TOPIC MAPS IN MANAGEMENT OF STUDY RESOURCES

Kamila Olševičová and Peter Mikulecký
    Katedra informačních technologií, Fakulta informatiky a managementu
    Univerzita Hradec Králové

Abstract
Information overload which we witness in all the areas of intellectual work has its negative influence at the universities preventing students from effective selection and understanding of most important stuff and its relationships in studied disciplines. The objective of this paper is to describe an innovative solution – an ontology-driven Topic Maps portal. Our pilot application was designed to help students of computer science to obtain more complex and deeper insight in the Artificial Intelligence area.

Keywords
Topic Maps, ontology, e-learning, knowledge management

1. Introduction

Information infoglut that recently threatens to drown knowledge workers all over the world creates a number of challenges for its overwhelming and has its negative consequences also in the educational process at universities. The vast of different resources of study materials and the stuff included in them are preventing students from understanding all the studied content correctly and in an effective way. In order to contribute to the infoglut overwhelming, a new solution is presented in the paper. The objective of this paper is to describe an innovative solution - an ontology-driven Topic Maps (TM) portal that was designed to help students to obtain more complex and deeper insight in the Artificial Intelligence (AI) area.

The paper is organized in the following way. Firstly we describe the structure of AI courses that are taught the Faculty of Informatics and Management, University of Hradec Kralove (FIM UHK). Then we introduce basics of TM technology and of creation TM application. In the last sections we explain how to apply TM technology to build the portal, useful to students, teachers as well as to other staff of the faculty. We also present snapshots of our experimental application here.

2. AI Courses at the FIM UHK

At the Faculty of Informatics and Management, University of Hradec Kralove six courses of AI are provided to students: courses of Logic Programming 1 and 2 (LP1, LP2) and courses of Knowledge technologies 1 to 4 (KT1, KT2, KT3, KT4), that are offered to students from the second to the fourth year of study programs of Applied Informatics and Information Management. As it can be seen from brief syllabuses (Tab 1.), the contents of all six courses are closely related and partly overlaps, because of the same area of interest (AI) and because of the limited palette of didactically interesting problems and tasks that can be understood and solved by most students.

Courses LP1 and LP2 are focused on basic principles of declarative programming and then on advanced programming techniques, used in solving the most classic AI tasks (towers of Hanoi, 8 queens' problem, farmer's dilemma etc.). To explore relevant topics practically, students pass exercises of programming in LPA-WinProlog [15] environment. In consecutive courses KT1 and KT2, parts of AI theory, expert and knowledge-based systems are presented.

Practical experiments are realized in CLIPS and Jess rule-based expert system shell [2],[7]. Courses KT3 and KT4 are offered to students with deeper interest either in the knowledge management area with related concepts of knowledge sharing, knowledge management systems (KT3), or in advanced techniques of AI like soft computing basics and applications (KT4).

Tab. 1. Artificial Intelligence courses content

	LP1 - Logic programming 1	*LP2 -Logic programming 2*
2nd year	Introduction to Prolog Syntax of Prolog Simple queries Backtracking List data structure Recursion Control predicates Database predicates	State space search Graph problems Problems on chessboard Two players games Small expert system
	KT1 – Knowledge technologies 1	*KT2 - Knowledge technologies 2*
3rd year	Introduction to AI Overview of AI areas of interest Knowledge representation Problem solving methods Rule-based systems	Introduction to knowledge-based systems Experts and expertise Knowledge engineering Knowledge-based system life cycle
	KT3 - Knowledge technologies 3	*KT4 - Knowledge technologies 4*
4th year	Knowledge management Knowledge sharing Knowledge management systems Communities of practice Communities of interest	Soft computing Fuzzy systems Genetic algorithms Genetic programming Neural networks

The following obstacles are related to our six courses:

- The LP2 course is optional, so students subscribed to KT1 may have different levels of previous knowledge and experiences. Those, who passed both semesters of the Logic programming courses, are able to go forward more quickly and some lectures may seem to them to be trivial, while other students suffer from the lack of experiences with certain programming techniques so the Knowledge technologies courses may seem difficult for them. Also two last courses (KT3 and KT4) were optional up to now and therefore at least introductory notes to themes like genetic programming or neural networks could not be omitted in lectures of KT1 to ensure that all students will know them at least a bit.

- Due to the study programmes, the audience of courses varies from future financial managers to technically oriented students with very good programming skills. Therefore, certain customization of courses' contents with the aim to better address individual students' needs is necessary and could make education more effective and attractive.

- Last, but not least, different courses are taught by different teachers and teaching assistants, some of exercises are managed by PhD students. Therefore it is difficult to

ensure that contents of courses are complementary, smothly connected, and that all parallel exercise groups progress with a same rapidity.

In previous years, all the six courses were supported by e-courses realized in the WebCT environment. These e-courses were used for delivery of digital educational resources in different forms, i.e.:

- common documents (text files, powerpoint presentations, web pages),
- in special formats (glossaries, image databases, self-tests, quizzes),
- management of the educational processes:
- students' agenda processing,
- management of assignments,
- asynchronous communication via email and discussions,
- planning events using calendar etc.

But for six courses that partly overlap in content and where the student audience is heterogeneous in sense of previous knowledge, level of interest etc., it was ineffective and clumsy to prepare and maintain six independent e-courses. Therefore we had looked for a way how to integrate all study resources of these courses together and how to provide more natural navigation mechanisms, than what was possible in the WebCT, where the sequential reading of chapters in the study content module is preferred. An appropriate solution was inspired by new ISO standard of Topic Maps, built above ontologies that capture domain vocabularies.

3. Topic Maps Technology

Topic maps, 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 can be expressed by domain ontologies, or by other modelling formalisms, e.g. subject categorizations, classifications or schemas, relational or object-oriented schemas, indices and thesauruses.

3.1. Topic Maps Model

Topic maps model defines three basic building blocks: topic, association and occurrence that together form "TAO of topic maps", as [12] 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 exists in this world. Through the topics, subjects transforms to computerized objects, about that any statements can be assigned to the TM. Particular structure of topics depends on the application domain, authors and users of the TM application. Each topic has got three kinds of characteristics: names, occurrences and roles in associations.

- Occurrences of topics are relevant information sources about the topics. Occurrences are characteristics of topics and are true in context of the particular scope. The occurrence is string value - either statement about the topic, or link to a resource (monograph, article, image, sound file etc.). The occurrences are inserted to TM using suitable identifiers (URI, HyTime addressing etc.). Through occurrences, information resources are included in the TM, although they are not changed, i.e. no special tags are necessary to be added to the resources for description of their content. This is one of the main advantages of TM: authors of documents are not enforced to respect any additional rules and all kinds of contributions are assignable to the TM.

- Associations are essential for establishing the network structure of topics. Associations simulate associative reasoning. Two or more topics are in an association, if there is some relation between them. Each topic plays some role in a particular association, what helps to interpret associations.

- Scopes express the range of validity of topics' characteristics, it establishes context for the validity of topics' names and their roles in associations. Scopes are used for capturing different points of view on the same reality. Through scopes, TM can be divided into semantic slices - filter language subversions, restrict view on resources according access rights or level of knowledge. It is important to say that scopes make visible or invisible characteristics of topics, but can not hide the whole topic, therefore always the whole TM is presented and user's view on the structure of the TM is not limited.

- Themes are defined as sets of topics used to specify the scope. Themes are topics used to limit the validity of a set of assignments in the TM.

- Published subjects are entities, placed and maintained at an advertised place, to standardize and reuse non-addressable subjects in various TM and to facilitate merging different TM. Agreement is up to the humans defining and using published subjects.

The TM model, described above, is applicable in different areas [13]. For our purpose, TM-based web portals are interesting. These portals apply TM principles with the aim to make navigation easier and more intuitive by providing subject-based categorization of resources. The current e-courses provided in WebCT (and in other courseware systems) are web-based applications, so the shift from the rigid structure of individual e-courses to the presentation of integrated content through TM application do not bring new obstacles in sense of making users to change their behaviour or to learn new working patterns. That is, the access via web browser do not change, only the application is built on the more sophisticated technology, closer to the reality of the upcoming semantic web era.

3.2. Creation of Topic Maps

No methodology of TM application creation is presented in the ISO/IEC standards, so the consecution totally depends on authors of particular TM and on software used for implementation of TM application. [13] recommends the main steps on the way towards the TM application:

- Definition of application area.
- Definition of functional requirements and the purpose of the future TM.
- Definition of schema of the TM - list of subjects and relationships.
- Selection of the tool for implementation of the TM solution.
- Development of the TM, especially population of instances, including evaluation of fulfilling previously defined restrictions and constraints and revision of the schema of the TM (if necessary).
- Maintaining TM and related applications with respect to newly coming requirements.

Using these steps, we created TM application for our AI courses, as it is shown in details in the next section. In general, the steps from the previous list were fulfilled in this way:

- Application area: information and knowledge resources of AI related courses taught at the FIM UHK.

- Requirements and purpose: defining all potential user groups (students, teachers, management of the faculty, administrative staff, technical staff) and defining relevant tasks and operations realized by each user group.
- Schema of TM: preparation of ontologies (general ontology of educational reality and domain ontology of AI).
- Tool: for experimental application we chose free version of TM software toolpack called Omnigator [11], provided by Ontopia .
- Development of TM: the structure of the map was defined in XTM syntax, different kinds of information and knowledge resources were identified, described by metadata and assigned to the TM either as string occurrences or as hyperlinks to resources.
- Maintaining TM application: In case of real application, it would be necessary to update occurrences, to check the consistency of the map, to defined suitable naming conventions etc.

4. Topic Maps for AI Courses

In the light of up-to-date e-learning terminology, all information and knowledge resources relevant to the course are called learning objects, digital resources are called digital learning objects. The learning objects can be described by metadata, similarly as assets stored in the library. Part of each metadata schema is focused on subject description of the resources. Our idea is to build a general model of the domain (ontology), then to reuse subjects descriptors of resources to assigning resources to relevant parts of ontology to obtain the TM of resources.

4.1. AI Courses at the FIM UHK Revisited

Primarily we focused on merging study resources used in different AI courses. Our effort was aimed on providing a unified view on them, and we wanted to point out the relations among them: not for all students it is simple to notice how LP and KT courses are related.

In parallel, our effort was focused on an integration of different information and knowledge resources and repositories that are accessed by students of our courses:

- website of the faculty, the source of current organizational and administrative information,
- students' agenda information system of the faculty, the source of information about courses, exams, teachers etc.,
- digital library DILLEO [3], where articles, papers, digital versions of bachelor and master theses etc. are stored,
- university library that offers hardcopies of books and journals,
- private or shared directories on the university computer network, where documents and applications are available.

The involving of meta-information from these repositories and information systems in our TM solution is explained in the following paragraph.

4.2. Preparation of ontologies

The process of TM application development can be accelerated if suitable schema of topics is available and can be reused. In our application that is focused on AI related information and knowledge resources, we decided to use:

- ACM Computer Classification Schema [1] that defines the tree of concepts of computer science, including categories of themes and subthemes studied in AI.

- WordNet lexical database [15], a representative of linguistic ontologies that conceptualize natural languages and help to understand terms and eliminate eventual misunderstandings.
- Different reusable, public available domain ontologies, e.g. ontology of activities around university [5], ontology of computer science community [4], personal knowledge ontology [9] etc. related to the educational reality.

The decision to use ACM Computer Classification Schema [1] is rationalized by the fact that the ACM CCS is generally accepted by computer science community. Therefore the TM application that reuses part of this schema could be easily merged with other TM, based on other parts of the ACM CCS. It means that e.g. our TM of AI would be merged with TM dedicated to computer graphics, information storage and retrieval etc. This aspect is very important, because it implies how our proposed TM solution should be applicable not only on few related courses, but on whole study programs of computer science.

The ACM CCS involves a four-level tree that has three coded levels and one un-coded level of subject descriptors (usually appearing at the fourth level). This tree comprises the categories and subject descriptors. For our purpose, the top level category "I. Computing methodologies" and its subcategory "I.2 Artificial Intelligence" are relevant. On the 3rd level, there are 13 subcategories defined, each of them has got own descriptors.

Having this classification scheme, the main structure of the TM ontology is established:

- The "main themes" of AI correspond with subcategories of AI on the 3rd level of the ACM CCS.
- The "subthemes" correspond to descriptors on the 4th level of the ACM CCS.
- Relations between main themes and subthemes can be defined by associations "is subtheme of".
- Of course, slight modification of the ACM CCS was necessary, e.g. we missed the category "miscellaneous" that was not practical for our purpose.

The terms used as categories labels or as descriptors are understandable for AI researchers, but we can not expect that all students will be familiar with all of them, so we defined four types of internal occurrences that explain the proper meaning of each term. Formulations of these explanations come either from WordNet lexical database or from classic AI resources. The origin of each explanation is clear from the TM, because the resource name appears in the occurrence label. Available occurrence types are:

- Wordnet definition from [15],
- Nilsson's definition from [10],
- Russell - Norvig's definition from [14],
- Luger - Stubblefield's definition from [8].

Through these occurrence types, the TM becomes a kind of encyclopaedia that presents terminology of AI discipline. Educational domain ontologies capture concepts and relations referring to information and knowledge resources and processes at the university, parts of individual courses, kinds of involved learning objects, instructional strategies, learning styles, student profiles, educational goals, but without direct relation to particular courses content.

The concepts define types of both internal and external occurrence in the TM. The relations between concepts define natural, rational or recommended navigation paths. E.g. for each instance of the concept "lecture", the relevant study materials are assigned through occurrences "presentation", "assignment" and others. The relation "is previous to" connects

instances of "lecture" concept and can be used for browsing presentations from particular lectures given during the semester.

4.3. Population of Topic Maps

Ontologies are static models of domains; here these models capture AI concepts and their relations, together with concepts and relations from the field of education and learning. The ontology-driven TM portal is web-based application that reuses ontologies for the navigation in resources. Our TM application was realized in freeware version of Ontopia Omnigator software [11]. For better explanation of how the application looks and works, we present few snapshots with comments.

The fisr shansphot (Fig. 1) presents the page of topic "Artificial intelligence". The page contains names of the topic in two languages (English, Czech), list of associations (i.e. hyperlinks to the topic pages that catch lower and upper categories in the ACM CSS, and a hyperlink to a topic page of the course dedicated to AI), two internal occurrences (strings), where definition of the topic are presented, list of external occurrences (hyperlinks) to relevant online resources.

Fig. 1. - The page of topic "Artificial intelligence"

The second snapshot (Fig. 2) shows the page of topic "Prolog", i.e. particular programming language and also important concept from AI. The association "ACM descriptor" catches the relation of Prolog to the ACM CCS. Course associations show what courses at the FIM UHK are related to Prolog. Association "has implementation" informs about Prolog programming environments (topics pages representing these environments are linked with official websites of their vendors). Internal and external occurrences refer to Prolog resources (books in the university library, licensed software installed on the university computer network etc.).

Fig 2. - The page of topic "Prolog"

The third snapshot (Fig. 3) is dedicated to information about particular course – here it is Logic programming. The page presents name of the course, study programs where the course is taught, name of the guarantee of the course and of teaching assistants, prerequisites and recommended semester, list of resources.

The last snapshot (Fig. 4) is focused on a particular lecture in the Logic programming 1 course. Different information and resources are available at this page, e.g. brief syllabus of the lecture, note about a PowerPoint presentation file stored in shared folders (direct url-addressing was not possible from our TM tool), hyperlinks to html-pages stored in the WebCT virtual study environment server Oliva, where traditional e-courses are provided, hyperlinks to topic pages of previous and following lectures and exercises where content of current lecture is reused, etc.

Fig. 3. - The page of topic "Logic Programming 1 course"

Snapshot 4. - The page of topic "Lecture 5 in LP1 course"

5. Conclusion

In the paper we suggested to apply the Topic Map technology for integration of study resources, used in different courses of AI offered to different audiences at the same educational institution.

Our proposed solution adopts part of the ACM Computer Classification Schema to define the hierarchy of themes that are expected to be studied by AI researchers. Other parts of the ACM schema would be applicable in description of resources dedicated to other subareas of computer science. Moreover, as individual Topic Maps can be merged, it would be possible to integrate AI-related study resources of different educational institutions, or computer science study resources of whole study programmes etc. Here it is important to note that not only online resources, but also other kinds of information and knowledge resurses, including contacts to human experts, best practices records etc. can be involved in a Topic Map. Then our TM application can be understood as a knowledge management solution designed for the purpose of educational institutions.

The idea of the solution is illustrated on pilot Topic Maps application implemented in an academic version of the Ontopia Omnigator Topic Maps software. Although this software was limited in certain functionalities, it helped us to present main features of our solution. More sophisticated implementation, interconnected with other information systems of the university, would increase the practical usability of the application. In relation to the dynamic development of TM technology that is primarily focused on design of TM Query Language, TM Constraint Language etc., numerous new opportunities to expand the idea of TM applications in the area of education promise to appear soon.

Acknowledgements

Resources

[1] ACM Computer Classification System website. Website: http://www.acm.org/class/1998/ (Accessed October 10, 2005)

[2] CLIPS Website. Website: http://www.ghg.net/clips/CLIPS.html (Accessed October 10, 2005)

[3] Dilleo website. https://e-dilema.uhk.cz/dilleo/ (Accessed October 10, 2005)

[4] Gibbins N. A broad ontology for describing an academic computer science community. Website: http://www.daml.org/ontologies/322 (Accessed: October 10, 2005)

[5] Heflin J. An ontology for describing universities and the activities that occur at them. Website: http://www.daml.org/ontologies/63 (Accessed 20 May 2005)

[6] Jess website.Website: http://herzberg.ca.sandia.gov/jess/ (Accessed 20 May 2005)

[7] LPA WinProlog Website. Website: http://www.lpa.co.uk (Accessed: October 10, 2005)

[8] Lugger G.F., Stubblefield W.A. Artificial intelligence: Structures and strategies for complex problem solving. The Benjamin/Cummings Publishing Company, Inc., Redwood, 1993

[9] Milton N. Personal Knowlege Ontology, PhD thesis, 2002. Website: http://www.epistemics.co.uk/staff/nmilton/ontology/standardframe.htm (Accessed 20 May 2005)

[10] Nilsson N.J. Artificial intelligence: A New Synthesis. Morgan Kaufman Publishers, San Francisco, 1998.

[11] Ontopia Omnigator website. Website: http://www.ontopia.net/omnigator/models/index.jsp (Accessed October 10, 2005)

[12] Pepper S. The Tao of Topic Maps. Website: http://www.ontopia.net/topicmaps/materials/tao.html (Accessed 20 May 2005)

[13] Rath H. The Topic Maps Handbook - white paper. Website: http://www.empolis.com/downloads/empolis_TopicMaps_Whitepaper20030206.pdf (Accessed 20 May 2005)

[14] Russell S., Norvig P. Artificial Intelligence: a modern approach. Prentice-Hall International, London., 1995

[15] WordNet Website. http://wordnet.princeton.edu/ (Accessed 20 May 2005)

Kontaktní adresy:

Mgr. Kamila Olševičová, Ph.D.
Fakulta informatiky a managementu
Univerzita Hradec Králové
Rokitanského 62
Hradec Králové 500 03
e-mail kamila.olsevicova@uhk.cz
tel. +420 493 331 225

prof. RNDr. Peter Mikulecký, PhD.
Fakulta informatiky a managementu
Univerzita Hradec Králové
Rokitanského 62
Hradec Králové 500 03
e-mail peter.mikulecky@uhk.cz
tel. +420 493 331 207