Analyze Me: Open Learner Model in an Adaptive Web Testing System

Fotis Lazarinis, Symeon 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. This paper presents the open learner modelling capabilities offered by an adaptive web testing tool. Adaptation decisions are based on user customizable compound rules requiring the identification of several learner characteristics which are then displayed graphically and textually to learners and educators. The open learner modelling capabilities include anonymous access to models of peers and restricted named access to friend peers. Opening the learner profiles supports learner and educator involvement in decision making so as to create a more credible learner model. The educational advantages of the open learner modelling capabilities are discussed on the basis of an evaluative study.

Keywords. Adaptive hypermedia testing system, user modelling, interoperability, learning standards

INTRODUCTION

Assessment and self-assessment are inseparable parts of the instructional process. Tests should be utilized in such a way that the learning process is enhanced, aiding both learners and educators to obtain a deeper understanding of their learning strengths and difficulties (Jensen & Feuerstein, 1987; Martinez & Lipon, 1989). Adaptive Hypermedia Testing systems provide a personalised assessment environment of a learner's knowledge and can be used either in conjunction with a complete adaptive hypermedia learning environment or as a stand-alone application for assessment or self-assessment, adding promising new dimensions in learning (Brusilovsky et al., 2004). Most of the existing adaptive testing tools are based on the computerized adaptive testing technique (van der Linden & Glas, 2000) and Item Response Theory (Hambleton et al., 1991) and are used mainly as limited skill meters basically presenting the learner's overall score on a subject and a pass/fail indication. A few adaptive web or stand-alone testing systems based on these techniques have been implemented (Collins et al., 1996; Romero et al., 2004; Conejo et al., 2004) or integrated into adaptive hypermedia educational systems (Arroyo et al., 2001; Guzmán & Conejo, 2002) offering certain advantages for learning.

Another adaptation technique, which is however mainly used in computer-assisted surveys, is adaptive questions, as defined by Pitkow and Recker (1995). This method causes the generation of a dynamic sequence of questions depending on the learner's responses. All these tools base their adaptation strategy primarily on the learner's performance, posing easier or more complex questions depending on the learner's answers. From a pedagogical point of view these approaches are limited as they do not allow educators to adequately diagnose all the possible learning weaknesses and misconceptions of their learners. More specifically, questions are not linked to particular misconceptions or bugs nor are they linked to specific concepts and therefore educators cannot understand the difficulties of their learners in the case of erroneous answers.
A possible solution so as to overcome these limitations is to create an open adaptive hypermedia testing system where adaptation decisions and learning objectives are adapted and set by educators and learners based on their personal needs and goals. This multicriteria approach can benefit the learning process as it grants educators the freedom to apply their own teaching intelligence and philosophy. Learners get different questions based on their performance and most importantly on their special abilities and weaknesses. Tailoring several criteria to achieve specific learning objectives is incontestably a complex procedure where different decisions are taken at various stages during assessment. These decisions imply identification of certain attributes, abilities and difficulties of learners which could be utilized to enhance the learning process.

A research trend in learning environments is open learner modelling. Opening the model to the learner is expected to yield pedagogic gains in providing the means for reflective learning (Kay, 1997; Bull & Nghiem, 2002). Learners will be able to analyze their learning process through the system's beliefs about them and, possibly, to argue about these beliefs. It has also been claimed that assessment might be facilitated by providing learners access to their learner model (Mitrovic & Martin, 2002). In an adaptive assessment system, especially in our tool which employs several user adapted rules, this is even more important as decisions on the adaptation could be taken with the aid of learners and educators after the externalization of the system's viewpoints and conclusions. Possible system misjudgements could be amended and potential inaccuracies in the design of a test could be revealed, leading to the alleviation of the adaptation strategy.

Our tool allows educators to create both formative and summative assessments (Angelo & Cross, 1993; Black & Wiliam, 1998) based on their beliefs. Formative assessment is often performed at the beginning or during a program, thus providing the opportunity for immediate evidence for student learning in a particular course or at a particular point in a program. Formative assessments could benefit from a personalized system with open learning modelling techniques, as the detailed descriptions of the assessment procedure and the inferences made could aid learners to view their understanding and misconceptions on certain concepts and eventually improve their learning. Summative assessment is comprehensive in nature, provides accountability and is useful for checking the level of learning at the end of the program. In our system, summative assessments could be faster and refinements made according to the aims of educators and the learners' portfolio.

This paper first overviews the modules and services provided by our adaptive web testing system, called CosyQTI. Next, the open learner modelling techniques are presented and analyzed. The potential educational advantages of the methods employed will be discussed on the basis of an evaluative study. Conclusions and future research directions are discussed at the end of the paper.

**DESCRIPTION OF CosyQTI**

CosyQTI is a web based adaptive hypermedia testing system where adaptation is based on prior accumulated knowledge about learners and on their learning progress. The component based architecture of the system consists of a learner model, a domain model and an adaptation model (see Figure 1).

These models are structured and stored using learning technology standards and they are editable through the system or they can be imported from other educational tools. The learner model contains demographics, goals, preferences, knowledge and usage data about a learner. The adaptation procedure relies on these characteristics (Brusilovsky, 2001). The attributes that structure the learner
model have resulted from a selection and combination of XML elements from IEEE PAPI (2002) and IMS LIP (2005) standards. This combination of elements serves our key objective for interoperability without compromising the attributes and services required (Dolog et al., 2003). The learner model is initialized either by the learner or the educator or it can be imported from a Learning Management System with which the learner has previously interacted.

![High level component architecture of CosyQTI](image)

**Authoring environment**

Editing the assessment and customization of the adaptation rules are core operations of CosyQTI offered through a homogenous interface (see Figures 2 and 4). Educators can create or re-use items (questions) of various types and group them into sections. The different types of items supported at the moment are (i) True/False, (ii) Multiple/single choice, (iii) Ordering, (iv) Fill in the blanks, (v) Multiple Image choice, (vi) Image hot spot. Items have a set of additional attributes which are difficulty level, hints, number of attempts, penalty for using the hint, and minimum and maximum score and the educator is able to alter the default values. Each section is associated with a concept which in turn is associated with a domain. Educators are free to associate the same concept with many sections and possibly structure successive sections with questions of increased difficulty weights.

CosyQTI is a domain independent web based adaptive hypermedia testing system, meaning that a mechanism has been developed which allows automatic integration of additional domains following the IEEE/ACM vocabulary structure (2001). Additionally, the domain model contains a series of learning objectives such as 'understand 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 and criteria for satisfying them 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 (De Bra et al., 2004) in the learner model based on the concepts and the learning objectives of an assessment.
Formally an assessment corresponding to a domain is denoted as $A_D$ and is the union of multiple sections $S$ each corresponding to a concept $C$ of the domain $D$ and a set of rules $R$ (see Figure 3a). Each section is an association of questions $Q$ of varying difficulty weights $w$ which have a set of attributes $a$ and a type $t$ (see Figure 3b).

\[(a) \ A_D = \bigcup \{ S(C_D), R \} \]
\[(b) \ S(C_D) = \bigcup \{ Q(t, a, w) \} \]

Fig.3. Semantics of an adaptive assessment.

The assessment data are structured in IMS QTI XML (2005), so that they can be exported and used by other IMS compliant applications. The educator is able to adapt the assessment to the requirements of an individual or the aims of a class by adapting a set of event-condition-action rules (see Figure 4). During the authoring phase trigger points can be set and actions can be specified based on the aims of the teacher. Thus educators are given the opportunity to apply their own experiences to the benefit of the learning process and of their learners. Compound adaptation decisions are also possible with the aid of Boolean operators.
Run time environment

During test execution the usage data collected updates the overlay learner model (De Bra et al., 2004). The update mechanism is based on the rule model. The knowledge level is constantly computed and updated and the subsequent sequence of questions is determined based on their difficulty weight and on the learner's performance. Scoring is normalized with respect to the expected maximum score of the questions posed. Assessment of the items contributes to the estimation of the knowledge level of a learner on a specific concept. These estimations propagate up to the domain level and contribute proportionally to the domain's knowledge level estimation.

\[ M_S = w_1 * M_{Q1} + w_2 * M_{Q2} + \ldots + w_n * M_{Qn} \]

\( n \): number of section questions
\( w_i \): difficulty weight of each question
\( Q_i \): mark in each question.

Fig.5. Parameterized computation of a learner's mark in a section.
The total mark $M_S$ in each section is computed with the aid of a parameterized linear equation (see Figure 5). As mentioned, a section is normalized with respect to the expected maximum score of the questions posed. This normalization is necessary as different learners may be presented with a diverse number of questions and complex weights in the same section. So the final computed mark of every learner in a section is calculated as a percentage in the same scale (e.g. 0%-100%). The final total score of a learner is the average of the individual section scores.

**Visualizing an assessment scenario**

Figure 6 visualizes an excerpt from an actual assessment scenario. Darker gray implies questions with increased difficulty weight. As can be seen, different learning paths are possible, based on the previous knowledge and on the adaptation model.

![Assessment Scenario Diagram](image)

Fig. 6. Visualization of an assessment scenario.

For instance, Rule 1 states that if the learner has some knowledge on the concept then testing should start with question 3. Rule 2 defines the transition to section 2 if some criterion based on the current estimated knowledge level is met. Following the non adaptive testing path would require learners to reply to all the questions 1 to 10 in each section. This simple graphically depicted assessment scenario shows that in CosyQTI non-linear content access based on definite rules is possible. Several adaptive paths can be dynamically formed based on the successful transitions. For example, based on the conditions met a learner may be presented with the questions $Q_1(S_1)$ to $Q_6(S_1)$, then with the questions $Q_4(S_2)$ to $Q_6(S_2)$ and finally with the questions $Q_6(S_3)$ to $Q_{10}(S_3)$. Another learner may also start on question $Q_3(S_1)$ but if rules $R_2$ and $R_3$ are false while $R_4$ is true then the sequence of questions would be $Q_3(S_1)$ to $Q_{10}(S_1)$, $Q_1(S_2)$ to $Q_{10}(S_2)$, $Q_1(S_3)$ to $Q_4(S_3)$ and last $Q_4(S_3)$ to $Q_{10}(S_3)$.

Presently the dynamically formed paths have a linear structure but it would be interesting to explore a graph oriented formation of questions which would possibly allow backtracking in some
cases. Also a module which would visualize the adaptive testing pathways producing graphs similar to the one of Figure 6 could help educators realize their testing strategy and could also be utilized when opening the learner model.

OPEN LEARNER MODELLING IN CosyQTI

Opening the learner model adds pedagogic value to the instructional process and helps both educators and learners in performing their tasks (Kay, 1997; Zapata-Rivera & Greer, 2001; Bull & Nghiem, 2002; Mitrovic & Martin, 2002). Both adult learners and children could benefit from inspecting their models (Bull, 2004; Bull & McKay, 2004). Additionally, Kay (1997) suggests that learners may wish to compare their progress to that of their peers. In a recent survey, undertaken to discover students' wishes concerning the contents, interaction and form of open learner models in intelligent learning environments, it was found that students indeed wish to inspect their model and compare it to the models of peers (Bull, 2004). The instructor may use their students' learner models as a source of information to help them adapt their teaching to the individual learner, or to the group (Zapata-Rivera & Greer, 2001).

CosyQTI is an open adaptive hypermedia testing system supporting different teaching strategies with the aid of multiple parameterized rules. The tests are designed to elicit and represent current knowledge and to compare the findings with the previous knowledge of the learner. Based on the findings of the studies in open learner modelling and on the capabilities of CosyQTI, opening the learner model in such a system is quite helpful as it will show the progress of learning and the learner's deficiencies. It will also provide educators with ample data to understand their learners and to review and possibly redesign their teaching and testing strategy.

Inspecting the learner's and a peer's learner model

Learner model attributes are displayed in alternative modes at various stages of the testing procedure. The achievements and misconceptions of learners are analyzed during the test or at the end of the assessment. Figures 7 and 8 show the results of a simple test based on multiple choice and true/false items only. As seen, textual descriptions and graphical information are available to learners and to educators. Analysis of the learning progress is both per assessment and per section. In this way learners realize explicitly which topics they know and which topics they need to study harder. Conclusions presented to learners are positive so as not to discourage learners from participating in the testing procedure. The information is a mixture of general feedback on performance: the presented data includes the overall performance and a brief textual explanation of the result, the number of questions posed and their average difficulty level; and data inferred from the learner model (the open learner model information) concerning known or problematic topics. The information presented to users results from comparing the previous knowledge level and the current estimated knowledge level as well. Also the formal education and portfolio of learners are taken into account. For example, if a learner achieves a score of 70% in a test then if her/his previous knowledge is high the learner is informed that a greater score would be expected since her/his starting knowledge level was high. We should mention here however, that in short tests or in cases for which the learner models are not complete the presented results tend to match the performance statistics as the system is unable to base the inferences made on the learner's previous knowledge.
Messages shown to users and inferences made, take into consideration the average query difficulty level and the number of questions posed. For example, in Figure 7, a student has successfully passed the test but s/he is advised to study harder for more demanding tests, since the specific assessment was of low difficulty and the achieved score was only 70%. Additionally, learners may see which question type better suits them based on their correct replies and on the number of attempts of each question type. This could be considered as a light attempt to choose the question types that better suit a learner. Of course, sometimes no safe conclusions can be drawn based on these data only and especially when assessments are short and use limited alternative question types.

Since each section is associated with a concept, which belongs to a domain, inferences are made also on a concept basis. Figures 7 and 8 textually and graphically depict the misconceptions and problematic areas. For example, it is easily realizable that the learner has performed better in the "software" related questions than in the "hardware" associated questions. The graphical depiction could be enhanced by using tree like representations of the domain with unknown concepts appearing in a darker colour. This way learners and educators could easily and promptly realize for which concept and in which domain their knowledge is not as solid as it should be.

The final classification of the student's performance (poor, average, good or excellent) and the decision about which topic is considered known relies on the performance of a learner (expected to reflect their current understanding as a snapshot of their knowledge identified through adaptive testing), and on the thresholds set by educators. For instance, an educator may consider that a score of 80% is adequate to classify a learner in the highest rank whereas another may believe that 90% is the appropriate lower limit for the final class. This depends on the goals of the assessment and on the pedagogical beliefs and experiences of educators.

CosyQTI attempts to identify problematic areas and misconceptions. Similar to Bull and McEvoy (2003) who use varying hues of gray to depict misconceptions and problematic topics, dark gray to portray the problematic topics and light gray to represent misconceptions in CosyQTI charts. Problematic topics are considered to be those topics which are related to many wrongly answered questions which have been tried more than once, and possibly after hints were used. Unanswered questions contribute to identifying problematic concepts as well. Discovering misconceptions is a more challenging task though. In some systems a number of already identified misconceptions for a specific domain have been modelled by the system so as to match to the students' replies (Bull & McKay, 2004). However, since our system is domain independent, modelling misconceptions is a quite difficult, if not impossible, task. Presently, only multiple choice questions which were answered wrong and were tried only once without using the available hints indicate possible misconceptions. Although this method may be considered somewhat crude, it is based on the assumption that although the user was confident, the given reply was wrong. Educators are also informed that incorrect options in multiple choice questions will need to be designed carefully to elicit misconceptions. In general, in this first version of the system the preference is to indicate errors as possible misconceptions rather than classifying them simply as problems. This tactic is adopted to direct both educators and learners into examining the given answers more thoroughly. A more sophisticated but more demanding misconception identification strategy would be to ask educators to classify wrong answers as either problems or misconceptions during the design of the assessment. This could be further enhanced by asking educators to provide adequate explanations and correct examples, especially in relation to possible misconceptions.
Fig. 7. Textual presentation of the learner model.
Anonymous access to the model of their peers is given to learners (see Figure 9). Our web testing system collects and aggregates the available data of all the learners who have completed a specific assignment and presents them upon request. The data of the peers are presented aggregated in diagrammatic form. This aggregated form is also available to educators who are able to compare the learning progress of their learners, to spot potential design problems in their adaptive test and to enhance their teaching practices. For instance, if many learners fail to answer correctly many questions in a section then it may be the case that these questions are ambiguous. Presently educators have to study the graphs and spot these problems on their own but we are in the process of developing a module which would automatically detect and present potential testing faults to them. Display of models to peers is a technique applied in previous studies as well (Bull & Nghiem, 2002; Bull & McKay, 2004). However, in our work the "average" learner model is also presented and since the data are in XML format other forms such as the best learner model or excerpts of the learner profiles could be presented on request. In this way educators can easily create customized views of the data in order to fit their needs or to form subgroups with similar learning characteristics (Mühlenbrock et al., 1998).

Another option of CosyQTI is named access to the profile of a learner from specified peers. This protected inspection should be permitted when the learner model is initialized by the learner itself. It is included as an option for contest purposes between friend-learners (i.e., friends who are learners) as a means for stimulating learning. This permissible "learning espionage" on friends is even possible during assessment as in most sports or other types of contests or as in oral examinations of pairs of students. Active open learner modelling is a method for generating the learner's model for specific purposes when a particular need arises (Hansen & McCalla, 2003). In CosyQTI although we maintain
full learner models, we present only the pieces of information needed. For instance, the named inspection of the learner model of friends restricts the presented conclusions during assessment to prevent undesirable collaboration of learners which could lead to cheating and eventually to wrong conclusions.

Storing all the available information in standardised learning formats allows for alternative externalizations of the learner models and sharing of the information with other systems. Learner profiles are thus reusable by different teaching systems and other applications. Different applications could interpret and portray the available data differently.

**Actively involving learners and educators**

The purpose of opening the learner profile is not only to make learners reflect on their own performance and to compare their progress to the one of their peers, but also to utilize it on specific learning decisions. At the end of each section of questions a learner is classified and, based on the adaptation rules, the difficulty level of the subsequent questions is decided. At this point learners are informed about this decision and they are shown their model. A short dialogue is then performed between the system and the learner to determine whether this decision will be finally accepted. This interactive open learner modelling approach (Dimitrova, 2003) relies on the ability of learners to compare their data to their peers which is helpful in taking the appropriate decisions. Granting learners with the ability to amend the difficulty level resembles diving competitions where a diver may take more demanding challenges to earn a higher ranking. Similarly, learners using the CosyQTI tool have the ability to earn higher grades by trying questions of increased difficulty or they are able to test their knowledge in an augmented number of simpler questions. Naturally a learner's intervention is possible only in cases of self-assessment or in case the educator who designed the test permits it.

Another way that opening the learner model may facilitate decision making is via information forwarded to the educator. At the end of an assessment, if the system alters the knowledge level of a learner or other significant data, an email is sent to the educator, or a dialog box is shown if s/he is connected to the system, with the facts which led the system to act accordingly (see Figure 10). The teacher, who may have a different opinion based on previous assessments or teaching experiences with a specific learner, will be able to increase or decrease the knowledge level and modify other attributes as well. These subtle alterations may increase the pedagogical value of the tool to the benefit of the following assessments. This type of intervention by the educator is by all means legitimate especially if it is the first time the system is used or if the adaptive web testing tool is used sporadically. In these cases it is wise to at least grant the educator with the option to alter the system's beliefs so as to create a much more credible user profile.

Involving learners in diagnosis has already been proposed by Self (1990) and utilized in other studies as well (Dimitrova et al., 2000; Dimitrova, 2003). It has been proposed as a feasible research direction that could accommodate the dynamics of student's behaviour and make student modelling more tractable. Our approach is essentially the same, although at the moment it is more limited.
EVALUATION

The CoSyQTI web based adaptive assessment system presented above supports learning by adapting the testing procedure to the performance, goals and preferences of learners and to the teaching experiences and preferences of educators. Additionally, various open learner modelling methods are employed. Thorough evaluation of such a system is a complex procedure involving compound criteria, several users with different backgrounds and different adaptation approaches. However in this evaluation experiment our primary aim was to see how learners and educators value the externalization of their data, how they interpret and exploit them and what they believe or additionally desire. Therefore we performed a focused evaluation in an adult class.

For evaluating the open learner modelling techniques, a two-phase experiment was carried out. Two adaptive assessments were created and run with the same learner population. In the first assessment round the five participants were shown only their final result, i.e. their score and a pass/fail indication, whereas in the second assessment the additional information was displayed. With field observation and discussions in the lab we tried to realize and record learners’ satisfaction and in general to identify tendencies in using the available information and the pedagogical gains resulting from opening the model.

Both adaptive assessments consisted of three sections and were used in an introductory training program on information technology. Each section related to a different concept and contained 8 questions. Since our primary aim was to see how users handle inspection of their models we employed only one adaptive rule which is "move to the next section if you achieve a score of 80% in the first 6 queries". During test execution two educators were present so as to record their opinions as well. The
passing score for each assessment was set to 75%. Table 1 presents the number of learners who took each test and the number of learners who passed or failed in each test.

In the first assessment learners were presented with only their score and a pass/fail indication. At the end of the assessment we observed that the three learners (3/5) who had failed could not understand exactly what went wrong. One of the other two learners who passed the test with only 75% was pleased; s/he did not think that there were a lot of concepts that s/he did not know and that s/he very nearly failed. The educators with whom we discussed these problems argued that it would be beneficial for the learning procedure to inform learners in detail about their mistakes.

In the second adaptive assessment the information was displayed only at the end of the assessment. Learners were able to see the textual and diagrammatic form of their performance, their misconceptions and the problematic areas. They could monitor the performance of their peers in aggregated form and they could anonymously access the model of their peers. During the inspection of their models and with the aid of the focus interview that followed, we made some very important observations.

Table 1
Number of students participating in each test and passed or failed

<table>
<thead>
<tr>
<th>No of learners</th>
<th>Participated</th>
<th>Passed</th>
<th>Failed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st assessment</td>
<td>5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>(no open learner modelling)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd assessment</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>(open learner modelling)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 shows the most important data recorded during the evaluation of the system. As can be seen, all the learners inspected their model which shows that the information presented was useful to them. They realized why they failed or passed and which topics they did not know or they had not understood well. The learners who failed realized exactly which questions led them to fail, a point which, as it was concluded from the previous assessment, was not known to them. Learners who passed the test realized that they have not fully comprehended some concepts, which could possibly lead them to failure in future assessments. By asking each learner individually, we realized that indeed they were aware of the reasons behind their failure or their success. For example, learners with low marks told us which questions specifically they did not ask and they linked these questions to the section of their workbooks that they had not studied thoroughly.

All the learners inspected all the available forms to find out what their score was and which questions they failed. All of them were quite interested in viewing the models of their peers as an informal competition is always present and, in our opinion, desirable since it is motivating in terms of learning. They first tried to understand what each screen presented, so they quickly scanned all the screens and then they started all over again looking at each form more carefully. The time spent in viewing the model varied from approximately 4 minutes to approximately 11 minutes. We noticed that the two participants with the weaker performance spent more time in inspecting their models trying to understand the main reasons behind their low score.
Although they found graphical information more appealing and easily understandable, learners spent more time on the textual description as it is more informative. They liked the varying hues of gray used to depict misconceptions and problematic topics in charts as well as the positive and encouraging messages shown even if they had failed. Some of the learners asked for a tighter coupling between textual and graphical representations. For example, they said that it would be helpful if they could click on certain points on the diagrams and be led to the respective part of the textual description. That way they could better and more rapidly reason about their learning progress.

The inspection module for the friend's model was inactive during this assessment test. However, we noticed that participants showed their models to each other and they compared and discussed their results. These actions are indications that the named inspection of the model of a friendly peer is a useful module and could assist the learning procedure. This module could be considered as a special case of collaborative learning as it drives learners to share opinions and knowledge after inspecting each other's model and trying to clarify their misconceptions.

An important point made clear to learners was that they were not given the same number of questions which eventually helped them realize the adaptive nature of the assessment and the benefits resulting from this fact (e.g. shorter testing procedure). Nevertheless, they were not sure why their testing routes were different. Some of the learners wanted to know, for example, why they were given only 6 questions in the first section. Thus they asked for an externalization of the adaptive rules in an understandable mode. While we are not fully convinced that this externalization would assist them they argued that they would realize what they had done that made them proceed faster than the others.

Learners also asked for a more precise report of the items that they had failed to answer correctly. More specifically they asked for a report of the wrong and the unanswered questions with the right option ticked. This could be an option if a test is used for self-assessment. However if it is used for actual testing on which formal grading is based it is preferable to avoid such reporting in order to
prevent the circulation of the testing material. Both the educators who participated in the evaluation experiment agreed on this point.

Moving on to the educators’ opinions, we can say that they are equally positive as in the case of learners. The educators were Computer Science teachers who were particularly interested in the use of technology in teaching. They had used automated non adaptive testing systems and simplistic learning management systems in the past. They both liked the fact that they could see an elaborated view of their learners' achievements and difficulties. Diagrams amassed the results and assisted them to draw general conclusions. However they stated that the textual analysis was more informative and permitted them to draw more specific conclusions. They found the email sent to them, summarising the performance of the learners and the decisions of the system, to be quite practical. This option allows them to relate the system’s behaviour with their approach to teaching and to their experiences with specific learners. Intervention is confronted positively and considered as a tool which enables avoidance of hasty conclusions on behalf of the system. Finally, they observed that in one section all the learners had a lower performance which is possibly an indication that this section contained ambiguous items. They would have liked to be shown more details so as to become conscious of the exact problems.

The main limitation of this work in terms of evaluating the potential of an open learner model is that, strictly speaking, the learner model data concerning knowledge, problems and misconceptions were presented together with more general feedback information and statistics. There was no systematic investigation of exactly which information learners were viewing. Nevertheless, such a combination of information types may be considered useful by learners.

**DISCUSSION AND FUTURE WORK**

In this paper we presented an adaptive web testing system and the open learning modelling techniques it supports. In the CosyQTI tool, learner models are displayed to learners and to educators in textual and graphical mode. Learners are given anonymous access to the model of their peers and eponymous access to the model of friend-learners. Educators can study the models of their students and spot potential problems with their teaching and adaptation strategies. Also they are able to alter certain learner model characteristics to create more credible and robust models. Learners have the option to discuss some of the system’s decisions and form a more "fitting" testing sequence. Apart from the open learner modelling potentials of the system, CosyQTI is a domain independent system conforming to the established learning technology standards in an attempt to promote interoperability. Since learner models are structured using IEEE PAPI (2002) and IMS LIP (2005) standards the information gathered could be utilized in other learning tools with open learner modelling support, allowing educators to have an elaborate view of the learning background of a user prior to her/his interaction with their tool. Its component based architecture makes it possible to expand the services of CosyQTI by integrating new modules or by using the existing modules with other learning tools.

An evaluation regarding the open learner modelling module of the system revealed that learners and educators benefited when the learners' performance was analyzed in detail. They realized their learning achievements and difficulties and they compared their progress to one of their peers initiating collaborative solving of the misconceptions presented in their models. Educators were automatically informed about the decisions of the system via an email forwarded to them. They found this automatic synopsis quite useful. Diagrams were understandable but presented the learning progress in condensed
form. Therefore learners and educators preferred the textual mode over the graphical as it tended to be more informative. The adaptive nature of the test was clear when the learner models were opened out.

This initial evaluation study produced new research directions as well. Participants asked for tighter coupling between graphical and textual modes and for more details on their adaptive testing routes. Visualization of the testing paths had the potential to be practical both in the authoring phase of the adaptive tests and at the end of each assessment. More elaborate reporting on the wrong questions was requested although this demand is arguable. Collaborative clarification of the misconceptions among learners is not supported by the system at the end of the assessment, even though it seems to be a desirable feature. This could be achieved by automatic learner matching based on their achievements and difficulties. Learners who answered some questions could be matched to learners who had not answered them so as to assist each other. Detection of ambiguous questions based on massive learner difficulties could be an option for the system to help educators to improve their testing material.

REFERENCES


