Innovations in Education and Teaching International
Vol. 45, No. 1, February 2008, 77–89

Routledge
Taylor & Francis Group

Diagnosing student learning problems based on historical assessment records

Gwo-Jen Hwang[a*], Judy C.R. Tseng[b] and Gwo-Haur Hwang[c]

[a]*Department of Information and Learning Technology, National University of Tainan, Taiwan, ROC;* [b]*Department of Computer Science and Information Engineering, Chung Hua University, Taiwan, ROC;* [c]*Information Management Department, Ling Tung University, Taiwan, ROC*

In the past decade, researchers have attempted to develop computer-assisted learning and testing systems to help students improve their learning performance. Conventional testing systems simply provide students with a score, and do not offer sufficient information in order to improve their learning performance. It would be of more benefit to students if the test results could be critically analysed and hence learning suggestions could be offered accordingly. This study proposes an algorithm for diagnosing students' learning problems and provides personalised learning suggestions for Science and Mathematics courses. An intelligent tutoring, evaluation and diagnosis system has been implemented based on the novel approach. Experimental results on a Mathematics course have demonstrated the feasibility of this approach in enhancing students' learning performance, making it highly promising for further study.

Keywords: Computer-based testing; learning diagnosis; science education; computer-assisted learning; concept-effect relationships

Introduction

Rapid advances in computer and communication technologies have encouraged researchers to apply the technologies to the development of computer-aided tutoring and testing systems (Antao et al. 2000; Chou 2000). For example, Vasandani et al. proposed a system which could assist in organising system knowledge and operational information to enhance operation performance (Vasandani and Govindaraj 1991, 1995; Vasandani et al. 1989); moreover, Gonzalez and Ingraham (1994) presented a system that automatically determined exercise progression and remediation during a training session based on past student performance. Meanwhile, various techniques and tools for developing intelligent tutoring systems were also being proposed. For example, Harp, Samad, and Villano (1995) employed neural network techniques to model students' behaviours in the context of intelligent tutoring systems; Rowe and Galvin (1998) demonstrated the use of planning methods, consistency enforcement, objects and structured menu tools to construct intelligent simulation-based tutors for procedural skills. Moreover, several technologies for detecting on-line status of students to establish interactive intelligent tutoring systems were also proposed (Giraffa et al. 1999; Hwang 1998). It can be observed that the development of intelligent tutoring systems and learning environments has recently become a key issue in both computer science and education (Ozdemir and Alpaslan 2000).

In conducting an intelligent tutoring process, the learning status of each student must be assessed, and examination is a typical evaluation method for identifying the learning status of

*Corresponding author. Email: gjhwang@mail.nutn.edu.tw

ISSN 1470-3297 print/ISSN 1470-3300 online

DOI: 10.1080/14703290701757476
http://www.informaworld.com

students. Researchers have shown that paper-administered and computer-administered exams are equivalent in terms of testing quality, encouraging the development of computer-based testing systems and relevant techniques (Olsen et al. 1989). For example, Fan, Tina, and Shue (1996) presented a system capable of changing the numeric part of test items during testing to prevent students from memorising the answers; Wainer (1990) proposed a form of computerised adaptive testing which applies some forecasting methodologies to shorten the test without compromising accuracy. In addition to the online testing results, an investigation for assessing students based on their group discussions and portfolio was reported by Rasmussen et al. (1997).

In conventional testing systems, students are assigned a score or grade as a test result to represent their learning status. This approach allows students to know their score (or grade) in reference to their learning status, but suggests that students might be unable to improve their learning status without further guidance. In 2003, researchers proposed a *concept-effect model* to represent the prerequisite relationships among concepts in a course (Hwang 2003a). The model demonstrated a systematic procedure to guide teachers to compute students' learning status for each concept semi-manually. Additionally, a computer-based approach was proposed to assist teachers in constructing the concept-effect relationships (Hwang et al. 2003). However, previous experiences also demonstrated the problems of applying the existing learning diagnosis models. For example, the learning diagnosis model is not applicable if no relevant test item of a concept was tested. Moreover, in the existing models, only subject opinions given by the teachers are adopted when constructing concept-effect relationships, while objective data (e.g. historical assessment records) are ignored.

To cope with these problems, this investigation proposes a novel approach for constructing concept-effect relationships and diagnosing students' learning problems. In the new approach, relationships among subject concepts and test items are reorganised and a conditional probability model is proposed to construct concept-effect relationships and to diagnose student learning problems, based on current test and the historical assessment records. Furthermore, an intelligent tutoring, evaluation and diagnosis system has been developed based on the proposed approach to provide objective evaluation and personalised suggestions for each student.

Related work

Although many researchers have identified testing and assessment as important issues in computer-based instruction and have suggested appropriate design strategies and techniques, few systems have attempted to diagnose relevant student learning problems or provide personalised learning guidance. Most conventional testing systems assign a score or status indicator to each student after testing, thus determining the learning status of that student, but do not consider how to improve upon it.

Salisbury (1998) indicated that the concept of learning information, including facts, names, labels or paired associations, is frequently a prerequisite to the efficient performance of a more complex and higher-level skill, particularly in Science and Mathematics courses. For example, thorough knowledge of the names and abbreviations of chemical elements and their atomic weights is required to comprehend scientific writing or chemical formulae. Hence, to learn a scientific concept effectively, people generally require knowledge of some basic concepts. Researchers have defined such relationships among concepts as the *concept-effect model* (Hwang 2003a).

Consider two concepts or skills, C_i and C_j. If C_i is a prerequisite for the efficient performance of the more complex and of higher-level concept C_j, then a *concept-effect relationship* $C_i \rightarrow C_j$ is said to exist. Notably, a concept may have multiple prerequisite concepts, and a given concept can also be a prerequisite concept of multiple concepts. For example, to learn the concept 'multiplication', one might first need to learn 'addition', while learning 'division' might require first learning 'multiplication' and 'subtraction'. To help teachers to construct concept-effect relationships, a computer-assisted system has been developed to analyse test answer sheets in order to provide statistical information to the teachers (Hwang et al. 2003).

Following the construction of concept-effect relationships, the main problem is the methodology used to diagnose students' learning problems. Notably, previous investigations employ a semi-manual approach to diagnose students' learning problems that forecast the following steps (Hwang 2003a; Hwang et al. 2003):

Step 1: Calculate the incorrect answer rate for each student and each concept.
Step 2: Define a threshold for identifying poorly learned concepts.
Step 3: Record the poorly learned concepts and give feedback to the students.

Although the previously proposed approaches have been applied to several subject units and have achieved encouraging results, there are several problems that need to be acknowledged (Tseng and Hwang 2004):

- Those approaches will not be applicable in determining the learning status of a concept if not enough test items related to the concept have been tested. That is, the learning diagnosis function will not be applicable when employing those approaches unless a sufficient number of test items have been tested for each concept in the specified subject unit.
- In those previous approaches, only subject opinions given by the teachers are taken into account in constructing concept-effect relationships, while objective data, i.e. historical assessment records, are ignored; therefore, the learning diagnosis results may be affected by personal bias.
- As historical assessment records are not taken into consideration in conducting the learning diagnosis, students' learning problems caused by the poorly learned concepts in other subject units could not be detected.
- Since the semi-manual approach involves several manual operations, it can only be applied to narrow-ranging subject units containing a limited number of concepts (so far the approach works for subject units containing less than 15 concepts). However, most science courses involve numerous and interrelated concepts, meaning that the semi-manual approach may not be applicable.

To deal with these problems, the following sections present an algorithm for diagnosing students' learning problems as well as the implementation of a tutoring, evaluation and diagnosis system based on the algorithm. Moreover, some experimental results are given to evaluate the performance of the system.

Method for diagnosing student learning problems

In order to diagnose the learning status of a student, a conditional probability-based algorithm for diagnosing student learning problems is proposed in this section. This has been applied to the development of a tutoring, evaluation and diagnosis system. In the following subsections, an Assessment Record Accumulation Model which describes cross-unit concept-related failure ratios according to historical assessment records of students is given, and the conditional probability approach to deriving concept-effect relationships and diagnosing student learning problems is proposed.

Assessment Record Accumulation Model of concept-related failure ratios

To compute accurately the concept-effect relationships among concepts, it is important to accumulate the test results and compute the concept-related failure ratios after performing each test. Figure 1 shows the graphical representation of the Assessment Record Accumulation Model. In

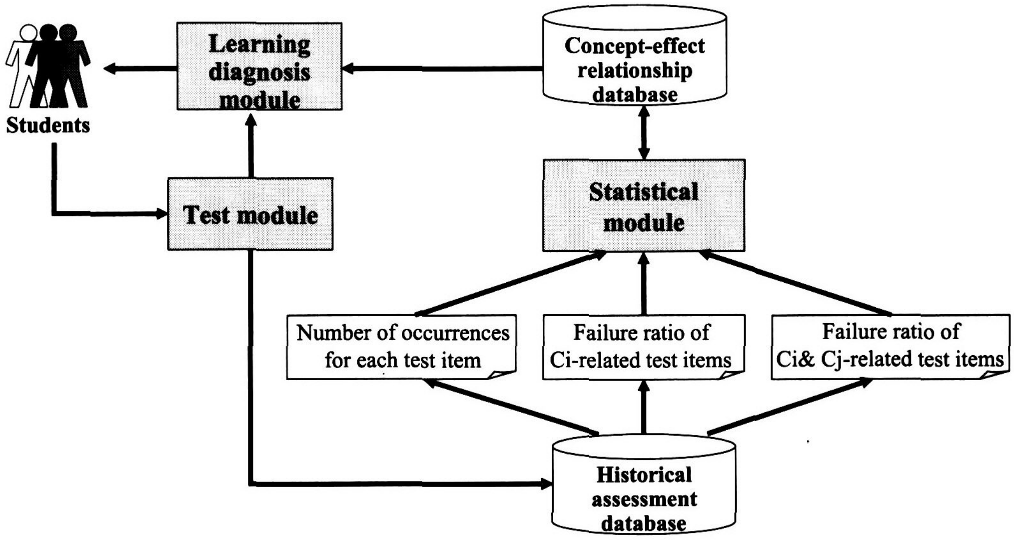

Figure 1. Graphical representation of the Assessment Record Accumulation Model.

this model, the historical test records of each student (i.e. student identification, data and time of receiving the test, the answer to each test item, the correctness of the answer) as well as the parameters of the relevant test items (i.e. the list of concepts that are relevant to each test item, the degree of relevance for each concept to each test item) are recorded in the historical assessment database.

The accumulated test records are then retrieved by a statistical module to find the information for determining concept-effect relationships by conditional probability, including the number of occurrences for each test item, the accumulated failure ratio for the test items that are related to each concept (say C_i), and the accumulated ratio for the students who failed to answer C_i-related test items who also failed to answer the test items related to another concept (say C_j).

According to the definition of conditional probability, $P(F \mid E)$ represents the probability of the occurrence of F under the premise that E occurs. Let $P(C_i)$ represent the failure ratio for students to answer correctly the test items related to concept C_i, $P(C_i \mid C_p)$ represents the failure ratio for students to answer correctly the test items related to concept C_i under the premise that the students fail to answer correctly the test items related to C_p. That is, $P(C_i \mid C_p)$ denotes the effect degree for concept C_p to concept C_i, that is, the degree of concept-effect relationship for C_p to C_i (Hwang 2003a; Tseng and Hwang 2004). For example, assume that C_i represents 'multiplication' and C_p represents 'addition', $P(C_i \mid C_p)$ denotes the failure ratio for students to answer correctly test items related to concept 'multiplication' under the premise that the students fail to answer correctly the test items related to 'addition'.

Conditional probability algorithm for deriving concept-effect relationships

Let $D(C_p, C_i) = P(C_i \mid C_p)$ denote the effect degree for the student learning status of concept C_p to that of C_i, which implies the failure of learning C_p will affect the learning of C_i, a directed link will be created from C_p to C_i in the graphical representation of the concept-effect relationships, called the *Concept Effect Graph*. The algorithm for constructing the Concept Effect Graph is given as follows:

Step 1: For each concept C_i, find two concepts, say C_p and C_k, with highest $D(C_i, C_p)$ and $D(C_i, C_k)$ values; that is, C_p and C_k are the two most influential concepts for students to learn C_i.

Step 2: Create directed links from C_p and C_k to C_i.
Step 3: For each concept C_i, find two concepts, say C'_p and C'_k, with highest $D(C'_p, C_i)$ and $D(C'_k, C_i)$ values; that is, without learning C_i well, the students are most likely to fail to learn C'_p and C'_k.
Step 4: Create directed links from C_i to C'_p and C'_k.
Step 5: Eliminate cycles among two concepts by removing the less influential relationships.
Step 6: If any cycle exists among three or more concepts in the Concept Effect Graph, remove the link with the smallest effect degree from the cycle.

By finding the effect degree for each pair of concepts, the concept-effect relationships are obtained and stored in the concept relationship database for diagnosis purpose. When students receive a test, their learning problems can be diagnosed based on their answers to the test items and the relevant concept-effect relationships in the database. Assuming that the failure ratio for student S_k to concept C_i is higher than the given threshold, C_i's parent concepts in the Concept Effect Graph are traced in sequence. If the failure ratio for student S_k to a parent concept C_p of C_i is higher than the threshold, C_p is added to the to-be-enhanced path of S_k. The tracing process is repeated until a low-failure-ratio parent concept is found, or the root of the Concept Effect Graph is reached.

Illustrative example

A Test Item Relationship Table (TIRT) represents the degree of association between test item Q_t and concept C_i. An illustrative example of a TIRT comprising 10 concepts and 12 test items is listed in Table 1, where TIRT $(Q_t, C_i) = 1$ represents 'relevant' and TIRT $(Q_t, C_i) = 0$ represents 'irrelevant'.

After performing a test, the answers of the students are recorded in an Answer Sheet Table (AST). Table 2 shows an illustrative example of an AST, where each entry AST (S_k, Q_t) is a value ranging from 0 to 1; 0 indicates that student S_k answered test item Q_t correctly; 1 indicates that S_k failed to answer Q_t correctly; and a value between 0 and 1 indicates a partially correct answer. Notably, for true/false questions and multiple-choice questions, AST (S_k, Q_t) value is either 0 or 1; while for short-answer questions, AST (S_k, Q_t) value can range from 0 to 1.

Table 1. Illustrative example of a Test Item Relationship Table (TIRT).

	Concept C_i					
Test item Q_t	Coefficient	Exponent	Addition of polynomial	Multiplication of polynomial	Division of polynomial	Distributive law
Q_1	1	1	0	0	0	0
Q_2	1	1	0	0	0	0
Q_3	0	1	0	0	0	0
Q_4	1	0	1	0	0	0
Q_5	1	0	1	0	0	0
Q_6	1	0	1	0	0	0
Q_7	1	0	1	0	0	0
Q_8	0	1	0	1	0	0
Q_9	0	0	0	1	0	1
Q_{10}	1	1	0	0	1	0
Q_{11}	0	0	0	1	0	1
Q_{12}	1	1	0	0	1	0
Q_{13}	0	1	0	0	1	0

Table 2. Illustrative example of an Answer Sheet Table (AST).

Student	Test item												
	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8	Q_9	Q_{10}	Q_{11}	Q_{12}	Q_{13}
S_1	1	1	1	0	0	0	0	0	1	1	0	1	0
S_2	0	0	0	1	1	0	0	1	0	0	0	0	0
S_3	0	0	0	0	0	0	0	0	0	1	0	0	1
S_4	1	1	1	0	0	0	0	1	0	1	0	1	0
S_5	0	0	0	0	0	0	0	0	1	0	0	0	0
S_6	0	0	0	0	0	1	0	1	1	0	0	0	0
S_7	0	0	0	1	1	0	0	0	0	0	1	0	1

By applying the Conditional Probability Approach, the concept-effect relationships in Figure 2 can be obtained.

If the failure ratio for student S_k to concept C_p (the parent concept of C_i) is unknown (i.e. none of the test items are relevant to C_p in the current test), the simple linear regression method is performed to find coefficients a and b of the least squares line $P(C_p) = a + b\,P(C_i)$, and hence, $P(C_p)$ can be estimated if $P(C_i)$ is known. Consider the concept-effect relationships in Figure 2, let A to F denote 'Coefficient', 'Exponent', 'Distributive law', 'Addition of polynomial', 'Division of polynomial' and 'Multiplication of polynomial', respectively. By applying a simple linear regression method to find least squares regression functions for A→D, B→E, B→F and C→D, we have A = −0.031 + 0.819D, B = 0.017 + 0.897E, B = 0.035 + 0.906F and C = 0.0021 + 1.131F.

If a student's failure ratio of 'Multiplication of polynomial' is greater than the given threshold (e.g. 0.4), the possible learning problem can be detected by taking the failure ratios of 'Exponent' and 'Distributive law' into consideration:

- If both the failure ratios of 'Exponent' and 'Distributive law' are higher than 0.4, the to-be-enhanced learning paths are 'Exponent' → 'Multiplication of polynomial' and 'Distributive law' → 'Multiplication of polynomial'.

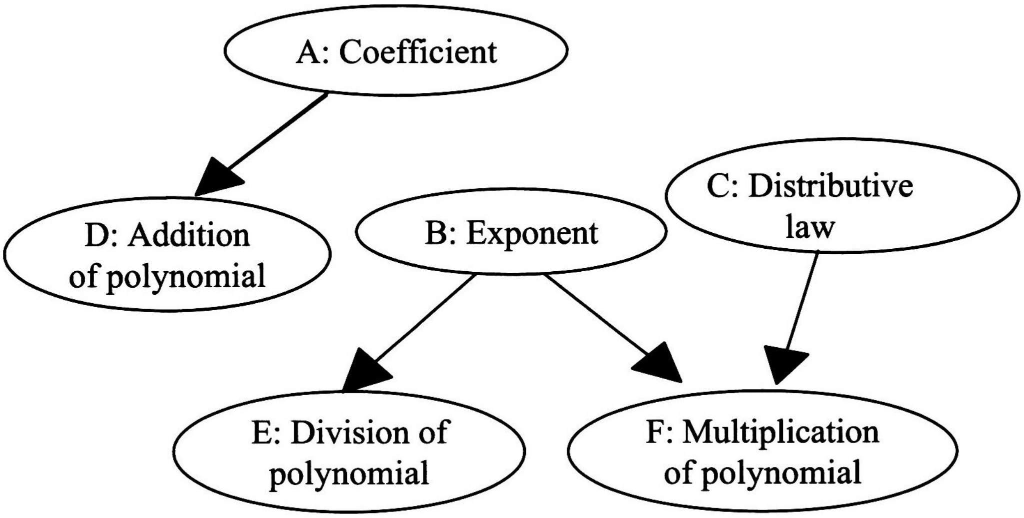

Figure 2. Illustrative example of derived concept-effect relationships.

- If both the failure ratios of 'Exponent' and 'Distributive law' are lower than or equal to 0.4, the only to-be-enhanced concept is 'Multiplication of polynomial'.
- If the failure ratio of 'Exponent' is higher than 0.4 and the failure ratio of 'Distributive law' is lower than or equal to 0.4, the only to-be-enhanced learning path is 'Exponent' → 'Multiplication of polynomial'.
- If the failure ratio of 'Exponent' is lower than or equal to 0.4 and the failure ratio of 'Distributive law' is higher than 0.4, the only to-be-enhanced learning path is 'Distributive law' → 'Multiplication of polynomial'.

In addition to those ideal cases shown above, it is possible that the failure ratios of part of the concepts are unknown, owing to the lack or insufficiency of relevant test items in the test. For example, it is possible that the student's failure ratio of 'Multiplication of polynomial' is known to be greater than the given threshold, while the failure ratios of 'Exponent' and 'Distributive law' are unknown. In this case, the failure ratios of 'Exponent' and 'Distributive law' can be predicted by applying the least squares regression functions. Assuming that the failure ratio of 'Multiplication of polynomial' is 0.43, the predicted failure ratios of 'Exponent' and 'Distributive law' are 0.48843 and 0.42458, respectively. As both the predicted failure ratios are greater than 0.4, the to-be-enhanced learning paths are therefore 'Exponent' → 'Multiplication of polynomial' and 'Distributive law' → 'Multiplication of polynomial'.

Moreover, it is possible that the test records may contain some answers with 'lucky guess' or 'careless fault'; therefore, the derived concept-effect relationships need to be examined by the teachers to eliminate incorrect relationships caused by such noisy data. That is, in implementing the testing and diagnosis system, it is necessary to provide a tool to assist the teachers in examining and modifying the derived concept-effect relationships.

System implementation

In 2000, the National Science Council of Taiwan funded the project of developing the ITED II (Intelligent Tutoring, Evaluation and Diagnosis system), which consists of three subsystems (see Figure 3): a Multi-Expert Tutoring Strategy Acquisition System, which supports multiple experts to work together for building subject materials and an item bank (Hwang 2002); a Personalised Tutoring System which analyses students' profiles and online records to generate personalised information for the tutoring module; and a Testing and Diagnosis System which diagnoses students' learning problems and gives personalised learning suggestions to the students by analysing their test results (Hwang 2003b).

ITED II has been implemented on the Windows NT environment using the Java language. In addition to the use of the Assessment Record Accumulation Model for deriving concept-effect relationships, it also provides a user interface for the teachers to examine and modify the derived concept-effect relationships. Moreover, a test interface is provided for the students to perform self-assessment, or participate in various group tests. To conduct a group test for a class of students, the teacher only needs to define some relevant parameters, such as the course identification, the class identification, the difficulty level of the test and the test time. The testing system will generate a test sheet accordingly. The students undergoing testing are requested to log on to computers located in the dedicated computer rooms, and during the test period their access rights are restricted to the test items only. Figure 4 displays an illustrative example of the group-test interface.

Following the submission of student answers to the test items, the system gathers the answers from clients and analyses the students' learning problems based on our novel approach. Based on the learning diagnosis results, learning guidance for individual students (see Figure 5) is generated,

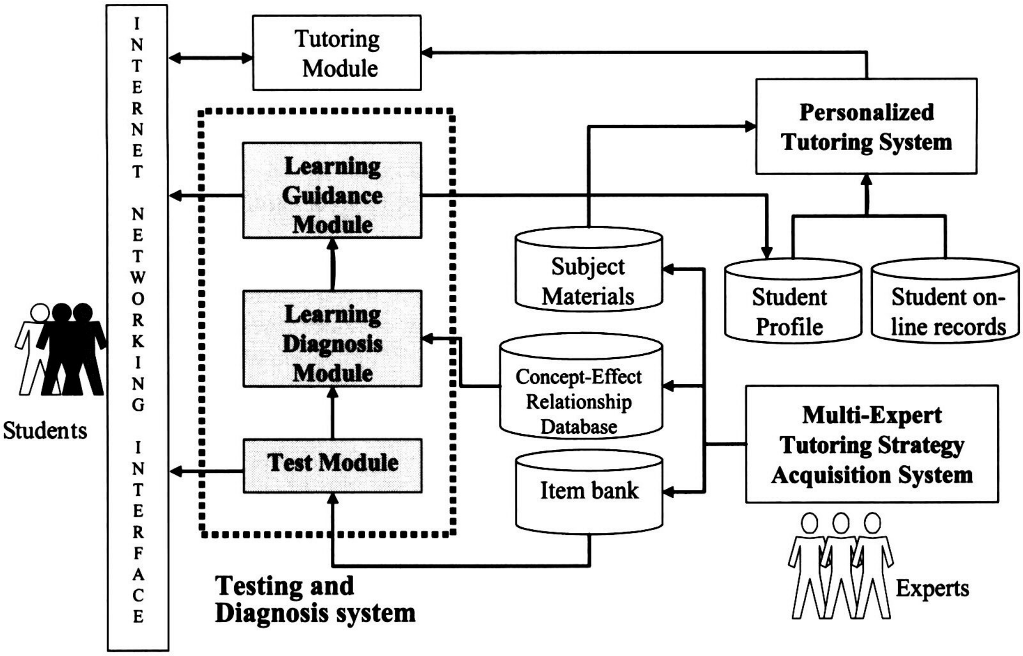

Figure 3. Structure of the ITED II.

as well as a summarised test report for the teacher. Moreover, those data are saved as a part of students' portfolios, which will be employed by the tutoring system to arrange personalised subject contents for each student.

Experiments and evaluation

To evaluate the efficiency of the presented approach, an experiment was conducted involving 76 junior high-school students enrolled in a Mathematics course. The experiment involved 40 concepts (as shown in Table 3), all of which were previously unknown to the students. The students were randomly separated into two groups, Group A (control group) and Group B (experimental group), each containing 70 students. The students in Group A (V1) received the learning guidance and relevant homework given by the teacher, while those in Group B (V2) received learning suggestions given by ITED II and relevant homework given by the teacher following each online test. Within a semester all of the students took two group tests (i.e. a pre-test and a post-test) and three self-assessments. The following presents the statistical results obtained from applying SPSS to analyse the group tests.

Pre-test

The pre-test aimed to ensure that both groups of students had the equivalent mathematics basis required for learning the course. Table 4 presents the t-test results of the pre-test. Notably, the means and standard deviations of the pre-test were 77.71 and 15.80 for V1 (control group), and 78.55 and 12.50 for V2 (experimental group). As the p-value (significant level) = 0.288 > 0.05 and $t = -0.258$, we can infer that in the pre-test, V1 and V2 do not significantly differ at a confidence interval of 95%. From the above, it was evident that the two groups of students have equivalent abilities in learning the Mathematics course.

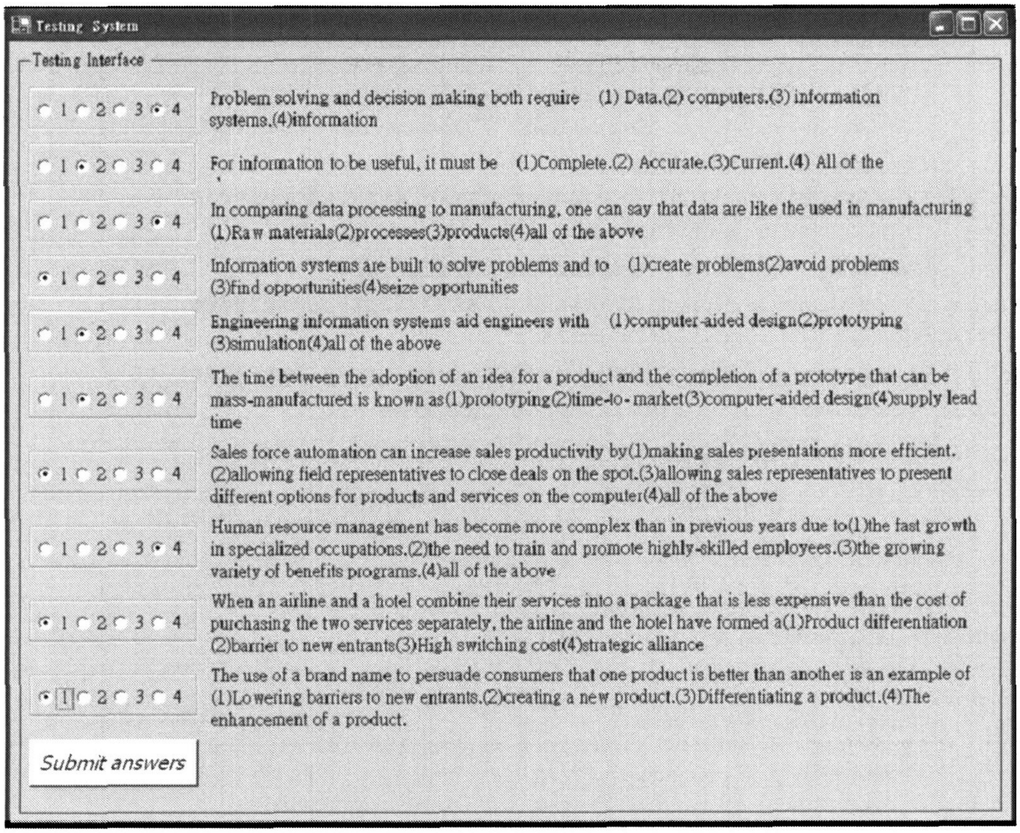

Figure 4. Illustrative example of the group-test interface.

Post-test

The post-test was intended to compare the learning achievements of the two groups of students after studying the Mathematics course. Table 5 lists the *t*-test values for the post-test results. Notably, the means and standard deviations of the post-test were 75.03 and 17.07 for V1, and 85.82 and 10.91 for V2. From the mean of the post-test, V2 at first observation achieves better performance than V1. As the p-value $= 0.015 < 0.05$ and $t = -3.284$, we can conclude that V2 achieved significantly better performance than V1 after implementing the subject approach.

Conclusions

This investigation presents a web-based intelligent testing and diagnostic system in a networked environment. A conditional Probability model is also proposed to diagnose poorly learned, partially learned and well-learned concepts, and to provide learning suggestions. An experiment on a Mathematics course was conducted to evaluate the performance of this approach. Based on the results of the pre-test, the two groups of students had equivalent basic knowledge before studying the Mathematics course. The students in both groups then received the same subject materials, an equivalent amount of homework, the same tests and were taught by the same teachers. The two groups differed only in terms of the source of learning suggestions. For the control group, the teacher gave some learning suggestions and assigned relevant homework to the students after each test. Meanwhile, for the experimental group, the learning suggestions were

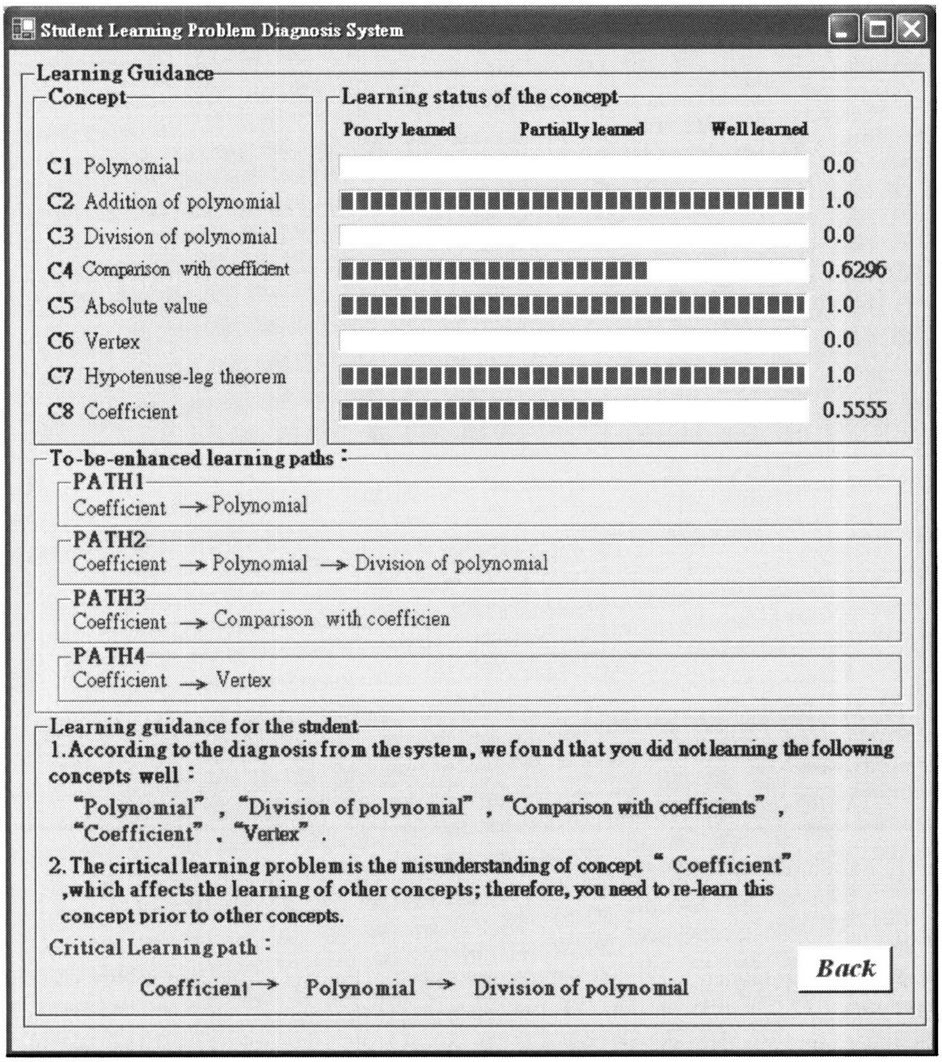

Figure 5. Illustrative example of computer-generated learning guidance.

given by applying the novel approach, and the teacher was asked to assign homework accordingly. From the results of the post-test, the students in the experimental group had progressed more significantly than the students in the control group, since more detailed and personalised suggestions were given by the system instead of the general suggestions given by the teacher.

Although the presented approach was adopted successfully in a Mathematics course, it could be applied to most Science courses. Our future studies will focus on applying the ITED II to several courses, including Physics and Chemistry courses in a junior high school, and Database Systems and Expert Systems courses in a university.

Acknowledgement

This study is supported by the National Science Council of the Republic of China under contract numbers NSC 95-2524-S-024-002 and NSC 95-2520-S-024-003.

Table 3. Concepts of junior high-school Mathematics course.

Concept number	Concept name	Concept number	Concept name
C1	Hypotenuse-leg theorem	C21	Elimination by substitution
C2	Calculate square	C22	Symbols of addition and subtraction
C3	Equivalence theorem	C23	Reduction of a fraction
C4	Polynomial	C24	Addition and subtraction between integer and root
C5	Coefficient	C25	Absolute value
C6	Exponent	C26	Comparison with coefficients
C7	Addition of polynomial	C27	Discriminant
C8	Multiplication of polynomial	C28	Axis of symmetry
C9	Division of polynomial	C29	Vertex
C10	Distributive law	C30	Standard expression
C11	Complete square of subtraction	C31	Convex or concave
C12	Squares difference	C32	Find the common factor
C13	Square root	C33	Extreme value of quadratic function
C14	Constant function	C34	Comparison with fraction and integer
C15	Linear function	C35	General expression
C16	Complete the square method	C36	Coordinates shift
C17	Complete square of sum	C37	Relationship between quadratic equation and x-axis
C18	The value of linear function	C38	Relationship between quadratic equation and y-axis
C19	Make a quadratic equation	C39	Solution of quadratic equation
C20	Symbol application	C40	Slope of quadratic function

Table 4. t-Test of the pre-test results (13 February 2003).

	N	Mean	SD	t	p
V1	38	77.71	15.80		
V2	38	78.55	12.50	−0.258	0.288

Table 5. t-Test of the post-test results (15 May 2003).

	N	Mean	SD	t	p
V1	38	75.03	17.07		
V2	38	85.82	10.91	−3.284	0.015

Notes on contributors

Gwo-Jen Hwang is currently Professor of the Department of Information and Learning Technology and Dean of the College of Science and Engineering at the National University of Tainan in Taiwan, Republic of China. His research interests include computer-assisted learning, computer-assisted testing, ubiquitous learning and expert systems. He has published over 200 academic papers in those fields.

Judy C. R. Tseng is currently Associate Professor of the Department of Computer Science and Information Engineering at Chung-Hua University in Hsinchu, Taiwan, Republic of China. She is also the chairman of that department. Her research interests include database systems, e-commerce, data mining and e-learning.

Gwo-Haur Hwang is currently Assistant Professor of the Information Management Department at Ling Tung University in Taichung, Taiwan, Republic of China. He is also the chairman of that department. His research interests include e-learning, e-commerce and medical expert systems.

References

Antao, B.A.A., A.J. Brodersen, J.R. Bourne, and J.R. Cantwell. 2000. Building intelligent tutorial systems for teaching simulation in engineering education. *IEEE Transactions on Education* 35(3), 222–25.

Chou, C. 2000. Constructing a computer-assisted testing and evaluation system on the world wide web – the CATES experience. *IEEE Transactions on Education* 43(3), 266–72.

Fan, J.P., K.M. Tina, and L.Y. Shue. 1996. Development of knowledge-based computer-assisted instruction system. Paper presented at the International Conference on Software Engineering: Education and Practice, Dunedin, New Zealand.

Giraffa, L.M.M., M.daC. Mora, and R.M. Viccari. 1999. Modeling an interactive ITS using a MAS approach: From design to pedagogical evaluation. Paper presented at the 3rd International Conference on Computational Intelligence and Multimedia Applications.

Gonzalez, A.V., and L.R. Ingraham. 1994. Automated exercise progression in simulation-based training. *IEEE Transactions on System, Man, Cybernetics* 24(6): 863–74.

Harp, S.A., T. Samad, and M. Villano. 1995. Modeling student knowledge with self-organizing feature maps. *IEEE Transactions on System, Man, Cybernetics* 25(5): 727–37.

Hwang, G.J. 1998. A tutoring strategy supporting system for distance learning on computer networks. *IEEE Transactions on Education* 41(4): 1–19.

———. 2002. On the development of a cooperative tutoring environment on computer networks. *IEEE Transactions on System, Man and Cybernetics: Part C* 32(3): 272–78.

———. 2003a. A concept map model for developing intelligent tutoring systems. *Computers & Education* 40(3): 217–35.

———. 2003b. A test sheet generating algorithm for multiple assessment requirements. *IEEE Transactions on Education* 46(3): 329–37.

Hwang, G.J., J.L. Hsiao, and Judy C.R. Tseng. 2003. A computer-assisted approach for diagnosing student learning problems in science courses. *Journal of Information Science and Engineering* 19(2): 229–48.

Olsen, J.B., D.D. Maynes, D. Slawson, and K. Ho. 1989. Comparisons of paper-administered, computer-administered, and computerized adaptive achievement tests. *Journal of Educational Computing Research* 5: 311–26.

Ozdemir, B., and F.N. Alpaslan. 2000. An intelligent tutoring system for student guidance in Web-based courses. Paper presented at the 4th International Conference on Knowledge-based Intelligent Engineering System and Allied Technologies.

Rasmussen, K., P. Northrup, and R. Lee. 1997. Implementing Web-based instruction. In *Web-based instruction,* edited by B.H. Khan. Englewood Cliffs, N.J.: Educational Technology.

Rowe, N.C., and T.P. Galvin. 1998. An authoring system for intelligent procedural-skill tutors. *IEEE Intelligent Systems* 13(3), 61–69.

Salisbury, D.F. 1998. Effect drill and practice strategies. In *Instructional designs for microcomputer courseware,* edited by David H. Jonassen. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Tseng, J.C.R., and G.J. Hwang. 2004. A novel approach to diagnosing student learning problems in e-learning environments. *WSEAS Transactions on Information Science and Applications* 1(5): 1295–300.

Vasandani, V., and T. Govindaraj. 1991. Intelligent diagnostic problem solving tutor: An experimental evaluation. *IEEE Transactions on Systems, Man and Cybernetics* 3: 1739–44.

———. 1995. Knowledge organization in intelligent tutoring systems for diagnostic problem solving in complex dynamic domains. *IEEE Transactions on Systems, Man and Cybernetics* 25(7), 1076–96.

Vasandani, V., T. Govindaraj, and C.M. Mitchell. 1989. An intelligent tutor for diagnostic problem solving in complex dynamic systems. Paper presented at the IEEE International Conference on Systems, Man and Cybernetics.

Wainer, H. 1990. *Computerized adaptive testing: A primer.* Hillsdale, N.J.: Lawrence Erlbaum Associates.