Using Semantic Web technologies to create a learning management system based upon Mixed Diagnostic Tests

Anna E. Yankovskaya, Vladimir V. Razin, Yury N. Dementyev, Valentin V. Popov

Abstract—An approach to the construction of learning management systems based on combined usage of Semantic Web and Mixed Diagnostic Tests that are a compromise between unconditional and conditional components is considered. The main concepts and definitions are provided. Ontology models that describe the academic subjects and test assignments are proposed. Construction of learning management systems that use the developed models as knowledge base is proposed. Directions of follow-up research are provided.

Index Terms—learning management systems, Mixed Diagnostic Tests, didactic units of academic subject, Semantic Web, ontology.

I. INTRODUCTION

Learning management systems have been widely used in the worldwide educational practice during more than 50 years [1-6]. Currently many automated, including intelligent, systems for managing various aspects of educational process have been developed [6-12]. Such systems are usually intended not only for forming and maintaining the structure of educational process, but also for organization and support of distance learning. As a rule, learning management systems consist of sets of modules that can be grouped into libraries and used in real-time mode. An important component of the process of learning management is a control of quality of mastering the learning material. That control is carried out according to educational standards in the field of education of one or another state. Further discussion will be based on Federal State Educational Standards of the Higher Professional Education of Russian Federation (FSES HPE, further FSES) that include a list of general and professional competences (GC and PC accordingly) that should be achieved by an University graduate. However, these competences do not presume a concrete list of questions and tasks that should be included into state examinations. So, both everyday control of the education quality, and final attestation of the graduate are, in fact, completely under control of an educational institution. Majority of institutions of higher education develop and maintain their own education quality control systems, including by means of modules for construction and editing the test questions that are included in the majority of modern learning management systems. Nevertheless, currently the activities are carried out in order to create united state-level system of education quality control in Russian institutions of higher education. Nowadays Russian Ministry of Education is conducting an experiment. Since 2005 Federal examination in the field of higher professional education (FEHE) has been introduced to the Russian universities. This experiment is based on computer testing of students and their skills to meet the requirements of state educational standards (SES). Testing students' knowledge is carried out using the Internet (off-line or online mode). Concept base of the student training quality evaluation model is an assessment of the student’s success in mastering of all didactic units of a given course. Test questions, which the examination FEHE consists of, are selected from common base of questions randomly for each didactic unit of a given subject. Number of students that have learnt all didactic units of given subject is an index of its overall mastering by a group. Training students for passing the FEHE examination is an important task for each institute of higher education. Most of existing and widely used learning management systems include a set of tools to create and/or edit test questions in order to maintain everyday and final education quality control. Unfortunately, in most cases the evaluation of passing the test by a student is binary: «the test is passed» or «the test is failed». Such evaluation method is barely flexible and, besides, it does not allow assessing the degree of mastering one or another section of the course. One of methods that allow solving that problem involves using the Mixed Diagnostic Tests (MDT) for the education quality control. MDT was proposed in [13-16] and are a compromise between unconditional and conditional components. These tests use fuzzy logic and allow making a differential evaluation of the quality of mastering the learning material. Majority of existing learning management systems are focused mainly on using the web to provide access to learning courses and test tasks. Besides, as it has been already mentioned above, the currently developed examination FEHE is supposed to be an internet-based testing and, therefore, it is also carried out by using the web. Modularity and ability to re-use separate elements of existing learning management systems are among their strengths. However, the data presentation in the web that is used by developers because of considerable focusing of these systems to online access is poorly suitable both for computer analysis and processing, as well as for standardization. Existing standards in this field, such as SCORM [17] require, as a rule, either substantial revision of existing system, or initial focusing of the developed system to the educational standard. To solve problems that are caused by difficulties of both web data structuring and its suitability for the computer processing, a
global initiative is carried out in order to reorganize the Internet data structure and convert it into the so called Semantic web [18] that gives much more capabilities to perform effective search and analysis of data, both by human and software agents. Modern Web, or World Wide Web (WWW), is based on HTML pages that contain information that can be extracted by human by means of a browser program. Semantic web stores the information in form of semantic network that provides a possibility to extract facts and make a logical deduction based on the facts by using a software client. Moreover, Semantic web operates in parallel with WWW and on its basis, by using HTTP protocol and unique resource identifiers URI. Semantic networks, on which the Semantic Web is based, are used as an addition to ordinary web-network and store metadata about both context of a web-page (name, authors, date of creation and/or last editing), and its content (main concepts that are used in text on the page, knowledge domain etc.). The initiative to create such a large-scale addition to World Wide Web has repeatedly been questioned. Two major obstacles to its implementation are the need to make authors and / or administrators of web-pages create semantic networks with metadata, and the need to create top-level semantic network describing the world as a whole, dividing it into separate concepts. However, both the idea itself, and existing Semantic Web technologies can be used for the creation and development of information systems within a single organization and, in particular, to create a learning management system. Semantic networks, which store metadata in the Semantic Web, are based on the description logic, are very flexible model of the information representation (and, strictly speaking, allow us to move from working with data to working with knowledge), and can be easily re-used. The main approaches to the use of Semantic Web technologies in education [19-23] suggest the use of domain models, carefully designed in advance, either for structuring teaching material, or for annotating existing documents (teaching materials, visual aids etc.). However, the main efforts of the experts fall on the creation of a domain model, and not on supporting the learning process. In addition, these approaches typically do not include methods to control the quality of education that differ from using tests based upon the above-mentioned binary evaluations. In this paper, an approach that involves the use of Semantic Web technologies to create a knowledge base of software learning management system using MDT is discussed, the basic definitions and concepts are given, the problem being solved is stated and formulated. Semantic models for describing the learning courses are proposed, illustrative example of the description of the course "Information technology" is provided, and directions for further research are given.

II. BASIC CONCEPTS AND DEFINITIONS.

PROBLEM STATEMENT

Let us consider the learning management system as a management system in which the object is a student, and the control action is a change of the structure of the course or repeated studying of the courses included in the curriculum. Sensors of such a system can be regarded as sets of test tasks, and the parameters of the control object – as numerical estimation of success of passing the tests. The approach to the evaluation of the quality of education, based on MDT and described in [10-15], is based on fuzzy logic and allows the use of probability estimation both for the complexity of the test, and for the result of its passing by student. This approach allows achieving flexibility in the parameter description of the control object as well as the possibility to use fuzzy control methods for the control object.

Student is a person that is passing some course, a control object. The main problem to be solved by means of MDT is the problem of classification of students. In this case, the degrees of mastering of a particular teaching course act as classes ("does not master the material", "satisfactory masters the material", etc). Teaching course in this paper is a set of educational materials combined by one educational subject (subject field).

Diagnostic test is a set of group characteristic features that distinguish any pair of objects belonging to different classes. In this paper we discuss two types of diagnostic tests: conditional and unconditional diagnostic tests. Unconditional diagnostic test assumes the simultaneous presentation of all characteristic features of a given object that are contained in a test, and the conditional diagnostic test - the sequential presentation of characteristic features depending on the values of characteristics that were presented earlier. Conditional diagnostic test in this case includes the obligatory calculation of the current estimation of the respondent, depending on the complexity of the test that has been predetermined by expert. Mixed diagnostic test (MDT) includes both unconditional and conditional components. After answering each question, the test score received by respondent is calculated according to a complexity of the question predetermined by test author.

Theme (module) is a component of the structure of the learning course devoted to the relevant section of the discipline. In this paper it is assumed that the theme as part of the discipline corresponds to some didactic unit. Lesson is an element of the theme dedicated to the consideration of certain issues included in that theme. Educational materials being discussed during one lecture (two academic hours), correspond to a lesson. Lesson is the minimal element of the course structure. The theme may include either other themes (subthemes) or lessons, but not both at the same time.

Ontology is a conceptual model that defines the terms used to describe and represent a domain of knowledge, and includes computer-readable (understandable for the software)
The main components of ontology are:

- **concepts** (classes) of the knowledge domain being described,
- **instances** of these concepts,
- **Properties**, which include object properties (that link concepts and their instances to each other) and data properties or attributes (that link concepts and their instances to some values).

A set of concepts and relations between them determines the overall scheme of storing data represented as a set of assertions about instances of concepts, or ontology axioms. These statements, or *triples*, have the form of triples "subject-predicate-object". Ontology is a central concept of the Semantic Web technology stack (or Semantic Web Stack). In fact, ontology is a semantic network storing any data described by means of Semantic Web. The main languages, by which the ontologies are described, are RDF [24] and OWL [25] languages standardized by the W3C. During the construction of MDT for some subject, the set of its sections is divided into conditional and unconditional parts. Then, using the algorithms described in [12], a sequence of questions that correspond to didactic units of each part is generated. We will give an example of MDT construction for the course "Information technology" for students of "Industrial and civil constructional engineering" specialization. The main sections of this discipline are:

1) the concept of information;
2) a general description of the data collection, transmission, processing and accumulation of information;
3) hardware and software implementation of information processes;
4) decision models of the functional and computational tasks;
5) algorithms and programming;
6) high level programming language;
7) databases;
8) the software and programming technology;
9) Computer graphics.

Sections of the discipline are associated with grouped characteristic features and didactic units are associated with the characteristic features. Each grouped characteristic feature contains a different number (from 4 to 8) of didactic units (characteristic features). In this case, it is acceptable to include the sections from the first one to Section 3.1 in the unconditional component of the test, and sections from 3.2 to 9 inclusive – in conditional component. It also makes sense to include tasks in sections 6-9 in the final part of the test. Figure 1 shows an example of the sequence of the MDT passing for this discipline, presented in the form of the search tree. Nodes of the tree represented by rectangles correspond to the sections of discipline, solid edges - to unconditional transitions, dotted edges - to transitions performed in case of the successful passing of the unconditional component of the test for a specific section. An answers weight matrix A with dimensions $n \times k$ corresponds to each MDT, where $n$ is number of test questions, and $k$ is maximum number of possible answers to the test question. Elements $a_{ij}$ of the matrix A correspond to the weight of answer $j$ to question $i$. Each $a_{ij}$ is pre-set by an expert in accordance with the complexity of the test and the proximity of response $j$ to the correct answer to the question $i$, passing the unconditional MDT is considered successful when a total weight of the student's answers exceeds a predetermined threshold. Passing of the conditional component of MDT is only possible after successful completion of its unconditional component.

![Fig 1. Transition tree for the process of MDT passing.](image)

**Problem statement.** A discipline is given, which is divided into sections, each of which consists of didactic units. The course has to be described as Semantic Web ontology suitable for use with the methods of constructing Mixed Diagnostic Tests.

**III. DESCRIPTION OF THE DISCIPLINE BY MEANS OF SEMANTIC WEB ONTOLOGIES**

An approach to the description of academic subjects using Semantic web ontologies proposed in this paper is based on a system of concepts presented in Section 2. According to the proposed approach, teaching course corresponds to a discipline, and themes correspond to its sections. The approach involves the use of three main ontologies: Course Level, Knowledge Level and System Level. These ontologies are supposed to be described in RDF [24] and OWL [25] languages. General scheme of Course Level ontology is shown in fig. 2. Concept Course corresponds to a discipline. Its attributes include name, number of hours and description. Concept Course is linked with concept Theme that corresponds to theme by property „includes”, and with Knowledge Domain concept by property “refers To”. Property “refers To” is, in fact, a one-sided association and sets a reference from one class to another. Concept Knowledge Domain contains information about a knowledge domain. Attributes of the concept include name and description. As it can be seen in the fig. 2, the knowledge domain may be a part of any larger domain, or may include other knowledge domains. Concept Theme is designed to store data about the themes and can include sub-themes and lessons.
Theme attributes include name and description. Concept Theme is also linked to Knowledge Domain concept by property “refersTo” that allows unambiguous identifying of the knowledge domain that theme is related to. Concept Lesson stores data about a particular lesson, and has the name and summary attributes. Concept Class Element, “atomic element of a class”, is essentially an abstraction for child concepts described in the following ontology. Concepts Test Question and Test Answer describe test questions on a particular Class Element. A test question can be associated to any Class Element by means of property “checking”. Test Question is described by attributes “body” (the answer body) of type string and “weight” of type double. By means of a „has Answer” property each question can be associated with the possible answer, Test Answer, described by fields “body” of type string and “truthValue” of type double. Concepts Knowledge Element, Person and Event are not direct components of this ontology, but they are mentioned in this context to link it to another ontology discussed below. Knowledge Level ontology that is shown in fig. 3 describes possible "knowledge elements" present in the disciplines. The central concept of this ontology is the Knowledge Element concept, which is, in fact, abstract, and generalizes a number of other concepts. Concept Person is used to store data about a person and is associated with the concept Knowledge Element by property “author”. Thus, a person may act as an author for any instance of the concept that inherits Knowledge Element. Concept Event describes a random event. The concept includes two child concepts to specify the type of event — interval one and instant one (Interval Event and Instant Event). Concept Person is associated with Event participant, describing the possibility of a person to participate in an event or in any way be related to it. Concept Theorem formally is used to describe theorems, but in the wide sense corresponds to any statement for which a formal proof can exist. Concept Proof describes the proof for theorem (Theorem concept). Concepts Theorem and Proof are linked by properties „has Proof” and “proof For” respectively. Concept Pronouncement is used to describe the statement made by some Person. Concept Viewpoint corresponds to a point of view that differs from Pronouncement by the fact that it has property „viewpoint On” linking it to Knowledge Element concept, linking author, his point of view and the subject in question. Concept Term corresponds to a concept of term or definition and can be described as term in the wide sense (for example, the concept or object in mathematical reasoning). Concept Classification is intended to describe the classifications of some term. Classification and Term concepts are linked to each other by properties “of Term” and “belongs To” respectively. Concept Classification Attribute is responsible for defining criteria for classifying terms. Concept Law describes an arbitrary law or rule. It can be represented in the form of “if - then” statement. This concept is directly related to the class Term, as it contains a variety of terms and concepts and operates them. Concept Artwork is used to describe the metadata of various resources.
Its subclasses are Picture, Sculpture, Architecture, Music, Invention, and Text. All instances of these concepts are not intended for storing image files, music, etc., but only for describing the data about these objects. Concept Text is divided into two sub-classes - Book and Article. The third of the developed ontologies is called System Level. It is used primarily to describe sequences of elements for disciplines, themes and lessons. The principle of precedence is not submitted in a separate concept, and is represented as an attribute with an integer type xsd:int. The main resources are the audio and video materials, images, books and articles, as well as websites. Corresponding concepts are shown in Figure
4. Concepts Theme with Priority, Lesson With Priority and Class Element With Priority are used to store data about the sequence order or priority of themes of the course, each lesson theme and the concepts discussed in class, respectively. Concepts Video, Audio, Book, Article, Image, Web-resource describe resources corresponding to the basic elements of the course, for example, the image of a human (Person), recording of an event (Event), an article about the theorem and its proof (Theorem and Proof respectively). Attributes of such classes are links to related materials storage. Classes Resource and Class Element are linked by “describes” property.

An example of ontologies describing part of the previously mentioned course "Information technology” is shown in fig. 5 and fig. 6. In this example, the course itself is presented in the ontology Course Level (fig. 5) by instance InformationScience270102.65 of concept Course. It refers (property refers To) to knowledge domain "Information Science” represented by an instance Information Science of concept Knowledge Domain.

Part of the course theme "Information" (an instance Information of concept Theme, which is associated by property “includes” with the course InformationScience270102.65) includes two lectures (instances of concepts Lesson, Lecture_1 and Lecture_2), entitled (data property “name”) “The concept of information” and “General characteristics of the collection, transmission, processing and storage of information.” Each of these lectures includes the term "information", to which corresponds an instance Information of concept Term, which is described in ontology Knowledge Level (fig. 6). Ontology Knowledge Level includes a classification description of information (instance Information of concept Term) in the form of presentation and method of perception (instances Inf ByRep and Inf By Perc of concept Classification, respectively). Classification criteria are presented in the ontology by instances of concept Classification Attribute - Representation Method for classification according to method of presentation and Perception Type for classification in form of information presentation - and are associated with instances representing classifications by property defined By. Finally, the types of information included in these classifications are associated with them by property “belongs To”. Each of these types has its own term (instance of concept Term), for example, Graphic Information for the information presented in graphical form, Tactile Information for information perceived by tactile receptors, etc.

Fig. 5. Partial Description of «Information Technology» Course in Course level Ontology.
IV. CONCLUSION

Described ontologies form a common knowledge base, which allows to store and edit ontological models describing the academic subjects in the form of disciplines, divided into themes, which correspond to didactic units. In turn, the themes are divided into lessons, including descriptions of a wide range of concepts, phenomena and events that may be part of any discipline. An important element of the ontological model of discipline is a description of a test question that can be associated to any of the concepts discussed in the learning process. Indications of the complexity of test tasks allow selecting the whole base of test tasks according to the concept, theme and section of the course, with a single query to the ontology, and creating a set of Mixed Diagnostic Tests for quality control of mastering of learning materials by students. Ontologies described in this paper can be supplemented by ontologies of lower level describing specific domains. Furthermore, these models can be reused in the future, both for creating new courses using the elements that are already described, and for creating the knowledge domain ontology of higher level. The presence of descriptions of audio and video materials and network resources in ontological models allows to supplement the discipline by visual teaching materials and to structure the lists of additional literature on related elements of the course. In the future, we plan to develop architecture of a learning management system based on the approach described in this paper, and a faculty-scale educational portal based upon such system. That learning management system will be used for supporting various educational courses, including courses on road climatic zoning, diagnostics and intervention of organizational stress and diagnostics and prevention of depression.

V. ACKNOWLEDGMENT

This research is funded by grant from the Russian Foundation for Basic Research (project no. 13-07-00373a, project no. 13-07-98037-r_sibir_a and project no. 14-07-00673) and by grant from the Russian Humanitarian Scientific Foundation (project no. 13-06-00709).

REFERENCES

AUTHORS PROFILES

Professor Anna Yankovskaya obtained her DSc in Computer Science from the Tomsk State University in Russia. She is currently a head of the Intelligent Systems Laboratory and a professor of the Applied Mathematics Department at Tomsk State University. She is a founder of the Intelligent Systems Laboratory and a professor of Tomsk State University. She is the author of more than 600 publications and 6 monographs. Her scientific interests include mathematical foundations for test pattern generation, artificial logic, and advanced techniques in education. E-mail: ayyankov@gmail.com

Vladimir Razin is currently a Professor Assistant at the Department of theoretical basics of Information technology of the Computer Science Department at Tomsk State University. He has graduated from the Computer Science Department at Tomsk State University in 2010. His scientific interests include artificial intelligence, Semantic Web, digital libraries, natural language analysis, and situational management. E-mail: vladimir.moonbreeze@gmail.com

Yuriy Dementiev received Dipl.Eng., Dr.Sc. techn. in electrical engineering from Tomsk Polytechnic University, Tomsk, USSR, in 1975 and from Budapest University of Technology and Economics, Budapest, Hungary, in 1984. He is presently head of Department of Electrical Drives and Equipment of National Research Tomsk Polytechnic University. He has authored more than 200 published papers on valve inverter control in variable-speed electrical drives and synchronous motors with self-excited inverter. His research interests include electromechanical systems and electrical drives, mathematical modeling, expert systems, discipline diagnostic testing. E-mail: dementiev@tpu.ru

Valentin Popov is a Master student of the 2nd year of studies of the Computer Science Department at Tomsk State University. His scientific interests include artificial intelligence, Semantic Web, management systems. E-mail: vvp_ii@sbmail.com.