A Cognition Assessment Authoring System for E-Learning

Jason C. Hung* and L. J. Lin**

**Department of Information Management*
Kuang Wu Institute of Technology
***Institute for Information Industry Education & Training Division*
Taiwan, R.O.C.

Wen-Chih Chang, Timothy K. Shih, Hui-huang Hsu, Han-Bin Chang, Hsuan-Pu Chang, Kuan-Hao Huang
Multimedia Information NEtwork (MINE) Lab
Department of Computer Science and Information Engineering
Tamkang University, Tamsui, Taipei Hsien
g8190239@tkgis.tku.edu.tw ;
tshih@cs.tku.edu.tw

Abstract

With the rapidly development of distance learning and the XML (Extensible Markup Language) technology, metadata becomes an important item in an e-learning system. Today, many distance learning standards such as SCORM, AICC CMI, IEEE LTSC Learning Object Meta-data (LOM), and IMS Learning Resource Metadata XML Binding Specification, use metadata to tag learning materials, shareable content objects, and learning resources. However, most metadata is used to define learning materials and test problems. Few metadata is dedicated for assessment in learning. In this paper, we proposed an assessment metadata model for e-learning operations. With the support from the assessment metadata, we can collect information at the question cognition level, Item Difficulty Index, Item Discrimination Index, questionnaire style, and question style. The assessment analysis model provides individual questions, summary of test results, and analytical suggestions. The suggestions and results can tell teachers why a question is not suitable and how to correct it. Teachers can see the analysis of test result and fix problematic questions. With the cognition level analysis, teachers can avoid missing items in teaching. The mechanism developed also suggests an e-learning system, with adaptive learning content and individualized tests, as well as good advice for the teachers.

Keyword: Cognition Level, Item Discrimination Index, Item Difficulty Index, Assessment Analysis Model, distance learning

1. Introduction

As the popularity and importance of distance learning increase, learning materials and group communications are widely established on Internet and wireless infrastructures. Whether in distance learning programs, e-learning portals, or the traditional education environment, instruction and assessment operates together as a complete learning cycle. But, how could a teacher realize the blind spot of a learner? This is the weakness of learning cycle. Assessment provides a suitable method to gather student feedback. A good assessment not only offers test, but also analysis test results for a teacher. With the interaction and analysis, teachers can fix their teaching strategies, and reedit or reorganize learning materials. In addition, the teacher can derive benefit from the assessment; the students can also realize what is the key point of learning materials. Assessment responses to the learners in terms of what is the major and most important part in each subject and course.

2. Related work

There are many e-learning standards exists. However, each standard emphasizes on different topics. We make a simple summary of these e-learning standards.

2.1 Learning Object Metadata

Metadata provides a common nomenclature for learning resources to communicate and exchange with the others in a common way. A good metadata need completeness, carefulness, and flexibility. The most famous metadata is the IEEE LTSC's Learning Object Metadata (LOM). It provides nine categories to describe learning resource.

2.2 Course hierarchy and structure

In an e-learning environment, course structure will effect on the learning resource transformation and educational knowledge constitution. About course hierarchy, the previous idea is content-block-sco. With the AICC [1] nomenclature, the course structure is divided into two elements.

2.3 Assessment Specification from IMS

IMS Question &Test Interoperability (Q&TI) [3] specification allows systems to exchange questions and tests. This standard is a powerful standard for complex assessment. The other distance learning standard (such as ADL SCORM [2], ULF offers item difficulty index and question style.

2.4 Environment

An environment to support e-learning operation, such as course management, student management, learning resource delivery, tracking service, integration of course and teaching methods and common format for exchange. DoD's ADL [2] proposed SCORM which has a clear concept and environment. In the Run-Time Environment, there are data model, SCO, Asset, API, Launch mechanism and LMS.

3. The MINE SCORM Meta-data

We defined an assessment metadata for E-learning in this section. We use SCORM as our reference. We hope the assessment metadata is conductive to E-learning. We call the assessment metadata MINE SCORM Meta-data Model. MINE SCORM Meta-data is designed specially for assessment in distance learning. Including assessment record, assessment analysis, and questionnaire and cognition level. The whole MINE SCORM Meta-data is represented in a tree-like structure in Figure 1.

3.1 Cognition level:

Instruction objective plays a very important role in teaching progress. If the instruction objective is clear, it guides teaching activities and evaluation precisely and properly. Bloom proposed the taxonomy of educational objectives into three domain, they are cognitive domain, psychomotor domain and affective domain. In cognitive domain, it includes knowledge, comprehension, application, analysis, synthesis, and evaluation.

3.2 Question style:

I. Essay: Defines the text of an open-ended essay question. You can also use it to represent shorter fill-in-the blank. Two elements are Question and Hint.
II. True False Item: Defines a question whose answer is either true or false. Two elements are Question and Hint.
III. Multiple Choice: Defines a question with multiple choice answers
IV. Match Item: Define a question with proper matched choice
V. Completion Item: Design a question like fill-in blank or cloze.
VI. Questionnaire:
   A. Question: The content could be text, graph, and draw a picture. In this metadata, we focus on text.
   B. Resumable: True means resumed and false means paused at a later time.
   C. Display Type: Fixed Order — for tests with a fixed number and order of questions. Random Order — for tests with a random order.

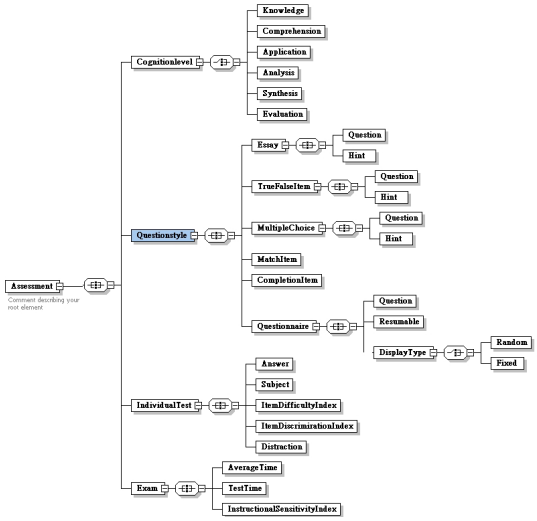

Figure 1. Our proposed assessment tree consists of ten sections with Cognition and Questionnaire Metadata illustrated.

3.3 IndividualTest:

I. Answer: Correct answer for explaining and query.
II. Subject: Define each question a main subject.
III. Item Difficulty Index: A simple explain is below.
P=R/N (100), which P: Item Difficulty Index, R: The number which people have right answer. N: Sum
For example, R=800, N=1000, then P= R/N=800/1000=0.8 (80%)
Generally speaking, the more Item Difficulty Index increase, the question is easier.
IV. Item Discrimination Index: An index for judging a question's discrimination.
V. Distraction: With the analysis, define students' distraction.

3.4 Exam:

I. Average Time: Each people take different time answering questions, we use average time for operation.
II. Test Time: A default time limit for testing.
III. Instructional Sensitivity Index: With the

comparison between the test result before teaching and the test result after teaching to analysis Instructional Sensitivity Index.

4. The Analysis Model

A completed teaching could divide into three parts. First part is teacher's teaching strategy (such as, game, direct, discussion and experimentation). Second part is the learning content (learning material, for example, handbook, music score and textbook). Last part is assessment (such as, questionnaire, test, exam and quiz). Assessment plays an important part in this model. A teacher uses proper teaching strategy and good learning content to teach students. However, we don't know if students receive the information or not. The only way is hold a test. With the test result and analysis, teacher may know what the students need, how the students received, what the learning content should add or delete.

4.1 Single question statistic and analysis

4.1.1 Number representation.

Teacher can see each question's status. Also it will provide some suggestions from the test question.

No	P_H	P_L	$D=P_H-P_L$	$P=(P_H+P_L)/2$
…	…	…	…	…

No: The question's Number

$P_{H:}$ the higher 25% of total student as the higher group

$P_{L:}$ the lower 25% of total student as the lower group

$D=P_H-P_L$ $P=(P_H+P_L)/2$

Prof. Kelly said that the best percentage is 27%, and the acceptable percentage is 25%-33% (Kelly, 1939). We tried to define the percentage 25% in this paper.

(1) 1st step: according to score height arrange the examination paper

(2) 2nd step: we define P_H the higher 25% of total student as the higher group and then P_L the lower 25% of total student as the lower group.

(3) 3rd step: calculate the people answer correct and his percentage in higher group and lower group in each question.

(4) 4th step: Calculate each question Item Difficulty Index $P=(P_H+P_L)/2$

(5) 5th step: Calculate each question Item Discrimination Index $D=P_H-P_L$

4.1.2 Signal representation.

With signal presentation, the advice to teacher becomes more easy and simple. In table 1, we defined a single problem item attribute. HA means the number of students in high score group select option A. The other HB, HC, HD, HE, LA, LB, LC, LD and LD are the same meaning.

Table 1: Problem Attribute

	Option A	Option B	Option C	Option D	Option E
High Score Group	HA	HB	HC	HD	HE
Low Score Group	LA	LB	LC	LD	LE

We defined four rules to analysis problems. These rules are introduced following:

Rule 1: If (LA|LB|LC|LD|LE)=0 then the option's allure is low

Example1	Option A	Option B	Option C	Option D	Option E
High Score Group	12	2	0	3	3
Low Score Group	6	4	0	5	5

Assume that the high score group=20, and the low score group=20. There are 6 people choose option A, 4 people choose option B, 0 people choose option C, 5 people choose option D and 5 people choose option E in the low score group. The answer of this question is option A. The option C didn't attract any one of the low score group. (LA|LB|LC|LD|LE)=0 Then we can find that the option's allure is low.

Rule 2: N={A, B, C, D, E}

If Option N is correct AND HN < LN Then the option is not well-defined.

If Option N is wrong AND HN > LN Then the option is not well-defined.

Example2	Option A	Option B	Option C	Option D	Option E
High Score Group	1	2	10	0	7
Low Score Group	2	2	13	1	2

Assume that the high score group=20, and the low score group=20. The correct answer is option C and the option E is wrong. We can see the people who choose option C in low score group is greater than high score group. Then we can find the option C has some problem. Also in this case, the option E is wrong, but the people in high score group is greater than low score group. Also we find that the option E is not a good option.

Rule 3: LM=MAX(LA ,LB ,LC ,LD ,LE)

Lm=min(LA ,LB ,LC ,LD ,LE)

LS=LA+LB+LC+LD+LE

If $|LM-Lm| \leq LS*20\%$ Then people in low score group lack concept

Example3	Option A	Option B	Option C	Option D	Option E
High Score Group	15	2	2	0	1
Low Score Group	5	4	5	4	2

Assume that the high score group=20, and the low score group=20. In this case, we can see the LM=5, Lm=2, and LS=20. $|LM-Lm|=3 \leq 4=LS*20\%$. We can find that low score group people choose every option equally. They don't know the answer in this problem. Then the information is very important to instructors to give the remedied course to low score group students.

Rule 4: HM=MAX(HA ,HB ,HC ,HD ,HE)

Hm=min(HA ,HB ,HC ,HD ,HE)

HS=HA+HB+HC+HD+HE

If |HM-Hm| ≤ HS*20% AND |LM-Lm| ≤ LS*20%

Then people in low score group and high score group lack concept

Example4	Option A	Option B	Option C	Option D	Option E
High Score Group	4	4	4	2	6
Low Score Group	5	4	5	4	2

Assume that the high score group=20, and the low score group=20. In this case, we can see the LM=5, Lm=2, LS=20, HM=6, Hm=2 and HS=20. |LM-Lm|=3 ≤ 4=LS*20% and |HM-Hm|=4 ≤ HS*20%. We can find that low score group and high score group people choose every option equally. They don't know the answer in this problem. Then the information is very important to instructors to give the remedied course to all students.

From each rule we can identify what kind of status in our test. Some of the information is useful for correcting the improper questions given in the exam, and the others are useful for instructors to realize students' learning. (Table 2) Table 3 shows the signal representation status. We may also see the prototype in Figure 2.

Table 2: Every status in four rules

	The option's allure is low	The option meaning is not clear	Careless	Not only one exact answer	Low score group lack concept	High score group lack concept
Rule 1	V	X	X	X	X	X
Rule 2	X	V	V	V	X	X
Rule 3	X	X	X	X	V	X
Rule 4	X	X	X	X	V	V

Table 3: Some advice and different suggestions about questions.

Status	Light signal	D	Rule 1	Rule 2
Good	Green	Higher 0.3		
Fix	Yellow	0.2-0.29	match	match
Eliminate or fix	Red	Lower 0.19		

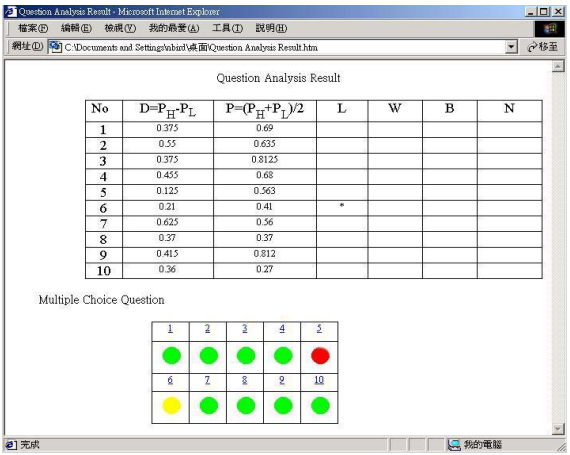

Figure 2. Signal represent interface for whole test.

Let's see an example of no.2 question. Assume that the class size is 44 students, the high score group and low score group is 11.

Class	A	B	C	D
High score group	0	0	10	1
Low score group	3	2	4	2

P_H=10/11=0.909 ≅ 0.91 P_L=4/11=0.36

D=P_H-P_L=0.91-0.36=0.55 D>0.3 The signal is green.

P= ($P_{H+}P_L$)/2=(0.91+0.36)/2=0.635

Let's see another example of no. 6 question.

Class	A	B	C	D
High score group	1	1	4	5
Low score group	0	2	4	4

P_H=5/11=0.45 P_L=4/11=0.36

D=P_H-P_L=0.45-0.36=0.09

P= ($P_{H+}P_L$)/2=(0.45+0.36)/2=0.41

Rule1: (LA|LB|LC|LD|LE)=0 The allure of option A is low.

4.2 Total Test statistic and analysis

The assessment analysis should be presented in different aspects. A total test analysis result could show the whole status of students.

4.2.1 Figure representation.

(1) Time (cross axle) and Number of answered question (vertical axle) figure: The figure shows the test time is enough or not.

(2) Test score (cross axle) and degree of difficulty (vertical axle) figure: The figure shows the distribution of score and difficulty.

(3) Cognition level (cross axle) and learning content subject (vertical axle) figure shows the cognition level, question number and subject. (See Table 4)

Table 4: Two-way specification table

	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation
Concept 1	A1	B1	C1	D1	E1	F1
...	...	...	...	...	...	...
Concept i	Ai	Bi	Ci	Di	Ei	Fi
	SUM(A1-Ai)	SUM(B1-Bi)	SUM(C1-Ci)	SUM(D1-Di)	SUM(E1-Ei)	SUM(F1-Fi)

4.2.2 Definition.

(1) Cognition level divided into six level, each named from A to F. Assume X is universal set, X={A, B, C, D, E, F} Ex.

Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation
A	B	C	D	E	F

(2) Concept in the test would be named from 1 to i, initial i=1 ex. Concept 1

(3) From concept 1, we write a question belongs to Knowledge cognition level. Then A1 is set [TRUE]. If over one question belongs to Knowledge cognition level exist in concept 1. A1 is [TRUE] to represent there is a question of knowledge level in concept 1 at least. If A1 is [FALSE], there is no question of knowledge level in concept 1 at least.

(4) SUM(Xi) is the question's sum of cognition level X in concept i.

ex. SUM(F3)=3, there are 3 questions of evaluation level in concept 3.

(5) SUM(Ai-Fi) is the question's sum in concept i. SUM(A10-F10)=8, there are 8 questions (From Knowledge to Evaluation level) in concept 10.

(6) SUM(B1-Bi) is the question's sum of Comprehension (From Concept 1 to Concept i). ex.SUM(C1-C7)=7, there are 8 questions (From Concept 1 to Concept 7).

4.2.3 Analysis.

(1) Concept Lost
If (A1|B1|C1|D1|E1|F1)=FALSE, Concept 1 lost in the exam.

(2) Cognition level and question's sum relation
SUM(A1-Ai) ≥ SUM(B1-Bi) ≥ SUM(C1-Ci) ≥ SUM(D1-Di) ≥ SUM(E1-Ei) ≥ SUM(F1-Fi)

(3) Distribution of cognition level and question (paint algorithm)

5. Assessment Authoring System Architecture

The assessment authoring system architecture is divided into several sections. (Figure 3) There are two databases, one is internal problem and exam database, and another one is SCORM compatible external repository. Assessment authoring system includes problem search, exam authoring, problem authoring and SCORM format output service. Another one is on-line exam monitor subsystem for capturing picture.

Authors, instructors and tutors use the assessment authoring system to edit problems or exam. They can search similar or specific subject or related problems from problem & exam database. In addition to search the database, instructors may edit the problem and exam by themselves. After authoring the problems, instructors can combine their own problems with the problems from database. In order to share the material of our problem and exam, our system provides SCORM format package output service. The service can package the original problem and exam files to SCORM compatible files. Other instructors may reuse the problem and exam files from SCORM compatible external repository. Administrator control the database and learning management (LMS) monitor function. Learners take the exam or the problems with Internet browser. When learners take the exam, monitor function captures the client picture for monitoring the exam progress.

5.1 Problem Type

Problem authoring provides several problem types, and there are choice problem (Figure 3), fill-in blank problem and true-false choice problem.

5.2 Problem Attribute

Problem in our system has two sections, one is metadata information, and another one is problem content. Metadata is used to describe the object information.

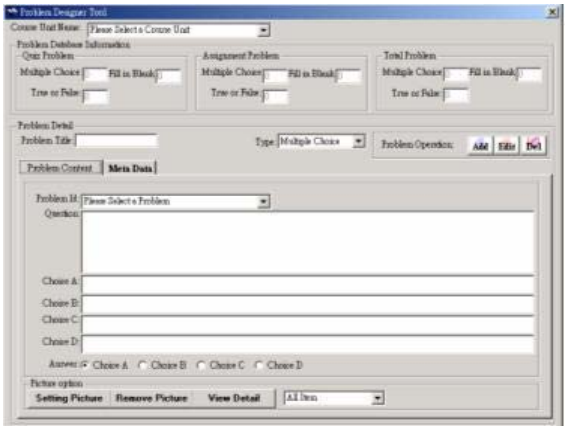

Figure 3. Choice Problem authoring interface

5.3 Template

We can put a picture in a problem, it is allowed to set the picture's position (x axis; y axis). Besides, we can set the question description and question selection items. In Figure 4, We set the presentation style by moving each item .

When instructor edited the problem, he wanted to copy the problem structure for reuse. He can add a new template in the exam. Also, he can delete an existed template.

Figure 4. Edited problem presentation style

5.4 Exam Authoring

There are various kinds of exam presentation style. It is hard to design all possible exam presentation styles. In order to solve the problem, instructors can use group

service to make all possible presentation style. (Figure 5)

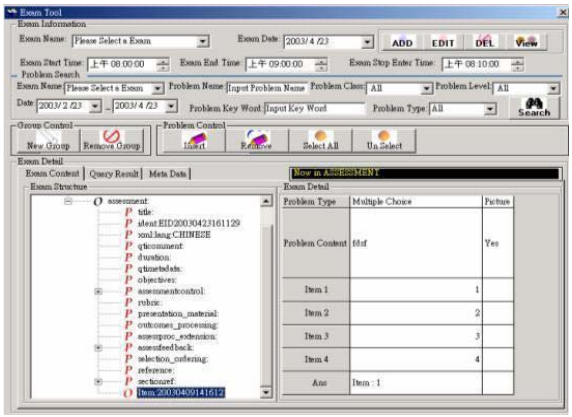

Figure 5. Exam authoring interface

5.5 SCORM Format Output

In SCORM standard, each file (ex. Html file, jpeg file, gif file, asf file, doc file, ppt file) has a descriptive xml file with the same level in the course structure. In addition to these descriptive xml files, a main description is an xml file called imsmanifest.xml. With this imsmanifest.xml, we can parse the whole course structure. Thirdly, java script files to communicate with API and learning management system are necessary to SCORM standard. Without these java scripts, the learning management can't find the API to communicate. Some API functions are used to set value (ex. learner record, learner progress, learner status), get value, error handler (ex. error message transfer, error status record, error dialog) and course beginning and ending (ex. course initial and course finish).

6. Conclusion

In this paper, we proposed the MINE Assessment Metadata to support e-learning operation. The assessment metadata strengthen SCORM assessment metadata. Also it offers an interaction to students and teachers.

In this paper we also implemented the assessment system provided instructors to design, edit, authoring the exam and problems easily by themselves. These files manufactured by the system are compatibly with SCORM standard and the authoring concept is also referenced IMS QTI. Besides, we provide a monitor function for capturing learner's picture during exam. In the near future, we will add the adaptive test algorithm and assessment feedback in our assessment system.

7. Reference

[1] Aviation Industry Computer Based Training Committee. Web site at http://www.aicc.org/ [Last access August 20th, 2003]

[2] Dodds. P. "*Sharable Content Object Reference Model (SCORM)*". Version 1.2. Technical report, Advanced Distributed Learning Initiative, January 2001. Electronic version available at http://www.adlnet.org/Scorm/downloads.cfm#spec [Last access August 20th, 2003]

[3] Smythe C., Shepherd E., Brewer L. and Lay. S. "*IMS Question & Test Interoperability Overview*". Version 1.2. Technical report, Feburary 2002. IMS Global Learning Consortium, Inc, Electronic version available at http://www.imsproject.org/question/qtiv1p2pd/imsqti_oviewv1p2.html [Last access August 20th, 2003]