

A group decision approach to developing concept–effect models for diagnosing student learning problems in mathematics

Gwo-Jen Hwang, Patcharin Panjaburee, Wannapong Triampo and Bo-Ying Shih

Gwo-Jen Hwang is a chair professor in the Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan. His research interests are mobile and ubiquitous learning, digital game-based learning, computer-assisted testing and learning diagnosis, and artificial intelligence in education. Patcharin Panjaburee is a lecturer of Science and Technology Education in the Institute for Innovative Learning at Mahidol University, Thailand. Her main research interests are mathematics education, computer science education, and computer-assisted testing and learning diagnosis. Wannapong Triampo is an associate professor in the Department of Physics at Mahidol University, Thailand. He is a director of the Institute of Innovative Learning at Mahidol University. His main research interests are biophysical, computational and theoretical modeling of biological and environmental systems, mathematics education, and computer-assisted testing and learning diagnosis. Bo-Ying Shih is a graduate student in the Department of Information and Learning Technology, National University of Tainan, Taiwan. His research interests are computer-assisted testing and learning diagnosis, mathematics education and web-based learning. Address for correspondence: Dr Gwo-Jen Hwang, Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, 43, Sec.4, Keelung Road, Taipei 106, Taiwan. Email: gjhwang.academic@gmail.com

Abstract

Diagnosing student learning barriers has been recognized as the most fundamental and important issue for improving the learning achievements of students. In the past decade, several learning diagnosis approaches have been proposed based on the concept–effect relationship (CER) model. However, past studies have shown that the effectiveness of this model heavily depends on the concept relationship knowledge provided by the domain experts (eg, experienced teachers or educators for a specified subject); ie, the performance of the developed learning diagnosis systems could be significantly affected by subjective opinions, ignorance or insufficient knowledge if those concept relationships are derived from a single domain expert. To cope with this problem, this study proposes a group decision approach for developing the CER model with the cooperation of multiple domain experts. Based on the proposed approach, a testing and diagnostic system has been implemented; moreover, an experiment has been conducted to evaluate the effectiveness of this new approach. The experimental results show that this approach is able to develop quality CER models, and hence the low-achievement students who received the generated learning suggestions had significantly better learning achievements than those who learned with the previous approach.

Introduction

Adaptive learning systems provide a learning facility that maintains the appropriate context to accommodate a diversity of student personalization (Magoulas, Papanikolaou & Grigoriadou, 2003; Wang & Liao, 2011). While developing an adaptive learning system, researchers need to consider the provision of meaningful and personalized feedback to or learning support for students, which has been recognized as one of the most important issues for assisting students in improving their learning achievements (Barbera, 2009; Draper, 2009; Hsu, Hwang & Chang,

Practitioner notes

What is already known about this topic

- Diagnosing student learning barriers is a fundamental and important issue for improving the learning achievements of students.
- Several learning diagnosis approaches have been proposed based on the concept–effect relationship (CER) model.
- The effectiveness of the learning diagnosis model heavily depends on the concept relationship knowledge provided by the domain experts.

What this paper adds

- A group decision approach for developing the CER model with the cooperation of multiple domain experts is proposed.
- A testing and diagnostic system was implemented based on the proposed approach.
- An experiment was conducted to investigate whether the students guided by the proposed approach had better learning achievements than those guided by the original CER model approach.

Implications for practice and/or policy

- The experimental results show that the low-achievement students who received the learning suggestions generated by the new approach had significantly better learning achievements than those who learned with the previous approach.
- Further studies of developing more effective tools or leading in other group decision strategies to help domain experts cooperatively determine the prerequisite relationships are needed.

2010; Jordan & Mitchell, 2009; Yeh, Chen, Hung & Hwang, 2010). In the past decades, many studies have been conducted to develop an effective model for analyzing the learning barriers of students such that helpful learning suggestions or guidance can be provided based on the analysis results (Chen & Bai, 2009; Hwang, 2003, 2007). In the meantime, researchers have developed various computer-assisted testing and diagnostic systems for diagnosing students' learning problems and providing appropriate learning guidance for individual students on the Internet (Casamayor, Amandi & Campo, 2009; Chen & Bai, 2009; Chiou, Hwang & Tseng, 2009; Hwang, 2003; Sieber, 2009). For example, Chen (2008) developed a genetic-based personalized learning system, in which a genetic algorithm was employed to generate appropriate learning paths based on the incorrect answers given by individual learners. However, such an approach ignores the relationship between the prior and subsequent knowledge while planning the personalized learning paths (Chen, 2010; Hwang, 2003; Jong, Lin, Wu & Chan, 2004; Lee, Lee & Leu, 2009).

The concept–effect relationship (CER) model is a concept map-oriented method for diagnosing student learning problems based on the prerequisite knowledge structure (ie, the CER between concepts) defined by domain experts. With the CER model, the adaptive learning systems are able to generate learning paths for individual students for improving their learning achievements (Hwang, 2003). This model has been shown to be effective in analyzing student learning problems and providing them with personalized suggestions by several researchers in various applications, including natural science, mathematics, physics and computer courses (Chu, Hwang & Huang, 2010; Hwang, 2007; Panjaburee, Hwang, Triampo & Shih, 2010; Tseng, Sue, Su, Weng & Tsai, 2007). Moreover, several studies have demonstrated the effectiveness of the learning

diagnosis model via using the diagnosis results to determine personalized learning paths or learning content (Jong *et al.*, 2004; Kwasnicka, Szul, Markowska-Kaczmar & Myszkowski, 2008; Tseng *et al.*, 2008).

Although the CER model has been shown to be beneficial in helping students improve their learning achievement, past experiences have also revealed a critical problem when applying it (Jong *et al.*, 2004; Hwang, Tseng & Hwang, 2008; Lee *et al.*, 2009); ie, the effectiveness of this model heavily depends upon the concept relationship knowledge provided by the domain experts. Therefore, subjective opinions, ignorance or insufficient knowledge could significantly affect the performance of the developed learning diagnosis systems if the concept relationship knowledge is obtained from a single domain expert.

Researchers have indicated that domain experts are likely to have different expertise or understandings for solving problems in the same domain owing to the working environments in which they have been situated, the cases they have experienced or knowledge they have constructed (Chu & Hwang, 2008; Hwang, Chen, Hwang & Chu, 2006; Léger & Naud, 2009; Panjaburee *et al.*, 2010). That is, different qualities of concept relationships might be provided by different domain experts, which could significantly affect the students' learning diagnosis results. This also implies that the quality of the learning suggestions given to the students could be unstable. That is, it remains a research issue to improve the reliability of using the CER model for developing adaptive learning systems.

In this paper, a group decision approach is proposed. With this new approach, the concept relationship knowledge is elicited and integrated from multiple domain experts such that more reliable and accurate learning suggestions can be given to the students. Moreover, a testing and diagnostic system has been implemented based on this innovative approach.

Literature review

The popularity of computer and communication technology has attracted researchers from various fields to develop or utilize computer-based assessment methods or tools. For example, Yin, Chang, Hwang, Hwang and Chan (2006) developed a computer-assisted testing system that enabled teachers to compose serial test sheets based on multiple assessment criteria; Marriott (2009) proposed an online summative assessment approach for evaluating the learning performance of students in an undergraduate financial accounting course; Cooner (2010) employed a formative evaluation strategy in a technology-enhanced blended learning environment to engage students in developing reflective skills; and Li, Liu and Steckelberg (2010) further used a peer-assessment strategy to investigate how it affected the quality of student projects in a technology application course.

Among the existing educational tools, concept maps have been recognized as being an important tool for assisting students to perform higher-order thinking (Chiou, 2008; Hwang, Shih & Chu, 2011); moreover, they have been used to evaluate the cognitive degree of the concepts (ie, the degree of understanding the meanings of and the relationships between the concepts) for individual students (Liu, Don & Tsai, 2005). Researchers have indicated the effectiveness of representing knowledge as concept relationships in learning processes; therefore, many cognitive learning tools or models have been proposed in the past decades (Peng, Su, Chou & Tsai, 2009; Roth & Roychoudhury, 1994). For example, the Cognitive Tutor is a computer-assisted learning system that interprets student problem-solving behaviors using a cognitive model in the form of production rules (Alevan & Koedinger, 2002).

The CER model is a concept map-oriented method that provides a systematic procedure for diagnosing students' learning problems and generating personalized learning guidance based on the assumption that prerequisite relationships exist between the concepts to be learned (Hwang,

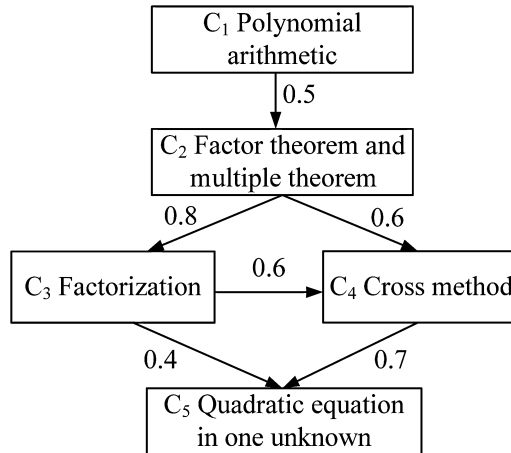


Figure 1: Illustrative example of the concept-effect relationship model

2003). Such relationships can be found in most science, engineering or mathematics courses, in which learning information, including facts, names, labels, concepts or paired associations, is often a prerequisite to efficiently learn more complex, higher-level concepts or skills. That is, the learning status of one concept could significantly affect the learning status of another if such a prerequisite relationship exists (Hwang, 2003). In this study, the term “concept” represents a notion of an idea that “responds to some class of entities and the features of the class” or “expresses how something can be accomplished.”

In the CER model, if concept C_i is a prerequisite to efficiently learn the more complex and higher-level concept C_j , a CER $C_i \rightarrow C_j$ is said to exist. Moreover, the domain expert needs to determine the corresponding weighting value, a real number ranging from 0 to 1, for representing the strength with which the learning status of the parent concept will affect that of the child concept. For example, to learn the concept “factorization,” one might need to learn “factor theorem and multiple theorem” first; before learning the concept “quadratic equation in one unknown,” one might need to learn “factorization” and “cross method.” Such learning sequence relationships can be represented as a CER model as shown in Figure 1.

Let weight value W_{C_i, C_j} represent the strength of the relationship $C_i \rightarrow C_j$. In Figure 1, $W_{C_1, C_2} = 0.5$, implying that the strength of the relationship $C_1 \rightarrow C_2$ is 0.5. When the relationships between the concepts are identified and the students’ learning status for each concept is known, a remedial learning path, which represents the suggested learning sequence of the concepts that the students have learned poorly, can be generated for individual students based on the following procedure:

Step 1: Calculating the error ratio (ER) of each concept for individual students. Considering the test items listed in Table 1, $ER(C_j)$ is the ratio of the weighting values of the C_j -related test items that the students failed to correctly answer; ie, $ER(C_j) = \text{ERROR}(C_j)/\text{SUM}(C_j)$, where $\text{ERROR}(C_j)$ is the sum of the weighting values of C_j -related test items that the students failed to correctly answer, and $\text{SUM}(C_j)$ is the sum of the weighting values of all of the C_j -related test items. For example, assuming that a student failed to correctly answer Q_1 , Q_4 and Q_{10} , we have $ER(C_1) = (0 + 1 + 0)/5 = 0.2$ and $ER(C_2) = (0 + 2 + 0)/5 = 0.4$.

Step 2: Determining the thresholds (θ_1 and θ_2) for categorizing the learning levels of the students. $ER(C_j) \leq \theta_1$ means that the students have learned concept C_j well; $\theta_1 < ER(C_j) \leq \theta_2$ means that the students have partially learned concept C_j ; otherwise, the students are said to have learned concept C_j poorly. Usually, the thresholds are given by the teachers; alternatively, they can be

Table 1: Illustrative example of the ER of each concept

		Concept C_j				
		C_1	C_2	C_3	C_4	C_5
Test items	Q_1	0	0	2	1	0
	Q_2	3	0	0	0	0
	Q_3	0	0	0	0	1
	Q_4	1	2	0	0	5
	Q_5	0	0	0	5	0
	Q_6	0	0	1	0	0
	Q_7	1	0	0	0	0
	Q_8	0	3	0	0	0
	Q_9	0	0	2	0	0
	Q_{10}	0	0	0	2	0
SUM(C_j)		5	5	5	7	6
ERROR(C_j)		1	2	2	3	5
ER(C_j)		0.2	0.4	0.4	0.43	0.83

determined by calculating the average ERs of the high/medium-achievement students and the medium/low-achievement students respectively (Hwang, 2003).

Step 3: Establishing the remedial learning paths for individual students based on their ER for each concept and the relationships between the concepts. The remedial learning paths are constructed from those partially learned or poorly learned concepts without any child concept (ie, the most difficult or complex concepts). For a poorly learned concept C_j , the parent concept of C_j , say C_p , will be added to the remedial learning path if the $C_p \rightarrow C_j$ relationship exists. If there is more than one parent concept C_p of C_j , for those C_p s with the strongest $C_p \rightarrow C_j$ relationships, different remedial learning paths will be generated by adding each parent concept to the corresponding learning path. The construction process is performed repeatedly until all of the remedial learning paths include the most basic concepts (ie, the concept without any parent concept).

Step 4: Removing the well-learned concepts from the generated remedial learning paths.

Consider the illustrative example given in Table 1. Assume that $\theta_1 = 0.3$ and $\theta_2 = 0.5$, concept C_5 with $ER(C_5) = 0.83$ (ie, a poorly learned concept without any child concepts) is selected as the starting point for constructing the remedial learning paths. The parent concepts with the strongest prerequisite relationships to C_5 are then added to the remedial learning paths until the most basic concepts are all included. From Figure 1, the prerequisite relationships of C_3 and C_4 to C_5 are 0.4 and 0.7 respectively; therefore, C_4 is added to the remedial learning path. The same process is then performed on concept C_4 . As the parent concepts of C_4 (ie, C_2 and C_3) have the same prerequisite relationship (ie, 0.7), two remedial learning paths are constructed by adding C_2 and C_3 to each path as follows:

$$\text{PATH1: } C_2 \rightarrow C_4 \rightarrow C_5$$

$$\text{PATH2: } C_3 \rightarrow C_4 \rightarrow C_5$$

Repeatedly, by checking the parent concepts of C_2 and C_3 , we have the following remedial learning paths:

$$\text{PATH}_1: C_1 \rightarrow C_2 \rightarrow C_4 \rightarrow C_5$$

$$\text{PATH}_2: C_1 \rightarrow C_2 \rightarrow C_3 \rightarrow C_4 \rightarrow C_5$$

As C_1 is the most basic concept in Figure 1, the construction process stops. From Table 1, it is found that $ER(C_1) = 0.2$, which is less than $\theta_1 = 0.3$; ie, concept C_1 is a well-learned concept. Consequently, C_1 is removed from the remedial learning paths:

$$\text{PATH1: } C_2 \rightarrow C_4 \rightarrow C_5$$

$$\text{PATH2: } C_2 \rightarrow C_3 \rightarrow C_4 \rightarrow C_5$$

It can be seen that the student has difficulty in learning concepts C_2, C_3, C_4 and C_5 ; furthermore, C_2 is the most basic concept for both learning paths. Therefore, the learning system will suggest that the student relearn those concepts in the order of $C_2 \rightarrow C_3 \rightarrow C_4 \rightarrow C_5$.

In the past decades, various applications of employing the CER model to identify student learning problems and give them personalized learning suggestions have produced positive results. For example, Jong, Chan and Wu (2007) developed a learning behavior diagnosis system based on the CER model. Their experimental results on a computer course showed that the approach was able to improve the learning achievements of the students. In the same year, Tseng *et al* (2007) employed the CER model in the Physics course of a junior high school and derived satisfactory results. Later, Hwang *et al* (2008) reported the effectiveness of using the CER model in the Mathematics course of an elementary school. Moreover, Panjaburee *et al* (2010) also showed the effectiveness of using the CER model in developing a testing and diagnostic system for a junior high school Mathematics course.

However, in the CER model, the quality of the learning suggestions given to the students highly depends on the concept relationship knowledge provided by the domain experts; therefore, subjective opinions, ignorance or insufficient knowledge could significantly affect the quality of the suggestions (Hwang *et al*, 2008; Lee *et al*, 2009). Such unreliable or low quality suggestions are generated because the knowledge is usually acquired from a single expert (consider the example given in Figure 2; the two domain experts have different opinions about the relationship between concepts C_2 and C_3). In some extreme cases, the domain experts could even have different opinions about the concepts to be taken into account; ie, the CER graph structures could be different. Such different opinions could significantly affect the learning suggestions given to the students. Therefore, in the following, a knowledge integration approach is proposed by eliciting and integrating concept relationship knowledge from multiple experts to cope with this problem.

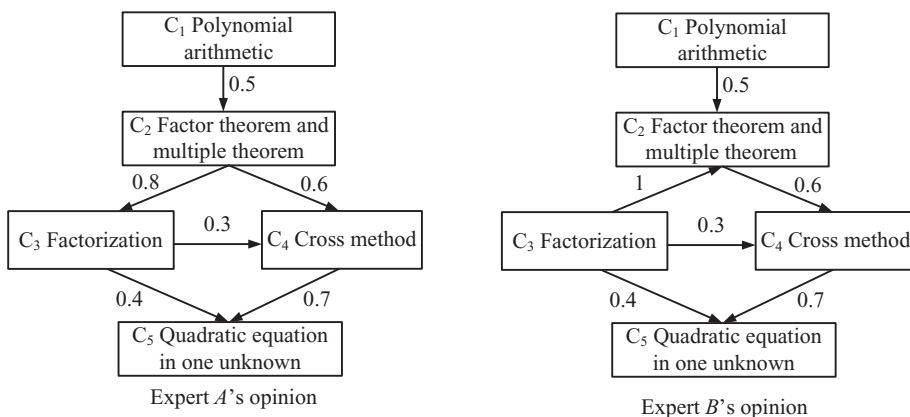


Figure 2: Comparison of the CER models constructed by expert A and expert B

A group decision approach for developing CER models

A group decision support system combines communication, computing and decision support technologies to facilitate the formulation and solution of unstructured problems by a group of people (DeSanctis & Gallupe, 1987). Past studies concerning group decision or knowledge integration have shown the effectiveness of such an approach in making quality and reliable judgments for complex problems (Chu & Hwang, 2008; DeSanctis & Gallupe, 1987; Hiltz, Johnson & Turoff, 1986). Moreover, researchers have indicated that, when eliciting knowledge from multiple experts, it is difficult to have a panel of experts constantly discussing the issues because it usually takes weeks or months to complete the group decision or knowledge acquisition process; ie, it is necessary to assume that most of the experts have difficulty in participating in synchronous or highly interactive discussions (Hwang, 1994; Panjaburee *et al*, 2010).

In this study, a group decision approach is proposed for developing quality CER models with the cooperation of multiple domain experts. It consists of three phases: the “concept eliciting and integrating” phase, the “relationship-eliciting” phase and the “relationship-integrating” phase. In the following, each phase of the multi-expert CER construction procedure is described in detail.

Step 1. Eliciting and integrating concepts from domain experts

In this phase, the domain experts are asked to provide a list of concepts to be taken into account for a course. Moreover, they are asked to use a statement to describe each concept they provide. After collecting the concepts from the experts, the integrated concept list is presented to them. To validate or verify the concepts given by these domain experts, based on the concept of group decision, they are asked to brainstorm to remove the redundant concepts (eg, “polynomial calculation” and “calculation of polynomials”). This implies that all domain experts could understand the concepts in their specific area in the same way.

Step 2. Eliciting <prior-concept, concept> relationships from individual experts

In this phase, the experts are asked to determine the relationships between concepts. The strength for C_x to be the prerequisite of C_y is represented as W_{c_x,c_y} , and the certainty degree for determining the strength is represented as $Certaintyc_{x,c_y}$. The strength of the prerequisite relationship is an integer ranging from 0 to 5, where “0” represents “no relationship” and “1” to “5” represent “very weak relationship” to “very strong relationship.” The certainty degree could be either “S” or “N,” where “S” and “N” represent “Sure” and “Not sure” respectively. Moreover, $W_{c_x,c_y}(E_i)$ and $Certaintyc_{x,c_y}(E_i)$ represent the W_{c_x,c_y} relationship and its corresponding certainty degree given by expert E_i . Furthermore, $Relationc_{x,c_y}(E_i)$ is used to represent the bidirectional relationship between C_x and C_y . $Relationc_{x,c_y}(E_i) = “X”$ if expert E_i has assigned “0” to both $W_{c_x,c_y}(E_i)$ and $W_{c_y,c_x}(E_i)$.

Table 2 shows an illustrative example in which the expert has determined the relationships between concepts C_1 , C_2 , C_3 , and C_4 . From this table, we have $W_{c_1,c_2}(E_A) = 4$ with $Certaintyc_{1,c_2}(E_A) = S$, $W_{c_1,c_3}(E_A) = 5$ with $Certaintyc_{1,c_3}(E_A) = S$, etc.

Table 2: Illustrative example of the relationships between concepts given by an expert

		Parent concept C_y			
		C_1	C_2	C_3	C_4
Child concept C_x	C_1	–	4,S	5,S	X,S
	C_2	0,S	–	0,N	X,N
	C_3	0,S	2,N	–	5,S
	C_4	X,S	X,N	0,S	–

Step 3. Integrating corresponding <concept, concept> weighting values from multiple experts

A set of knowledge-integrating rules is defined to check and integrate the corresponding <concept, concept> weighting values from multiple experts. There are four categories of rules (as shown in Table 3): (1) rules for integrating the weightings given by multiple experts for the same prerequisite relationship, (2) rules for integrating the relationships between two concepts with different prerequisite directions and degrees of confidence, (3) rules for integrating the relationships between two concepts with different prerequisite directions but the same degree of confidence and (4) rules for integrating the “X” relationships.

Table 3: Summary of the knowledge-integrating rules

Rule#	Condition	Integrated weighting	Certainty degree
1A	All of the domain experts agree on the same prerequisite relationship direction by assigning the weighting values, and <i>most of them show high confidence</i> in giving the values.	Formula (1) $W_{C_y, C_x} = 0$	“S”
1B	All of the domain experts agree on the same prerequisite relationship direction by assigning the weighting values, and <i>most of them show low confidence</i> in giving the values.	Formula (1) $W_{C_y, C_x} = 0$	“N”
2	There are opposite opinions, ie, the domain experts have assigned the weighting values to different prerequisite directions for two concepts with different degrees of certainty. Moreover, the number of domain experts who support $C_x \rightarrow C_y$ with <i>high confidence is greater than</i> that of the experts who support $C_y \rightarrow C_x$ with <i>low confidence</i> .	Formula (2) $W_{C_y, C_x} = 0$	“N”
3	There are opposite opinions, ie, the domain experts have assigned the weighting values to different prerequisite directions for two concepts with the same degree of certainty. Moreover, the number of domain experts who support $C_x \rightarrow C_y$ with <i>high confidence is equal to</i> the number of domain experts who support $C_y \rightarrow C_x$ with <i>the same degree of confidence</i> .	Ask the experts to check and reconsider their ratings	
4.1	There are some experts who have assigned the weightings to “X.” The number of domain experts who have assigned the $C_x \rightarrow C_y$ relationship with <i>high confidence is greater than</i> the number of domain experts who have assigned the $C_y \rightarrow C_x$ and the weights to “X.”	Formula (2) $W_{C_y, C_x} = 0$	“N”
4.2A	There are some experts who have assigned the weightings to “X.” Furthermore, <i>most of the domain experts have assigned the $C_y \rightarrow C_x$ relationship and “X” with high confidence</i> .	Ask the experts to check and reconsider their ratings	
4.2B	There are some experts who have assigned the weightings to “X.” Furthermore, <i>most of the domain experts have assigned the $C_x \rightarrow C_y$ with high confidence</i> .	Ask the experts to check and reconsider their ratings	
4.3A	There are some experts who have assigned the weightings to “X.” Furthermore, <i>most of the domain experts have assigned “X” with high confidence</i> .	“X”	“S”
4.3B	There are some experts who have assigned the weightings to “X.” Furthermore, <i>most of the domain experts have assigned “X” with low confidence</i> .	“X”	“N”

Along with the rules, the following formulas are used to determine the integrated weighting and certainty degree. Note that in the formulas $S_i = 2$ if $\text{Certainty}_{C_x, C_y}(E_i) = S$ and $S_i = 1$ if $\text{Certainty}_{C_x, C_y}(E_i) = N$.

$$\text{Formula (1): } W_{C_x, C_y} = \frac{\sum_{i=1}^n (W_{C_x, C_y}(E_i) \times S_i)}{\sum_{i=1}^n S_i}$$

where n is the number of domain experts.

$$\text{Formula (2): } W_{C_x, C_y} = \frac{\sum_{i.i \in \{p \vee q\}} (W_{C_x, C_y}(E_i) \times S_i)}{\sum_{i.i \in \{p \vee q\}} S_i}$$

where p and q represent the domain experts who support $C_x \rightarrow C_y$ with $\text{Certainty}_{C_x, C_y}(E_i) = S$ and $\text{Certainty}_{C_x, C_y}(E_i) = N$ respectively.

Step 4. Showing the resulting map to the domain experts for validation

In this step, the resulting map is shown to the domain experts for validation. Those who do not agree with the map can note, discuss and resolve their concerns. In that case, Steps 2–4 will be conducted repeatedly by focusing on the questioned relationships. After all of the domain experts agree with the integrated CER in the map, the learning diagnosis mechanism will analyze the learning problems of individual students and give them suggestions based on those relationships.

System development

Based on the proposed approach, a web-based testing and diagnostic system is implemented. The system provides an interface for developing the CER model, as shown in Figure 3.

After all of the domain experts have determined the relationships between concepts, the system integrates the weightings based on the proposed approach. If some unsolvable conflicts exist, the conflicted <concept, concept> weighting list is presented to all domain experts, and the system will require them to brainstorm to check and reconsider their weighting and certainty values. The procedure is repeatedly conducted until no further checking and considering of the <concept, concept> weighting information is needed. Based on the concept of group decision, all domain experts are asked to brainstorm to check the final CER model for validation. The final CER model is then used to analyze the learning problems of individual students and provide learning suggestions to them accordingly.

Figure 4 shows an illustrative example of the learning suggestions for a student. The suggestions include the test items, the correct answers to the items, the students' answers, the remedial learning path and the link to the supplementary materials for each poorly learned concept. The student is instructed to relearn those concepts following the sequence given in the remedial learning path by clicking the corresponding link to browse the supplementary materials and to do the relevant exercises. After completing the remedial course and passing a test for a concept, the student can proceed to learning the next concept.

Evaluation and analysis

To evaluate the performance of our approach, an experiment was conducted on the “computations and applications of quadratic equations” unit, which is one of the mathematics units that

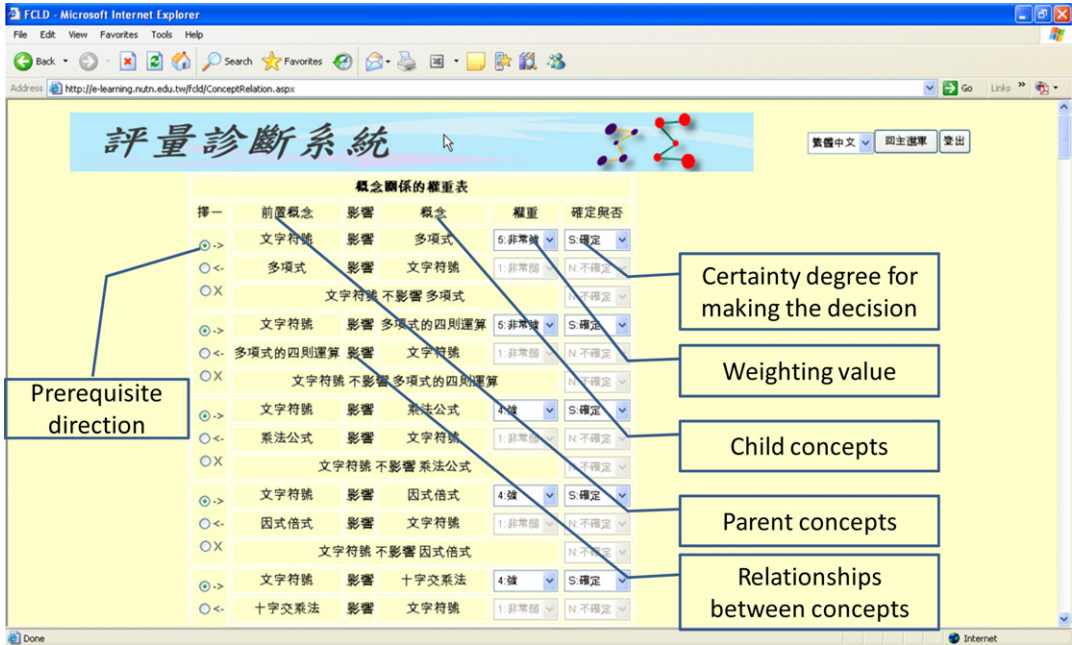


Figure 3: Interface for developing the CER model

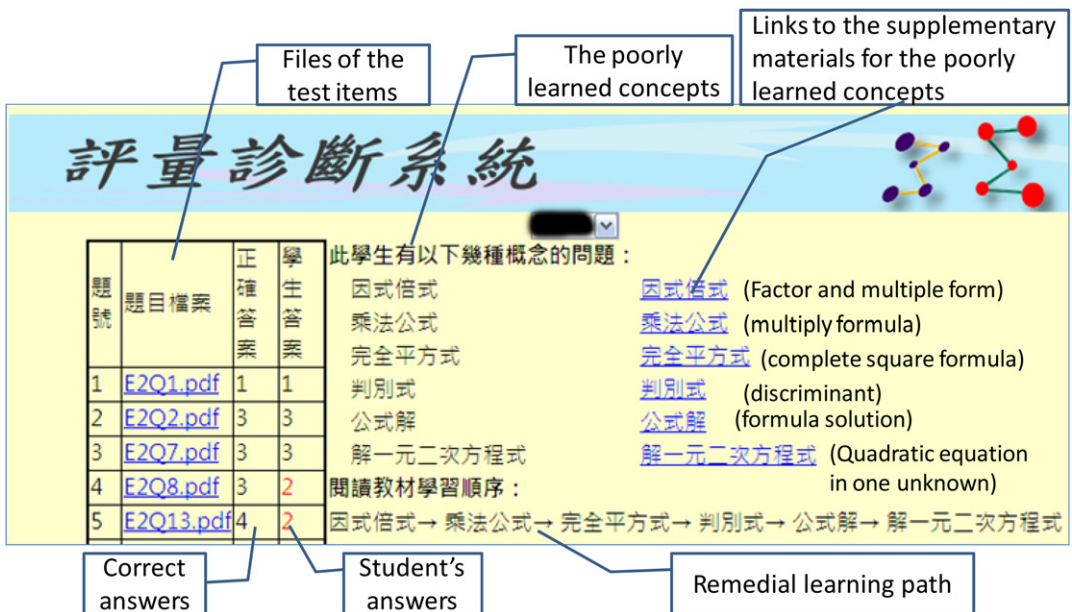


Figure 4: Interface for providing learning suggestions

most high school students have difficulties learning (Mathematical Association, 1962). In this subject unit, the students learned to solve problems via setting variables, determining the polyamines that represent the relationships between the variables and finding the answers. Such procedural knowledge for problem solving has been recognized as being a higher-order cognitive

process in Bloom's taxonomy of educational objectives, ie, "analyze," "evaluate" and "create" (Anderson *et al*, 2001). Moreover, it has also been categorized as a kind of intellectual skill (ie, the critical, analytical, synthesizing and problem-solving skills) defined in Gagne's taxonomy (Gagne, Briggs & Wager, 1992).

Participants

The participants were 104 grade eight students and three domain experts (D1, D2 and D3) who had a number of years' experience of teaching mathematics courses. In each group decision phase, the domain experts were guided by the group decision support system to provide or modify the list of concepts or the relationships between concepts in a specified period of time (eg, 1 week), and then had a synchronous discussion on the due date of that phase.

To compare the performance of the multi-expert approach and the original CER model in enhancing the learning achievement of the students, the students were divided into four groups (ie, an experimental group and three control groups):

1. Experimental group E1: in this group, the students received learning suggestions based on the CER elicited and integrated by employing the group decision approach proposed in this study.
2. Control groups CG1, CG2 and CG3: in the three control groups, the students received learning suggestions based on the CER models provided by D1, D2 and D3 respectively.

Experimental procedure

Before participating in the learning activity, the students were asked to take a pretest, which aimed to evaluate their prior knowledge for taking the subject unit. The entire learning activity lasted four weeks, during which the students took three tests on what they had learned in the time period. The test results were analyzed by applying the learning diagnostic system for generating learning suggestions to individual students and guiding them to learn with corresponding subject materials. After finishing the learning activity, all of the students took a posttest to compare the learning achievements between the experimental group and the three control groups.

Experimental results

The test scores of the four groups were analyzed by applying analysis of covariance (ANCOVA). It was found that there was no significant difference between the experimental group and the control groups, as shown in Table 4.

By examining the remedial learning paths generated for the students in different clusters of the pretest scores (ie, high, medium and low achievements), it was found that the students in the high- and medium-achievement clusters received few suggestions from the learning system because they had learned most of the concepts in that subject unit well. In contrast, the students in the low-achievement cluster required more assistance from the learning system, implying that the low-achievement students could be the ones most likely to benefit from the proposed approach (Lee *et al*, 2009).

Table 4: ANCOVA of the posttest results

Group		N	Mean	SD	F	Pairwise comparisons
(a)	E1	26	74.04	13.80	3.018	
(b)	CG1	25	67.04	20.38		
(c)	CG2	26	59.15	21.52		
(d)	CG3	27	57.78	29.92		

Table 5: ANCOVA of the posttest results for students with different prior knowledge

Group		N	Mean	SD	F	Pairwise comparisons	d
High-achievement cluster	(a)	8	89.17	3.90	12.44		
	(b)	6	90.56	0.00			0.50
	(c)	7	83.33	2.63			1.75
	(d)	7	95.24	3.45			1.65
Medium-achievement cluster	(a)	9	77.04	4.23	10.37		
	(b)	13	69.74	9.18			1.07
	(c)	12	62.78	9.41			1.95
	(d)	13	57.18	18.00			1.51
Low-achievement cluster	(a)	9	57.41	4.34	71.27*		
	(b)	6	37.78	7.20		a > b*	3.30
	(c)	7	29.05	5.68		a > c*	5.61
	(d)	7	21.43	7.90		a > d*	5.64

* $p < .05$.

(a) experimental group E₁; (b) control group CG₁; (c) control group CG₂; (d) control group CG₃.

Consequently, we further compared the learning performance of the students in different clusters of the pretest scores. An ANCOVA on the posttest scores, with the pretest as the covariate, was used to explore the learning achievement between the experimental group and the three control groups of each cluster, as shown in Table 5. It was found that the low-achievement students in the experimental group achieved significantly better test results than those in the three control groups, but there was no significant difference found in the high- and medium-achievement clusters, implying that the proposed approach is able to develop a better quality CER model than the ones developed by employing the original approach. After examining the CER models used in the control groups, it was found that several factors might have affected the effectiveness of those models:

1. The models of two control groups CG₁ and CG₂ contained additional concepts that were removed in the first step; therefore, the suggested learning paths generated by these two models contained some less relevant concepts, which could lead the students to spend additional time on irrelevant practicing; therefore, the low-achievement students might be affected because they did not have time to fully concentrate on facing their most critical learning problems. On the other hand, the high- and medium-achievement students were not seriously affected because their learning performance was relatively high.
2. The models of the three control groups contained additional links (non-zero weightings) because the domain experts were not sure about the relationships between some concepts; therefore, additional learning suggestions could be given and some important learning suggestions could be missed. For those low-achievement students, it was important to receive precise suggestions because they needed more time to practice each of the to-be-enhanced concepts. In that case, their learning performance might be significantly affected.

As the number of students in each cluster in the experiment is small, we further computed the effect size of the test results based on the Cohen's *d*-value (Cohen, 1988) as researchers have suggested reporting the effect size when the sample is small (Ge, Chen & Davis, 2005; Ge & Land, 2003). Cohen tentatively defined effect size as "small, $d = 0.2$," "medium, $d = 0.5$ " and "large, $d = 0.8$ " (Cohen, 1988); usually, a test result is said to have a large effect size if its Cohen's *d*-value is greater than 0.80. As shown in Table 5, for the posttest results between the experimental group E₁ and Control groups CG₁, CG₂ and CG₃, the *d*-values are 3.3, 5.61 and 5.64 respectively,

indicating that the finding (ie, the innovative approach is helpful to low-achievement students in improving their learning achievement) has a very large effect size.

Conclusions

This paper presents an innovative approach for integrating the CER models from multiple domain experts. Based on the proposed approach, a testing and diagnostic system has been implemented. The system can be used to work with an online learning system by detecting students' learning problems according to their test answers and giving them personal guidance based on their online learning performance. It can also be used to analyze the learning problems of students for an in-class course.

To evaluate the performance of this innovative approach, an experiment was conducted on a mathematics course in a junior high school. Three domain experts were asked to participate, and 104 junior high school students were recruited to compare the performance of the original CER model and our enhanced model. The experimental results show that the low-achievement students who received the learning suggestions provided by the knowledge integration approach made significantly better progress than those who received the suggestions generated by the existing CER model. This implies that the group decision approach is effective; moreover, it is found that the low-achievement students require more support or guidance in comparison with those medium-achievement or high-achievement students. Therefore, it is concluded that this group decision-making approach can cope with the key problem of constructing the CER models for diagnosing student learning problems via guiding the domain experts to work cooperatively. Such a finding not only plays an important role in helping students improve their learning performance but also provides valuable references for those researchers who are engaged in the study of learning diagnosis models (Hwang *et al*, 2008; Panjaburee *et al*, 2010) or the development of testing and diagnostic systems (Chen & Bai, 2009; Chu, Hwang, Tseng, Judy & Hwang, 2006; Hwang, 2007; Lee *et al*, 2009).

The finding of this study can be generalized to other applications in which prerequisite relationships exist, such as science courses (eg, physics and chemistry) and other mathematics courses. It also implies that further studies to develop more effective tools or to develop other group decision strategies to help domain experts cooperatively determine prerequisite relationships are needed. In addition, the following lead-in procedure is recommended to those who intend to introduce the proposed approach into the classroom:

Step 1: Give a brief about the meanings of concepts and relationships between concepts to the instructors by showing some illustrative examples. This step usually takes 30–60 minutes.

Step 2: Demonstrate the functions of the learning diagnosis system to the instructors. Usually it takes 20–30 minutes for the demonstration.

Step 3: Guide the instructors to determine the concepts to be taken into account in the selected subject unit (ie, the extent of conducting the learning activity). This step can be carried out by conducting online or face-to-face discussions.

Step 4: Guide each instructor to determine the relationships between the test items and concepts. This step can be carried out by invoking the proposed system.

Step 5: Guide individual instructors to determine the CER and integrate the derived relationships. This step can be conducted by invoking the proposed system. Individual instructors only need to provide the weight of the relationships between concepts (ie, 1, 2, 3, 4 or 5). Accordingly, the proposed system automatically calculates the integrated weight from those given by multiple instructors.

Step 6: For each scheduled test to be conducted during the learning activity, ask the instructors to provide a set of test items based on the scope of that test. Remove the redundant or similar test

items under the agreement of the instructors. This step can also be carried out online or face to face.

Step 7: Conduct the online tests or paper-and-pencil tests. For the students who receive paper-and-pencil tests, the test results need to be uploaded to the learning diagnosis system.

Step 8: Invoke the learning diagnosis system to analyze the test results and generate the learning guidance for individual students. In this step, some personalized paths or learning content can be provided if the learning diagnosis system is linked to an adaptive learning system. Alternatively, the learning suggestions can be printed out and given to individual students, and the instructors can ask the students to restudy the relevant part in the text book or provide additional homework to individual students based on the suggestions.

Acknowledgement

This study is supported in part by the National Science Council of the Republic of China under contract numbers NSC 99–2511-S-011-011-MY3 and NSC 100–2631-S-011-003, and the Mahidol University research fund.

References

- Aleven, V. W. M. M. & Koedinger, K. R. (2002). An effect metacognitive strategy: learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science*, 26, 2, 147–179.
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruickshank, K. A., Mayer, R. E., Pintrich, P. R. *et al* (2001). *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives* (Abridged ed.). New York: Longman.
- Barbera, E. (2009). Mutual feedback in e-portfolio assessment: an approach to the netfolio system. *British Journal of Educational Technology*, 40, 2, 342–357.
- Casamayor, A., Amandi, A. & Campo, M. (2009). Intelligent assistance for teachers in collaborative e-learning environments. *Computers & Education*, 53, 4, 1147–1154.
- Chen, C. H. (2010). Ontology-based concept map for planning a personalized learning path. *British Journal of Educational Technology*, 40, 6, 1028–1058.
- Chen, C. M. (2008). Intelligent web-based learning system with personalized learning path guidance. *Computers & Education*, 51, 2, 787–814.
- Chen, S. M. & Bai, S. M. (2009). Learning barriers diagnosis based on fuzzy rules for adaptive learning systems. *Expert Systems with Applications*, 36, 8, 11211–11220.
- Chiou, C. C. (2008). The effect of concept mapping on students' learning achievements and interests. *Innovations in Education & Teaching International*, 45, 4, 375–387.
- Chiou, C. K., Hwang, G. J. & Tseng, J. C. R. (2009). An auto-scoring mechanism for evaluating problem-solving ability in a web-based learning environment. *Computers & Education*, 53, 2, 261–272.
- Chu, H. C. & Hwang, G. J. (2008). A Delphi-based approach to developing expert systems with the cooperation of multiple experts. *Expert Systems with Applications*, 34, 4, 2826–2840.
- Chu, H. C., Hwang, G. J. & Huang, Y. M. (2010). An enhanced learning diagnosis model based on concept effect relationships with multiple knowledge levels. *Innovations in Education and Teaching International*, 47, 1, 53–67.
- Chu, H. C., Hwang, G. J., Tseng, J. C. R. & Hwang, G. H. (2006). A computerized approach to diagnosing student learning problems in health education. *Asian Journal of Health and Information Sciences*, 1, 1, 43–60.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cooner, T. S. (2010). Creating opportunities for students in large cohorts to reflect in and on practice: lessons learnt from a formative evaluation of students' experiences of a technology-enhanced blended learning design. *British Journal of Educational Technology*, 41, 2, 271–286.
- DeSanctis, G. & Gallupe, B. (1987). A foundation for the study of group decision support systems. *Management Science*, 33, 589–609.
- Draper, S. W. (2009). What are learners actually regulating when given feedback? *British Journal of Educational Technology*, 40, 2, 306–315.
- Gagne, R. M., Briggs, L. J. & Wager, W. W. (1992). *Principles of instructional design*. Fort Worth, TX: Harcourt Brace Jovanovich.

- Ge, X., Chen, C. H. & Davis, K. A. (2005). Scaffolding novice instructional designers' problem-solving processes using question-prompts in a web-based learning environment. *Journal of Educational Computing Research*, 33, 2, 219–248.
- Ge, X. & Land, S. M. (2003). Scaffolding students' problem-solving processes in an ill-structured task using question prompts and peer interactions. *Educational Technology Research and Development*, 51, 1, 21–38.
- Hiltz, S. R., Johnson, K. & Turoff, M. (1986). Experiment in group decision making: communication process and outcome in face-to-face versus computerized conferences. *Human Communication Research*, 13, 2, 225–253.
- Hsu, C. K., Hwang, G. J. & Chang, C. K. (2010). Development of a reading material recommendation system based on a knowledge engineering approach. *Computers & Education*, 55, 1, 76–83.
- Hwang, G. H., Chen, J. M., Hwang, G. J. & Chu, H. C. (2006). A time scale-oriented approach for building medical expert systems. *Expert Systems with Applications*, 31, 2, 299–308.
- Hwang, G. J. (1994). Knowledge elicitation and integration from multiple experts. *Journal of Information Science and Engineering*, 10, 1, 99–109.
- Hwang, G. J. (2003). A conceptual map model for developing intelligent tutoring systems. *Computers & Education*, 40, 3, 217–235.
- Hwang, G. J. (2007). Gray forecast approach for developing testing and diagnostic systems. *IEEE Transactions on System, Man and Cybernetic Part C*, 37, 1, 98–108.
- Hwang, G. J., Shih, Y. R. & Chu, H. C. (2011). A concept map approach to developing collaborative Mindtools for context-aware ubiquitous learning. *British Journal of Educational Technology*, 42, 5, 778–789.
- Hwang, G. J., Tseng, J. C. R. & Hwang, G. H. (2008). Diagnosing student learning problems based on historical assessment records. *Innovations in Education and Teaching International*, 45, 1, 77–89.
- Jong, B. S., Chan, T. Y. & Wu, Y. L. (2007). Learning log explorer in e-Learning diagnosis. *IEEE Transactions on Education*, 50, 3, 216–228.
- Jong, B. S., Lin, T. W., Wu, Y. L. & Chan, T. (2004). Diagnostic and remedial learning strategy based on conceptual graphs. *Journal of Computer Assisted Learning*, 20, 5, 377–386.
- Jordan, S. & Mitchell, T. (2009). e-Assessment for learning? The potential of short-answer free-text questions with tailored feedback. *British Journal of Educational Technology*, 40, 2, 371–385.
- Kwasnicka, H., Szul, D., Markowska-Kaczmar, U. & Myszkowski, P. B. (2008). Learning assistant—personalizing learning paths in e-learning environments. In *Proceedings of the 7th International Conference on Computer Information Systems and Industrial Management Applications*, Ostrava, Czech Republic, June 26–28.
- Lee, C. H., Lee, G. G. & Leu, Y. (2009). Application of automatically constructed concept map of learning to conceptual diagnosis of e-learning. *Expert Systems with Applications*, 36, 2, 1675–1684.
- Léger, B. & Naud, O. (2009). Experimenting statecharts for multiple experts knowledge elicitation in agriculture. *Expert Systems with Applications*, 36, 8, 11296–11303.
- Li, L., Liu, X. & Steckelberg, A. L. (2010). Assessor or assessee: how student learning improves by giving and receiving peer feedback. *British Journal of Educational Technology*, 41, 3, 525–536.
- Liu, C. C., Don, P. H. & Tsai, C. M. (2005). Assessment based on linkage patterns in concept maps. *Journal of Information Science and Engineering*, 21, 5, 873–890.
- Magoulas, G. D., Papanikolaou, K. & Grigoriadou, M. (2003). Adaptive web-based learning: accommodating individual differences through system's adaption. *British Journal of Educational Technology*, 34, 4, 511–527.
- Marriott, P. (2009). Students' evaluation of the use of online summative assessment on an undergraduate financial accounting module. *British Journal of Educational Technology*, 40, 2, 237–254.
- Mathematical Association (1962). *The teaching of Algebra in schools*. London: G. Bell & Sons.
- Panjaburee, P., Hwang, G. J., Triampo, W. & Shih, B. Y. (2010). A multi-expert approach for developing testing and diagnostic systems based on the concept-effect model. *Computers & Education*, 55, 2, 527–540.
- Peng, H. Y., Su, Y. J., Chou, C. & Tsai, C. C. (2009). Ubiquitous knowledge construction: mobile learning re-defined and a conceptual framework. *Innovations in Education and Teaching International*, 46, 2, 171–183.
- Roth, W. M. & Roychoudhury, A. (1994). Science discourse through collaborative concept mapping: new perspectives for the teacher. *International Journal of Science Education*, 16, 4, 437–455.
- Sieber, V. (2009). Diagnostic online assessment of basic IT skills in 1st-year undergraduates in the Medical Sciences Division, University of Oxford. *British Journal of Educational Technology*, 40, 2, 215–226.
- Tseng, S. S., Su, J. M., Hwang, G. J., Hwang, G. H., Tsai, C. C. & Tsai, C. J. (2008). An object-oriented course framework for developing adaptive learning systems. *Educational Technology & Society*, 11, 2, 171–191.

- Tseng, S. S., Sue, P. C., Su, J. M., Weng, J. F. & Tsai, W. N. (2007). A new approach for constructing the concept map. *Computers & Education*, 49, 3, 691–707.
- Wang, Y. H. & Liao, H. C. (2011). Adaptive learning for ESL based on computation. *British Journal of Educational Technology*, 42, 1, 66–87.
- Yeh, Y. F., Chen, M. C., Hung, P. H. & Hwang, G. J. (2010). Optimal self-explanation prompt design in dynamic multi-representational learning environments. *Computers & Education*, 54, 4, 1089–1100.
- Yin, P. Y., Chang, K. C., Hwang, G. J., Hwang, G. H. & Chan, Y. (2006). A particle swarm optimization approach to composing serial test sheets for multiple assessment criteria. *Educational Technology & Society*, 9, 3, 3–15.