studied as well as in the e-course's study content module in current virtual study environments. Also, all other parts of the e-course (students' agenda, evaluation tools, communication tools), can be integrated into the

TM portal using the occurrence elements. All these integrations are motivated by the effort to unify access mechanisms to information. E.g. while using WebCT 4.1 Campus Edition, students have to uncomfortably click through individual e-courses and check e-mail boxes or assignments evaluation records in each of the e-courses separately. The TM solution enables unified access to all information of the same type, e.g. to all e-mails, all assignment scores etc.

4. Development of the portal

No methodology of TM applications creation is presented in the basic ISO/IEC Topic Maps standard, so the consecution totally depends on authors of particular TM applications and on software used for the implementation of the TM application. General TM developers' guides suggest starting either top-down (i.e. by defining the application area (Rath, 2003), or bottom-up, i.e. by summarizing the available information and knowledge resources to be covered by the TM application (Vatant, 2001). The first approach helps to reduce the space of documents and resources to be considered with respect to the application domain, while the opposite approach promises not to omit any currently available resources and to enable the TM application to access these resources and repositories.

In both approaches, the next TM application development contains the following steps:

- definition of functional requirements and the purpose of the future TM,
- definition of schema of the TM portal - the ontology,
- selection of the tool for implementation of the TM solution,
- population of instances, including evaluation of fulfilling all restrictions and constraints,
- optional revision of the schema of the TM.

The fundamental task is the selection of reference vocabularies and ontologies of types, categories, relationships etc. which help to coordinate the effort of the team of TM portal developers, especially to avoid misunderstandings and misinterpretations of the skeleton of the TM portal and to ensure further extension and sharability of the portal.

4.1 Reusable ontologies

For our purpose, that is using ontologies in the description of educational resources and processes in the context of virtual study environments, three types of reusable ontologies are of interest:

Linguistic ontologies (also lexical databases) which conceptualize natural languages and help to understand terms, which seem to be common, but in practice can become sources of misunderstanding. The most prominent ontology of this type is Word Net (WordNet, 2005), a large lexical database in English based on psycholinguistic theories. WordNet attempts to organize lexical information in terms of word meanings rather than word forms, though inflectional morphology is also a criterion. WordNet 2.0 contains words organized in sets of synonyms (synsets), each representing one underlying lexical concept, with a brief explanation of the intuitive sense in English. Synsets are interlinked via relations such as synonymy, antonymy, meronymy (part-of relation), holonymy (has-a relation) hyponymy (subclass-of relation), hyponymy (super class-of relation). The lexicon is divided into five categories: nouns, verbs, adverbs, adjectives and function words. WordNet database is under continuous development.

Domain ontologies which model concepts and relations in particular areas, so we can use them to describe disciplines which are studied in courses. It means that domain concepts are applicable as subject descriptors of resources; the relations between concepts make relations between resources visible. If we consider using ontology as an integrative element that would allow us to interconnect educational resources related to the course or the whole study programme, optionally prepared by various authors from different institutions, it is clear that such an ontology must not only to cover the relevant domain (area, discipline), but must be of a good quality and must be accepted by all authors of resources, especially authors of applications that make these resources available to users (i.e. to students, teachers, other staff of the faculty). Unfortunately, at the moment there are not suitable ontologies available for all domains, but only for those areas, where the critical necessity to manage large repositories of documents has already prompted the development of ontologies. Such areas are e.g. medicine, chemistry or law. In disciplines, where ontological engineering is not so advanced, it is possible to think about adopting some of developed taxonomies or categorizations and reusing them for the development of a new ontology. E.g. in the area of computer science education, we propose to use the ACM Computer Classification System, which is recommended to be used for description of content of all ACM publications (ACM CCS, 2005). Reusable domain ontologies are available in web libraries of ontologies.

Knowledge management ontologies which are included among the knowledge-based techniques for building organizational memory systems. For our broader objective, i.e. description of information and knowledge resources at the university, such an approach is of interest. (Abecker et al., 1998) distinguishes three types of knowledge management ontologies that may partly overlap in certain concepts and relations:

- domain ontologies which model the content of the information sources (as previsouly discussed),
- information ontologies which describe the different kinds of information resources, their structure, access rights, format properties,
- Enterprise ontologies which model the context of an organization, business processes, organization of the enterprise etc.

4.2 Creation of the ontology

The process of ontology construction is critical, mainly in complex domains. There are two main approaches that aid a large-scale ontology construction from scratch:

- The first one facilitates the manual process of ontological engineering by providing editors, consistency checkers, natural language processing tools, etc. Protégé-2000 is an example of this category (Protégé, 2005).
- The second approach relies on machine learning and automated language processing techniques to extract concepts and ontological relations from given databases and texts. This approach, usually called ontology learning, is explained in great detail in (Maedche, 2002). One example of such system is OntoLearn, presented in (Navigli et al., 2003).

Our proposed ontologies for the description of educational reality are expected to be created manually. The development methodology would partly depend on the chosen ontology editor. If we use Protégé-2000, the following steps (formulated in Protégé-2000 terminology) are recommended (Noy and McGuiness, 2001, Noy et al., 2001; Gómez-Pérez et al., 2004):

- Determine the domain and scope of the ontology, i.e. answer basic questions “What is the domain that the ontology will cover?”, “For what types of questions should the information in the ontology provide answers?”, “Who will use and maintain the ontology?” A good way to define the scope is to start with so-called competency questions; this means a set of queries that will be used above the knowledge base, based on the ontology. The list of competency questions can help to check if the ontology contains enough information.
- Consider reusing existing ontologies. The process of ontology creation can be accelerated by refining and extending existing ontologies from public libraries, web repositories etc.
- Enumerate important terms in the ontology, without worrying about any overlap between concepts they represent, distinguishing properties and relations.
- Define the classes and the class hierarchy either using the top-down or bottom-up approach, or a combination of both approaches. One question is whether to start with the most general class and then generate more specialized concepts or vice versa. Class hierarchy automatically assumes inheritance, which means that if a class A is a super class of class B, then every instance of B is also instance of A.
- Define the properties of classes – slots. Properties of classes can be intrinsic, extrinsic, parts (if concept is structured), relationships with other individuals. All subclasses of a class inherit slots of the class. A slot should be attached to the most general class that can have that property.
- Define the facets of the slots. A slot can have different facets describing the value type, allowed values, cardinality and other features.
- Create instances. Deciding whether a particular concept is a class or an individual instance depends on what the potential applications of the ontology are and takes into consideration the lowest level of granularity in the representation. If concepts form a natural hierarchy, then we should represent them as classes.

4.3 Ontology of educational environment

For the description of a university environment and its information and knowledge resources, three kinds of ontologies are needed:

- a general ontology describing the reality of the educational institution,
- a course ontology that defines the structure of the course,
- Domain ontologies that conceptualize individual disciplines.
- The general ontology of the educational institution contains concepts and relations referring to information and knowledge resources and processes at the university, but without direct relation to particular courses. It contains:

- concepts such as address, article, bachelor program, bachelor thesis, conference, consultation, course, credit, database, department, email, entrance test, enrolment, examination, faculty etc.,
- And relations such as (course)-is-previous-to-(course), (course)-is-recommended-to-(study programme/semester), (person)-is-supervisor-of-(course), etc.
- Course ontology defines concepts that describe parts of individual courses, kinds of involved learning objects, instructional strategies, learning styles, student profiles, educational goals, etc. There are:
- concepts such as active material, answer, assignment, difficulty, discussion, evaluation, exam, exercise, feedback, glossary, grade, homework, image, interactivity, item, learning style etc.,
- And relations such as (lecture/exercise etc.)–Is-previous-to–(lecture/exercise etc.), (prerequisite)–is requested–in (course), (presentation/assignment/question)–assigned-to–(lecture/exercise) etc.

Domain ontologies describe terms and their relations which are valid in a certain discipline, taught in courses. For each field, a different ontology is demanded. If possible, existing domain ontology can be adopted. The usefulness of these ontologies lies in the fact that they can help to structure courses' contents logically according to structures of disciplines. For students, it is highly important to gain an insight into the terminology of the studied area. If concepts of ontologies are used for defining keywords or subject descriptors of educational resources, while searching repository of resources, students internalize the discipline terminology in a natural way.

These three kinds of ontologies (general, course and domain) can be merged together. Therefore, we get a large collection of concepts and relations, describing the university with all its information and knowledge resources and its study programmes on a very detailed level. Above such a large ontology (or above its semantic slices), it is possible to implement an application which serves as a universal gate to all resources. Practical realization of this application may reuse TM technology.

The presented extracts of general, course and domain ontologies are illustrative only. Real applications would be based on ontologies, which should result from a detailed analysis and broad understanding of all involved concepts and their relations among users of particular ontology application. We expect that these ontologies for the description of an educational reality, courses and particular disciplines would be reusable for the purpose of different educational institutions, so the efforts of educational experts and university management to develop the ontologies would be very fruitful. Knowledge engineers and ontology engineers, skilful in the construction of knowledge bases, will have to participate on the development of the proposed TM e-learning portal.

5. Pilot application

For better explanation of our proposed TM solution, we developed a pilot application using Omnigator, a free version of TM software provided by (Ontopia, 2005). Our application integrates information and knowledge resources related to several courses on Artificial intelligence (AI), sub areas which are taught in the Faculty of Informatics and Management at the University of Hradec Kralove. These resources are traditionally separated in different information systems and repositories, e.g. website of the faculty, student's agenda information system of the faculty, digital library of the faculty, university library, private and shared folders on university computer network etc. The underlying ontology of our application reuses parts of ACM CCS and the WordNet lexical database. The following snapshots illustrate what the application looks like and how it can be used. (Our version of TM software was limited in graphic functions; therefore the design of the application is not as attractive and user-friendly as the full solution would be.)

The first snapshot (Fig. 1) shows how the TM application presents information related to particular AI course (here it is Logic programming 1.). The page contains:

- Name of the course,
- Relevant study programs (with hyperlinks to their topic pages),
- Name of the supervisor of the course and of teaching assistants (with hyperlinks to their topics pages),
- String information about prerequisites and recommended semester,
- List of resources (different kinds of online resources – e-courses, scheduling application and student agenda system – again with hyperlinks),
- Hyperlinks to 4 of 94 scoped occurrences, i.e. particular resources that are associated with the scope of the topic “LP1“.

Figure 1: Topic page of Logic programming course

The second snapshot (Fig. 2) presents a particular lecture in the Logic programming 1 course. Different information and resources are available on this page, e.g.:

- Brief syllabus of the lecture,
- Note about a PowerPoint presentation file stored in shared folders,
- Hyperlinks to html-pages stored in the WebCT virtual study environment, where traditional e-courses are still provided,
- Hyperlinks to topic pages of previous and following lectures and exercises where the content of the current lecture is reused.

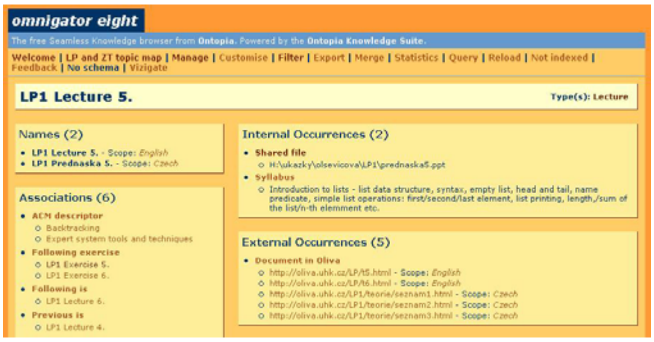

Figure 2: Topic page of 5th lecture of Logic programming course

The last snapshot (Fig. 3. shows the page of the topic "Agent", i.e. particular important concept in AI.

- Associations "ACM CSS category", "Has subcategory" and "Is element of" explain the position of the topic in the ACM CCS and relation to other close topics.
- The course association „Studied in“ shows what courses (lectures) are related to Agent topic.
- Internal and external occurrences refer to topic resources: e.g. information about books available in the university library, hyperlinks to online web resources, string value information – here, quotations of definitions from the WordNet lexical database. The scopes assigned to occurrences inform about the level of prerequisited knowledge, i.e. that a particular book is aimed mostly at experts, or about language of resources etc.

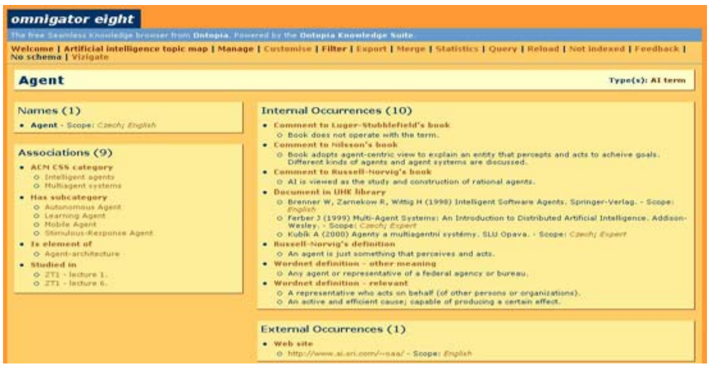

Figure 3: Topic page of the term "Agent"

6. Conclusion

In this paper we have proposed understanding virtual study environments as a kind of knowledge management system and integrating them with other information and knowledge resources, both digital and non-digital, which are available at the educational institution. We see the Topic Maps technology as a suitable tool for this proposed integration and we offer a general framework for the development of the Topic Maps-driven e-learning portal. The main idea is to interconnect different information systems, used by students and staff of the faculty and to reuse domain ontologies for the subject classification of the content. See (Olsevicova, 2005), where the presented idea is explained in detail. See also (Dicheva et al. 2004; Dichev and Dicheva, 2005) for other TM applications in the educational context, mainly a novel architecture improving the effectiveness of searches in concept-based digital course libraries. (Schwotzer, 2004) explains how Topic Maps can be used for knowledge representation in distributed knowledge management systems and understands the exchange of Topic Maps to be the exchange of explicit knowledge. In this perspective, we can consider the interlinking of our proposed Topic Maps portal of a particular educational institution with similar portals of other educational institutions.

Acknowledgements

This contribution was partially supported by the Czech Science Foundation Project No. 406/04/2140 KNOMEDIAS

References

Abecker A, Bernardi A, Hinkelmann K, Kühn O, Sintek M (1998) *Toward a Technology for Organizational Memory*, IEEE Intelligent Systems, Vol. 13, No. 3, pp.40-48.
ACM CCS (2005) *ACM Computer Classification System website*, [online], http://www.acm.org/class/1998/.
Dichev C., Dicheva D., Aroyo L. (2004). Using Topic Maps for Web-based Education, *Int. J. of Advanced Technology for Learning*, Vol. 1, No. 1, pp1-7.
Dichev C., Dicheva D. (2005). Contexts in Educational Topic Maps. *12th International Conference on Artificial Intelligence in Education, AI-ED 2005*, July 18-22, 2005, Amsterdam, The Netherlands (to apear).
Gómez-Pérez, A., Fernández-López, M., Corcho, O. (2004) *Ontological Engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic web*, Springer-Verlag, London.
ISO 13250 (2000) *ISO/IEC 13250:2000 Topic Maps*, [online] http://www.y12.doe.gov/sgml/sc34/document/0129.pdf.
Maedche A (2002) *Ontology Learning for Semantic Web*. Kluwer Academic Publishers, Boston.
Mondeca (2005) *Mondeca webiste*, [online] http://www.mondeca.com.
Navigli, R., Velardi, P., Gangemi, A. (2003) *Ontology Learning and Its Applications to Automated Terminology Translation.* IEEE Intelligent Systems, Vol. 18, No. 1, pp22-31.
Noy, F.N., McGuinness, D.L. (2001) *Ontology Development 101: A Guide to Creating Your First Ontology*. KSL Technical Report KSL-01-05, Knowledge Systems Laboratory, Stanford University, Stanford, [online] http://protege.stanford.edu/publications/ontology_development/ontology101.html.
Noy, N.F., Sintek, M., Decker, S., Crubézy, M., Fergerson, R.W., Musen, M.A. (2001) *Creating Semantic Web Contents with Protégé-2000*. IEEE Intelligent Systems, Vol. 16, No. 2, pp60-71.
Olsevicova, K. (2005) Knowledge Management in Virtual Study Environment, *PhD dissertation*, University of Hradec Kralove.
Ontopia (2005) *Ontopia website*, [online] http://www.ontopia.net.
Pepper, S. (2000) *The Tao of Topic Maps*, [online] http://www.ontopia.net/topicmaps/materials.
Protégé (2005) *Protégé-2000 website*, [online] http://protege.stanford.edu.
Rath, H. (2003) *The Topic Maps Handbook - white paper*, [online] http://www.empolis.com.
Schwotzer, T. (2004) Modelling Distributed Knowledge Management Systems with Topic Maps. *Proc. of the 4th International Conference on Knowledge Management*, Graz, Austria, June 30 – July 2 2004, pp. 52-59.
Vatant, B. (2001) *Managing complex environments with Topic Maps*, [online] http://mondeca-publishing.com/s/anonymous/title12140.html.
WordNet (2005) *WordNet lexical database website*, [online] http://www.cogsci.princeton.edu/~wn/index.shtml.