Integrating Ontologies into Learning Management Systems — A Case of Czech

Pavel Smrž

Faculty of Informatics, Masaryk University Brno
Botanická 68a, 602 00 Brno, Czech Republic
E-mail: smrz@fi.muni.cz

Abstract. The paper deals with the role of ontologies in e-learning. We report our research on the application of ontologies in intelligent platforms for self-learning and computer-mediated education that are developed at the Faculty of Informatics, Masaryk University, Brno, Czech Republic. The use of ontologies in the organization of knowledge about the area of learning management systems is tackled first. The next part discusses a new direction of our research aiming at automatic generation of e-learning self-tests based on information extracted from ontologies. A very new topic deals with the preliminary experiments on an ontology-based automatic evaluation of student essays.

1 Introduction

The importance of ontologies is widely recognized, especially in relation to the Semantic Web vision. Ontologies provide common conceptualization, i.e. identify and properly define concepts and relations among them. Although the role of ontologies is often presented only in the context of machines that will understand information on the future web, they are obviously helpful in present-day systems to define shared understanding and unification of conceptual framework.

The role of ontologies in e-learning and particularly in today's learning management systems is often neglected. This paper reports our work in this area. The first part stresses the fact that ontologies can be extremely important not only for machine processing but also for organization of knowledge in e-learning, as means to facilitate communication and cooperation among people and to enable interoperability among learning management systems (LMS). Also system engineering benefits — reusability, reliability, and specification — are briefly discussed.

The formal specification of the concepts and relations between them takes usually advantage of the new XML-family standards, RDF and OWL. The latter serves as a base for our recent research on automatic generation of exercise questions asking for relations between concepts given by an ontology. As the number of available ontologies for various subject domains will surely increase this direction of research has significant potential.

A completely new domain of ontology application discussed in the paper is their use for self-testing. As far as we know, this is the first application of ontologies for such a task. Relations among concepts extracted from a domain ontology are used in automatic generator of e-learning tests (and an automatic evaluation of students' answers). Even more visionary seem to be experiments with the inverse process. The content of students' essays is compared to the concepts and their relations extracted from an ontology, and the automatic evaluation of (information) quality of the text is performed. Although we are not able to present ultimate results on this topic, the preliminary experiments are very promising.

2 Shared Understanding as a Necessary Condition for Integration of Learning Management Systems

The range of LMS used or tested at Masaryk University, Brno, Czech Republic (MU) is rather broad. The most important ones are ILIAS (`http://www.ilias.uni-koeln.de/`) and MOODLE (`http://moodle.org`). Both systems are developed and distributed under the term of the GNU General Public License and provide open platform appropriate for the integration of NLP solutions. The actual project at MU aims at unification of the used e-learning platforms and their integration with the administrative information server IS MU [1]. Although separate systems would be more modular, easily maintainable and extendable, we opt for the integrated solution that will benefit from the permanent technical support and personal assistance of the administrative information server team.

The discussion of the needs of the unified e-learning platform at our university showed that participants not only differ in their opinions but also in the conceptualization of the subject area. Such a situation is well known in other areas and software engineering techniques exist that can facilitate the search for consensus. The necessary prerequisite of the consensus however is the sharing of concepts participants are speaking about. Ontologies are natural devices for the task.

An advisory group of current LMS users have been formed to be a partner for the developers of the IS MU. The members successively form the ontology of the area of LMS at MU. Of course, the core is based on the general conceptualization of e-learning that is connected to the international standards in the area. The support of open e-learning standards is rather crucial here as all new learning objects developed at MU should be developed in such a way to facilitate sharing and interchange. As the field of LMS is rather hot and many commercial as well as public-domain systems emerged recently, the newly produced learning object should also reflect the possibility of a switch between platforms without an information loss. The currently developed part of the ontology is compliant with LOM/SCORM standareds. It could help to put the platform independent standards into wide use.

What is specific for our understanding of LMS is the must to cover area of the education administration functions supported by IS MU. For example, the ontology has to deal with the mapping the LMS data model to that of IS

MU. Also, the concepts have to reflect the multilinguality of the systems at MU. Other important issue is the need for strong authentication. The Kerberos authentication protocol is commonly used in large systems with many users. Main advantages are security, reliability and transparency for the users.

The developed ontology provides also conceptualization for:

- Teacher-friendly editing and semi-automated publication of study materials prepared in DocBook and its derivatives (DocBook Slides).
- Integration with standalone system for semi-automated testing of students' programs in Java and other languages.
- NLP support for e-learning enabling to go beyond simple multiple-choice tests.
- Automatic linking of additional study materials available at MU based on intelligent text analysis.

3 Self-test Generation and Evaluation Based on Ontologies

The promising results of automatic question answering (QA) reported recently [2] led us to the idea to automatically generate questions and perhaps whole tests on the base of e-learning courses. It is usually easy to extract "what is" questions or ask about a particular fact explicitly stated. In some cases, the structure of the documents itself helps to identify the important knowledge; sometimes, the keyword extraction algorithm can be employed.

The tricky part of the process is therefore to create a procedure that would be reliable enough for the automatic checking of answers. Again, the basic QA module is satisfactory for the questions that expect factoid as the answer. However, it is much more difficult to automatically evaluate more elaborate answers. Although we are not able to present final results for this part yet, the preliminary ones show that the application of the same method as for the comparison of student essays with the content could provide at least a good approximation of the assessment that would be given by human teachers.

The weak point of the described automatic generator of exercises is its focus on the factography and impossibility to verify whether students really understand the content, whether they got the heart of the matter and are able to apply the obtained knowledge. Obviously, it is still far from today, when computers will be able to substitute human teachers in this respect. An interesting step, that at least simulates such a function and that is investigated by our current experiments, is the employment of standard ontologies for this purpose.

Table 1 presents the relations from SUMO (`http://www.ontologyportal. org/`) that have been used in our experiments. A set of test-question templates is defined for each type of the relations and natural language questions (in Czech) are generated.

The attempts to make the assessment of student knowledge as objective as possible and to reduce the teacher's work led to the spread of multi-choice

Table 1. Relations from SUMO covered by semantic patterns

beforeOrEqual	geometricDistance
causes	geometricPart
causesSubclass	geopoliticalSubdivision
completelyFills	greaterThan
component	greaterThanOrEqualTo
conditionalProbability	hasPurpose
confersObligation	immediateInstance
confersRight	immediateSubclass
connected	increasesLikelihood
connectedEngineeringComponents	independentProbability
connects	lessThanOrEqualTo
connectsEngineeringComponents	material
containsInformation	meetsSpatially
contraryAttribute	meetsTemporally
decreasesLikelihood	member
disjointDecomposition	occupiesPosition
disjointRelation	origin
domainSubclass	overlapsPartially
employs	overlapsSpatially
experiencer	overlapsTemporally

tests in last decades. It is a well-known fact that this form of testing has many disadvantages too and that the focus on them can easily produce test-experts rather than people understanding of the subject. The current technology offers various means to implement intelligent tests and to escape the trap of multi-choice tests. Our research concentrates on the integration of ontologies into the evaluation module of LMS. The experience shows that it is relatively easy to provide such functionality for short answers in the form of phrases described by simple grammatical patterns. The results of the first limited experiment are very promising as the method reduced the number of answers that needed to be processed manually significantly (only 31 % of the cases remained for the manual check). However, there are also many open questions concerning scaling-up the method of the answer patterns. The longer the answers are, the less matches between the pre-defined applications are found. A more general solution of the analyzing mechanism in the form of a general grammar is obviously needed. However, the available grammar of Czech is developed for a robust analyzer, so the ambiguity of the analysis tends to be rather high. Also it is much more difficult for authors that do not specialize in the computational linguistics to define the acceptable forms of answers. The assessment of open questions is definitively one of the actual topics of our research.

Although the current state-of-the-art in NLP does not allow full computerization of the general test assessment, there are areas, where the technology can already take over the role traditionally falling upon the teacher. We have re-

cently implemented the automatic evaluation of programs in the Java programming course at FI MU and also the phrase-structure grammars developed by students of the "Introduction to the computational linguistics" course. Students are obliged to provide their assignments in the designated form. Pre-defined tests are applied to verify the correctness of the provided solutions. We are to provide a similar function for a grammatical pre-checking of students' essays. Even though only a limited number of error types can be identified automatically, the method can still reduce a significant portion of the tedious teachers' work.

4   Conclusions

We currently work on a module that will be able to summarize the content of related documents based on the ontological information. The immediate target of the tool is the summarization of messages in course discussion groups. Sometimes, students are very active (the discussion is occasionally off-topic) and even a simple browsing through the discussion threads from previous runs of a course could present a tedious work. The automatic extraction of the most important points can significantly help newcomers.

Ontologies can serve as a launching point for the LMS personalization. Some e-learning courses already include special mechanisms to offer a path through the teaching material adapted to the needs of a particular user. Such an enhancement can improve the acceptance of the content by users which often face learning materials that are too in-depth in some parts and too sketchy in others. Also, the speed of presentation of the teaching material can vary according to the user's needs. Besides the above-mentioned assignment of essays, the simple statistical technique has been adopted that automatically evaluates the match of students' answers and determines knowledge levels of the students compared to relations extracted from ontologies. The solution is now ready to be applied generally to all e-learning courses provided by MU. It can be generalized to allow semi-automatic generation of tests for each part of the subject matter that would "send" students (back) to the parts they should go through again.

Acknowledgements

References

1. Pavlovic, J., Pitner, T., Smrz, P., Verner, J.: Customization of ILIAS and its integration with the university information system. In: ILIAS 2003, Cologne, Germany (2003)
2. Smrz, P.: Integrating natural language processing into e-learning — a case of czech. In Lemnitzer, L., Meurers, D., Hinrichs, E., eds.: Proceedings of the Workshop on eLearning for Computational Linguistics and Computational Linguistics for eLearning, COLING 2004, Geneva, Switzerland (2004) 1–10