

Support for Knowledge Tests: Brief Summary of Regulations and Software

Bošnjaković, M., Andrija; Tartalja, I., Igor; and Protić, Ž., Jelica

Abstract—Intention of this paper is to brief the reader with overall status of the support for knowledge tests in contemporary software. Need and multiplicity of ways to ask a question and to administer a knowledge test have lead to development of numerous systems with different output formats, thus pushing the need for standardization to front. In order to classify relevant terms, two main criteria are taken into account – whether a term refers to a regulation or to a software product and term's generality. Additional criterion is applied to classify products: free of charge (even open source) or commercial. Examples are given for all of the classes constructed using these criteria. Also provided are list of further readings and brief comparison of relevant features in most prominent software tools, based on the proposed question-type taxonomy.

Index Terms—CAA, IMS QTI, knowledge test, LMS, question, quiz, SCORM, VLE.

1. INTRODUCTION

REGARDLESS of the field, need for assessment of human resources is ubiquitous within academia and industry. During this process, psychological, especially knowledge, tests are often used as a time-saving tool. There are three common phases during their life cycle: assembly, administration, and scoring. The term assembly denotes creation and/or selection of items for a test, administration is presentation of some or all items contained in the test to the examinees, while scoring is associating some, most often numerical, data with examinees' responses. The concept has been basically the same for decades, what changes over years is the increase in use of computers to automate the process.

One of important parts of e-learning is knowledge assessment, which belongs both to education and psychology. Within this context, tests are often referred to "knowledge tests" or "quizzes", items are referred to as "questions" or "problems", and examinees are referred to as "students" or "learners". In this paper, the first of the two options will be used for each of three terms.

This paper is created as a first part of broader survey conducted at School of Electrical Engineering at University of Belgrade during research and development that resulted in the assessment

software called *test* [1], which is made up of modules that help automate described phases of knowledge test's life cycle.

2. PROBLEM STATEMENT

There are many ways to create and to ask a question. There are many ways to select questions for a test and decide on the order of their appearance. There are numerous related software systems in existence, only simple listing of which would take a lot more than one journal paper. That is why there are many question types available in the existing software, some of which were impossible to use in a paper-and-pencil test, mostly due to their interactivity. Software products that support creation of diverse question types and various test assembly and/or administration methods are producing custom output, which can hardly be used by some other similar tool. This plethora of output formats caused the situation where most tools are not interoperable, which further lead to introduction of standards and specifications that regulate format of questions and tests (e-learning content in general).

Primary goal of this paper is to briefly inform the reader about key products and regulations in the field of e-learning and/or computer-based knowledge testing, rather than to provide in-depth analysis of subtopics. Additional goal is to compare support for most frequent question types in related software. Although important, test administration features will be only mentioned, while scoring will not be discussed at all in this paper, as it is most often trivial for knowledge tests.

3. ORGANIZATION OF THE PRESENTATION

This section lists criteria used to classify terms in this presentation, and then lists appropriate classes under meaningful combinations of criteria values, used to organize the paper.

3.1. Criteria

One has to distinguish a term that refers to a regulation (a standard or a specification) from a term that refers to particular software that is compliant or not to a regulation (criterion: *term nature*). Furthermore, both standards and tools can be focused only at questions and tests or to learning content in general (criterion: *generality*). Finally, software tools and systems can be either commercial or free to use and modify (criterion: *price*).

Manuscript received December 15, 2006.

A.M.Bosnjakovic, I.I.Tartalja, and J.Z.Protic are with the School of Electrical Engineering, University of Belgrade, Serbia (e-mails: andrija@etf.bg.ac.yu, tartalja@rcub.bg.ac.yu, jeca@sezampro.yu, respectively).

3.2. Classification

The first criterion classifies term in question as either *software* or *regulation*. These two classes are further divided using the other two criteria.

Applying the generality criterion to regulations (standards and specifications) classifies them to those specialized only to question and test representation (class: *specialized regulations*) and to those that cover more general area of describing learning content (class: *general regulations*). Applying the generality criterion to software products introduces other two classes. Tests can be created, previewed and/or administered through a Computer Aided Assessment (CAA) tool, specialized for one or more of these actions (class: *CAA software*); frequently, such tools have support for item parameters and their analysis, which are not as important to educators as to psychologists and psychometricians. Another common way is to use assessment component of e-learning software systems, which are most often called Virtual Learning Environment (VLE) or Learning Management System (LMS). These systems are designed to help instructors (teachers or skilled professionals) in the management of educational content for their learners (students or company employees) and help learners to use the content accordingly to their needs. Within scope of this paper, the term VLE will be used when referring to this kind of software (class: *VLE software*).

Finally, the price criterion is applicable only to software products, as regulations are inherently open to public and ones described in this paper are freely available. Software products can be further classified as *commercial* and *free*, which yields total of 4 classes pertaining to software.

4. REGULATIONS AND EXAMPLE SOFTWARE

This section introduces the most prominent relevant regulations and software.

4.1. Specialized standards and specifications

Regulations that focus only on question and test representation are almost always a subset of more general documentation.

4.1.1. IMS QTI

IMS Question and Test Interoperability [2] is a specification that was introduced and is constantly developed by IMS Global Learning Consortium (IMS GLC). It is a specification for a standard way of sharing assessment data, such as description of questions and tests and reporting of results. QTI deals with the structure of assessment related information rather than the way it is going to be rendered using some particular software. The QTI specification is now being implemented within a number of assessment systems and virtual learning environments. IMS GLC is maintaining a lot of other specifications related to learning content [3]. Although this specification is quite mature (the current version is 2.1), support for it in current software products is at a relatively early stage, as most products that support

QTI are compliant with previous version (1.2) and many of the tools are still under development.

4.2. General standards and specifications

There are several regulations that refer to format of e-learning content. It is worth noting that these are neither complementary nor mutually exclusive.

4.2.1. SCORM

Sharable Content Object Reference Model (SCORM) [4] was created in 2000, as a result of cooperation between government, academia, and industry and is maintained by Advanced Distributed Learning (ADL), which is a part of U.S. Ministry of Defense. It consolidates other regulations, such as IMS specifications [3], AICC recommendations [5], and IEEE 1484 standards (LOM, ECMAScript API, Data model) [6], into unified reference model that enables interoperability, accessibility, and reusability of Web-based learning content. Among other things, it covers packaging of e-learning content (Content Aggregation Model) and the way data are exchanged at runtime between the content and software using it (Run-Time Environment). Similarly to QTI, it does not imply content rendering (unless the content itself contains such information), as it only defines rules on e-learning content storing and exchanging. ADL certified 178 products to be compliant to SCORM 1.2 or SCORM 2004 [7]. This growing number and agreement between ADL, IMS and IEEE [8], further establish SCORM as the industry standard.

4.3. Commercial VLE software

This class includes VLE software that is not freely available, i.e. that one (an institution or an individual) must pay for in order to use it.

4.3.1. Blackboard and WebCT products

Blackboard Inc. is a company that offers several commercial software products. Besides VLE software, the company offers tools that manage transactions related to functioning of an educational institution. After acquisition of the rival WebCT company, Blackboard Inc. became the leading provider of e-learning solutions in the USA, in terms of total users. Having in mind that the given software's e-learning features do not differ substantially from the ones found in other similar software [9], Blackboard software stands out mostly with its transactional capabilities.

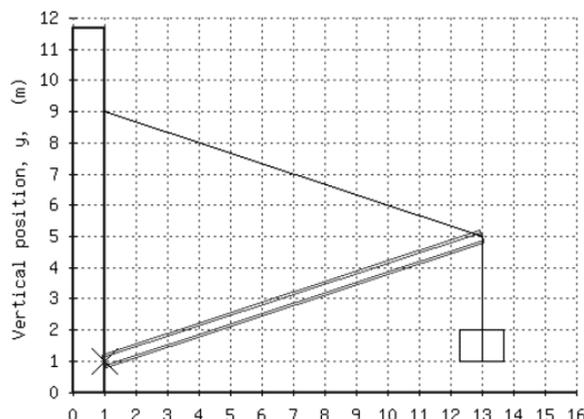
4.4. Free VLE software

Global trend of more and more really usable free and/or open source software impacts this kind of software like all other types. That is why majority of examples presented are in this section. It is worth noting that, besides the examples given here, there are several other solutions in this class, each having respectable list of implemented features and growing user community.

4.4.1. Moodle

Designed by Martin Dougiamas, this open source system has been under constant development for several years, with participation from

A crate with a mass of 155.5 kg is suspended from the end of a uniform boom with mass of 89.5 kg. The upper end of the boom is supported by a cable attached to the wall and the lower end by a pivot (marked X) on the same wall. Calculate the tension in the cable.



A crate with a mass of 177.5 kg is suspended from the end of a uniform boom with mass of 88.5 kg. The upper end of the boom is supported by a cable attached to the wall and the lower end by a pivot (marked X) on the same wall. Calculate the tension in the cable.

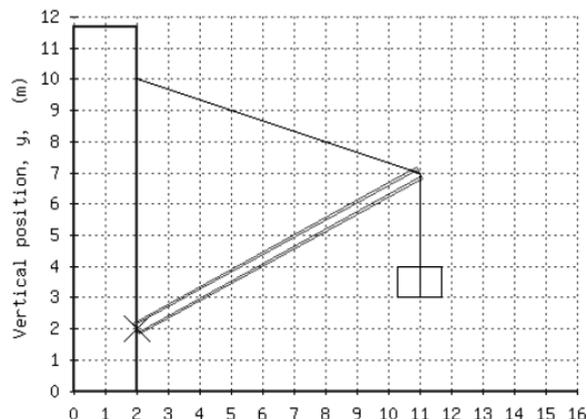


Figure 1. Two versions of the same question. The values of parameters differ in each version of a same question, as they are randomly selected from predefined ranges. Question text and associated image are consistent with these parameters in each version. System uses predefined formulas to automatically generate the result based on the concept that a student must comprehend in order to answer correctly and, optionally, results based on frequent concept misunderstandings. It compares student's response with these result(s) and provides appropriate feedback. This approach "encourages collaboration between students on a conceptual level, but prevents blind copying of answers" [13].

numerous community members and supported by Moodle.com company. Original idea beyond this system is constructivism – approach to learning where knowledge is actively constructed by the learner, not passively received from the environment [10]. The name of the system stands for Modular Object-Oriented Dynamic Learning Environment, which reflects its key attributes – modularity and ease of extension and/or customization. There are more than 100 modules which add some functionality to the system, and the usage statistics [11] are impressive: approximately 50,000 downloads during one month and nearly 20,000 known installations with total of nearly 8 million users and nearly 6 million quiz questions. Information on best practices can be found in [12] and at the Moodle.org community discussion groups. Ease of use and constant improvements make this system the first choice for most educators.

4.4.2. LON-CAPA

Nearly two decade long development of this system is lead by LON-CAPA Academic Consortium, headed by Michigan State University and University of Illinois at Urbana-Champaign. LON-CAPA stands for the LearningOnline Network with Computer-Assisted Personalized Approach, which denotes software's key features: distributed content repository and personalized questions (Figure 1). Moreover, system is utilizing "one source-multiple target" concept, rendering the questions appropriately for the given output format (desktop computer, PDA, bubble-sheet, etc.). Parts of this system are in use for more than a decade, and the distributed repository holds almost 250,000 resources, 80,000 out of which are questions (as of Fall 2006). LON-CAPA related work has produced a long list of publications [14].

Due to its personalized approach, it is best suited for sciences such as physics and mathematics, where concept comprehension is more important than plain memorization of facts.

4.4.3. ILIAS

The name stands for Integrated Learning, Information and Co-Operative Work System, which is being developed since year 1997 by University of Cologne. ILIAS supports a lot of different authentication methods and course design based on learning objectives. Data in different formats (L^AT_EX, OpenOffice.org, etc.) can be imported into the learning content database. ILIAS is also known as the first free SCORM conformant software [15]. Also interesting are extended forms of most frequent question types; for instance, short answer questions may have associated a set of terms that helps automate scoring – student's answer is compared against this set. This system is interesting to users having lots of existing content in popular formats; also, institutions having user authentication system for other purpose can use it to authenticate ILIAS users.

4.5. Commercial CAA Tools

The history of specialized assessment software is long in terms of computer science, and there are many examples. Here are presented only two representative products.

4.5.1. Questionmark Perception

Questionmark is a privately owned company from Norwalk, CT, USA, and have started developing CAA software in year 1988. The company offers a set of test assembly, administration, and reporting tools. Perception Authoring Manager, specialized in question creation and test assembly with support for most frequent question types, has feature-rich question editor and an option to

Question type	Brief description	1	2	3	4	5	6	7	8
multiple choice single response	question text (stem) and several answers (alternatives), one of which is correct, one can be selected	ü	ü	ü	ü	ü	ü	ü	ü
multiple choice multiple response	question text (stem) and several answers (alternatives), zero or more are correct, more than one can be selected	ü	ü	ü	ü	ü	ü	ü	ü
term comparison short answer	student provides a short text answer, and that text is compared against one (or more) possible correct answers	ü	ü	ü	ü	ü	ü	ü	ü
regular expression short answer	student provides a short text answer, correct answer is formed using regular expressions	ü	ü	ü	ü	ü	ü	ü	ü
cloze (fill in the blanks)	question text contains empty text fields for the user to fill in (fields are most often treated as short-answer questions)	ü	ü	ü	ü	ü	ü	ü	ü
cloze (general)	contains several questions of various formats embedded within the text of one composite question	ü	ü	ü	ü	ü	ü	ü	ü
matching items	student is presented with two lists of terms, and should match terms correspondent in a given context	ü	ü	ü	ü	ü	ü	ü	ü
numerical	short-answer question that expects a numerical value instead of a word, correct answer is defined by range	ü	ü	ü	ü	ü	ü	ü	ü
calculated (simple)	similar to numerical type, values in question text and correct answer are parameters that vary in each instance	ü	ü	ü	ü	ü	ü	ü	ü
calculated (complex)	extension of simple calculated type, values in question text and associated pictures consistently vary in each instance	ü	ü	ü	ü	ü	ü	ü	ü

Table 1. Overview of most often supported question types with their brief description. Numbers in the heading row pertain to: (1) Blackboard (all products), (2) Web-CT (all products), (3) Moodle 1.6.3, (4) LON-CAPA 2.1, (5) ILIAS 3.7.0, (6) Perception 4, (7) FastTEST Professional 2, and (8) Hot Potatoes 6.

import questions in IMS QTI format. Along with other tools, such as connectors to SCORM conformant, Blackboard or WebCT systems, Perception represents a rounded set of tools, serving more than 15,000 users who author content and millions of users who use it in some way. The company is providing a list of white papers [16] that contain information useful both for introductory and in-depth reading on assessment. Perception is set of feature rich software tools, useful to all users who want to create and/or modify SCORM, QTI and MS Office content.

4.5.2. ASC products

Assessment Systems Corporation (ASC) was founded in 1979 by David Weiss, psychometrics professor at the Minnesota State University; company's products of most relevance to this paper are the historical DOS software MicroCAT, and its successors, FastTEST Professional and FastTEST. The first two support both computerized adaptive testing (CAT) and paper-and-pencil tests, while FastTEST is focused only at the latter. CAT is the kind of testing where the sequence of questions adapts to the examinee's previous responses. Along with other features, MicroCAT has support for MCATL (Minnesota Computerized Adaptive Testing Language), an authoring language specifically designed for specifying adaptive tests. MicroCAT was the first software to support CAT. ASC provides software and books that cover many areas of computerized testing [17]. FastTEST Professional offers different ways and precise control of test administration, as well as support for most frequent question types.

4.6. Free CAA Tools

Although they are not yet mature to be compared with commercial systems, some free assessment software tools are worth discussing, even at this stage of their development.

4.6.1. Hot Potatoes

Since 1997, development team affiliated with the University of Victoria Humanities Computing and Media Centre is working on this set of tools, use of which is free for most non-commercial purposes [18]. All the tools produce interactive educational exercises viewable in most browsers. The package contains several separate tools for different types of learning material (specific purpose enclosed in parenthesis): JCloze (fill in the gap exercises), JCross (crosswords), JMix (jumbled-sentence exercises), JMatch (matching and ordering exercises), and JQuiz (quizzes made of some of most frequent question types). All of the exercises use JavaScript and HTML for their functionality, and the content can be made to comply with the Dublin Core metadata standard [19]. Having in mind supported types of exercises, these tools are most suitable for language learning and purposes where assessments should preferably be in some presentational format attractive for younger examinees.

5. FURTHER READINGS

The above list is only a top of an iceberg, when compared to constantly increasing list of relevant products. For more information on terms described in this paper, interested reader is first directed to their respective sites. Regarding standards and specifications, further readings include already mentioned AICC recommendations and

white papers [5] and IEEE 1484 standards [6]. This list is rounded with efforts conducted by CETIS [20] and JISC [21]. Detailed description of most contemporary VLE software, along with option for direct comparison of several systems' features at a time is available at [9].

6. DISCUSSION OF APPROACHES

Although it should not be the sole criterion when choosing a system for question authoring, it is good to know what question types are available in which software. Table 1 provides very coarse overview, as only the most frequent question types are listed. Most of the tools have some kind of expansion possibilities, so that new or existing question types can be easily added. Moodle and LON-CAPA are good example: both of them utilize Java molecular editor, a graphical editor where a student can construct molecular structure [22]. Once completely drawn, this structure is converted to a string, which is further processed as a response to short answer question.

Regarding support for test administration, most common features supported in CAA and VLE software are: defining the time during which a test will be available, randomization of questions within a test and answer options within a multiple choice question, feedback after a question is answered and/or a test is completed, etc.

Most of the described software products support Internet/Web based or computer/desktop based testing. There is also an approach that is somewhere in between this two concepts – use specialized desktop software to create questions and tests, while using Internet only to deliver the content to the public. One example of this approach is already mentioned *test* software [1]. An interesting feature of this software is automatic test assembly, based on predefined question selection criteria, such as question difficulty, number of questions in given category, etc.

7. CONCLUSIONS

Table 1 indicates that VLE software has become comparable with specialized CAA tools, in terms of diversity of supported question types. Communities gathered around open source solutions are constantly contributing to adding new features to these solutions, which further leads to the fact that most features in free VLE software are at least comparable to ones of commercial VLE software.

Free VLE software will probably substantially increase its part in the e-learning market share, as all the crucial components are beginning to fit into place: quantity of available content is increasing due to standardization, features are improving day by day, and there are a lot of small privately owned companies that support developers and offer user support and/or training for free VLE software.

Diverse question types supported in current software contribute to better assessment of stu-

dent's knowledge. However, these types are only a part of a much larger list given in [23], which opens horizons for future development. Having in mind easy extensibility of the most of relevant software, it is only a matter of time until all existing and, perhaps, some new question types will be implemented and made available, thus boosting the shift already present in education: from traditional testing towards Internet based testing.

ACKNOWLEDGMENT

The authors would like to thank Professor Veljko Milutinovic for his continuous support and encouragement to publish the results of their work.

REFERENCES

Unless otherwise noted, URLs given in this section are last retrieved on December 15, 2006.

- [1] Protic, J., Bojic, D., and Tartalja I., „*test*: Tools for evaluation of students' tests - a development experience“, *31st ASEE/IEEE Frontiers in Education Conference*, Reno, Nevada, USA, 2001, pp. F3A.6-F3A.12.
- [2] IMS Global Learning Consortium, "IMS Question & Test Interoperability Specification," <http://www.imsglobal.org/question/>, 29.9.2006.
- [3] IMS Specifications, Specifications, <http://www.imsglobal.org/specifications.html>, 12.12.2006.
- [4] Advanced Distributed Learning, SCORM 2004 3rd Edition, <http://www.adlnet.gov/scorm/index.cfm>, 12.12.2006.
- [5] Aviation Industry CBT Committee, Index of Available AICC Publications, <http://www.aicc.org/pages/down-docs-index.htm>.
- [6] IEEE Learning Technology Standards Committee, <http://ltsc.ieee.org/>.
- [7] Advanced Distributed Learning, SCORM Certified Products, <http://www.adlnet.gov/scorm/certified/index.cfm>, 5.12.2006.
- [8] Advanced Distributed Learning, "ADL Reaches Agreement with IMS and IEEE on Use of Specification/Standards Work in SCORM® 2004 3rd Edition," <http://www.adlnet.gov/news/articles/384.cfm>, 16.11.2006.
- [9] EduTools Course Management System Comparisons, <http://www.edutools.info/static.jsp?pj=4>.
- [10] Dougiamas, M., "A Journey into Constructivism," <http://dougiamas.com/writing/constructivism.html>, November 1998.
- [11] Dougiamas, M., et al., "Moodle statistics," <http://moodle.org/stats/>.
- [12] Cole, J., "Using Moodle," *O'Reilly*, Cambridge, MA, USA, July 2005.
- [13] LON-CAPA Academic Consortium, "Examples of Individualizing Problems," <http://www.lon-capa.org/problems.html>.
- [14] LON-CAPA Academic Consortium, "Publications connected to the LON-CAPA project," <http://www.lon-capa.org/publications.html>.
- [15] Kunkel M., et al., ILIAS History, <http://www.ilias.de/ios/info-e.html>, 6.9.2006.
- [16] Questionmark, White Papers, <http://www.questionmark.com/us/whitepapers/index.htm>.
- [17] Assessment Systems Corporation, <http://www.assess.com/>.
- [18] Hot Potatoes, <http://hotpot.uvic.ca/>.
- [19] Dublin Core Metadata Initiative, <http://www.dublincore.org/>.
- [20] Centre for Educational Technology & Interoperability Standards (CETIS), <http://www.cetis.ac.uk/>.
- [21] Joint Information Systems Committee (JISC), <http://www.jisc.ac.uk/>.
- [22] Ertl, P., "Java molecular editor," <http://www.molinspiration.com/jme/doc/index.html>, October 2004.
- [23] IMS Global Learning Consortium, IMS QTI Implementation Guide, http://www.imsglobal.org/question/qtiv2p1pd2/imsqti_implv2p1pd2.html, 14.10.2005.