Lin, J.-W., Lai, Y.-C., Szu, Y.-C., Lai, C.-N., Chuang, Y.-S., & Chen, Y.-H. (2014). Development and Evaluation of Across-Unit Diagnostic Feedback Mechanism for Online Learning. *Educational Technology & Society*, *17* (3), 138–153.

Development and Evaluation of Across-Unit Diagnostic Feedback Mechanism for Online Learning

Jian-Wei Lin[1], Yuan-Cheng Lai[2], Yu-Chin Szu[3], Ching-Neng Lai[4], Yuh-Shy Chuang[1] and Yen-Hung Chen[5]

[1]Chien Hsin University, Taiwan // [2]National Taiwan University of Science and Technology, Taiwan // [3]St. John's University, Taiwan // [4]Hsing Wu University, Taiwan // [5]Institute for Information Industry, Taiwan // jwlin@uch.edu.tw // laiyc@cs.ntust.edu.tw // belinda@mail.sju.edu.tw // Sebastian.lai@gmail.com // yschuang@uch.edu.tw // pplong@gmail.com

[*]Corresponding author

(Submitted March 23, 2013; Revised August 13, 2013; Accepted September 8, 2013)

ABSTRACT

Solving well-structured problems often requires using considerable related concepts which are usually scattered and introduced throughout different learning units of a subject. In addition, poor learning of related concepts of preceding units may block the learning of subsequent units, and eventually leads to the inability to solve well-structured problems of a subject. Thus, this work proposes using across-unit diagnostic feedback, which can identify weak concepts not only within a unit but also in different units. Furthermore, the provided feedback can be used to recommend remedial learning paths for students, and inform the students the priority of the paths to understand which weak unit should be remedied first and which weak concepts within a unit should be remedied first. Students can refer to the instructions and use the provided corresponding remedial materials to conduct remedial learning in a systematic way. To discriminate the learning effect among various feedback types, this project will compare the proposed system, the Across-Unit Diagnostic Feedback System (AUDFS), with two other systems, the Single-Unit Diagnostic Feedback System (SUDFS) and the Traditional Feedback System (TFS). Experiment results show that the proposed system significantly enhanced learning achievement and the ability to solve well-structured problems for students. The mean student retention time of the proposed system is significantly higher than that of other systems, indicating that the proposed system enables sustained connection between students and the system. Additionally, positive correlations exist between student retention time of the proposed system and student post-test scores. Through a questionnaire and interviews, most students expressed positive attitudes to the proposed system.

Keywords

E-learning, Diagnostic feedback, Remedial learning, Online assessments

Introduction

Traditional education breaks an entire course into units and requires that learners focus on each unit independently, such that learners are often unable to perceive the big picture before all units are presented (Wang, Peng, Cheng, Zhou, and Liu, 2011). Conceptual understanding of a domain is often lacking in books and learning materials. Moreover, after the author teaching computer science courses a number of times, it was clear that most students prepared for their midterm or final exam at the last minute. That is, students rarely reviewed the material of previous units. Without prior knowledge of preceding units, students typically had difficulty in understanding subsequent units. The vicious cycle may propagate until the semester ends, resulting in reduced learning efficiency for the students and leaving students poorly prepared for midterms and final exams. Prior knowledge plays an important role in learning new concepts, such that students typically require prerequisite knowledge to learn a new concept (Ausubel, 1968). Lin, Lin, and Huang (2011) developed a test-based diagnostic system to assist instructors and students in diagnosing prior knowledge before new instruction is undertaken.

Moreover, the most common problems for students to solve in scientific subjects at universities are well-structured problems that feature logical, story- and rule-based problems with pre-defined steps and correct solutions (Jonassen, 1997). Arithmetic problems test the scientific conceptual understanding of students, requiring comprehension of a finite number of regular concepts, rules, and knowledge which may span a subject (Laxman, 2010). Without comprehension of preceding units, this can be a hurdle to the acquisition of follow-up knowledge and further reduce a student's ability to solve well-structured problems.

Formative assessment plays an essential role in learning processes, through improving student learning efficiency by offering feedback rather than evaluation for course grades (Pachler, Daly, Mor, and Mellar, 2010). Students who have received feedback on their performance may take steps to remedy whatever weak concepts the assessment has exposed (Buchanan, 2000). Conducting one assessment for every unit throughout a course is very likely to stimulate students to practice and review preceding units, in turn benefiting learning in subsequent units.

However, learning feedback should show more information, instead of just simple results (correct/incorrect) or correct answers. That is, besides informing learners of their correctness, learners need to be told where their problem-solving process went wrong and coached from that point onward (Jonassen, 1997). An effective coaching method, "diagnostic feedback," contributes to learning effectiveness. There are a handful of literatures which applied data mining and fuzzy theories to diagnose learning barricades or to developing adaptive learning for each individual (Chen and Bai, 2009; Chen, Hsieh, and Hsu, 2007; Lee, Lee, and Leu, 2009). Chen and Bai (2009) presented a method to diagnose learning barriers of learners based on fuzzy rules. Chen et al. (2007) presented a learning diagnosis system based on association rules for learning mathematics in an elementary school, trying to discover any learner misconceptions according to the responses of incorrect testing items during the learning processes. Lee et al. (2009) developed an intelligent concept diagnostic system based on a technique of data mining, enabling teachers to diagnose the learning barriers and misconception of learners with constructed concept maps.

However, most related studies provide diagnostic feedback for one assessment (i.e., one specific unit) instead of for an entire subject (i.e., a series of related assessments). For example, Chen, Hsieh, and Hsu (2007) designed a learning diagnostic system based on association rules for the course unit of "Fractions" in a mathematics course at an elementary school. Hwang (2003) developed a method to identify remedial learning paths for an assessment in the unit called "Number expressions and operators" in a science course at an elementary school. Chen (2011) proposed a personalized diagnostic system for an assessment in the unit called "Oriented-object programming design concept." Feedback from only a small learning scope may hinder the acquisition of connected concepts across successive units and adversely affect the ability to solve well-structured problems. In other words, how to provide feedback for large learning scope (*e.g*., a series of online assessments) is still blurred at present.

This work provides diagnostic feedback across units in a subject and assigns one assessment to each unit. This work further developed the novel Across-Units Diagnostic Feedback System (AUDFS). The proposed diagnostic feedback scheme detects weak concepts within an assessment and across assessments. One assessment can solidify fundamental knowledge of the current unit, paving the way for the next unit. As arranged assessments in a course schedule increases, students can cement fundamental knowledge across units, gradually understanding correlations among units and the inner concepts of such units, and thus enhance their ability to solve well-structured problems. The recommended remedial learning path and material are regarded as scaffolding feedback for students, which motivates a learner's interest and reduces frustration (Shute, 2008).

The remainder of this paper is organized as follows. The proposed approach is presented, after which the system development is elaborated. Then the educational evaluation is presented and finally the conclusion is delivered.

The proposed approach

An explicit representation of the structure of conceptual knowledge is constructed by capturing key knowledge concepts and their relationships in a visual format. This visual knowledge structure is a cognitive roadmap that facilitates the knowledge construction and high-level thinking of online learners (Wang et al., 2011). According to Zhuge and Li (2006), sequential and subtype relationships are the two relationships between knowledge concepts. A sequential relationship indicates that one concept should be learned before another, and is based on Ausubel's assimilation theory (Ausubel, 1968), meaning that the role of prerequisite knowledge is essential for learning new concepts. A subtype relationship means that one child concept belongs to one parent concept, which conforms to Fisher, Gleitman, and Gleitman (1991) who states that such a relationship forms a tree structure, where the scope of the parent concept is broad, while that of a child concept is specific (Liu, Chen, and Chang, 2010). Hwang (2003) argued that a hierarchical structure offers an overall cognition of subject content from a top-down perspective.

Consequently, subject content can be presented in a tree structure, comprising of constituent units (e.g., chapters or sections) and key concepts within a unit. In reality, such a structure of conceptual knowledge is frequently used in a science textbook index. An assessment is assigned to each unit, which contains a number of key concepts. An assessment evaluates a student's familiarity with key concepts in a unit. According to Wang (2011), the proper number of key concepts in an assessment should not exceed ten, so that learners may have manageable and measurable milestones and not feel overwhelmed. This work considers only the unit level and does not consider the sub-unit level. Figure 1(a) shows a simple example of a conceptual knowledge structure for a subject with three units (i.e., U_1, U_2, and U_3) in which assessment 1 (A_1) belongs to U_1 and contains six key concepts, C_1–C_6. Within each assessment, the sequential relationship is set for inner key concepts (Fig. 1(a)). Additionally, sequential relationships should be established among units (Fig. 1(b)).

Notably, a subject typically contains numerous concepts. Without considering the subtype relationship (i.e., considering only concepts without units), all concepts will be mixed up and placed in Level 1, directly below the root (i.e., subject) of the tree. Thus, one assessment will contain all subject concepts, which contradicts the position of Wang et al. (2011) who suggested that the number of concepts in one assessment should not exceed ten. Additionally, the structure of conceptual knowledge will become disorganized and unreadable. Using the subtype relationship, one can organize and present numerous concepts in a systematic manner.

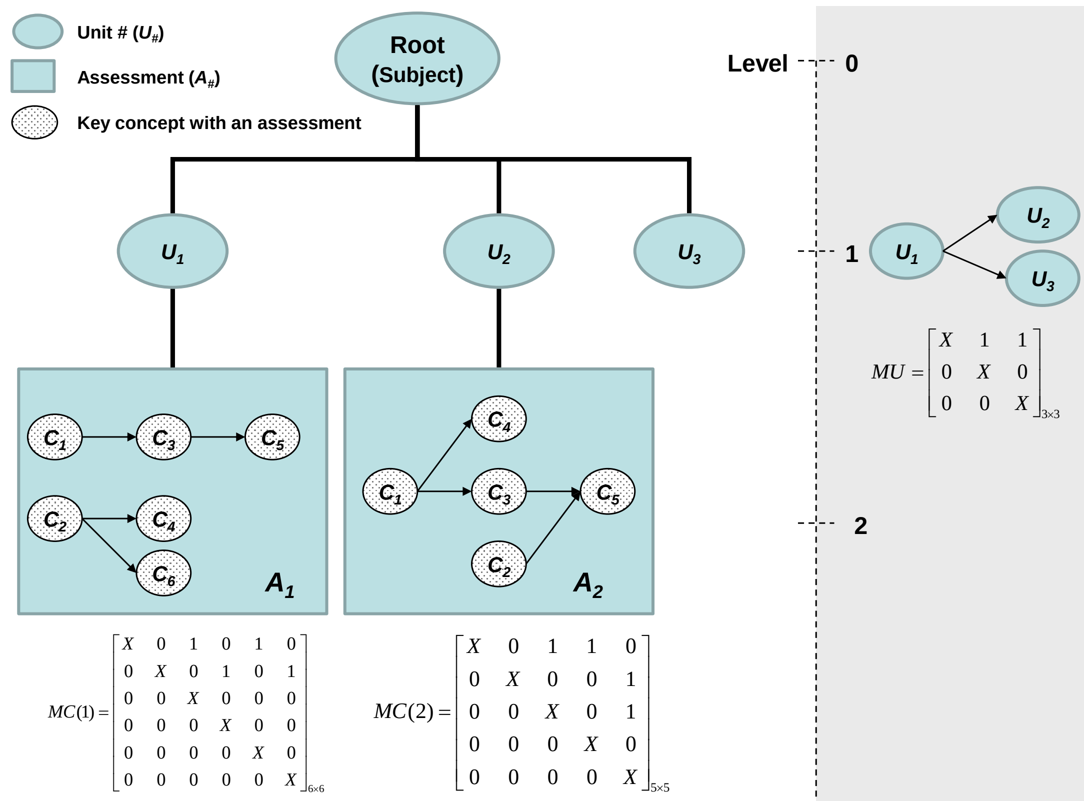


Figure 1. (a) Left part: A conceptual knowledge structure of a subject, which presents subject content in a tree structure and identifies sequential relationships among inner concepts in a unit (b) Right part: Identifying sequential relationships among units.

Instead of asking learners to construct the structure of conceptual knowledge alone, students can use expert knowledge to support their thinking and learning based on a solid foundation, reducing the likelihood of misconception (Wang et al., 2011). Experts have acquired a great deal of well-organized content knowledge, and their organization reflects a deep understanding of subject matters (Bransford, Brown, and Cocking, 2000). Thus, in this work, a teacher determines the subtype and sequential relationships to construct the structure of conceptual knowledge.

The proposed across-unit diagnostic algorithm

For facilitating algorithm computing in practice, sequential relationships can be represented in a matrix. $MU=[mu_{j,k}],\ 1 \le j,k \le M$ represents the sequential relationship between all units within a subject in which M is the number of units in a subject. An element is binary, indicating the presence (1) or absence (0) of a relationship. The element in the diagonal is not defined. If k is a succeeding unit of j (i.e., j is a preceding unit of k), $mu_{j,k} = 1$, else $mu_{j,k} = 0$. Similarly, $MC(i)=[mc(i)_{j,k}],\ 1 \le j,k \le N_i$ represents the sequential relationship between all concepts within the *i-th* unit in which N_i is the number of concepts in the *i-th* unit. Fig.1 (b) shows an example of *MU* and Fig.1 (a) shows the examples of *MC(1)* and *MC(2)*.

```
MU:  The matrix of the sequential relationships among all units within a subject
MC(i): The matrix of the sequential relationships among all concepts within the i-th unit
QU:  The queue of unit remedy path within a subject
QC(i) : The queue of concept remedy path within the i-th unit

GenerateQU (current_unit) {
        for (j = 1; j <= current_unit; j++) {
                if ( j is a weak unit ) and (Visited[j] is false) {
                        Add (j) to QU ;
                        GenerateQC(j) ;
                        for ( k = j+1; k <= current_unit; k++) {
                                If (k is a weak unit) and ( mu_{j,k}=1 ) {
                                    Add (k) to QU ;
                                    Visited[k] = true;
                                    GenerateQC(k) ;
                                }
                        }
                }
        }
}
GenerateQC (i-th_unit) {
        for (j= 1; j <= last_concept; j++) {
                If ( j is a weak concept ) and (Visited[j] is false) {
                        Add (j) to QC(i-th_unit) ;
                        for ( k = j+1; k <= last_concept; k++) {
                                If (k is a weak unit) and ( mc(i)_{j,k}=1 ) {
                                    Add (k) to QC(i-th_unit);
                                    Visited[k] = true;
                                }
                        }
                }
        }
}
```

Figure 2. Pseudo code of the across-unit diagnostic algorithm

Figure 2 shows the pseudo-code of the proposed across-unit diagnostic algorithm. The main aims of this algorithm are to identify weak concepts and units and further generate the across-unit remedial paths. *QU* is the queue of the unit remedy path within a subject. After the algorithm is executed completely, sequentially outputting the elements in the *QU*, which has the property of First In First Out (FIFO), can acquire the unit remedy path and priority. Similarly, *QC(i)* is the queue of concept remedy path within the *i*-th unit. Sequentially outputting the elements in the *QC(i)* can acquire the concept remedy path and priority for the *i*-th unit.

Both inner pseudo-codes in the *GenerateQU()* and *GenerateQC()* functions are similar. The difference between these two functions is that *GenerateQU()* is used to generate unit remedial paths between the first unit and the current testing unit while *GenerateQC()* is used to generate concepts remedial paths within a given unit.

The *GenerateQU()* function checks each unit from the first unit to the current testing unit. If the *j*-th unit is weak, the function first adds the unit into *QU* and then calls the *GenerateQC()* function to generate concept remedial paths within the *j*-th unit. Furthermore, it also checks the *j+1*-th element to the *current_unit*-th element in *j*-th row of the *MU* (i.e., $mu_{j,j+1}$, $mu_{j,j+2}$ … $mu_{j,current\_unit}$) to identify whether there are succeeding units of the the *j*-th unit which are weak. If existing (e.g. the *k*-th unit), these succeeding weak units are sequentially added into *QU* and the *GenerateQC()* function is individually called to generate concept remedial paths for every succeeding weak unit. A succeeding weak unit (e.g. the *k*-th unit) which has been visited will never be visited again. The above procedure processes repeatedly until the last unit has been finished.

The *GenerateQC()* function checks each concept from the first to the last within a given unit, the *i*-th_unit. If *j*-th concept is weak, the function first adds the concept into *QC(i)* and further checks the *j*+1-th element to the last element in *j*-th row of the *MC(i)* (i.e., $mc(i)_{j,j+1}$, $mc(i)_{j,j+2}$ … $mc(i)_{j,last\_concept}$) to identify whether there are existing succeeding concepts of the *j*-th concept which are weak. If existing (e.g. the *k*-th concept), these succeeding weak concepts are sequentially added into *QC(i)*. A succeeding weak concept (e.g. the *k*-th concept) which has been visited will not be visited again afterward. The above procedure processes repeatedly until the last concept has been finished.

The method of identifying weak concepts within a unit through an assessment

The assessment result is the main factor in determining a learner's performance (Zhuge and Li, 2006). Each question in an assessment may be associated with certain concepts (Lee et al., 2009; Zhuge and Li, 2006). The method to identify weak concepts within an assessment is based on the study by Hwang (2003), and has been slightly modified as follows. Herein, 1 represents a question associated with a concept, whereas 0 represents no association. Instructors define the association of questions with concepts through authoring interfaces. Suppose A_1 (Assessment 1 in Fig. 1) consists of six questions (Q_1–Q_6) and six concepts (C_1–C_6) that are associated with the questions (Table 1).

Suppose a student incorrectly answers Q_2 and Q_4. The student can get the Correct Rate (*CR*) for concepts 1–6 of 100%, 66%, 100%, 66%, 80%, and 75%, respectively, where *CR* is calculated as *CR* = Correct Amount / Total Amount.

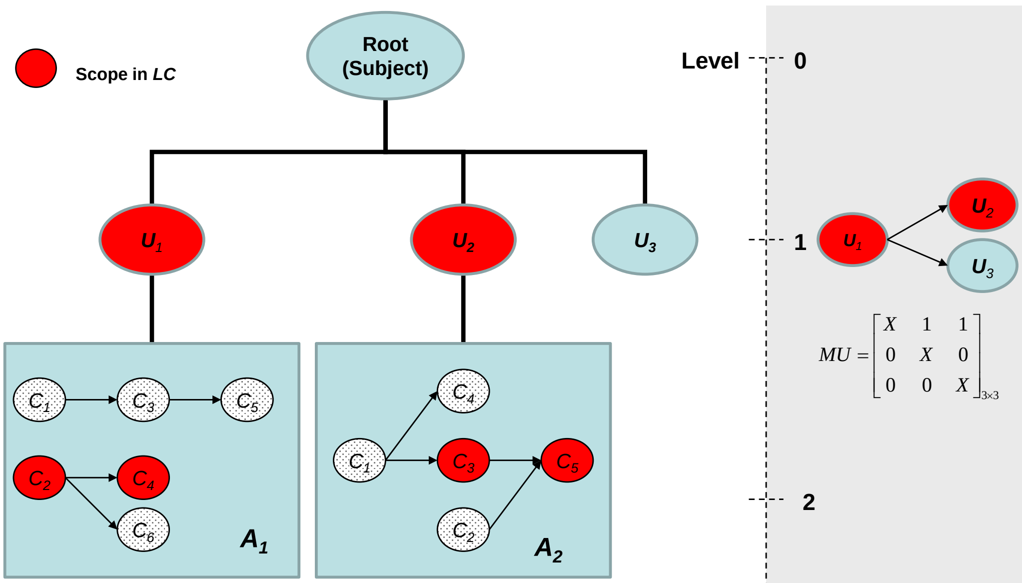
Table 1. The association between questions and concepts in Assessment 1 (A_1) in Figure 1

Questions	Concepts C_1	C_2	C_3	C_4	C_5	C_6
Q_1	1	1	1	1	1	0
Q_2	0	1	0	0	0	1
Q_3	1	1	1	0	1	1
Q_4	0	1	0	1	1	0
Q_5	1	1	1	0	1	1
Q_6	1	1	1	1	1	1
Correct Amount	4	4	4	2	4	3
Total Amount	4	6	4	3	5	4
Correct Ratio(CR)	100%	66%	100%	66%	80%	75%

A concept *CR* of a student below the average is classified as a Low-score Cluster (*LC*); conversely, that above the average is classified as a High-score Cluster (*HC*). After all participants finish an assessment, the average *CR* for C_1–C_6 can be calculated. By comparing a concept *CR* with the average, a student can obtain his/her learning status (i.e., which cluster a concept is located) for all concepts within an assessment. A concept located in *LC* is deemed as a weak concept and needs to be remedied. If more than two weak concepts are exposed in an assessment, the student should focus on weak concepts that precede other weak concepts. With A_1 (Fig. 3) as an example, if C_2 and C_4 are the two weak concepts, the student is advised that C_2 has a highest priority to be remedied because it precedes C_4 in a sequential relationship.

Across-unit diagnostic feedback

Using the above-mentioned methods, one can acquire the learning status for any key concept within a unit when the corresponding assessments are completed. For instance, once A_1 is completed, the learning status of key concepts C_1–C_6 and unit 1 (U_1) can be obtained. This is the same scenario for A_2. If a student completed A_1 and A_2 (Fig. 2) and the scores in the *LC* are marked in red, the suggested remedial learning paths should be addressed before entering the next unit, U_3. At this time, applying the proposed across-unit diagnostic algorithm can obtain the following feedback. First, the weak unit U_1 is added into *QU* and then generates the QC(*1*) for U_1 which contains two weak concepts: C_2 and C_4. Then, the weak unit U_2 is added into *QU* and then generates the *QC(2)* for U_2 which contains two weak concepts C_3 and C_5. Finally, sequentially outputting *QU* obtains U_1 and U_2, which means the unit remedy path is U_1->U_2. Sequentially outputting *QC(1)* obtains C_2 and C_4, which means the concept remedy path for U_1 is C_2->C_4. Sequentially outputting *QC(2)* obtains *C3* and *C5*, which means the concept remedy path for U_2 is C_3->C_5. More specifically, these results can be explained as follows. *U1* should be remedied first because a poor U_2 likely results from a poor U_1 according to their sequential relationship. Thus, the first recommended concept remedy path is C_2 -> C_4 within U_1 because C_2 precedes C_4. The second recommended concept remedy path is C_3-> C_5 within U_2 because C_3 precedes C_5.



$$MU = \begin{bmatrix} X & 1 & 1 \\ 0 & X & 0 \\ 0 & 0 & X \end{bmatrix}_{3\times3}$$

$$MC(1) = \begin{bmatrix} X & 0 & 1 & 0 & 1 & 0 \\ 0 & X & 0 & 1 & 0 & 1 \\ 0 & 0 & X & 0 & 0 & 0 \\ 0 & 0 & 0 & X & 0 & 0 \\ 0 & 0 & 0 & 0 & X & 0 \\ 0 & 0 & 0 & 0 & 0 & X \end{bmatrix}_{6\times6} \quad MC(2) = \begin{bmatrix} X & 0 & 1 & 1 & 0 \\ 0 & X & 0 & 0 & 1 \\ 0 & 0 & X & 0 & 1 \\ 0 & 0 & 0 & X & 0 \\ 0 & 0 & 0 & 0 & X \end{bmatrix}_{5\times5}$$

Figure 3. Identifying remedial learning paths and their priority from a top-down perspective

In addition to the textual information like above, the proposed system also provides the corresponding graphical information, as shown in Fig. 3. Miller and Miller (1999) explicated the focal tasks of a web-based learning system as follows: 1) present course content in a manner that hierarchically structures the sequence of information; and 2) obtain student feedback to insure accuracy of understanding. In our proposed framework, students can understand the structure of the entire course, perceive weaknesses, and conduct remedial learning from the bottom up. Once these weak key concepts are remedied by remedial learning, high-level units will be improved accordingly. Liu et al. (2010) stated that a learner must link and align the relationship of child concepts under the same parent concept to comprehend this parent concept; this knowledge construction process is bottom-up comprehension. Notably, as a course schedule is ongoing and assessments are gradually conducted, students can retrieve diagnostic feedback and do not have to wait until semester end. This allows students to identify relationships among concepts and units, generate semantic networks, and conduct high-level thinking for review and reflection (Wang et al., 2011).

System development

The operational procedure

Teachers must finish the following two tasks ahead of class. The first task is to establish the tree structure of a subject according to course content and build sequential relationships among units and within each unit (Fig. 1). The second task is to establish a Question Bank Database for each assessment. The assessments are based on textbook material and handouts. In this work, most questions are multiple choices. The operational procedure (Fig. 4) consists of the following steps.

- Step 1: Students learn a new unit in class.
- Step 2: Students take an assessment and obtain their results, including their answers, correct answers, and question restatements, to monitor their learning status for the unit.
- Step 3: Students check every incorrectly answered question. Meanwhile, students can retrieve diagnostic feedback.
- Step 4: If all units are finished, the operational procedure terminates; otherwise, return to Step 1 for a new unit and start again.

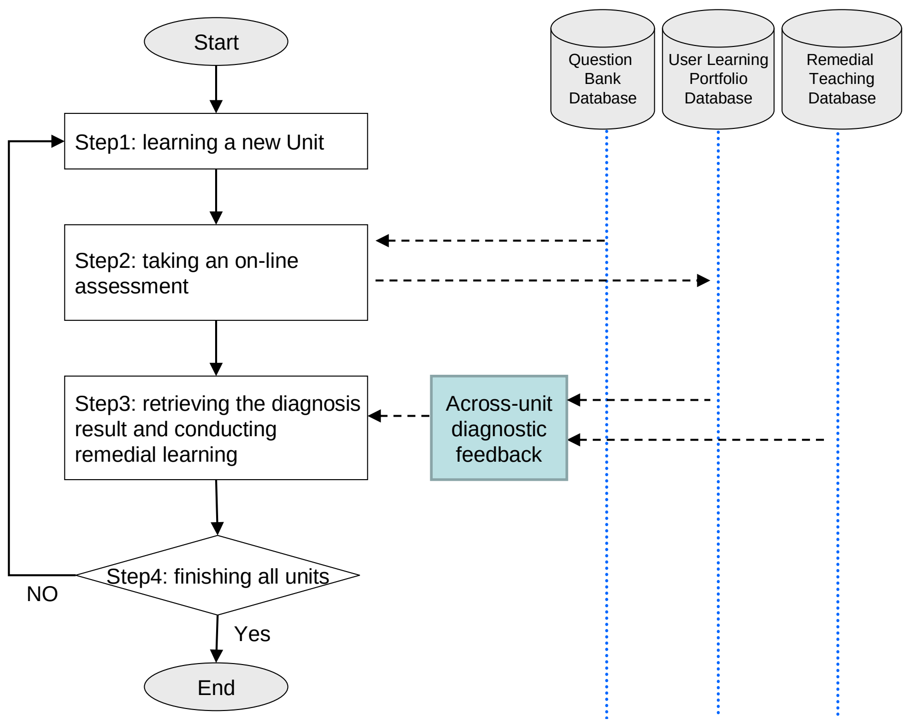


Figure 4. Flowchart of the operational procedure

Students can log in at will to review their current or historical assessments, receive diagnostic feedback, and complete an assessment that they did not complete previously. Providing students with across-unit diagnostic feedback may facilitate students with the acquisition of systematic knowledge, sustain the students' connections to the system, and help students obtain learning skills to cope with well-structured problems. Therefore, using the proposed approach, students can take advantage of the system in the following scenarios. Scenario 1: A student who fails to attend an assessment session will receive an email reminder, notifying the user to complete the missed assessments to keep up with the class. Scenario 2: When a student who missed preceding assessments finishes conducting the current assessment, he/she will only receive the diagnostic feedback for the current unit. To obtain complete remedial support (i.e., across-unit diagnostic feedback), he/she has to complete the missed assessments. This scenario may engender his/her interest to complete the assessments to remedy weaknesses. These two scenarios demonstrate how the system raises a student's awareness of his/her learning status, which may facilitate a student's sustained connection to the system.

The developed system

The proposed system has two parts: a front-end User Platform and a back-end Administrator Platform. The front-end platform is available to students for assessments and diagnostic feedback, while the back-end platform is only available to administrators (teachers) for managing student-related activities.

Back-end platform

The Administrator Platform has several functions for managing course structure, sequential relationships, and assessments. Other functions are remote user management and statistical analysis. Some graphical user interfaces are shown in Fig. 5. The top-left window shows the construction of the course structure while the bottom-left window is to construct sequential relationships among units. The top-right window shows the assessment edition. By using the multimedia editor, embedded by the free software component FCKEditor (http://www.fckeditor.net), a teacher can edit multimedia questions by inputting text, images, tables, flash objects, and hyperlinks and further arrange the layout. The bottom-right window sets the association between questions and concepts within an assessment (Table 1).

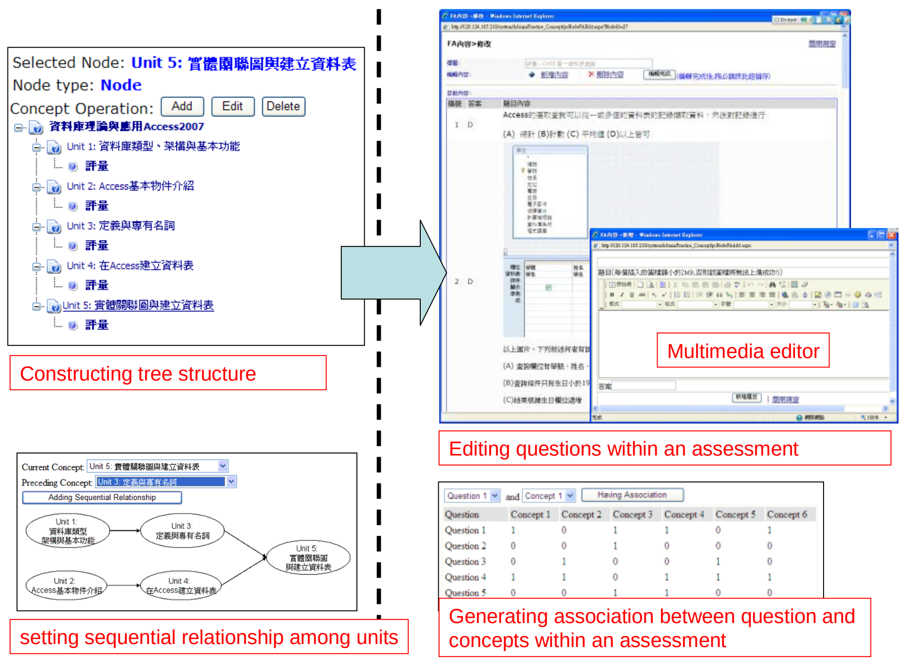


Figure 5. Snapshots of main functions of the back-end platform

Front-end platform

A student takes an assessment through the front-end platform. The assessment is generated by an administrator through the back-end platform. The student can review personal learning profiles from the assessment list and past results (Fig. 6)

Figure 7 shows snapshots of the across-unit diagnostic feedback that students can obtain. In the left window, the upper part shows the diagnostic result among units in a tree structure, while the lower part shows weak units and the remedial learning sequence for units. The top-right window shows weak concepts within a unit and the remedial learning sequence for inner concepts. Students can click a weak concept and receive the corresponding remedial material, as shown in the bottom-right window. The remedial material presented in a multimedia content format, including audio, video, PowerPoint presentations, and file downloads, and it is created using PowerCAM software (http://www.powercam.com.tw/). Students can select or replay the desired material through operational interfaces.

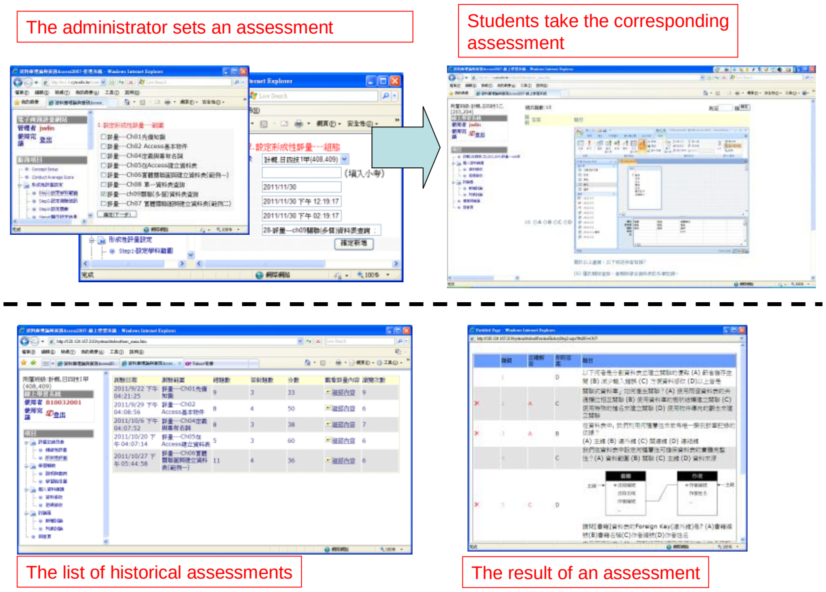


Figure 6. Snapshots of the main functions of the front-end platform

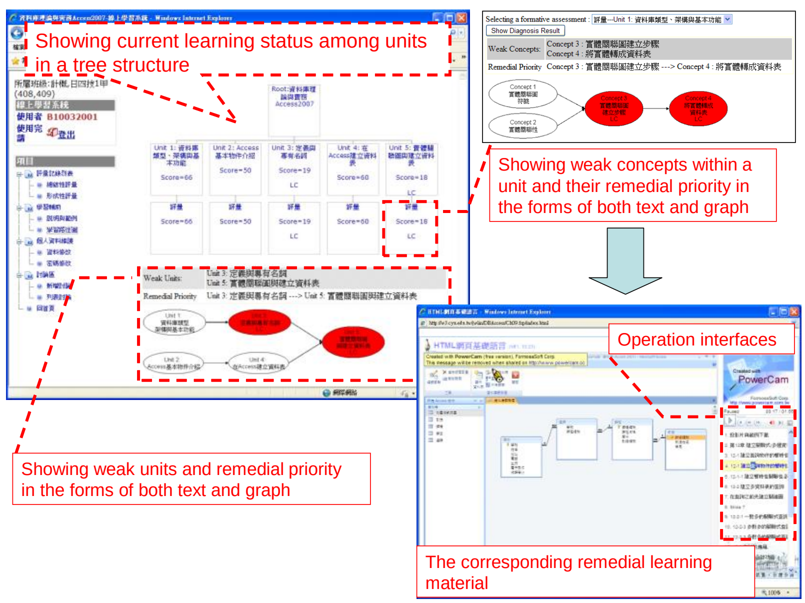


Figure 7. Snapshots of across-unit diagnostic feedback

Educational evaluation

Objectives

This evaluation compares the proposed system, the AUDFS, with two other systems, the Single-Unit Diagnostic Feedback System (SUDFS) and Traditional Feedback System (TFS). Once an assessment is completed, all three systems give students performance results, including questions, answers, correct answers, and scores. Notably, students using the AUDFS can receive across-unit diagnostic feedback (Fig. 7). Students using the SUDFS only receive diagnostic feedback for one unit. Students using the TFS do not receive diagnostic feedback, but only information about whether their answer was correct. The three systems were used by three first-year classes at a university, with an approximate enrollment of 160 students. The first class, consisting of 52 students, used the

AUDFS. The second class, comprising 52 students, used the SUDFS. The third class, consisting of 55 students, used the TFS.

This work (1) compares the learning achievements of these classes, and (2) compares student retention time among classes and within the AUDFS class. Finally, a questionnaire was administered and 12 students were interviewed.

Research tools and procedure

The experiment course is a computer science course called “Database Theory and Application—Microsoft Access 2007.” The course goal is to teach students about databases and Microsoft Access 2007. The experiment subject has five units: “Relational Database Concept Introduction”; “Access Object Introduction”; “Entity-Relationship Diagram (ERD)”; “Building a Relational Access Table”; and “Query for Relational Tables.” The upper part (a) of Fig. 8 shows the sequential relationships among the five units within the experiment subject while the lower part (b) of Fig. 8 shows an example of the sequential relationships among the concepts within a unit, the Unit 3. In total, 12, 13, 15, 14, and 13 multiple-choice questions were generated for units 1–5, respectively, to establish question banks. The test content in each assessment primarily originates from teaching materials.

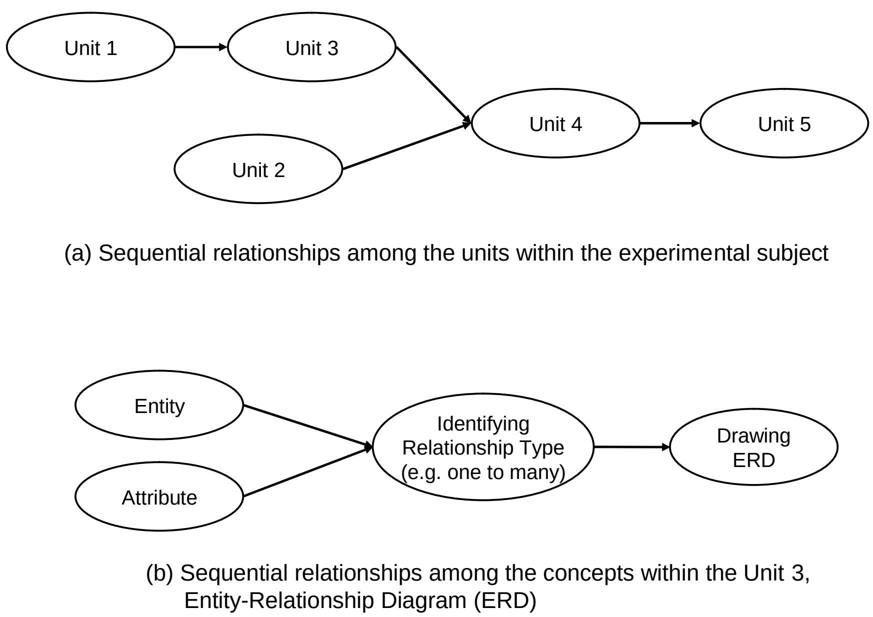


Figure 8. Sequential relationships among units and among concepts within a unit

This work adopted a quasi-experimental design. The experiment treatment was 2 months, 2 hours weekly. Before the experiment started, all classes were trained and practiced using their designated system for 1 week. Students in the AUDFS class were advised to utilize its novel functions (Fig. 7), especially its capabilities to identify weaknesses in current and prior units and provide corresponding remedial materials. During the experiment, all classes adhered to the same teaching and assessment schedules. Upon completion of one unit, all classes took the same assessment of that unit. The posttest content came from teaching materials and assessments. Students who did not complete assessments had no historical records in their designated system. Before the posttest, they also had no historical records of assessments in their personal learning profile to review. This motivated students to attend each assessment.

To verify whether pretest scores of students in these classes differed significantly, a pretest of background knowledge was conducted before the evaluation. To assure pretest validity and reliability, two experts reviewed the pretest, which was then tested by 56 students. Inappropriate questions were removed according to their difficulty and discrimination levels, leaving 16 multiple-choice questions with a Cronbach’s α of 0.83.

The validity and reliability analyses of the posttest were the same as those of the pretest, leaving 22 questions with a Cronbach's α of 0.82. Most questions in the posttest were well-structured problems; for instance, a question was: "How do you transfer an existing ERD of a small business to relational Access tables?" The question was followed by four options, of which only one was correct. To answer this question, one must have an understanding of connected concepts across several units, such as the concepts of "Access Objects" and "Table Object" in Unit 2, "Entity," "Attribute," and "Relationship" in Unit 3, and "Creating Tables," "Assigning the Primary Key," and "Establishing Relationships" in Unit 4. Thus, completing a posttest question required knowledge across different units.

Data collection and analysis

To analyze participant preferences, all systems recorded participant activities as logged data, including login time, activity types (i.e., reviewing or testing), source IP (i.e., locations such as home or school), and stay period (i.e., the time a visitor spends in the system). The SPSS program was used for statistical analysis.

One-way analysis of variance (ANOVA) was applied to examine the effects of these systems on each assessment (Table 2). No significant differences existed in mean scores for units 1 and 2 among the three classes. This is reasonable because units 1 and 2 are both preceding units (Fig. 8) of the subject. At this time, the function of the AUDFS is the same as that of the SUDFS, as the AUDFS can perform for only one unit. For Unit 3, although the result does not reach significance ($F = 1.39$, $p > .05$), the mean score of the AUDFS class is higher than that of the other two classes. At this time, the positive effect of the AUDFS starts to emerge. For Unit 4, the mean score of the AUDFS class is significantly higher than that of the TDS class. Additionally, the mean score of the AUDFS class is higher than that of the SUDFS class, even though the result is not significant. For Unit 5, the mean score of the AUDFS class is significantly higher than that of the SUDFS and TDS classes. This is because addressing Unit 5 demands a larger scope, requiring concepts from units 1–5, compared with Unit 4, which requires concepts from units 1–4. The AUDFS gradually showed its effectiveness in the latter units because understanding concepts in these latter units required more concepts than preceding units. The AUDFS identifies weak concepts within the current unit and within preceding units, further offering remedial material. These features benefit learners in solving well-structured problems, which require the comprehension of a series of concepts in different units. In contrast, the SUDFS identifies only weak concepts within one unit, which may limit learner capability to solve well-structured problems. The TFS is the weakest system because it does not provide feedback support.

Table 2. Results of one-way ANOVA

Assessment #	Class	Mean	*SD*	*F*	Post Hoc[a]
Unit 1	AUDFS	37.34	17.87	1.10	N/A
	SUDFS	33.38	15.12		
	TFS	36.93	12.22		
Unit 2	AUDFS	40.07	20.56	0.79	N/A
	SUDFS	40.51	17.02		
	TFS	44.24	20.16		
Unit 3	AUDFS	52.58	21.18	1.39	N/A
	SUDFS	48.31	19.20		
	TFS	45.88	20.15		
Unit 4	AUDFS	54.89	19.65	3.77*	AUDFS > TFS
	SUDFS	49.02	18.50		
	TFS	44.13	28.91		
Unit 5	AUDFS	54.42	19.19	3.88*	AUDFS > SUDFS
	SUDFS	45.33	20.18		AUDFS > TFS
	TFS	42.07	20.94		

[a] Multiple comparisons: LSD.
[*] $p < .05$.

Analysis of learning achievement among classes

The paired-samples *t*-test was applied to compare learning achievement among classes. All students in these three classes significantly improved on their posttest (Table 3). One-way ANOVA was further applied to the pretest to verify whether background knowledge of the three classes differed significantly. No significant difference existed in background knowledge among the three classes ($F = 1.25$, $p > .05$).

Table 3. Results of the paired-samples *t*-test

Class	Test	N	Mean	*SD*	*t*
AUDFS	Pretest	52	37.42	18.03	-16.73*
	Posttest		78.29	12.08	
SUDFS	Pretest	52	32.94	14.93	-14.14*
	Posttest		70.08	17.57	
TFS	Pretest	55	36.13	11.12	-9.77*
	Posttest		64.42	18.47	

[*] $p < .05$.

One-way Analysis of Covariance (ANCOVA) was also applied to compare learning achievement among classes. The analysis regarded the experiment treatment as the independent variable, posttest score was the dependent variable, and pretest score was the covariate. Before analyzing covariance, homogeneity of regression coefficients was tested to examine whether intra-group homogeneity existed. The SPSS analysis demonstrates that the *F* value of regression coefficients was 1.16 ($p > .05$); thus, the hypothesis of homogeneity was accepted. Thus, covariance analysis was further conducted.

Posttest scores were adjusted by removing the influence of the pretest from posttest scores. Pretest scores had a significant effect on the posttest scores ($F = 8.42$, $p < .05$) (Table 4). Learning achievement of the three classes differed significantly ($F = 9.61$, $p < .05$), indicating that a great difference in learning achievement existed among these three classes.

(1) Students in the AUDFS class performed significantly better than students in the SUDFS class and TFS class. For a well-structured problem, an error in one step can carry over to subsequent steps, and, consequently, to the final solution (Corbalan, Paas, and Cuypers, 2010). That is, a mistake can propagate over an entire problem-solving process. The concepts required to solve a well-structured problem usually span different units. The SUDFS can only identify misconceptions within one unit; in contrast, the AUDFS can identify all misconceptions in different units as well as the source of the misconceptions, and provide recommended remedial paths and material. Students using the AUDFS can ideally obtain a deeper understanding, progressively comprehend the correlation across different units, and acquire a better ability to solve well-structured problems.

(2) Students using the SUDFS performed better than those using the TFS. Wang (2008) stated that web-based assessment would be more effective if it could provide strategies and feedback to learners. The TFS only provides the correct or incorrect answer, which is insufficient for overcoming learning barriers, limiting students' problem-solving ability, and encouraging rote memorization.

Table 4. One-way ANCOVA on posttest scores

Variable	Class	Mean[a]	*SD*	*F*	Post Hoc[b]
Pretest				8.42*	N/A
Type of System	AUDFS	77.81	2.21	9.61*	AUDFS > SUDFS *
	SUDFS	70.71	2.22		AUDFS > TFS *
	TFS	64.26	2.15		SUDFS > TFS *

[a] Covariates appearing in the model are evaluated at Pretest = 35.51.
[b] Adjustment for multiple comparisons: LSD (equivalent to no adjustments).

Analysis of retention time for the three classes

Retention time of each student, which was the total time a student spent on his/her system during the evaluation period, was computed. This value was calculated by accumulating the staying time after every login. Table 5 shows that retention time result is significant (F = 4.06, p > .05). The mean retention time in the AUDFS class (mean = 150.02) is significantly higher than that of the TFS class (mean = 93.72). This is likely because the AUDFS offers more learning assistance and enables a sustained connection between students and the system than the TFS. However, the retention time of the AUDFS class (mean = 150.02) is not significantly higher than that of the SUDFS class (mean = 118.15). This is because students using the SUDFS also feel that their system help their learning.

Table 5. One-way ANCOVA on retention time

	Class	Mean	*SD*	*F*	Post Hoc[a]
Retention Time (in Minutes)	AUDFS	150.02	124.56	4.06*	AUDFS > TFS *
	SUDFS	118.15	101.35		
	TFS	93.72	83.85		

[*] p < .05.

In the AUDFS class, students whose posttest scores were above the average were allocated to the Posttest High-cluster (PostHC) group; otherwise, they were allocated to the Posttest Low-cluster (PostLC) group. The independent samples t-test was applied to comparison results (Table 6). For retention time, the mean of the PostHC group (mean = 262.14) is significantly higher than that of PostLC group (mean = 84.35) (t = -2.40, p < .05). Moreover, the Pearson coefficient was generated to assess the strength of the correlation between the posttest scores and retention time. A significant positive correlation existed (correlation coefficient = 0.36, p < .01), showing that retention time is moderately conducive to learning achievement.

Table 6. Results of the independent samples *t*-test

	Group [a]	N	M	*SD*	*t*
Retention time (Minutes) within the AUDFS class	PostHC	27	262.14	371.77	-2.40*
	PostLC	25	84.35	74.46	

[a] Because two students' posttest scores were the same, the two groups did not have the same number of people.
[*] p < .05.

Questionnaire and interview

To understand student satisfaction, a questionnaire with a Likert scale ranging from 5 for "*strongly agree*" to 1 for "*strongly disagree*" was given to all classes at study end. Among 52, 52, and 55 students in the AUDFS, SUDFS, and TFS classes, 49, 50, and 51 valid questionnaires were collected for data analysis. A questionnaire with all answers empty or all answers the same was deemed invalid. After completing the questionnaire, short interviews of 12 randomly selected students were conducted to elicit their feelings.

Table 7 and figure 9 show the questionnaire questions and results, respectively, revealing that all systems received positive feedback for most evaluated aspects. For questions 1–3, most students felt their designated system was stable and convenient. However, the remaining answers (question 4-10) show that the AUDFS class was more satisfied than the other two classes, likely because the AUDFS has more functions for learning assistance. The following focuses on questionnaire and interview results for the AUDFS class.

Table 7. Questionnaire questions

#	Question Content
1	The designated system provided a convenient environment (*e.g.*, user-friendly interfaces).
2	The designated system had good stability (*e.g.*, quick response).
3	You are satisfied with the operation of the designated system.
4	Assessments helped you identify your weaknesses.
5	The designated system assisted and reinforced your weaknesses.
6	The designated system benefited learning new concepts.
7	The designated system helped you understand the course scope and correlation among units.
8	You usually logged in to the designated system for review in your daily life.
9	You logged in to the designated system for review before the post-test.
10	The designated system significantly improved your learning.

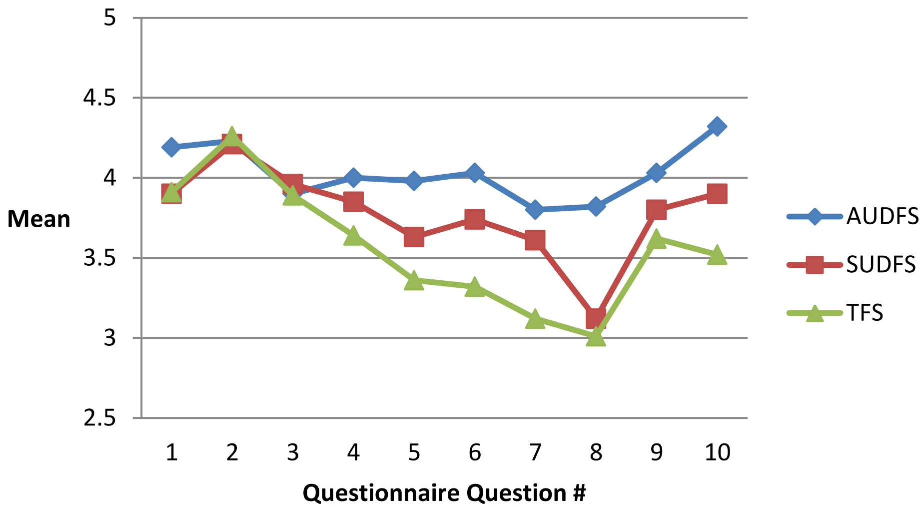


Figure 9. Questionnaire results among the three classes

Answers to questions 1–3 show that most students were satisfied with the AUDFS. For example, one interviewee stated, "*The user interfaces were simple and friendly. As a novice, it was easy for me to catch on by first or second use.*" Another interviewee expressed, "*The system was pretty stable and convenient and I could review at school or home.*"

Answers to Question 4 show that most students felt that the AUDFS is a practical auxiliary tool for identifying learning weaknesses. For example, one interviewee stated, "*The system helped me identify where my weaknesses were.*" Two interviewees said that they listened to material in class and the system told them which parts of unit they did not understand.

Answers to Question 5 show that the AUDFS identified weaknesses. Informing learners of their weaknesses and providing coaching are important (Jonassen, 1997). One interviewee stated, "*For every weakness, I can instantly look at the corresponding PowerCAM remedial material and replay it if necessary. The remedial material was easy to understand because it was illustrated by video.*" One interviewee said, "*Although the system can spot my weak concepts in each unit, it was hard for me to figure out the better remedial path, especially when there were too many weak concepts or the learning scope was huge. Thanks to the AUDFS, I can follow the provided guide to strengthen my weaknesses step by step.*"

Answers to questions 6 and 7 show that the AUDFS benefited learning new concepts, understanding the learning scope, and correlation among units. One interviewee commented, "*Sometimes I tried to strengthen my weak concepts in the latter unit (e.g., Unit 5) by directly reading the teaching material of that unit, but I was still confused. However, the AUDFS unveiled which of my weaknesses in the preceding units should be remedied first. I followed the instructions and they worked!"*

When investigating whether the AUDFS stimulated students to spend time in their daily life or before the posttest (questions 8 and 9), some said they reviewed in daily life, while some admitted that they reviewed only before the posttest. Although students welcomed the AUDFS, some suffered from excessive loads from other subjects, resulting in insufficient time for review. One interviewee stated, "*I regularly logged in to the system to review after school. I cannot stand reviewing at the last moment (*i.e., *only before the posttest), because I have to prepare too many subjects at that time.*" Another one stated, "*I reviewed only before the posttest because this strategy gave me a deeper impression and better memory.*" From answers to Question 8, one can see that students were willing to spend time on the system.

Question 10 obtained the highest score, directly reflecting the fact that the students in the AUDFS class felt that the overall system significantly improved learning.

Conclusions

This work provides diagnostic feedback for a large learning scope through a series of online assessments that detect weak concepts within an assessment and across assessments. The deployment of assessments is based on the conceptual knowledge structure for a subject. Through the recommended remedial learning path and material, students can solidify their foundations for concepts in the current unit and across other units, which can enhance the students' ability to solve well-structured problems.

Experiment results show that the AUDFS is conducive to learning subsequent units. The mean post-test score in the AUDFS class is significantly higher than that in the SUDFS and TFS classes, indicating that the AUDFS significantly enhances student learning and ability to solve well-structured problems. The mean retention time in the AUDFS class is significantly higher than that of the TFS class, indicating the AUDFS enables a sustained connection between students and the system. Within the AUDFS class, mean retention time of the Post-HC is significantly higher than that of the Post-LC. Retention time and post-test score were positively correlated. The questionnaire and interview results reveal that most students using the AUDFS have positive attitudes toward all evaluated aspects.

The proposed system should fit such subjects as mathematics and programming courses for the following two reasons. First, these courses have fixed learning sequences and a clear concept structure, which can represent course content into a tree structure (Fig. 1). Second, typical questions in these subjects are well-structured problems, requiring across-unit concepts to solve. However, this assumption needs further verification.

As the online learning environment is characterized as autonomous, the ability of learners to engage in self-regulated behaviors is construed as a crucial factor for successful online learning (Barnard, Lan, To, Paton, and Lai, 2009). A self-regulated learner can play an active role in learning, setting task-oriented and proper goals, taking responsibility for their own learning, monitoring their own learning, and maintaining their own learning motivation (Heikkilä and Lonkab, 2006; Wang, 2011). This ability for self-regulation is important, because online learning requires learners to be more disciplined; they also require considerable persistence and determination (Shea and Bidjerano, 2012). Kauffman (2004) and Wang (2011) also asserted that the learners in an online learning environment must be highly self-regulated; otherwise their learning effectiveness may be low. Thus, in the future, this study will further verify whether the self-regulation level of a student significantly influences his/her learning behavior and learning achievement in the AUDFS.

References

Ausubel, D. (1968). *Educational psychology: A cognitive view.* New York, NY: Holt, Rinehart, and Winston.

Barnard, L., Lan, W., To, Y., Paton, V., & Lai, S. (2009). Measuring self-regulation in online and blended learning environments. *Internet and Higher Education*, *12*(1), 1–6.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school* (expanded ed.). Washington, DC: National Academies Press.

Buchanan, T. (2000). The efficacy of a World-Wide Web mediated formative assessment. *Journal of Computer Assisted Learning, 16*(3), 193–200.

Chen, C. M., Hsieh, Y. L., & Hsu, S. H. (2007). Mining learner profile utilizing association rule for web-based learning diagnosis. *Expert Systems with Applications, 33*(1), 6-22.

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.

Chen, L. H. (2011). Enhancement of student learning performance using personalized diagnosis and remedial learning system. *Computers & Education, 56*(1), 289-299.

Corbalan, G., Paas F., & Cuypers, H. (2010). Computer-based feedback in linear algebra: Effects on transfer performance and motivation. *Computers & Education, 55*(2), 692-703.

Fisher, C., Gleitman, H., & Gleitman, L. R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology, 23*(3), 331–392.

Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology: Research and Development*, *45*(1), 65–94.

Heikkilä, A., & Lonkab, K. (2006). Studying in higher education: Students' approaches to learning, self-regulation, and cognitive strategies. *Studies in Higher Education, 31*(1), 99–117.

Hwang G. J. (2003). A conceptual map model for developing intelligent tutoring systems. *Computers & Education*, *40*(3), 217-235.

Kauffman, D. F. (2004). Self-regulated learning in web-based environments: instructional tools designed to facilitate cognitive strategy use, meta-cognitive processing, and motivational beliefs. *Journal of Educational Computing Research, 30*(1&2), 139–161.

Laxman, K (2010). A conceptual framework mapping the application of information search strategies to well and ill-structured problem solving. *Computers & Education, 55*(2), 513-526.

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.

Lin, Y. C., Lin Y. T., & Huang Y. M. (2011). Development of a diagnostic system using a testing-based approach for strengthening student prior knowledge. *Computers & Education, 57*(2), 1557-1570.

Liu, P. L., Chen, C. H., & Chang, Y. J. (2010). Effects of a computer-assisted concept mapping learning strategy on EFL college students' English reading comprehension. *Computers & Education, 54*(2), 436-445.

Miller, S.M., & Miller K.L. (1999). Using instructional theory to facilitate communication in web-based courses. *Educational Technology & Society, 2*(3), 106-114.

Pachler, N., Daly, C., Mor, Y., & Mellar, H. (2010). Formative e-assessment: Practitioner cases. *Computers & Education, 54*(3), 715-721.

Shea, P., & Bidjerano, T. (2012). Learning presence as a moderator in the community of inquiry model. *Computers & Education*, *59*(2), 316-326.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153-189.

Wang, M., Peng, J., Cheng, B., Zhou, H., & Liu, J. (2011). Knowledge visualization for self-regulated learning. *Educational Technology & Society, 14*(3), 28–42.

Wang, T. H. (2008). Web-based quiz-game-like formative assessment: Development and evaluation. *Computers & Education*, *51*(3), 1247-1263.

Wang, T. H. (2011). Developing web-based assessment strategies for facilitating junior high school students to perform self-regulated learning in an e-learning environment. *Computers & Education*, *57*(2), 1801-1812.

Zhuge H., & Li, Y. (2006). Learning with active e-course in knowledge grid environment. *Concurrency and Computation: Practice and Experience*, *18*(3), 333-356.