A Computer-Assisted Approach to Diagnosing Student Learning Problems in Science Courses*

GWO-JEN HWANG**, CHIA-LIN HSIAO+ AND JUDY C. R. TSENG++
**Information Management Department
National Chi Nan University
Puli, Nan-Tou, 545 Taiwan
E-mail: gjhwang@ncnu.edu.tw
+Information Management Department
Ling Tung College
Taichung, 400 Taiwan
E-mail: jialin@mail.ltc.edu.tw
++Department of Computer Science and Information Engineering
Chung Hua University
Hsinchu, 300 Taiwan
E-mail: judycrt@chu.edu.tw

The growing popularity of computer networks has led researchers to attempt to develop distance learning systems. However, students in network-based learning environments may need additional guidance and assistance when they encounter problems in learning certain concepts. Therefore, how to provide testing and diagnostic mechanisms in a computer-assisted learning environment is an important issue. This study presents a novel computer-assisted approach to diagnosing student learning problems in science courses. A network-based testing and diagnostic system is also implemented based on the proposed approach, which can diagnose science learning problems and offer students advice accordingly.

Keywords: learning diagnosis, concept effect relationship, computer-assisted testing, computer-assisted learning, science education

1. INTRODUCTION

In the last decade, researchers have exploited computer networks for educational applications, thus making the development of intelligent tutoring systems and learning environments an important issue in both computer science and education [2, 7, 11, 13, 15, 16].

During tutoring, the learning status of each student must be evaluated, and tests are a typical method for evaluation [14]. Computer techniques have been adopted for testing for over two decades. Early studies focused on whether computer-based tests were equivalent to paper-and-pencil tests, assuming that identical tests were administered in the two formats. Olsen et al. [7] compared the performance of paper-administered, computer-administered, and computer-adaptive tests in testing the mathematical abilities of

Received October 4, 2001; revised April 10, 2002; accepted June 14, 2002.
Communicated by Chuen-Tsai Sun.
* This study is supported by the National Science Council of the Republic of China under contract numbers NSC-90-2520-S-260-001 and NSC-90-2520-S-260-002. The original version of this paper was presented in the 5th Global Chinese Conference on Computers in Education, Taiwan, May 29-31, 2001.

third- and sixth-grade students. They found that paper-administered and computer-administered tests did not significantly differ in terms of testing quality. That study also identified the three types of tests to be equivalent in terms of score rank order, means, dispersions, and distribution shapes [7].

Due to the growing popularity of computer-assisted instruction (CAI), computer-based testing has received an increasing amount of attention. Jacob and Chase noted that computers can employ more varied methods of presentation than paper-and-pencil tests, including 3-D diagrams, motion effects, rotating geometric forms and so on [5]. Besides the traditional multiple-choice, fill in the blank, and short essay type questions, Rasmussen et al. [10] suggested that Web-based instruction could also allow students' progress to be evaluated through participation in group discussions and portfolio development. Furthermore, Khan [6] suggested that designers of Web-based instruction systems could create facilities to allow students to submit comments on courseware design and delivery. Finally, Chou proposed the CATES system, a collective and collaborative project intended to integrate an interactive testing system with theoretical and practical research on complex technology-dependent learning environments [1].

Although many researchers have identified testing and evaluation as important issues in computer-based instruction and have suggested design strategies and techniques, few systems have attempted to diagnose student learning problems. Most conventional testing systems assign a score or status indicator to each student after conducting a test, thus determining the learning status of that student, but without considering how to improve their learning status. In 1998, Hsu et al. [3] proposed a concept effect relationship (CER) model to demonstrate how the learning status of certain concepts can possibly be influenced by the learning status of other concepts. After the test results of students were analyzed based on concept effect relationships, the students were given guidance on areas needing improvement and on how to improve their learning status [3].

Although the concept effect relationship model appears desirable based on experimental results, its application is time-consuming for teachers unfamiliar with computer programming. Previous experience in applying the concept effect relationship model to tutoring has revealed that most teachers require assistance to define concept effect relationships [4]. To cope with this problem, this investigation proposes a computer-assisted approach to assisting teachers in defining and analyzing concept effect relationships. The following sections present this novel approach, along with practical experiences in applying it in several courses. Experimental results demonstrate that this novel approach can be used to diagnose student problems and help improve their learning status.

2. CONCEPT EFFECT RELATIONSHIP MODEL

To diagnose student learning problems, a concept effect relationship (CER) model has been proposed in [2]. The following subsections describe how this model finds the poorly-learned and well-learned concepts of a student, and how it generates learning guidance.

2.1 Concept Effect Relationships

Salisbury [12] indicated that learning information, including learning facts, names, labels, or paired associations, is often a prerequisite to achieving efficient performance in a more complex, higher level skill, especially in science courses. For example, the names and abbreviations of chemical elements and their atomic weights must be well learned in order to comprehend scientific writing or chemical formulas. That is, effectively learning a scientific concept normally requires first learning some basic concepts.

Consider two concepts or skills, say Ci and Cj. If Ci is a prerequisite to efficiently performing the more complex and higher level concept Cj, then a concept effect relationship $Ci \rightarrow Cj$ 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." Fig. 1 presents an illustrative example of the concept effect relationships among "positive integer," "negative integer," "addition," "subtraction," "multiplication" and "division." The effect relationships are important in diagnosing student learning problems. For example, if a student fails to answer most of the test items concerning "subtraction", the problem is likely that the student has not thoroughly learned "subtraction" or its perquisite concepts (such as "negative integer" or "addition"). Therefore, teachers can identify student learning problems by tracing the concept effect relationships.

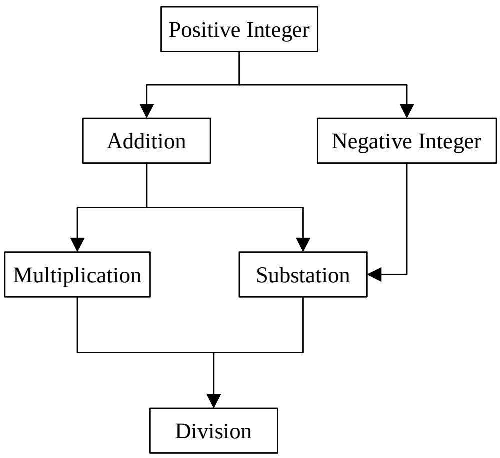

Fig. 1. Illustrative example of Concept Effect relationships.

2.2 Constructing a Concept Effect Relationship

Given a learning unit comprising 10 concepts (C_1, C_2, $C_3 \ldots C_{10}$) and a test sheet, containing 10 test items (Q_1, Q_2, $Q_3 \ldots Q_{10}$), a *test item relationship table* (TIRT) can be created as shown in Table 1. Each TIRT(Q_i ,C_j) entry represents the degree of association between test item Q_i and concept C_j. Meanwhile, the values of each entry range from 0 to L with 0 meaning no association and 1, 2… L representing an increasingly strong association, with a three-level (L = 3) or five-level (L = 5) rating scheme generally being employed. Table 1 shows a five-level (L = 5) rating scheme, where SUM(C_j) denotes the total strength of concept C_j on the test sheet. Furthermore, in Table 1, Q_1 and Q_6 are related to C_1; thus, SUM(C_1) = TIRT(Q_1, C_1) + TIRT(Q_6, C_1) = 5 + 1 = 6. ERROR(C_j) represents the total strength of the incorrect answers involving C_j, while ER(C_j) = ERROR(C_j)/ SUM(C_j) is the ratio of incorrect answers concerning concept C_j.

Assume that the student fails to correctly answer Q_3, Q_6 and Q_7; we have

ERROR(C_1) = TIRT(Q_3, C_1) + TIRT(Q_6, C_1) + TIRT(Q_7, C_1) = 0 + 1 + 0 = 1, and
ER(C_1) = ERROR(C_1)/ SUM(C_1) = 1/6 = 0.16.

Table 1. Illustrative example of a test item relationship table.

		C_j									
		C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}
Q_i	Q_1	5	1	0	0	0	0	0	0	0	0
	Q_2	0	4	2	0	0	0	0	0	0	0
	Q_3	0	0	3	1	2	0	0	0	0	0
	Q_4	0	0	0	5	0	0	0	0	0	0
	Q_5	0	0	0	0	5	0	0	0	0	0
	Q_6	1	0	0	0	0	4	1	0	0	0
	Q_7	0	0	0	0	0	0	5	0	0	0
	Q_8	0	0	0	0	0	0	0	3	1	2
	Q_9	0	0	0	0	0	1	0	0	4	0
	Q_{10}	0	0	0	0	0	1	0	2	0	5
	SUM	6	5	5	6	7	6	6	5	5	7
	ERROR	1	0	3	1	2	4	6	0	0	0
	ER(C_j)	0.16	0	0.6	0.16	0.28	0.66	1	0	0	0

Given that the teacher has already defined the concept effect relationships as shown in Fig. 2, each concept C_j in the Concept Effect relationship can be labeled with a corresponding ER(C_j) value.

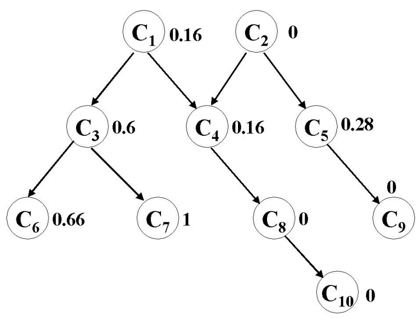

Fig. 2. Illustrative example of labeled Concept Effect relationships.

2.3 Finding the Enhanced Learning Paths

Diagnosing learning problems using CER requires identifying possible learning paths. For the example given in Fig. 2, five learning paths exist:

PATH1： $C_1 \rightarrow C_3 \rightarrow C_6$;
PATH2： $C_1 \rightarrow C_3 \rightarrow C_7$;
PATH3： $C_1 \rightarrow C_4 \rightarrow C_8 \rightarrow C_{10}$;
PATH4： $C_2 \rightarrow C_4 \rightarrow C_8 \rightarrow C_{10}$;
PATH5： $C_2 \rightarrow C_5 \rightarrow C_9$.

A threshold θ defines the acceptable error rate. If $ER(C_j) \leq \theta$, the student is said to have learned concept C_j; otherwise, the student has failed to learn the concept, and the concept is selected as a node of the *enhanced learning path*. Assume that θ is 0.3, and that $ER(C_3)$, $ER(C_6)$ and $ER(C_7)$ are greater than 0.3; then, the enhanced learning paths are as follows:

PATH1：$C_3 \rightarrow C_6$;
PATH2：$C_3 \rightarrow C_7$.

Clearly, the key problem in learning the subject unit is a lack of understanding of concepts C_3, C_6 and C_7, and the student should also learn concept C_3 before learning C_6 and C_7.

2.4 Problems With Employing the Concept Effect Relationship Model

A major problem encountered in applying the concept effect relationship model is eliciting concept effect relationships from experienced teachers or educational experts. To evaluate the performance of the concept effect relationship model, more than ten teachers were asked here to assist in defining concept effect relationships for certain science courses. However, few teachers were willing to participate in the experiment owing to the time consuming process of defining concept effect relationships. To cope with this problem, this study proposes a novel algorithm for finding concept effect relationships based on a set of testing results. Moreover, the following sections present a tool, called the CER Builder, that can help teachers define concept effect relationships.

3. CER BUILDER: A TOOL FOR FINDING CONCEPT EFFECT RELATIONSHIPS

While the Concept Effect relationship model appears useful, it is time-consuming for teachers to apply it unaided. For most teachers who are unfamiliar with computer programming, a computer system with a friendly user interface and intelligent analytic functions would be a more convenient source of aid. To cope with the above problems, the *CER Builder* is proposed (see Fig. 3). It comprises a CER Generator, a CER Analyzer and a CER Assistant.

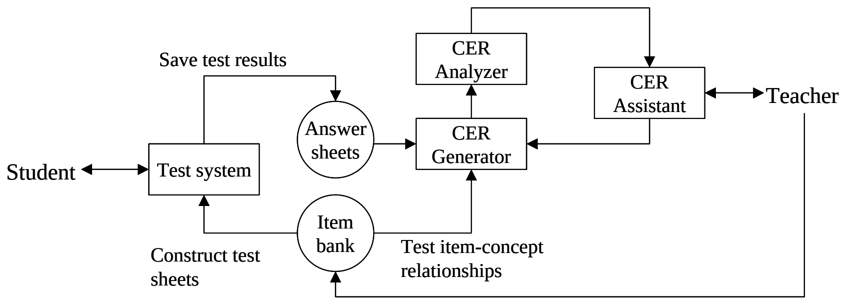

Fig. 3. Structure of the CER Builder.

Teachers can interact with the CER Builder via the CER Assistant, which is an interface for defining parameters and invoking the CER Generator for rapid prototyping and editing of concept effect relationships. Students can receive tests via the test system, and the test results are recorded in an answer-sheet database, which is later used by the CER Generator to generate possible concept effect relationships. The CER Analyzer is a long-term analysis system for identifying the relationships among the parameters used to generate concept effect relationships. The following subsections introduce each component in detail.

3.1 CER-Generator

The CER-Generator applies an algorithm to generate concept effect relationships from student test results. The algorithm is based on statistical prediction. It first finds the test item that most students failed to answer correctly, then finds the other test items that were incorrectly answered by those students, and finally uses this information to determine the relationships among the test items. The relationships among concepts can thus be determined based on the relationships among test items, and between test items and concepts.

An *Answer Sheet Summary Table* (ASST) and *Test Item Relationship Table* (TIRT) must be produced before applying the concept effect relationship-generating algorithm. An ASST is a collection of each student's answers for a test, and Table 2 presents an illustrative example. In the table, Si denotes the identity of an individual student and Q_i represents the test item number. If student Si correctly answered test item Q_i, we have ASST[Q_i, S_j] = 1; otherwise, ASST[Q_i, S_j] = 0. Entry Err_Count[Q_i] denotes the number of students who failed to correctly answer test item Qi.

Table 2. An illustrative example of an ASST.

	S_1	S_2	S_3	· · · · · · · ·	Err-Count[Q_j]
Q_1	1	0	1	· · · · · · · ·	
Q_2	0	0	0	· · · · · · · ·	
Q_3	1	0	0	· · · · · · · ·	
⋮	⋮	⋮	⋮	· · · · · · · ·	⋮

A TIRT records the relationships between each test item Qi and each concept Cj. Table 3 presents an example of an ITRT with a three-level rating scheme.

Table 3. Illustrative example of an ITRT.

	C_1	C_2	C_3	C_4
Q_1	3	0	1	0
Q_2	0	2	0	0
Q_3	0	0	0	3
Q_4	2	0	2	

Each test item Qi may have multiple related concepts, all with different relationship weights. A *Relationship Ratio Table* (RRT) is constructed to indicate the degree of association between concept Cj and test item Qi. Each RRT entry records the ratio of the concept in the total ratings against the test item. For example, $\sum$ITRT[Q_1, C_j] = 3 + 0 + 1 + 0 = 4; thus, RRT[Q_1, C_1] = 3/4 = 0.75 and RRT[Q_1, C_3] = 1/4 = 0.25, as listed in Table 4. An RPT[Q_i, C_j] value represents the possibility of a student failing to learn C_j if he (or she) answers Q_i incorrectly.

Table 4. Illustrative example of an RPT.

	C_1	C_2	C_3	C_4
Q_1	0.75	0	0.25	0
Q_2	0	1	0	0
Q_3	0	0	0	1
Q_4	0.5-	0	0.5	0

Symbols and Notations

The following symbols and notations are used in the concept effect relationship-constructing algorithm throughout this paper:

- *NS* Number of students tested
- E_{max} Set of test items that most students failed to answer correctly
- N_{max} Number of students who failed to correctly answer the test items in E_{max}
- Q_i i-th test item
- C_j j-th concept to be learned
- N_{Qi} Number of students who failed to correctly answer Qi
- $RC_{Qi}=\{C_1,C_2,C_3\ldots C_k\}$ Set of concepts that are related to Q_i
- $FAIL_{Qi}$ Set of students who fail to correctly answer Q_i
- $ES_{Qi}=\{Q_1,Q_2,Q_3\ldots Q\}$ Set of test items that the students in $FAIL_{Qi}$ failed to answer correctly
- $RRT[Q_i ,C_j]$ The ratio of concept C_j in the total ratings to test item Q_i
- n1 Number of accumulated test items
- n2 Number of new test items

• $R(C_i,C_j)$ Degree of certainty for $C_i \rightarrow C_j$ (if C_i is true then C_j is true). The value of $R(C_i,C_j)$ ranges from 0 to 1, indicating the relative degree to which "the learning status of C_i" influences "the learning status of C_j."
• $R(C_i,C_j)_{new}$ New degree of certainty for $C_i \rightarrow C_j$
• Support A threshold representing the proportion of students who failed to correctly answer test item Q_i. For $N_{Qi}/N >=$ support, a sufficient number of students is assumed to have failed to correctly answer Q_i such that Q_i is likely to be important in identifying concept effect relationships. For $N_{Qi}/N <$ support, since few students failed to answer Q_i, it may be difficult to find concept effect relationships using Q_i.
• Belief A threshold representing the lowest acceptable connection level for two concepts that students failed to correctly answer based upon the conditional probability. Assume that x% students who failed to correctly answer problems involving Concept X also failed to correctly answer problems involving Concept Y. If the defined *belief* is b% and x% ≥ b%, then the implication relationship "If one failed to answer X, then he/she might fail to answer Y" is accepted and recorded for future use.

Concept effect relationship generating algorithm

$step 1$: $E_{\max} = \{Q_{e1}, Q_{e2}, ...Q_{em}\}$
$step 2$: $N_{\max} = N_{Qei}$ for $Q_{ei} \in E_{\max}$
$step 3$: $while (N_{\max} / N \geq \text{support})$
{
  $for (i = 1; i \leq m; i++)$
  {
    Find RCQ_{ei} // * Set of concepts that are related to Q_{ei} * //
    Find $FAILQ_{ei}$ // * Set of students who fail to answer Q_{ei} * //
    Find ESQ_{ei} // * Set of test items which the ones in $FAILQ_{ei}$ fail to answer * //
    $while (ES_{Qei} \neq \Phi)$
    {
      $\exists Q_j \in ES_{Qei}$
      Find RC_{Qj} // * Set of concepts that are related to Qj * //
      $\forall C_k \in RC_{Qei}$
      $\forall C_l \in RC_{Qj}$
      $$R(C_l, C_k)_{new} = \frac{N_{Qj} * RRT(C_l, Q_j) * RRT(C_k, Q_{ei})}{RRT(C_l, Q_j) * N_{\max}}$$
      if $(R(C_l, C_k)_{new} > \text{belief})$ $then$
      $$R(C_l, C_k) = \frac{n_1 * R(C_l, C_k) + n_2 * R(C_l, C_k)_{new}}{n_1 + n_2}$$
      $ES_{Qei} = ES_{Qei} - \{Q_j\}$
    }
  }

```
    Remove the test records related to Qei
    Find a new Emax set
  }
```

In Step 1, E_{max}, the set of test items that most students failed to answer correctly, is constructed. In Step 2, for each test item in E_{max}, the number of students who failed to correctly answer that test item, say Q_{ei}, is counted. In Step 3, if the ratio of the number of students failing to correctly answer Q_{ei} equals or exceeds the support value, then the set of concepts related to Q_{ei} (i.e. RC_{Qei}), the set of students who failed to correctly answer Q_{ei} (i.e. $FAIL_{Qei}$) and the set of test items that the students in $FAIL_{Qei}$ failed to correctly answer are recorded for future use. The relationships among the concepts are then constructed based on the definition of $R(C_i, C_j)$.

3.2 CER-Assistant

The CER-Assistant provides a graphical user interface to help teachers adjust the parameters (i.e., *support* and *belief*) used in the CER-Generator and define the final concept effect relationships. If $N_{max}/N \geq$ support, then the concepts within that test item may be important in the diagnosis phase. The CER-Generator thus tries to find the relationships among the above concepts and others. A *Belief* value represents the lowest acceptable connection level for two concepts with related test items that were incorrectly answered by the students based upon conditional probability. For example, assume that 60% of students who fail to correctly answer "subtraction" problems also fail to correctly answer "division" problems; then, the connection level between "subtraction" and "division" is 0.6. Meanwhile, if the defined *belief* is 0.5, then the implication relationship "If one failed to answer subtraction problems, then he/she might fail to answer division problems" is recorded for further use in the learning diagnosis phase. Figs. 4 and 5 present the on-line interface of the CER-Assistant.

The teacher can use the CER-Assistant to efficiently define and analyze the concept effect relationships needed for diagnosing student learning problems.

3.3 CER-Analyzer

The *support* and *belief* values are clearly two thresholds for generating the Concept Effect relationships. If the thresholds approach 0, then too many noisy relationships may be generated, while if the thresholds approach 1, then some important relationships may be missed. The CER-Analyzer is used to determine the most appropriate *support* and *belief* values for generating concept effect relationships based on the outcome of previous applications. Each record in the database is composed of several attributes, including course name, the number of students, average score, score variation, the discrimination degrees of the test items, the difficulty degrees of the test items, belief value and support value. The CER-analyzer employs data mining techniques proposed in [9] to explore the relationships among these attributes, and the relationships identified are then employed to assist in determining belief and support values for future applications.

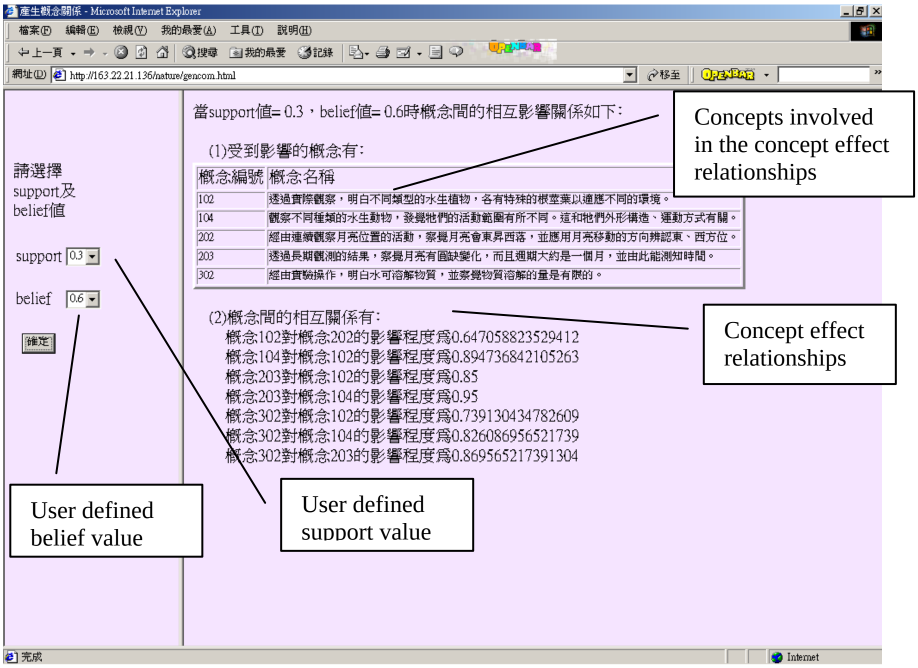

Fig. 4. Output of the CER-Assistant.

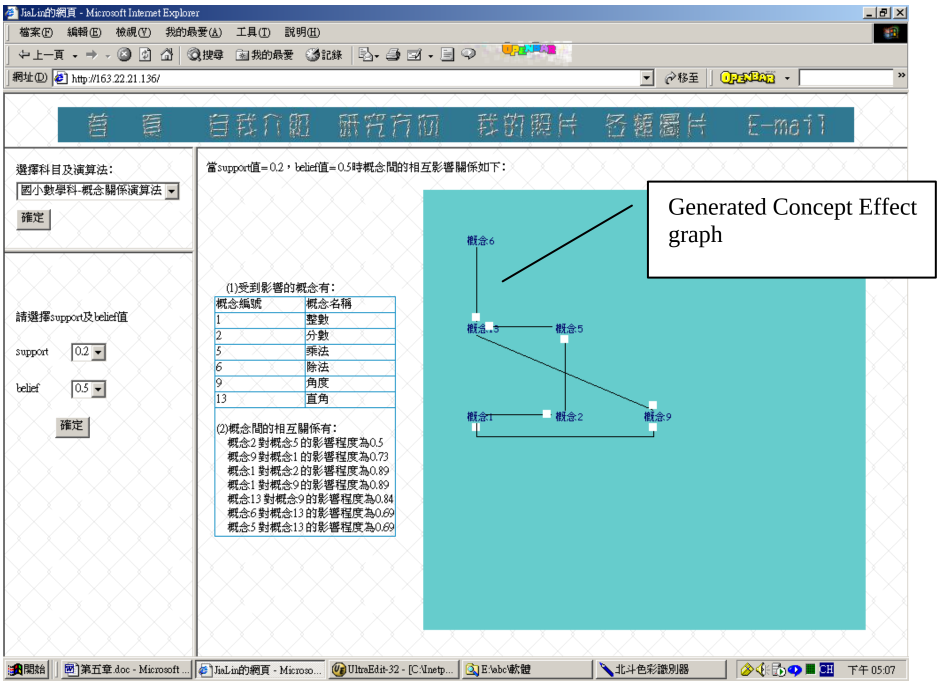

Fig. 5. Illustrative example of a constructed Concept Effect graph.

3.4 Testing and Diagnostic System

This study applies the proposed approach to implement a network-based testing and diagnostic system, which is composed of a student status database, item bank, Java-based interface and testing-and-diagnostic unit. The student status database contains general information on each student as well as his or her learning status. This database provides information necessary for the testing and diagnostic unit, which then constructs test sheets according to the learning status of each student. A tabular interface (see Fig. 6) is provided to help teachers define the relationships between individual concepts and test items. Following each test, the system identifies both the well-learned and poorly-learned concepts, and presents learning guidance (which is composed of a set of learning suggestions for each student based on the diagnostic results) as shown in Fig. 7.

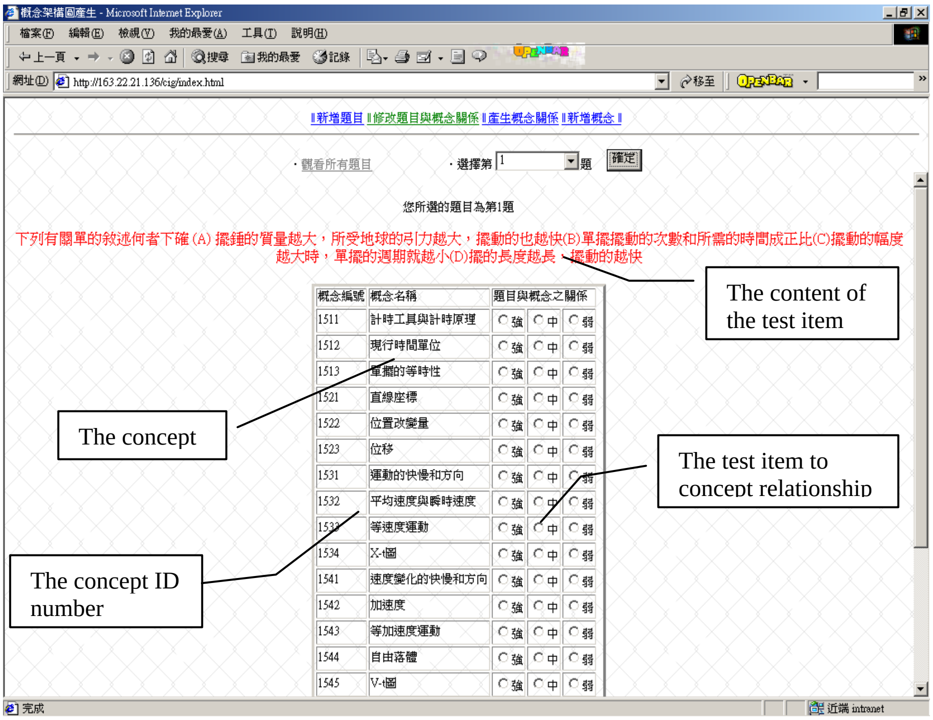

Fig. 6. User interface to construct the test item relationship table.

4. EXPERIENCES OF LEARNING DIAGNOSIS BASED ON CER APPROACH

This study applied the CER approach in several courses to evaluate its effectiveness. With the cooperation of several teachers, three sets of experiments were conducted to evaluate our novel approach. Table 5 lists the basic data of the three experiments in “Physics,” “Natural Science” and “Mathematics” courses, and the following subsections describe each experiment in detail.

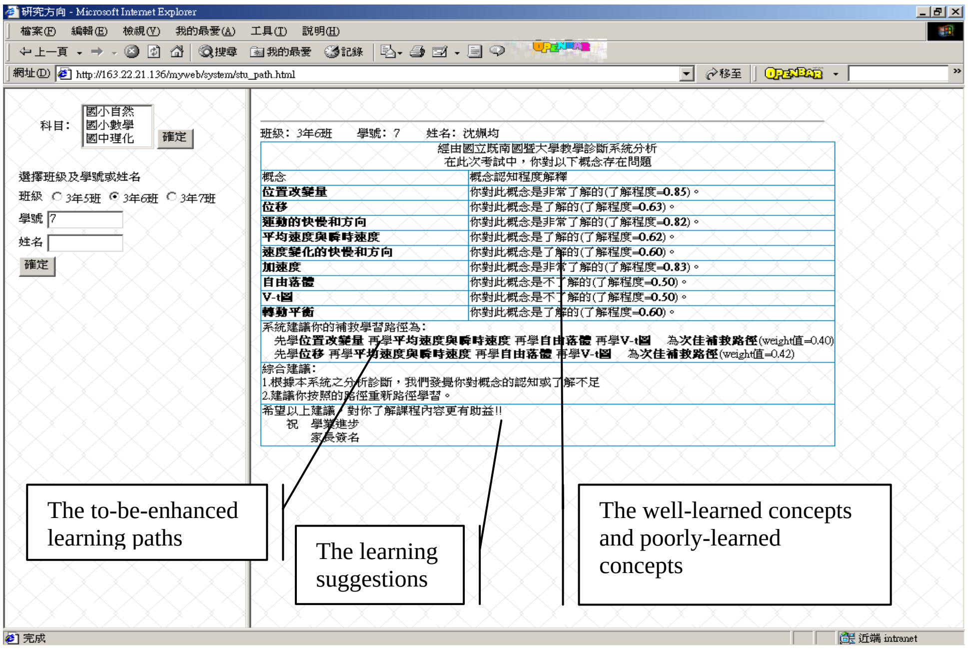

Fig. 7. Learning suggestions presented by the system.

Table 5. Statistics of three experiments.

Course	Physics	Natural Science	Mathematics
School & Grade	Junior High School K-9	Elementary School K-3	Elementary School K-4
Number of students	104	92	97
Average score	61.06	86.98	86.32
Standard deviation of scores	18.2	13.3	10.93
Average discrimination of the test items	0.395	0.187	0.222
Discrimination range of the test items	[0.08, 0.71]	[0.06, 0.32]	[0.03, 0.66]

4.1 Physics Course Experiment

The first experiment was based on a test in a "Physics" course at a junior high school from October 2000 to November 2000. One hundred and four K-9 students from two classes participated in the experiment, and their average test score was 61.06, while the average discrimination level of the test items was 0.395. Table 6 lists the concepts tested.

Table 6. Concept table of the Physics course.

Concept ID	Concepts to be learned	Concept ID	Concepts to be learned
1511	Tools and theories for timing	1542	Acceleration
1513	Equal time feature of swing	1543	Fixed acceleration
1522	Change of location	1544	Free fall
1523	Movements	1545	V-t diagram
1532	Average speed and instant speed	1612	Balance of forces
1534	X-t diagram	1621	Pushing position of a lever
1541	Change of speed and direction	1622	Balance of rotation

The teacher used the CER Builder to create a series of possible relationships among the concepts by adjusting the support and belief values (see Fig. 8). Fig. 9 displays the final concept effect graph with support = 0.2 and belief = 0.9.

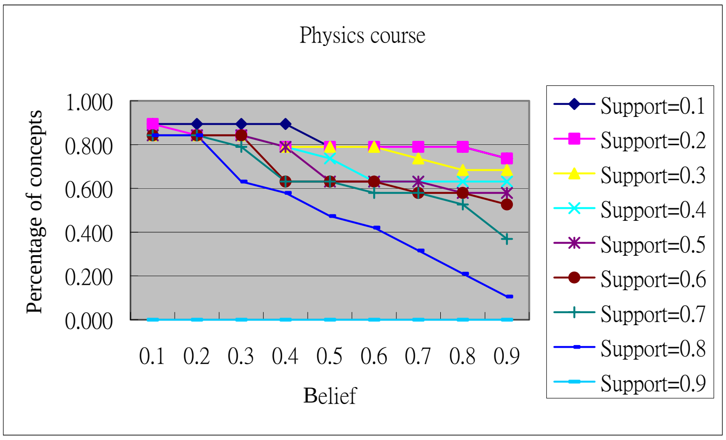

Fig. 8. Percentage of concepts involved in the concept effect relationships of the Physics course.

4.2 Natural Science Course Experiment

The second experiment was based on “Natural Science” tests administered at an elementary school from March 2000 to June 2000. Ninety two K-3 students from four classes participated in the experiment, and their average test score was 86.98, while the average discrimination level of the test items was 0.187. Table 7 lists the concepts tested.

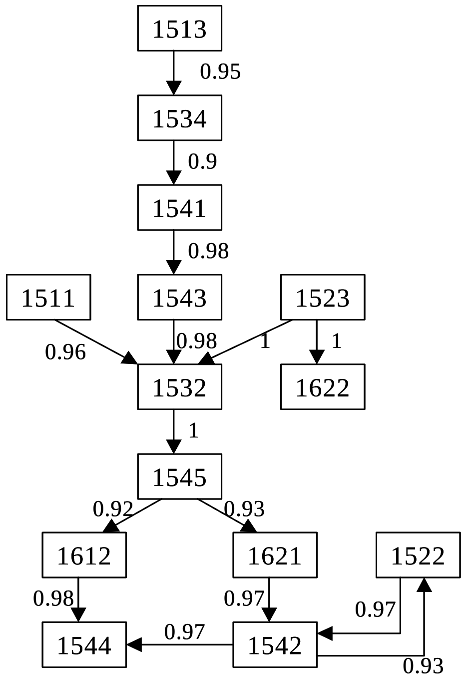

Fig. 9. Concept effect relationships of the Physics course.

Table 7. Concept table of Natural Science course.

Concept ID	Concepts to be learned
101	The locations of plants in water
102	The beam shapes of plants in water
103	How light affects the growth of plants
104	The active locations of animals in water
202	The change of the position of the moon
203	The relationship between time a period and the change of the position of the moon
302	Some objects can be melted in water
304	Different percentages of objects melted in water lead to different weights
305	The ability to observe and the operations in an experiment

Figs. 10 and 11 illustrate the concept inheritance relationships derived by applying the CER Builder.

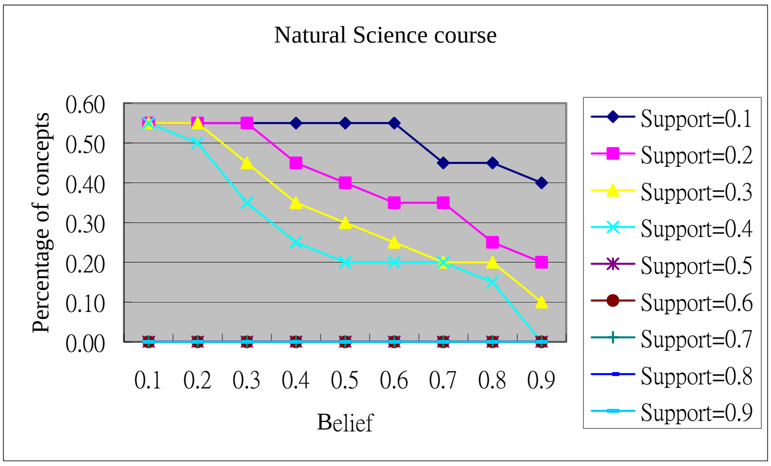

Fig. 10. Percentage of concepts involved in the concept effect relationships of the Natural Science course.

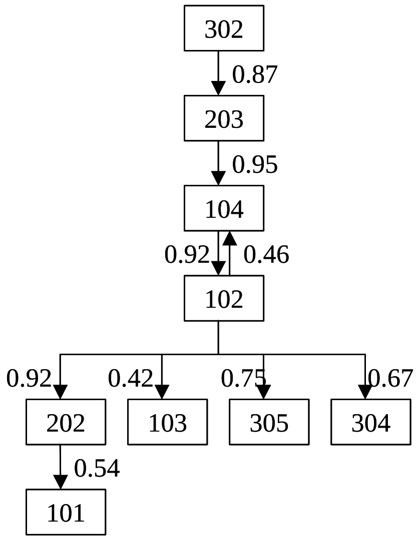

Fig. 11. Concept effect relationships of the Natural Science course with Support = 0.2 and Belief = 0.4.

4.3 An Experiment in a Mathematics Course

The third experiment was based on "Mathematics" tests administered at an elementary school from March 2000 to June 2000. Ninety-seven K-4 students from four classes participated in the experiment, and their average test score was 86.32, while the average discrimination level of the test items was 0.222. Figs. 12 and 13 illustrate the concept inheritance relationships derived using the CER Builder.

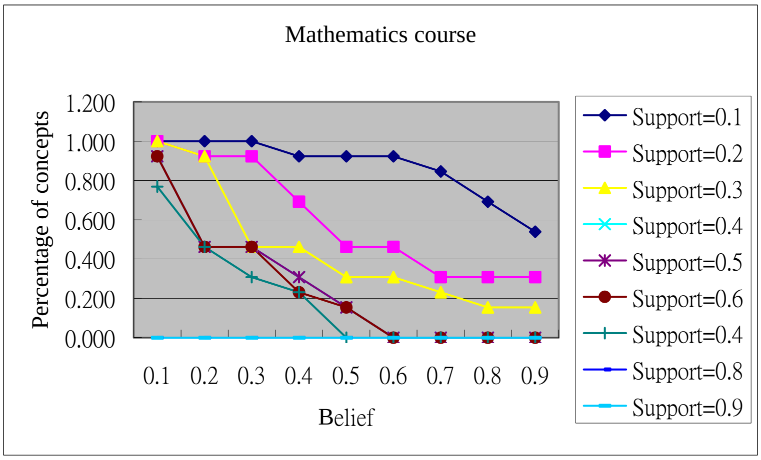

Fig. 12. Percentage of concepts involved in the concept effect relationships of the Mathematics course.

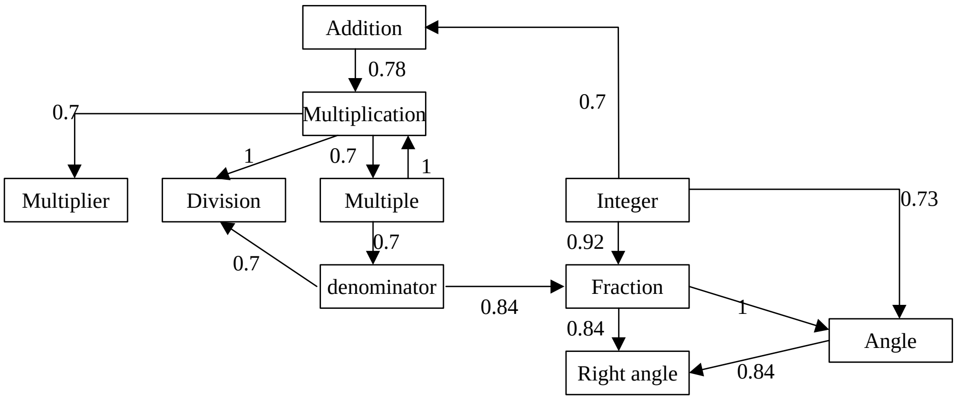

Fig. 13. Concept Effect relationships of Mathematics with Support = 0.1 and Belief = 0.7.

4.4 Evaluation of Student Attitudes

For each experiment, the learning diagnostic system analyzed the answer sheets and provided learning guidance to each student. Table 8 lists student feedback after the students received the learning guidance and clearly indicates that most of the students saw the learning guidance as valuable and were willing to continue receiving it in the future.

4.5 Evaluation of the Efficacy of the CER Builder

To evaluate the efficacy of our novel approach, an experiment in the Natural Science course at an elementary school was conducted from March 2001 to June 2001. Sixty K-3 students from two classes taught by the same teacher participated in the experiment and were separated into two groups, A (control group) and B (experimental group), each containing 30 students. The students in *Group-A* (V1) received regular on-line testing without learning guidance while those in *Group-B* (V2) received learning suggestions and relevant homework after each on-line test. All 60 students were given two tests within the space of one semester (including a pre-test and a post-test). The statistical results obtained by applying SPSS to analyze the two tests are presented below.

Table 8. Evaluation of student attitudes.

Course	Physics	Natural Science	Mathematics	Physics	Natural Science	Mathematics
Number of students	104	92	97	104	92	97
Questions / Answers	Is the learning guidance helpful?			Would you like to receive learning guidance in the future?		
Yes	88	49	83	92	79	96
Maybe	6	25	4	12	13	1
I don't know	10	18	10	0	0	0
Maybe not	0	0	0	0	0	0
No	0	0	0	0	0	0

• Pre-test

Table 9 lists the t-test values for the pre-test results. Since the p-value was 0.222, the t value of “Equal” variances was adopted. The p-value(Sig.2-tailed) = 0.024 > 0.01, implying that hypothesis H_0 can be accepted; that is, the performance of Groups A and B does not differ significantly in the pre-test. Therefore, we can conclude that in the pre-test, the mean score of Group A equaled that of Group B.

Table 9. t-test values of the pre-test results.

SCORE	t-test for Equality of Means						
	t	df	Sig. (2-tailed)	Mean Differ-ence	Std. Error Difference	99% Confidence Interval of the Difference	
						Lower	Upper
Variances equal	2.323	58	.024	4.00	1.72	-.59	8.59
Variances not equal	2.323	56.66	.024	4.00	1.72	-.59	8.59

• Post-test

Table 10 lists the t-test values for the post-test results. From the mean of the post-test, Group B performed better than Group A. Since the p-value was 0.022 (not significant), the t value of "Equal" variances was adopted. The p-value (Sig.2-tailed) = 0.009 < 0.01 and t = -2.719, implying that hypothesis H_0 should be rejected. Consequently, we can conclude that Group B performed significantly better than Group A because it benefited from the novel approach developed herein.

Table 10. t-test values of the post-test results.

SCORE	t-test for Equality of Means						
	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	99% Confidence Interval of the Difference	
						Lower	Upper
Variances equal	-2.719	58	.009	-9.57	3.52	-18.94	-.20
Variances not equal	-2.719	51.3	.009	-9.57	3.52	-18.98	-.15

5. CONCLUSIONS

This study has proposed a computer-assisted approach to assisting teachers in defining and analyzing concept effect relationships, thus helping them to diagnose student learning problems. Furthermore, a testing and diagnostic system has been implemented based on the novel approach. The system provides learning suggestions for each student by analyzing answer sheets and the relationships between subject concepts and test items. Experimental results indicate that our novel approach helps students making progress in learning. The authors are currently conducting further experiments on the Internet to accumulate more experience in diagnosing student learning problems, and to find possible relationships among the parameters used to describe test items and concepts.

ACKNOWLEDGEMENT

The authors would like to thank the National Science Council of the Republic of China for financially supporting this research under contract numbers NSC-90-2520-S-260-001 and NSC-90-2520-S-260-002 as well as Rei-Hsin Pon, Chin-I Hwei and Bau-Wen Hsu for their help in conducting the experiments.

REFERENCES

1. C. Chou, "Constructing a computer-assisted testing and evaluation system on the world wide web-the CATES experience," *IEEE Transactions on Education*, Vol. 43,

2000, pp. 266-272.
2. L. M. M. Giraffa, M. C. Mora, and R. M. Viccari, “Modeling an interactive ITS using a MAS approach: from design to pedagogical evaluation,” *Third International Conference on Computational Intelligence and Multimedia Applications*, 1999, pp. 153-158.
3. C. S. Hsu, S. F. Tu, and G. J. Hwang, “A concept inheritance method for learning diagnosis of a network-based testing and evaluation system,” *The 7th International Conference on Computer-Assisted Instructions*, 1998, pp. 602-609.
4. C. L. Hsiao, G. J. Hwang, and J. C. R. Tseng, “An interactive concept relationship construction assisted system for learning diagnosis,” *The 5th Global Chinese Conference on Computers in Education*, 2001, pp. 925-932.
5. L. C. Jacobs and C. I. Chase, *Developing and Using Tests Effectively: A Guide for Faculty*, San Francisco, CA: Jossey-Bass, pp. 168-177.
6. B. H. Khan, “Web-based instruction (WBI): What is it and why is it?” in *Web-Based Instruction*, B. H. Khan, ed., Englewood Cliffs, NJ, Educational Technology, 1997, pp. 5-18.
7. S. Labidi and J. S. Ferreira, “Technology-assisted instruction applied to cooperative learning: The SHIECC project,” *The 28th Annual Frontiers in Education Conference*, Vol. 1, 1999, pp. 286-291.
8. J. B. Olsen, D. D. Maynes, D. Slawson, and K. Ho, “Comparison and equating of paper-administered, computer-administered and computerized adaptive tests of achievement,” *The Annual Meeting of American Educational Research Association*, 1986.
9. J. S. Park, M. S. Chen, and P. S. Yu, “Using a hash-based method with transaction trimming and database scan reduction for mining association rules,” *IEEE Transactions on Knowledge and Data Engineering*, Vol. 9, 1997, pp. 813-825.
10. K. Rasmussen, P. Northrup, and R. Lee, “Implementing web-based instruction,” in *Web-Based Instruction*, B. H. Khan, ed., Englewood Cliffs, NJ, Educational Technology, 1997, pp. 341-346.
11. M. Rosic, S. Slavomir, and V. Glavinic, “Intelligent tutoring systems for asynchronous distance education,” *The 10th Mediterranean Electrotechnical Conference*, 2000, pp. 111-114.
12. D. F. Salisbury, “Effect drill and practice strategies,” in *Instructional Designs for Microcomputer Courseware*, D. H. Jonassen, ed., Lawrence Erlbaum Associates, 1998, pp. 103-124.
13. C. T. Sun and C. Chou, “Experiencing CORAL: design and implementation of distance cooperative learning,” *IEEE Transactions on Education*, Vol. 39, 1996, pp. 357-366.
14. H. Wainer, *Computerized Adaptive Testing − A Primer*, Hillsdale, NJ, Lawrence Erlbaum Associates.
15. L. H. Wong, C. Quek, and C. K. Looi, “TAP: a software architecture for an inquiry dialogue-based tutoring system,” *IEEE Transactions on Systems, Man and Cybernetics*, Part A, Vol. 28, 1998, pp. 315-325.
16. A. Yoshikawa, M. Shintani, and Y. Ohba, “Intelligent tutoring system for electric circuit exercising,” *IEEE Transactions on Education*, Vol. 35, 2000, pp. 222-225.

Gwo-Jen Hwang (黃國禎) was born on April 16, 1963, in Taiwan, Republic of China. In 1991, he received his Ph.D. degree from the Department of Computer Science and Information Engineering at National Chiao Tung University in Taiwan. He is now an associate professor of National Chi-Nan University. His research interests include computer-assisted learning, computer-assisted testing, expert systems and mobile computing.

Chia-Lin Hsiao (蕭嘉琳) was born on October 13, 1977, in Taiwan, Republic of China. In 2001, she received her Master degree from the Department of Information Management at National Chi-Nan University in Taiwan. She is now a Lecturer of Ling-Tung College. Her research interests include computer -assisted instruction, data mining, e-commerce and e-learning.

Judy C. R. Tseng (曾秋蓉) was born on August 21, 1965, in Taiwan, Republic of China. In 1992, she received her Ph.D. degree from the Department of Computer Science and Information Engineering at National Chiao Tung University in Taiwan. She is now an associate professor of Chung-Hua University. Her research interests include database systems, e-commerce, mobile computing and e-learning.