Engineering an interoperable adaptive hypermedia testing tool supporting user adaptable strategies

Fotis Lazarinis, Simeon Retalis
University of Piraeus
Department of Technology Education and Digital Systems
80 Karaoli & Dimitriou
185 34 Piraeus, Greece
lazarinf@unipi.gr, retal@unipi.gr

Abstract. In this paper we present a web based tool for authoring adaptive assessments based on IMS QTI, IMS LIP and IEEE PAPI learning standards. The tool allows authors of assessments to customize the applied adaptation strategies to meet their own teaching experiences and goals. During the assessment procedure the system accumulates knowledge about learners and adapts the exposure of items based on this knowledge and on the adaptable rule model. Design decisions and the employed user, domain and rule models are analyzed and discussed. Results from a usability evaluation are discussed at the end of the paper.

1 Introduction

Adaptive Educational Hypermedia Systems provide a learning environment adapted to the individual characteristics and goals of the learners [1]. Adaptive Educational Hypermedia Systems can be used both for assessment and self-assessment providing a tailored testing environment which results to a shorter and less demanding assessment.

Most of the current adaptive testing systems rely on the Computerized Adaptive Testing technique [2] and the Item Response Theory [3] and depending on the learner's performance they pose easier or more complex questions. This unique adaptation criterion limits the capabilities offered to educators for identifying the learning difficulties and misconceptions of their students. The majority of these systems represent their data in a custom way, i.e. in a custom made database or XML structure, preventing knowledge sharing between applications [4, 5].

CosyQTI has been designed taking into account two very important factors. Firstly, that the data models employed should comply with established learning

standards to support interoperability among e-learning systems. Secondly, adaptive decisions should not be fixed and thus pedagogically limited. Instead, the adaptivity decisions should be left to educators who could then integrate their own teaching experiences and intelligence into the system. This can be achieved through the customization of a broad set of predefined adaptive rules.

The paper is organized as follows. In section 2, a brief introduction to adaptation in assessment systems is provided. Section 3 presents the basic modules of CosyQTI and discusses the factors that influenced their design. Section 4 presents the results of usability experiments and section 5 summarizes and gives directions for further research.

2 Adaptation in Assessment Systems

The most well known adaptation technique is Computer Adaptive Testing (CAT). This involves a computer-administered test in which, the selection/presentation of each question and the decision to end the process are dynamically adapted to the learner's performance in the test [2]. Computer Adaptive Testing is based on Item Response Theory [3]. In the adaptive testing procedure, depending on the learner's answers to questions, easier or questions of increased difficulty are posed. The criterion for selecting questions is to match the question's difficulty level with the learner's estimated knowledge level. It has been shown that these questions are more "informative" in terms of conclusions that can be drawn on the learner's knowledge. The goal is to accurately estimate the learner's knowledge. Several approaches exploit the idea of adaptive testing. Huang in [6] describes the CBAT-2 adaptive testing algorithm that generates content-balanced questions and SIETTE [7] is an adaptive assessment authoring tool that employs this technique.

Another adaptation technique, which is however mainly used in computer-assisted surveys, is Adaptive Questions, as defined by Pitkow and Recker in [8]. This method causes the generation of a dynamic sequence of questions depending on learner's responses. In this case, questions are divided in categories based on whether they do not cause any adaptation at all, trigger one follow-up question, or trigger more than one follow-up question. It has been shown that web-based adaptive questionnaires can reduce the number and complexity of questions presented to users. Adaptive questionnaires have been used to assess web users' attitudes in CATES [9].

3 Overview of CosyQTI

CosyQTI is a web based tool allowing authors to create both adaptive and non-adaptive web based assessments. The component based architecture of the system, depicted in figure 1 using UML notation, consists of a learner model, a domain model and a rule model. The item selection engine is implicitly created based on these models. The learner model contains information such as the goals, preferences, qualifications, knowledge estimations and usage data of each learner. The domain

model follows the IEEE/ACM Computing Curricula structure [10] and additional domains can be easily integrated. The rule model is user adaptable meaning that educators adapt a set of predefined rules to meet their own pedagogical goals and beliefs.
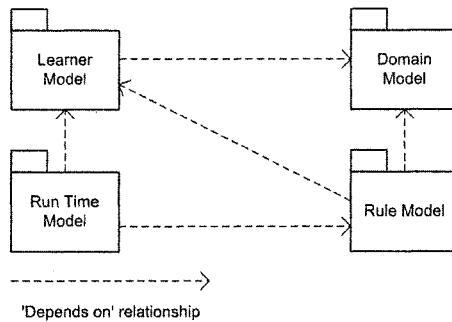

Fig. 1. High level component architecture of CosyQTI

3.1 Assessment authoring

After authentication educators select the domain of the assessment and create a new assessment or edit an existing one through a homogeneous interface (see figure 2). Educators can create or re-use items (questions) of various types and group them to sections. Each section is associated with a concept which in turn is associated with a domain.
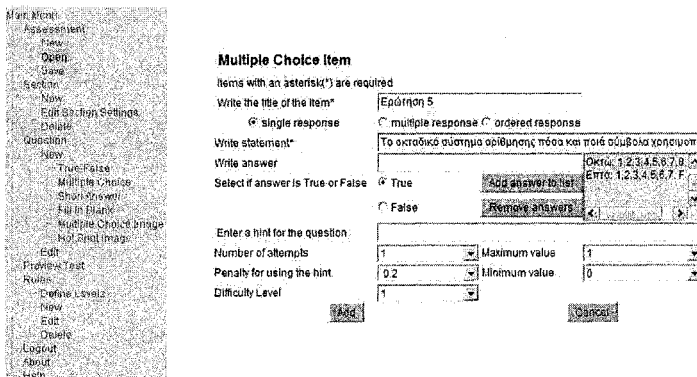

Fig. 2. Creation of multiple choice questions

For each assessment item the educator may alter the default values of the additional data which are difficulty level, hints, number of attempts, penalty for using the hint, and minimum and maximum score (see figure 2). The different types

of items supported are:
(i)True/False
(ii) Multiple choice (single, multiple or ordered response)
(iii) Fill-in-the-blanks
(iv) Multiple Image choice
(v) Image hot spot.

Interoperability is a significant design requirement in CosyQTI. Therefore the created assessments conform to IMS QTI (Question and Test Interoperability) XML standard [11], so that they can easily be exported and used by other IMS compliant applications. QTI structures material into assessments, sections, and items. An item is the formal name for a question and assessment is the terminology used for a test.

The structure of an assessment resembles a two dimensional array. Each row is a set of questions which are grouped in a section and associated with a concept of the domain. The number of items is diversified between sections and they are sorted based on their difficulty level. Testing paths differ between learners as in real world classes. Experienced teachers pose questions of increased complexity to students with a high knowledge level or an augmented number of simpler questions to students with lower performance. In both cases the ultimate goal of educators is to encourage all the students to actively participate in the learning procedure. Similarly in CosyQTI the testing path of a new learner may initiate in question 1 of section 1, while in other learners the assessment procedure may start in question 1 of section 2 or in a question of increased difficulty level of section 1.

3.2 Learner model

The adaptivity of the system relies primarily on the learner model which contains the goals, preferences, knowledge and usage data of each learner. This information is crucial as the adaptive rules are based on these data as we will see in the next sections. The information in the user profile has to be such that the system can better adapt to the user's individual needs [12]. To support the adaptive approach taken in CosyQTI we identified five categories of data in the learner profile:
(i) Demographic information: data that remains unchanged, such as age, gender, etc.
(ii) Learner goals: which are related to the long and short term learning goals (e.g. "to pass test X with a high score").
(iii) Learner preferences: preferences related to various aspects of the learning environment and procedure (e.g. the mode of delivery, accessibility requirements, or assessment).
(iv) Learner knowledge: records the knowledge level and the weaknesses and strengths related to the concepts to be learnt and the formal qualifications of a learner.
(v) Usage data: information like which pages were viewed, in what order, for how long, etc.

For satisfying the interoperability design principle we structured the data using learning standards. More specifically, the attributes that compose a learner's profile

have resulted from a selection and combination of elements from IMS LIP [13] and IEEE PAPI [14] standards. This combination of elements serves our key objective for interoperability without compromising the attributes and services required [15].

3.3 Domain model

Domain model is a set of associated concepts which form hierarchies of concepts. Concepts of the domain model may be atomic or composite containing other atomic or composite concepts. The IEEE/ACM vocabulary [10] is utilized for testing purposes in the current implementation of CosyQTI. However CosyQTI is an open domain web testing tool, meaning that a mechanism has been developed which allows automatic integration of domains following the IEEE/ACM vocabulary structure. Thus educators of various disciplines are able to utilize the system.

The domain model contains additionally a series of learning objectives such as 'learn concept X', 'describe the common characteristics of concept X', etc. Learning objectives are high level abstract learning goals which are associated with concepts at run time. Educators define learning objectives for each section or item of an assessment and the system automatically determines, based on the learner's performance, whether these learning goals are satisfied or not. The domain model is overlaid [16] in the learner model based on the concepts and learning objectives of an assessment.

3.4 Rule model

Adaptation decisions are set by the educators during the authoring phase. An educator is able to adapt the assessment to the requirements of an individual or to the aims of a class by adapting a set of *IF* <condition> *THEN* <action> rules. During the creation stage trigger points are set and actions are specified based on the learning preferences and pedagogical aims of the teacher (see figure 3). Rules are applicable either to new users or to existing users for whom the system, based on their previous interactions, has a complete learner profile. Adaptation rules may be combined with the aid of Boolean operators.

Events relate to the previous knowledge on the assessed subject, formal qualifications, previously estimated knowledge level or performance during the execution of an assessment. Conditions are set by the educators and are thresholds which should a learner pass an action is realised. Actions may be section advancement, knowledge level increase or decrease, assessment ending, ascent or descent of the difficulty level. For example an educator may impose section advancement and knowledge level increase if the score of a learner during the first five questions of a section is greater or equal to 75%.

3.5 Run time model

When learners log in to the system they are identified and described by their personal

model. The user model has to form an accurate image of the learner so it needs to be updated frequently to include any changes. The learner's knowledge is updated in our system, based on the section topics a learner is assessed on. The algorithm used to update the user's knowledge is implicitly formed based on the author's rules. Learners get immediate feedback and help on their choices.

System keeps the learner's usage data which includes historical information about the learner's performance in the test as well. It is very useful to keep track of the learner's actions and the sequence of sections and items visited. This information is required in cases where an author wants to apply advanced rules such as "If the user performed very well in section A but performed poorly in section B, then present a new section with intermediate difficulty level". This type of rule condition is currently under development. Also usage data can be used for opening the learner model during or at the end of an assessment [17].

Fig. 3. Rule adaptation options

5   Evaluation

CosyQTI is a web based adaptive assessment tool which can be used for self-assessment as well. Its main goal is to support learning by adapting the testing procedure to the performance, goals and preferences of learners and to the teaching experiences and intelligence of educators. Thorough evaluation of such a system is a complex and time consuming procedure involving compound criteria, several users and different approaches.

The initial evaluation performed regarded the system's added value offered to educators. The system's authoring capabilities and features were explained and

exhibited to five experienced IT instructors. Then we asked them to create a 4 section assessment consisting of 10 questions each. We asked the participants to supply their own queries and should they judge it appropriate and necessary to adjust some of the available adaptive rules.

During the test authoring phase we recorded their adaptation decisions. As seen in table 1 most of the educators defined section advancement if a learner achieves a specific score in the first 5 or 7 questions. However two of them created more specialised compound rules applied either at the start or at the end of each section. These rules depend on the estimated knowledge level of learners. If it is high then force the system to increase the knowledge level and thus the difficulty level of the subsequent questions.

Table 1. Adaptive rules created from the educators

Educator	Trigger point	Condition	Action
1	Question 5 of each section	Score > 75%	Move on to the next section
2	Question 7 of each section	Score >= 90%	Move on to the next section
3	1. Test initiation	1. Degree in computer science or certificate in IT	1. Pose questions of increased difficulty level
	2. Question 5 of each section	2. Score >= 100%	2. Move on to the next section
4	End of first section	Score >= 80%	Increase knowledge level and difficulty level of posed questions
5	Start and end of each section	High or low knowledge level in the domain	Increase/decrease difficulty level of posed questions

This short evaluation revealed that such a system is quite useful because it grants educators with freedom to apply their own intelligence and teaching philosophy. However enrichment of the rule model is necessary and possibly a mechanism for defining new rules instead of simply adjusting the predefined ones is necessary as well. In any case more exhaustive evaluation experiments are needed to realize the full potential of the system and to recognize design weaknesses.

5   Discussion and Future Work

This paper presents the basic components and design decisions of a web based adaptive assessment authoring tool. Data models conform to international standards making the system interoperable with other e-learning tools. Educators have the

ability to tailor multiple criteria to their goals and to the individual needs of their learners. The authoring environment makes the addition and modification of learning content straightforward and allows the educators to efficiently customize the execution of the assessment and assign various learning objectives to items and sections. Learners get immediate feedback on each question and an analytical explanation at the end of the assessment.

The usability experiment revealed that the capabilities and the non restrictive nature of the system are very useful to educators. Different educators applied different pedagogical approaches to test the knowledge of their students on the same domain. The evaluation produced new research directions with respect to the adaptation model. More adaptive rules and more adaptability freedom needed to simulate real teacher/student classes.

References

1. Brusilovsky, P: Adaptive hypermedia. User Modeling and User-Adapted Interaction 11 (2001) 87-110.
2. van der Linden, W. J., Glas, C. A. W.: Computerized Adaptive Testing: Theory and Practice. Netherlands: Kluwer Academic Publishers (2000)
3. Hambleton, R. K., Swamination, H., Rogers, H. J.: Fundamentals of Item Response Theory. Newbury Park: Sage Publications (1991)
4. De Bra, P., Aroyo, L., Chepegin, V.: The Next Big Thing: Adaptive Web-Based Systems. Journal of Digital Information, 5(1) (2004) http://jodi.tamu.edu
5. Retalis, S.: Usable and Interoperable E-Learning Resources Repositories. In: Sanjaya Mishra & Ramesh C. Sharma (eds): Interactive Multimedia in Education and Training, Idea Group Inc. 249-269 (2004)
6. Huang, S.: A Content-Balanced Adaptive Testing Algorithm for Computer-Based Training Systems. In: Frasson, C., Gauthier, G., Lesgold, A. (eds.): Intelligent Tutoring Systems, Third International Conference, Springer (1996) 306-314
7. Guzmán, E., Conejo, R., García-Hervás, E.: An Authoring Environment for Adaptive Testing Educational Technology & Society 8 (3) (2005) 66-76
8. Pitkow, J., Recker, M.: Using the Web as a Survey Tool: Results from the Second WWW User Survey. Computer Networks ISDN Systems 27(6) (1995) 809-822
9. Chou, C.: Constructing a Computer-Assisted Testing and Evaluation System on the World Wide Web – the CATES experience. IEEE Transactions on Education 43(3) (2000) 266-272
10. IEEE/ACM Computing Curricula 2001 – CS Body of Knowledge, http://www.computer.org/education/cc2001/final/appa.htm
11. IMS Question and Test Interoperability Specification, http://www.imsglobal.org/question/
12. Kobsa, A.: Generic User Modeling Systems. User Modeling and User-Adapted Interaction 11 (2001) 49–63
13. IMS Learner Information Package Specification, http://www.imsproject.org/profiles/
14. IEEE PAPI Learner http://edutool.com/papi/
15. Dolog, P., Gavriloaie, R., Nejdl, W., Brase, J.: Integrating Adaptive Hypermedia Techniques and Open RDF-based Environments. In: Proc. of 12th International World Wide Web Conference, Budapest, Hungary (2003)
16. De Bra, P., Aroyo, L., Cristea, A.: Adaptive Web-Based Educational Hypermedia. Web Dynamics, (2004) 387-410
17. Bull, S., Pain, H.: Did I say what I think I said, and do you agree with me? Inspecting and Questioning the Student Model. In: J. Greer (ed): Proceedings of World Conference on Artificial Intelligence in Education, AACE: Charlottesville, VA (1995) 501-508