

Designing a Diagnosis Mechanism of Misconceptions for Providing Learning Feedback to Students based on Knowledge Map

Shao-Chun Li¹, Jyh-Cheng Chang², Maiga Chang³

^{1,2}Dept. of Information and Computer Engineering, Chung-Yuan Christian Univ.

³National Science & Technology Program for e-Learning in Taiwan

^{1,2}#200, Chung-Pei Rd. Chung-Li, 32023

³#300, Jhongda Rd., Chung-Li, 32001

TAIWAN

http://maiga.dnsalias.org/maiga/resume_e.htm

Abstract: - There are many researchers applied either concept map or concept hierarchy to diagnose the learning effects of students via test. After a student completes the test, diagnosis mechanism can figure out what concepts the student may misunderstand according to the answer sheet. However, the diagnosis mechanism always only provides supplemental materials of the concept terms without a complete learning sequence. This paper integrates Knowledge Map and Information Retrieval to develop a three steps diagnosis mechanism in order to provide a supplemental learning sequence to students.

Key-Words: e-Learning, Concept Map, Knowledge Map, Learning Diagnosis, Learning Feedback

1 Introduction

With the growth and development of the computer and Internet, the teaching activities in the Internet are more convenient and popular and called distance learning. In recent years, many researchers mentioned the Computer-Assisted Instruction (CAI) [2] [17] [22]. The purpose of CAI is to assist the teacher to improve the instructions of course, however, it still has a lot of issues which are faced by teachers when they are applying technologies into their courses and classes. Additionally, lots of scholars are devoted to the development of on-line instant feedback test to replace the traditional paper pen test.

Learning process has their gradations, for example step by step or from easy to difficult, every gradation of learning is to prepare for the next gradation and the qualitative changes during the learning process will also happen [18]. In order to facilitate learning effectively, the main principle of adaptive learning is to diagnose students by their answers of the test and arrange the follow-up teaching materials in accordance with the learning achievement of students.

With the purpose of improve the learning effects of students on Internet, the methods of measurement, tracking, and evaluation are used to discover the learning states of students and decide the suitable teaching materials, it is so-called the performance technology of learning [6] [16]. According to Clark and Langdon's opinions, the suitable and effective teaching materials can facilitate learning more

efficiency. Moreover, interaction is also one of the indispensable characteristics in teaching and the main shaft designed of interaction system on Internet is the test and commenting of achievement [21].

There are two kinds of applications which apply the computer technology to the test [1]: (1) Build and construct a test (2) Use the computer as the media of the examination. The two kinds of applications can be saw as active and passive applications simply. The passive application is always recording the processes, preserving and reporting the statistics and learning results of students [5]. The active application will guide students learning, diagnose the misconceptions, and prevent students from making a mistake.

A good test system not only can assess learning effects of students, but also can diagnose the learning barrier of students and help students to overcome those obstacles. Therefore, a good test system must reach two major objectives: (1) analyzes the basic obstacles of students while they are learning, (2) settles a learning sequence to improve the academic achievement of students. In order to reach these two objectives, this paper applies the knowledge map to the distance learning environment in order to diagnose students' misconceptions and generate the learning sequence of supplemental teaching materials after test. Beside that, the diagnosis mechanism and the feedback system will also make the full-loop learning completely.

The knowledge map incorporates the concept map with concept schema and use for knowledge

representation is proposed by Kuo et al. in 2002. In education, the concept map often used to diagnose the learning achievement and the students' mental model [19]. The knowledge map has two advantages: (1) represent the relations between concepts clearly just like the concept map; and, (2) represent the related attributes of concept which are kept in the concept schema.

The problem proposed by this paper is formulated in Section 2, and Section 2 will also describe some research backgrounds such like Concept Map and Knowledge Map on WWW. Section 3 will analyze the diagnosis mechanism and give a complete example. Algorithms in the three-steps diagnosis mechanism are designed in Section 4. In Section 5, the experiment and the evaluation results are discussed. Section 6 makes a simple summary and describes the possible future researches.

2 Concept Map and Knowledge Map

2.1 Definitions of Concept Map

Concept maps are logic-based knowledge representation formalism, whose elements belong either to concepts or relations. The technique was developed by Professor Joseph Novak at the Cornell University in 1960s. Concepts within a designated domain are given as a concept set $C = \{c_i\}$. For example in Chemistry the concept set of elements, Celement, might be Celement = {SUBSTANCE, METAL, NONMETAL, IRON, MERCURY, OXYGEN, SULFUR, j-}.

The relation between two concepts, c_i and c_j , can denote as $\rho_k = \{c_i, c_j\}$, where k indicates the specific category in the relation category set and the relation set is $\rho = \{\rho_k\}$. The relation category is also a set, that is, $k = \{\text{part of, type of, leads to, analogy, characteristic, evidence}\}$ according to previous researches [8] [9] [12] or $k = \{\text{clustering, aggregation, linkage}\}$ proposed by Meyer in 1985 and so on. Proposition is usually used to represent the meaning of a sentence, it can correspond to a Boolean value, that is true or false.

A proposition describes the relation between concept could be written as $\phi_i(\rho_k(c_i, c_j, \dots)) \rightarrow V = \{\text{true, false}\}$. For example, "the ACCELERATION of GRAVITY is 9.8" could be written as $\phi_i(\rho_{is}(\text{ACCELERATION OF GRAVITY}, 9.8)) \rightarrow \text{true}$.

The knowledge not only can help students to understand the concepts in the specific domain, but

also can guide students during the learning process. We can find the misconceptions of the student by diagnosing the change of the concept map which is stored inside the student's mind [10] [13] [15].

2.2 Definitions of Knowledge Map

Knowledge map is a knowledge representation formalism comprising with concept hierarchy and concept schema [14]. Concept hierarchy represents knowledge in a hierarchical way. Collin and Quillian indicated that human organize knowledge in a structure way and store characteristics of the concept in different level of the structure [7]. The concept hierarchy is defined as $CH(C, \rho)$, where C is the concept set and ρ is the relation set. Fig. 1 is a example of concept hierarchy.

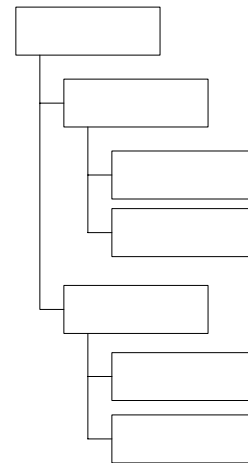


Fig. 1. Concept Hierarchy

In cognitive psychology, schema is used to identify messages from the real world and store the attributes of the message. Schema is a high level knowledge representation formalism integrate with nature, events and text [11]. Concept schema is used to describe the attributes of a concept. The concept schema of a concept set C is defined as: $CS(C) = \{\rho_k(c_i, c_j)\}$. The following Table 1 is the concept schema of elements in the Chemistry domain.

Table 1. Concept Schema

Element Concept Set	Status Concept Set
SUBSTANCE	SOLID
SUBSTANCE	LIQUID
SUBSTANCE	GASSEOUS
METAL	SOLID
METAL	LIQUID
METAL	GASEOUS
.....
OXYGEN	GASEOUS
MERCURY	LIQUID

Once we have the concept hierarchy and the concept schema, a knowledge map for the specific domain (or a concept set) can be integrated as Fig. 2 shown below.

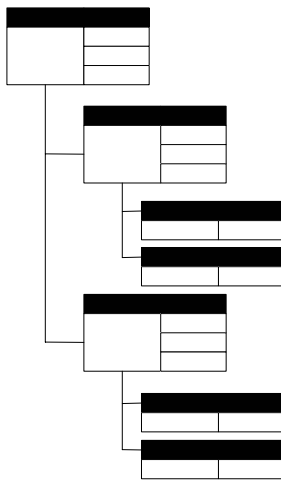


Fig. 2. A Knowledge Map for Specific Domain

3 Learning Diagnosis and Feedback

This research tries to diagnose the learning achievement of students on the Internet-based learning environment and offer the suitable learning feedbacks in sequence according the learning achievement. Before the analysis is made, there are three assumptions needed to list for this research.

- ASSUMPTION 1: In the test, student gets the wrong answer of an item means student has missing some relevant concepts in the item.
- ASSUMPTION 2: In the concept hierarchy of knowledge map, the concept in the upper (more abstract) level will involve those concepts in the lower (more specific) level.
- ASSUMPTION 3: When a concept becomes a barrier to the student, the relevant concepts will also become barrier to the student.

In order to find the student's misconception, the teaching materials and test items must be analyzed first. Then student can make test on Internet and the system collect the test result. Finally, system according to the test result and teaching materials selected suitable teaching material for remedy learning and represent as a learning path. The learning diagnosis and feedback step is described below.

3.1 Analysis in teaching materials with knowledge map

According to the knowledge schema and concept hierarchy, the teaching materials could be analyzed and reorganized. After analyzing the teaching materials (web pages), a matrix is created which is called the embedded concept matrix (*ECM*).

Every teaching material, no matter what representation format, contains concepts that related to the course, therefore, the embedded concepts (*EC*) of teaching material can be obtained via analyzing by the domain experts. The association between concept c_i and teaching material d_j is denoted as $EC(c_i, d_j) \in [0,1]$, high value means high association. Embedded concept matrix, $ECM = [EC(c_i, d_j)]$, represents the concept association between documents, and the value of $EC(c_i, d_j)$ is given by expert or calculated by content analysis method. The test items also can be analyzed and found the embedded concept and represent as *ECM* with the same way.

3.2 Misconception diagnosis via online test

Moreover, since there are association among concepts in both concept map and knowledge map, the cumulative embedded concept is then constructed for representing the concept hierarchy. Also the domain experts (teachers) will take responsibility to assign the concept thresholds and hidden concepts for a test. According to the ASSUMPTION 2 and ASSUMPTION 3, the cumulative embedded concept matrix (*CECM*) is used to introduce the relations between the concepts in the knowledge map by using the original *ECM* matrix and the concept hierarchy. Hence, *CECM* can come from *CHM* and *ECM*, it define as (1) and (2).

$$CHM(C, C) = [\rho(c_i, c_j)] \forall c_i, c_j \in C \quad (1)$$

$$CECM(C, Q) = (CHM(C, C) + I(C))^T \cdot ECM(C, Q) \quad (2)$$

According to *CECM*, the distribution of concepts in all items can be found by (3).

$$DCECM(i, j) = CECM(i, j) / \sum_{j=1}^{n} CECM(i, j) \quad (3)$$

3.3 Misconception diagnosis via online test

Once the student completes the test, the misconception diagnosis mechanism can compute the distribution of misconception for the student and judge which concept belongs to misconception according to the threshold. The value of threshold, which given by domain expert, is between 0 and 1, that, the misconception value of a student higher than the threshold means the student get miss correspondent concept.

The first step of the process of diagnosis is comparing *CECM* with the student's answer matrix $A = [a_i] \in [0,1]$ in test; $a_i = 0$ if the student's answer is

SOLID
NONMETAL LIQUID
GASEOUS

right and $a_i=1$ if the student's answer is wrong. After that the misconception degree matrix Φ_s of students can be found from the student answer matrix A and $DCECM$ as (4).

$$\Phi_s = A \cdot DCECM^T \quad (4)$$

3.4 Learning sequence generation

This research is trying to create a learning sequence based on the recommendatory weights of teaching materials. The feature vector for each teaching material will be retrieved at first, and then the similarity between misconception degree of student and ECM of teaching material is used for selecting supplemental materials. With the weight order of supplemental materials, a remedy learning sequence will be generated automatically.

The concepts are involved in a teaching material which similar to the students' misconception will be recommended as the supplemental materials. The calculation of similarity is defined as (5).

$$sim(\Phi_s, ECM(C, d_j)) = \frac{\Phi_s \bullet ECM(C, d_j)}{|\Phi_s| \times |ECM(C, d_j)|} \quad (5)$$

According to the threshold of misconception and the students' misconception, the system can select teaching materials for each student as supplemental materials. The learning path in this research is representation in a catalog way and sorted according to similarity. Students can learning sequence suggested by system or selected by themselves.

4 Evaluation Methodology

This paper uses the terms of statistics includes the Precision, the Recall, and the F-measure, as the measurement of evaluating the effects of recommendatory materials [20]. The Precision (P), the Recall (R), and the F-measure (F) between concepts in the misconception t and the teaching material c are:

$$P_{c,t} = \frac{N_t \cap N_c}{N_c} \quad (6)$$

$$R_{c,t} = \frac{N_t \cap N_c}{N_t} \quad (7)$$

$$F_{c,t} = \frac{2}{\frac{1}{P_{c,t}} + \frac{1}{R_{c,t}}} \quad (8)$$

The evaluation methodology takes the percentage of recommendation errors, that are, the Type I error and the Type II error as Table 2 listed below. In statistics the Type I error means data retrieved is neither expected nor full covered and the Type II error means data retrieved is more than needs. Therefore, in this work the Type I error is the system

provides concepts that belongs to the student's misconception but not exist in the recommendatory teaching materials, and the Type II error is the system provides concepts that exist in the recommendatory teaching materials however do not belong to the student's misconception.

Table 2. Error Type

	CONCEPT IN RECOMMENDATORY TEACHING MATERIAL	CONCEPT NOT IN RECOMMENDATORY TEACHING MATERIAL
CONCEPT BELONGS TO STUDENT'S MISCONCEPTION	CORRECT	Type I error (assume true)
CONCEPT NOT BELONGS TO STUDENT'S MISCONCEPTION	Type II error (assume false)	CORRECT

5 Experiment and Results

The experiment was done with the database course in the Dept. of Information and Computer Engineering, Chung-Yuan Christian University. There were 42 students participated this experiment, there were 2 graduate students, 15 senior students and 25 junior students among the participators.

The average value of Precision, Recall and F-measure is 0.8569, 0.9255 and 0.8868. The evaluation results showed that the recommendatory teaching materials are suitable for remedying the misconception of the students.

In this experiment, the Type I error, α , is the misunderstanding concepts of student but have not included in suggested materials. When the P-value = 1 and the R-value < 1, the Type I error occurs. Similarly, the Type II error, β , is there are some concepts in suggested materials not belongs to the misconception of students. When the R-value = 1 and the P-value < 1, the Type II error happens. In the statistical hypothesis the value of Type I error should be lower than 0.05 and the value of Type II should be also lower than 0.2.

The Type I error, α , in this experiment is 0.0476, it is quite low and it means the selected teaching materials can cover all concepts for the teaching course. Because in the experiment, the teaching materials can not cover all concepts of the examination, the α would be reduced if the teaching materials could cover every concepts of test. The Type II error, β , in this experiment is 0.4524 which is greater than 0.2. The reason causes the value is so high is because the each unit of teaching materials in

this experiment contains too many concepts. If the teaching materials could be subdivided furthermore, the Type II error might be reduced.

6 Conclusion

Different student has different knowledge, learning states and learning achievement in the Internet-based learning environment. The diagnosis mechanism is designed in this work and used to provide sequential remedy materials based on the adaptive learning according to students' answers of test. After this experiment, the evaluation of the recommendatory teaching materials has been made. The evaluation results are good but also reveal some limits. The limits are: (1) students do not follow the learning sequence which generated by the system, although the system is an adaptive system; (2) students do not interest in the Internet-based learning. Beside the limits, there are also some further researches should be done: (1) the system should also analyze the learning effect of the student after repeatedly learning; (2) the system should be able to provide feedback to the teacher in order to let the teacher changes his/her teaching strategies; and, (3) the system should be able to discover a more useful learning sequence based on the theories of traversal pattern and learning path.

References:

- [1] S. M. Alessi, S. R. Trollip, *Computer-based instruction: Methods and development*, Englewood Cliffs, NJ: Prentice-Hall, 1991.
- [2] C. Bloom, F. Linton, B. Bell, Using Evaluation in the Design of an Intelligent Tutoring System, *Journal of Interactive Learning Research*, Vol. 8, No. 2, 1997, pp.235-275.
- [3] P. Brusilovsky, Methods and Techniques of Adaptive Hypermedia, *User Modeling and User-Adapted Interaction*, Vol. 6, No. 2-3, 1996, pp. 87-129.
- [4] P. Brusilovsky, Adaptive Hypermedia, *User Modeling and User-Adapted Interaction*, Vol. 11, No. 1-2, 2001, pp.87-110.
- [5] C. Chou, A computer logging method for collecting user-reported inputs during formative evaluation of computer network-assisted distance learning, In *the Proceedings of the 7th World Conference on Educational Multimedia and Hypermedia*, (ED-Media 1996), Boston, USA, June 17-22, 1996, pp.143-153.
- [6] R. C. Clark, Hang up your training hat, *Training and Development*, Vol. 48, No. 9, 1993, pp.61-63.
- [7] A. M. Collins, M. R. Quillian, Retrieval time from semantic memory, *Journal of Verbal Learning and Verbal Behavior*, Vol. 8, 1969, pp. 240-247.
- [8] D. F. Dansereau, The development of a learning strategies curriculum, In O'Neill, H.F. Jr. (Ed.), *Learning strategies*, New York: Academic Press, 1978, pp. 1-26
- [9] D. F., Dansereau, K. W. Collins, B. A. McDonald, C. E. Holley, J. C. Garland, G. Diekhoff, S. H. Evans, Development and evaluation of an effective learning strategy program, *Journal of Educational Psychology*, Vol. 71, 1979, pp. 64-73.
- [10] B. Ferry, J. Hedberg, B. Harper, How Do Preservice Teachers Use Concept Maps To Organize Their Curriculum Content Knowledge? *Journal of Interactive Learning Research*, Vol. 9, No. 1, 1998, pp.83-104.
- [11] E. D. Gagne, C. W. Yekovich, F. R. Yekovich, *The Cognitive Psychology of School Learning*, New York: HarperCollins College Publishers, 1993.
- [12] C. D. Holley, D. F. Dansereau, B. A. McDonald, J. C. Garland, K. W. Collins, Evaluation of a hierarchical mapping technique as an aid to prose processing, *Contemporary Educational Psychology*, Vol. 4, 1979, pp.227-237.
- [13] D. H. Jonassen, T. C. Reeves, N. Hong, D. Harvey, K. Peters, Concept Mapping as Cognitive Learning and Assessment Tools, *Journal of Interactive Learning Research*, Vol. 8, No. 3-4, 1997, pp.289-308.
- [14] R. Kuo, M. Chang, D. X. Dong, J. S. Heh, Applying Knowledge Map to Intelligent Agents in Problem Solving Systems, In *the Proceedings of World Conference on Educational Multimedia, Hypermedia & Telecommunications*, Denver, Colorado, USA, Jun. 24-29, 2002, pp.1053-1054.
- [15] D. Kennedy, C. Mcnaught, Use of Concept Mapping in the Design of Learning Tools for Interactive Multimedia, *Journal of Interactive Learning Research*, Vol. 8, No. 3-4, 1997, pp.389-406.
- [16] D. Langdon, Performance technology in three paradigms, *Performance and Instruction*, Vol. 30, No. 7, 1991, pp.1-7.
- [17] C. Srisethanil, N. Baker, ITS-Engineering Providing Adaptive Teaching in the Engineering Tutor, In *the proceedings of the 25th IEEE Frontiers in Education*, Vol. 1, 1995, 2a3.22-2a3.26.
- [18] L. Sherry, B. Wilson, Transformative Communication as a Stimulus to Web

Innovations, *Web-based Instruction*, Englewood Cliffs, New Jersey: Educational Technology Publications, 1997, pp.67-73.

- [19] J. Turns, C. J. Atman, E. Adams, Concept maps for engineering education: a cognitively motivated tool supporting varied assessment functions, *IEEE Transactions on Education*, Vol. 43, No. 2, 2000, pp.164-173.
- [20] C. J. Van Rijsbergen, *Information Retrieval*, 2nd Edition, London: Butterworth-Heinemann, 1979.
- [21] B. Willis, J. Dickinson, Distance Education and the World Wide Web. *Web-based Instruction*, Englewood Cliffs, New Jersey: Educational Technology Publications, 1997, pp.81-84.
- [22] L. H. Wong, C. Quek C. K. Looi, TAP: A Software Architecture for an Inquiry Dialogue-Based Tutoring System, *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, Vol. 28, No. 3, 1998, pp.315-325.