Réka VAS

KNOWLEDGE REPRESENTATION AND EVALUATION – AN ONTOLOGY-BASED KNOWLEDGE MANAGEMENT APPROACH

Competition between Higher Education Institutions is increasing at an alarming rate, while changes of the surrounding environment and demands of labour market are frequent and substantial. Universities must meet the requirements of both the national and European legislation environment. The Bologna Declaration aims at providing guidelines and solutions for these problems and challenges of European Higher Education. One of its main goals is the introduction of a common framework of transparent and comparable degrees that ensures the recognition of knowledge and qualifications of citizens all across the European Union. This paper will discuss a knowledge management approach that highlights the importance of such knowledge representation tools as ontologies. The discussed ontology-based model supports the creation of transparent curricula content (Educational Ontology) and the promotion of reliable knowledge testing (Adaptive Knowledge Testing System).

Keywords: **competencies, ontology-based model, higher education**

The Bologna Declaration (Bologna Declaration, 1999) aims at providing solutions for the problems and challenges of European Higher Education. One of its main objectives is the introduction of a common framework of transparent and comparable degrees that ensures the recognition of knowledge and qualifications of citizens all across the European Union. Clearly, measuring knowledge of students in a reliable and objective way is a cornerstone in the implementation of the common framework. In the paper we will discuss a testing approach which incorporates an Educational Ontology and its application to certain knowledge areas.

The European Higher Education Area is structured around three cycles where each level has the function of preparing the student for the labour market, for further competence building and for active citizenship. Accordingly, the main focus areas of the paper are the following:

- One goal is to establish *ontological relation between competencies acquired during a training program and labour market requirements*. That means that the selected tool for building a bridge between the needs of the labour market, the content of curricula and competences are ontologies.
- Another critical aspect of this research is to provide support for the adaptive knowledge testing of students with the help of the developed Educational Ontology. Parts of the curricula of Business Informatics will be represented and analyzed in the Educational Ontology model.

This paper will focus on the higher education aspects of the approach, at the same time the model can be easily adapted to *business environment* as well. Life long learning has become a leading philosophy in corporate trainings. Employees are expected to be open to acquire new skills, attitudes and knowledge at any time. The discussed ontology-based approach provides an excellent opportunity for self-training and self-assessment that can be also integrated with corporate e-learning sites.
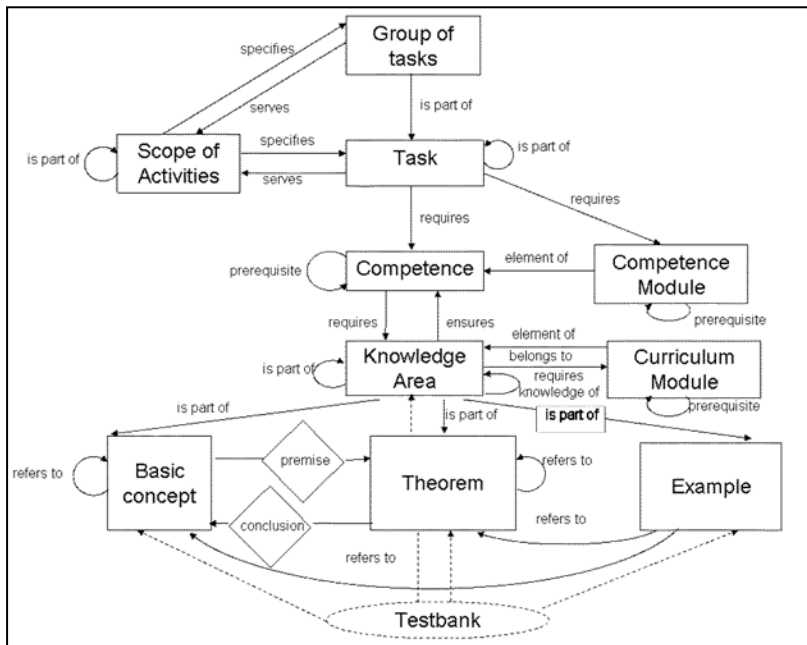
Educational Ontology

There is no doubt that ontology is a key concept in modern computer sciences. It is also a popular topic in the semantic web community and in the field of developing computer systems that are based on formal ontologies. The wide spectrum of application areas also proves that both business and scientific world has acknowledged that the detailed exploration of semantic relations must stand in the middle of knowledge mapping and proc-

Educational Ontology Model

esses description – beside the precise definition of concepts (Corcho – Gómez-Pérez, 2000; Gómez-Pérez – Corcho, 2002). That is what ontologies provide.

Corcho and his colleagues – by taking into consideration several alternative aspects – constructed a rather precise and practical definition of ontology: *„ ontologies aim to capture consensual knowledge in a generic and formal way, so that they may be reused and shared across applications (software) and by groups of people. Ontologies are usually built cooperatively by a group of people in different locations."* (Corcho et al., 2003: p. 44.). Characteristics highlighted in this definition also have a key role concerning the Educational Ontology.

The scope of curricula taught in Business Informatics training program is broad and curricula in general are substantively different in nature, which clearly poses a challenge on the ontology model building process. It should be also taken into consideration that the struc-

ture and content of a subject might be at least partly different in different institutions. Major classes of the ontology that were developed in the first cycle should meet these challenges. Taking all these into consideration the model of Educational Ontology is depicted by Figure 1 using the following notation:

- Rectangles sign classes.
- Arrows depict 0-N relations (so a competence may have several prerequisites, scope of activities may specify more tasks at the same time and it is also possible that a competence those not have any prerequisites). *(Figure 1)*

Figure 1

Knowing the amount of work required producing ontologies even for the simplest concepts, during development we focused on the provision of easily definable and applicable classes and precise determination of relations. This section gives a description of all of the classes in the ontology.

"Scope of Activities" Class

The "Scope of Activities" class contains all of those professions, employments and activities that can be successfully performed with the acquisition of those competencies that are provided by the given training program.

"Task" and "Competence" Class

A job consists of numerous tasks that should be executed in the course of everyday work. At the same time the employee must posses certain competences to be able to accomplish tasks relating to her position. So each task should be in "requires" relation with competences.

On the other hand one scope of activities should be in direct "specified by" – "served by" relation with tasks. This way the given scope of activities prescribes a number of concrete tasks that will define concretely required competences.

"Group of Task" and "Competence Module" Class

By defining separate classes for tasks and competences, not sets (competence modules, group of tasks), but their elements are connected to each other. At the same time the "Group of Tasks" and the "Competence Module" classes should be entered to the model to enable the definition of sets of tasks and competences as well and ensure further ways of comparison.

"Knowledge Area" Class

Knowledge areas and competences are connected directly with the "requires" and "ensures" connection. (A competence requires the knowledge of a given knowledge area and the good command of a knowledge area ensures the existence of certain competence(s).)

The class of "Knowledge Area" is an intersection of the ontology, where the model can be divided into two parts:

- One part of the model describes the relation of knowledge areas and labour market requirements with the help of the above-described elements.
- The other part will depict the internal structure of knowledge areas.

The internal structure of knowledge areas must be refined to allow of effective ontology construction and the efficient functioning of the adaptive knowledge testing system.

"Knowledge Area" is at the very heart of the ontology, representing major parts of a given curriculum. Each "Knowledge Area" may have several Sub-Knowledge-Areas through the "is part of" relation". Not only the internal relations, but relations connecting different knowledge areas are important regarding knowledge testing, too. This is described by the "is part of" relation. At the same time another relation has to be introduced, namely the "requires knowledge of" relation. This relation will have an essential role in supporting adaptive testing. If in the course of testing it is revealed that the student has severe deficiencies on a given knowledge area, then it is possible to put questions on those areas that must be learnt in advance.

For the sake of testing all of those elements of knowledge areas are also listed in the ontology about which questions could be put during testing. These objects are called knowledge elements and they have the following major types: "Basic concepts", "Theorems" and "Examples". In order to precisely define the internal structure of knowledge areas relations that represent the connection between different knowledge elements also must be described *(See Figure 1).*

Testbank

In order to provide adequate support by the Educational Ontology for the adaptive testing system several theoretical foundations and conceptions must be laid down in this model concerning the test bank. One pillar of the testing system is the set of test questions. Accordingly all test questions must have the following characteristics:

- All questions must be connected to one and to only one knowledge element or knowledge area. On the other hand a knowledge element or knowledge area may have more then one test question. This way the Testbank is structured by the Educational Ontology.
- All questions should be weighted according to their difficulty.
- Test questions will be provided in the form of multiple-choice questions. Therefore parts of a question are the following:
  – Question
  – Correct answer
  – False answers

The Testbank does not form an integral part of the ontology. This means that questions do not have to form a part of the ontology if we want to represent correctly a given curriculum. That is the very reason for connecting the Testbank to other elements of the ontology with dashed lines *(Figure 1).*

This way the ontology model that provides the base for the application (the adaptive testing system) is completed. The only task that must be accomplished before starting the construction of the adaptive testing system is to lay down the main principles of our own adaptive testing methodology and work out its process.

Principles of Adaptive Knowledge Testing

In contrast with traditional examination the number of test items and order of questions in an adaptive test is only determined during writing the test itself with the goal of determining the knowledge level of the test taker as precisely as possible with as low number of questions as possible (Linacre, 2000). Adaptive testing is not a new methodology and despite the fact that it has many advantages compared to traditional testing, its application is not widespread yet. This research has focused on the computerized form of adaptive testing; whose main characteristics – independently from the methodological approach – are the following:

- The test can be taken at the time convenient to the examinee; there is no need for mass or group-administered testing, thus saving on physical space.
- As each test is tailored to an examinee, no two tests need be identical for any two examinees which minimizes the possibility of copying.
- Questions are presented on a computer screen one at a time.
- Once an examinee keys in and confirms his answer, he is not able to change it.
- The examinee is not allowed to skip questions nor is he allowed to return to a question which he has confirmed his answer to previously.

- The examinee must answer the current question in order to proceed onto the next one.
- The selection of each question and the decision to stop the test are dynamically controlled by the answers of the examinee (Thissen – Mislevy, 1990).

We have elaborated a methodology of adaptive testing that provides help in determining the knowledge level of the student with asking as few questions as possible. The main principles and steps of this methodology are discussed in the next sections.

Goal of Examination, Problem Introduction

In the course of examination we select from a previously created set of questions. All test items are multiple-choice questions that have the following characteristics:

- One – exactly one – knowledge area is attached to one question.
- Text of the question.
- Set of answers in which the correct answer is indicated separately.
- Difficulty level of the test question (on a scale form 1 to 100).

The goal of knowledge evaluation and the ontology together determine the scope of questions that belongs to the given examination. The test taken by the student contains these questions. Based on the answers and their distribution on the ontology structure we are able to determine the knowledge level and the missing knowledge areas of the student. Accordingly the characteristics of the methodology are the following:

- If the candidate's performance exceeds a predefined pass level, meaning that she knows the given area on an adequate level, then she succeeded at the exam.
- If the candidate's performance is below the required level, meaning that she does not know the given area, then she failed.
- Finally, if the candidate is approaching the required knowledge level (from above or from below), then the final score can only determined after testing the sub-areas of the given knowledge area.

This means that even if the candidate knows the main knowledge area on a suitable level it must be assured that candidate's deficiencies on sub-areas can also be determined. In other words the methodology must ensure the possibility of examining the sub-areas of the given knowledge-area.

Functioning of the Adaptive Testing System

Before starting the exam several kick-off parameters must be provided. Beside the determination of these parameters the phases of the exam (the testing algorithm) is also discussed in this section.

(1) Start

Before testing can be started the following parameters must be defined in the system:

- a 'T' knowledge area, that we want to test
- the $0 \leq low \leq pass \leq high \leq 100$ values, and
- a true/false value, that describes whether the sub-areas of the given 'T' knowledge area must be tested or not.

The role of *low*, *pass* and *high* values are the following: In each step of the test the actual knowledge level of the candidate must be estimated. Pass level is that level which must be achieved to pass the exam. If the candidate constantly remains under the low level, then the exam stops quickly and the candidate fails. In this case it is unnecessary to examine the sub-areas, since the candidate does not know anything about knowledge area 'T'. High level means the opposite. If the candidate permanently remains above the *high* level her knowledge concerning knowledge area 'T' is much more then enough, so we take as the candidate knows the hole area and she has know deficiencies.

(2) Phases of Testing

Having defined the initial conditions the test can be started. The course of testing is based on the Rasch model (Rasch, 1992) and has the following steps:

- The system asks some question whose difficulty levels are around the pass level.
- Based on the answers the knowledge level of the candidate is estimated, we call this value *D*.
- Further questions are asked; whose difficulty level is always above the actually estimated knowledge level and based on the last answer the value of *D* is corrected. We continue until one of the stop rules is met. (Further details concerning stop rules are provided in the next section.)
- If D < *low*, then the knowledge of the candidate is absolutely missing, so testing can be finished.
- If D > *high*, then the knowledge of the candidate completely, so testing can be finished again.

If *low* < D < *high* then the knowledge of sub-knowledge areas must be tested too. All sub-knowledge areas are tested and results of each sub-area testing are summarized. If the summarized result is below the pass level then the candidate failed, otherwise succeeded.

(3) Stop Rules

Finally those rules have to be defined that determine when to terminate testing, ensuring that evaluation is accomplished by as few questions as possible. If there are no more questions attached to the knowledge area then obviously we do not ask more questions. Testing is also terminated if after a minimal number of questions (or after series of questions that have affected all sub-areas) D < low or D > high. At the same time if we ask 1,5 more then the minimal question number and low < D < high, then instead of knowledge area 'T' its sub-areas are examined more thoroughly. We use the same algorithm for each sub-area testing. If testing concerning all areas is accomplished, then the actual values of D is calculated based on questions and answers accomplished in the meanwhile. If D < low then the candidate failed, otherwise succeeded.

This testing system is built into Protégé-2000 ontology editor that was also used for building up the Educational Ontology. This way the efficient and effective cooperation of the two systems is ensured.

Implementation of the Adaptive Testing System and Further Experiences

Beside the evaluation of students' knowledge level, another goal is to measure the effectiveness of teaching. This has been kept in view during the whole course of development. It is practical to apply the top-down approach in the course of constructing the ontology model of curricula. This means that most general concepts were defined first and then more and more specialized concepts were only defined afterwards.

Let's take Knowledge Management curriculum for which we have already developed ontology model. In our case the main knowledge areas of Knowledge Management are the following:

• Basics of Knowledge Management
• Knowledge Management Strategies
• Technological Support of Knowledge Management
• Knowledge Creation and Development
• Knowledge Transfer
• Intellectual Capital
• Organizational Questions of Knowledge Management
• Human resource Management Questions of Knowledge Management

These knowledge areas are divided into sub-knowledge areas. In this case 28 sub-areas were identified. In some cases sub-areas are further divided into other sub-areas. The primary taxonomy is formed this way. The relations among knowledge areas must also be defined; one must examine whether a given knowledge area requires anticipatory acquirement of another knowledge area or not. In Knowledge Management there are several examples for such situation. For example understanding of the "Knowledge Management Strategies" area requires knowledge of some areas of Strategic Management (e.g.: resource based approach, core-competence approach). Finally basic concepts, examples, theorems and their relations must be defined.

In the next step the other "side" of the ontology must be developed, meaning that one must describe the competences that should be acquired by a student if she makes herself master of Knowledge Management. Competences are defined in the official accreditation documents of each training program. On BSc and MSc levels in Business Informatics 54 competences are distinguished. Furthermore 3 elements were uploaded into the "Scope of Activities" class (Knowledge Management Expert, Knowledge Management Project Leader and Chief Knowledge Officer) and tasks relating to the given job were also defined in the ontology.

When instantiation of the ontology model by the Knowledge Management curriculum was completed then students could be tested. As discussed earlier the system provides estimation for the knowledge level of the candidate concerning each knowledge area based on the difficulty level of the asked questions and answers. The results are evaluated as well as summarized; the result is shown to the candidates. This evaluation report also indicates how many questions were for each knowledge area. For example, we tested the knowledge of a candidate concerning to the 'Knowledge Transfer' knowledge area. One of the sub-areas of 'Knowledge Transfer' is called 'Cooperation'. Eight questions were asked concerning this sub-area and from these 8 questions 7 questions are connected to further sub-knowledge areas. In our case the knowledge level of the student on this knowledge area was 75%. But on other sub-knowledge areas (of the Knowledge Transfer main area) the performance of the candidate was very low (only 21%), this way the whole area (the Knowledge Transfer area) was not accepted and the student failed on this exam.

Beside the introduction of testing, this example has also shown how the ontology model of a given curriculum can be built up and the content of a given curriculum can be aligned with the needs of labor market through competences.

Exploring missing knowledge areas

The Educational Ontology can also be used for other applications as a base. In this section another approach is introduced, where the exact determination or estimation of the knowledge level is not a goal. The emphasis, however, is on the detailed exploration of missing knowledge areas and on the discovery of each knowledge area where the candidate lacks some knowledge (Borbásné Szabó, 2006).

This approach was worked out and applied in the Human Resource Development (HRD) – "Development of Knowledge Balancing, Short Cycle e-Learning Courses and Solutions" project. In this project 12 Hungarian higher education institutions were involved that all developed a curriculum and a relating ontology model as well. The developed e-learning curricula are the following:

- Databases
- Application Development
- Information Technology
- Linear algebra
- Mathematics
- OO programming and Java
- Operational Research
- System development I.
- System development II.
- Computer architecture
- Accounting
- Management

Twelve institutions and 291 students were involved in testing the curricula and the testing system. In total 74504 questions were responded by the participating students.

Table 1 summarizes the most important observations, average response time (in sec.) per knowledge areas, average effectiveness, standard deviation of results and number of responses. Average effectiveness is defined as a ratio of good answers given for a certain question and the total number of questions. There are other two characteristics of answers that are worth to be taken into consideration. One – concerning effectiveness – is the standard deviation of results. It can be seen that in most of the cases responses have spread the on the 50% of the range of possible responses. Only Information Technology showed significant difference, since results were born in a narrower interval. The "best" performance (average effectiveness) was produced in case of Information Technology curriculum. At the same time, the standard deviation is the lowest and the number of responses the highest on this knowledge area. Taking into consideration the effectiveness of responses this test was easier then the average. Another characteristic is response time; in which there is no subsequent difference between the subject, and solemn question what is the reason for students having only spent 26 seconds with answering the questions. It is especially interesting due to the complexity of questions, since in case of the Mathematics questions we suppose that more time is required for interpreting and answering the questions.

Table 1

Results of System Testing

Tests of Knowledge Areas	Average response time (sec)	Average effectiveness	Standard deviation of results	Number of responses
Databases	23.284	55.20%	49.73%	8175
Application Development	26.085	66.07%	47.35%	5730
Architectures	27.000	77.30%	41.90%	2982
Information Technology	24.545	85.28%	35.43%	23618
System Analysis and Development I.	27.000	76.33%	42.52%	1259
OO programming and Java	26.217	74.25%	43.73%	8854
Mathematics, Linear algebra, Operational Research	26.176	57.36%	49.46%	14201
Management Accounting	26.069	70.77%	45.49%	2145
Management and Organizations	27.000	61.96%	48.57%	1112
System Analysis and Development II.	26.875	65.93%	46.37%	6428
Average	*26.025*	*69.05%*	*45.06%*	*74504*

Conclusions

The Bologna Declaration set the objective of the introduction of a common framework of transparent and comparable degrees. The HRD project in Hungary concerned about the possible way of achieving this objective. It had the primary goal to develop a competitive training evaluation system that promotes the transition between the BSc and MSc levels of higher education. Business Informatics training program was selected to demonstrate the solution.

Ontologies, as an abstract and precise way of knowledge description were used to describe the curricula and the needs of the market in the project. This pilot demonstrated and proved the feasibility of using such tools.

The main obstacle of mobility is that the outputs of different bachelor programs do not provide homogeneous input for a given master program. For that very reason such methodology must worked out with which it is possible to explore what a student does not know. In other words those knowledge areas must be determined that the student must learn before starting the given master program. In order to enable closing up of a student, a dedicated personal teaching material must be assigned to the student. After having acquired the newly assigned material the candidate has to be evaluated again.

This approach proved to be fruitful and we are going to extend the scope of this research to wider academic community.

References

Bologna Declaration (1999): The Bologna Declaration of 19 June 1999. [online], http://www.bologna-berlin2003.de/pdf/bologna_declaration.pdf

Borbásné, Szabó (2006): Educational Ontology for Transparency and Student Mobility between Universities. in: Proceedings of ITI 2006 (under progress)

Corcho, O. – Gómez-Pérez, A. (2000): Evaluating knowledge representation and reasoning capabilities of ontology specification languages. in: Proceedings of the ECAI 2000 Workshop on Applications of Ontologies and Problem-Solving Methods, Berlin

Corcho, O. – Fernández-López, M. – Gómez-Pérez, A. (2003): Methodologies, tools and languages for building ontologies. Where is their meeting point? in: Data & Knowledge Engineering, Vol. 46, p. 41–64.

Linacre, J.M. (2000): Computer-adaptive testing: A methodology whose time has come, in: Chae, S. – Kang, U. – Jeon, E. – Linacre, J. M. (eds.): Development of Computerized Middle School Achievement Tests, MESA Research Memorandum No. 69., Komesa Press, Seoul, South Korea

Gómez-Pérez, A. – Corcho, O. (2002): Ontology Languages for the Semantic Web. IEEE Intelligent Systems, Vol. 17, No. 1, p. 54–60.

Rasch, G. (1992): Probabilistic Models for Some Intelligence and Attainment Tests. MESA Press, Copenhagen and Chicago

Thissen, D. – Mislevy, R.J. (1990): Testing Algorithms. in: Wainer, H. : Computerized Adaptive Testing, A Primer. Lawrence Erlbaum Associates, Publishers, New Jersey, p. 103–135.