Proc. Natl. Sci. Counc. ROC(D)
Vol. 9, No. 1, 1999. pp. 1-9

Development of an intelligent Testing and Diagnostic System on Computer Networks†

GWO-JEN HWANG

Information Management Department
National Chi Nan University, Pu-Li Nan-Tou
Taiwan, R.O.C.

(Received March 5, 1998; Accepted May 10, 1999)

ABSTRACT

In recent years, researchers have tried to employ the techniques of artificial intelligence to the development of intelligent computer-assisted instruction system. They also tried to develop more effective programs to test and improve the learning performance of students. However, as conventional testing systems are usually developed for a single user, students seldom have chances to discuss with others during self-assessment. Moreover, it seems to be of little use to present only a score as a test result to a student who really needs learning guidance. In this paper, we propose a network-based intelligent testing and diagnostic system that can give dynamic tests according to the degree of difficulty, discrimination, and other parameters. The system also permits several students to take a test cooperatively through a computer network. With proper discussions, students can solve their problems and will be more willing to learn. Moreover, by applying the techniques of fuzzy reasoning, the system is able to provide learning suggestions that can be used to guide the students to raise their learning status.

Key Words: computer-assisted instruction, world wide web, fuzzy logic, expert systems

I. Introduction

As computer techniques progressed rapidly in recent years, researchers have tried to enable computers to perform intelligent behaviors, which has led to the advance of artificial intelligence (AI) techniques. The issue of developing Intelligent Computer-Assisted Instruction (ICAI) systems has therefore attracted attention from both AI and CAI researchers. However, the problem that conventional computer-assisted instruction systems fail to play the roles of learning partners has encouraged some researchers to turn their attention to other studies. Recently, the popularity of computer networks has enabled the development of distance instructional environments, which solve the problems of making contact with learning partners by means of network communications (Sun & Chou, 1996).

During the instruction process, it is important to evaluate the learning status of each student. Tests are a typical method of evaluation, and it is important to select proper test items, those which fit the characteristics of students accurately and efficiently. In a conventional testing system, the test items for a unit are seldom changed; that is, no matter how many times examinations on the same unit are given, the students always confront the same test items. In that case, students may receive good grades just because they are familiar with the test items, rather than actually learning the unit well. Besides, giving students nothing but their scores on a test will not help them to improve their studies. In this paper, we propose a new system to solve these problems. According to the degrees of difficulty, discrimination, and some other parameters, the system can construct test sheets dynamically; that is, different test items will be given when an identical unit is tested repeatedly. Moreover, we try to give objective evaluations and suggestions to the students by analyzing their answers sheets. Therefore the students are able to understand their

†This study is supported by the National Science Council of the Republic of China under contract number NSC87-2511-S-260-001-ICL.

areas of difficulty and obtain learning suggestions. To offer such feature, a fuzzy expert system approach has been applied. Some features of expert systems, including reasoning ability and the advantage of easy maintain, can be very usefully applied to the development of more intellectual computerized testing systems.

Within the on-line conversation window of the system, students can sort out many concepts through network discussions. While traditional CAI systems only offer one user on a single machine, this system allows students from everywhere to discuss the same test sheet jointly through the network, which can help them to find their blind spots and increase their interests of learning. We have also applied database technology to the implementation and management of item banks. By combining the mechanisms of expert systems, CAI systems and database systems, test sheets and learning guidance can be generated according to the learning status of the students.

II. Relevant Research

Instructional applications of computers have increased rapidly in recent years, so have the computerized testing systems. Taking the GRE as an example, people have taken the test through computers since 1992. The GRE has been offering a computerized form since 1993 and plans to abandon the paper-and-pencil form in 1999. The IBM Co. and Arthur Andersen Co. have begun to work on the development of a computerized testing system. Such systems, which change the form of tests from paper-and-pencil to on-line, are proliferating rapidly.

In New Zealand, three researchers propounded a “Knowledge Based Computer Assisted Instruction System” which can change the numeric part of items when the test is in progress to prevent the students from memorizing answers (Fan, Tina & Shue, 1996). Another branch of relevant researches is the Computerized Adaptive Testing, which applies some prediction methodologies to shorten the length of the test without losing the preciseness (Wainer, 1990; Ho, 1997).

With the fast development of computer networks, people can access information and communicate with others without being constrained by space and time. Through network communications, people can make discussions with others to solve their problems. Therefore, it becomes an important issue to implement testing and practicing systems on computer networks (Hsu, Tu, Yeh, Chu. & Hwang, 1997).

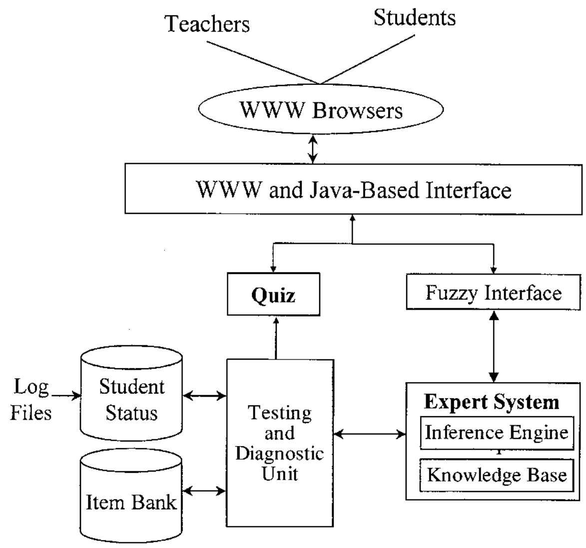

Fig. 1. Structure of the intelligent testing and diagnostic system.

III. The Intelligent Testing and Diagnostic System

As computer networks have become popular, various relevant studies have been conducted, including research on network-based banks, digital libraries, distance education, network-oriented advertisements, etc. In this section, we shall present a network-based intelligent testing and diagnostic system as shown in Fig. 1.

The system includes a student status database, an item bank, a Java-based interface, a testing and diagnostic unit, and a fuzzy interface. In the following, we shall describe each component in detail:

1. The Student Status Database

This contains general information and the learning status of each student. It provides necessary information to the testing and diagnostic unit, which then invokes the fuzzy expert system to construct test sheets by analyzing the learning status of each student. Elements of the student status database include (Hopper, 1992):

A. The basic information of students
B. Records concerning concentration, patience, and willingness which are derived by analyzing the previous learning behavior of the students
C. Records concerning learning path and browsing time for each student
D. The grade level of each student
E. The answer records of each student

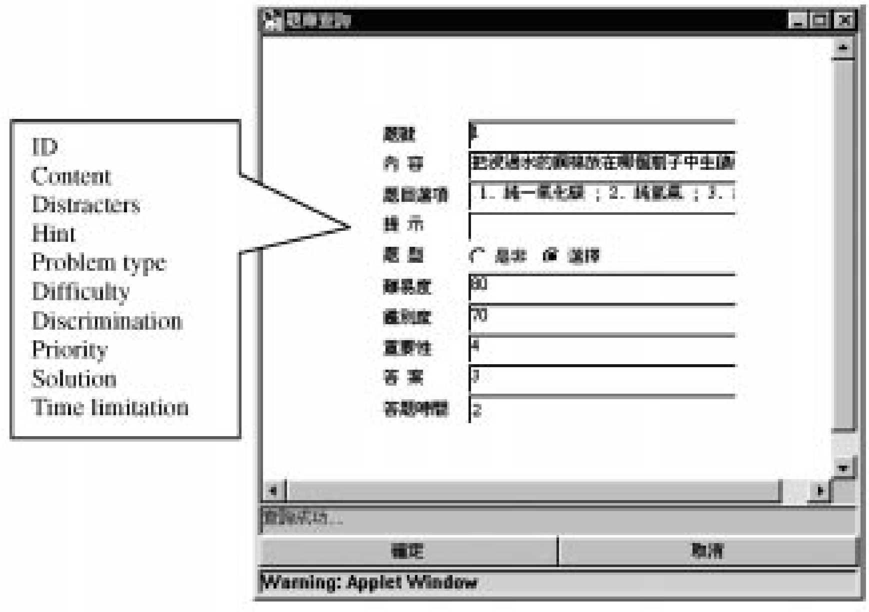

Fig. 2. Parameters of a test item.

1. The Item Bank

This contains the test items and the corresponding parameters as follows:

A. Content: the question of the test item
B. Distracters: a set of candidate choices separated by “;”
C. Hint: a hint to the students concerning the test item
D. Type: an indicator of a Yes/No question or a multiple-choice question.
E. Difficulty: an indicator of degree of difficulty of the test item.
F. Discrimination: an indicator of the degree of discrimination of the test item.
G. Priority: the priority of the test item, which is used by the system to determine when the test item should be selected. A test item with higher priority will be selected more frequently. The legal value of priority ranges from I to 5.
H. Solution: the answer of the test item
I. Time: the estimated time for answering the question.

An example of a test item and its corresponding parameters is given in Fig. 2..

3. The Java-based User Interface

A Java-based user interface is provided so that the users can access the system through WWW browsers. Both registration and login processes are controlled by the user interface. After successfully logging in, a user can adjust the parameters of the test items to take a test or join the on-line discussions. An explanation of the test results is presented to guide the user to further learning programs after doing the test.

4. The Testing and Diagnostic Unit

This unit is used to allocate suitable test items for each student and record students’ basic information. By invoking the inference engine of the fuzzy expert system, a feasible test sheet can be constructed according to the given parameters and the records of the students.

5. The Fuzzy Interface

This is an interface for fuzzifying the user-requested parameters of the test items. The fuzzified parameters, along with a set of fuzzy rules, are then sent to an expert system shell to perform the inference process.

Since these components work independently, it is possible to substitute any component to perform specified tasks. For example, we can link the system to different expert system shells or database systems to perform special functions. Moreover, as each function of the system can be either enabled or disabled, we can adapt the system to fit various needs. For example, a teacher can initiate a test with network communication functions disabled; while in the self-

Table 1. A Table to Illustrate the Evaluation Balance Table

	Key concepts-1	Key concepts-2	Key concepts-3	...	Key concepts-n
$[Q_1]$	e_{11}	e_{12}	e_{13}	...	e_{1n}
$[Q_2]$	e_{21}	e_{22}	e_{23}	...	e_{2n}
$[Q_3]$	e_{31}	e_{32}	e_{33}	...	e_{3n}
⋮	⋮	⋮	⋮	⋮	⋮
$[Q_m]$	e_{m1}	e_{m2}	e_{m3}	...	e_{mn}
	E_1	E_2	E_3	...	E_0

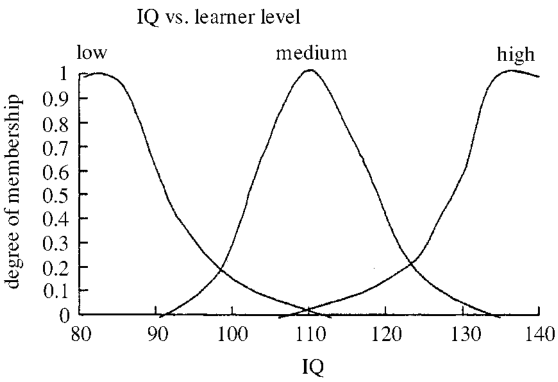

Fig. 3. The membership function of IQ vs. learner's level.

assessment mode, a student not only can communicate with others, but also can initiate the test with desired parameters.

IV. Test Sheet Constructing Rules

We have applied an Evaluation Balance Table (Hsaio, Tseng & Hwang, 1993; Hwang, Chen & Tseng, 1995) method as the theoretical basis of constructing test sheets. To apply this method, we suppose that each subject unit contains several key concepts to be learned. In Table 1, *n* key concepts (KC-1, KC-2... KC-n) are put across the top of the table and each row of the table represents a test item $[Q_i]$.

The value of e_{ij} ranges from 0 to 5, which signifies one of the following meanings:

5: Test item $[Q_i]$ is very important to concept KC-j.
4: Test item $[Q_i]$ is important to concept KC-j.
3: Test item $[Q_i]$ is more or less important to concept KC-j.
2: Test item $[Q_i]$ is related to concept KC-j.
1: Test item $[Q_i]$ has a weak relationship with concept KC-j.
0: Test item $[Q_i]$ has no relationship with concept KC-j.

Accordingly, a correlative evaluation rule for each test item $[Q_i]$ is generated. An n-array vector $\overline{SCOP}=(S_1, S_2, S_3 \ldots S_n)$ can be used to denote the learning status of a student, where S_i denotes the learning effect of KC-i. At the beginning, we have $\overline{SCOP}=(0, 0, 0, \ldots, 0)$ and the diagnostic rule for test item $[Q_i]$ is defined as

IF the answer of $[Q_i]$ is right

THEN $\overline{SCOP}=\overline{SCOP}+(e_{i1}, e_{i2}, e_{i3} \ldots e_{in})$ for $i=1, 2, 3, \ldots m$.

As test item $[Q_i]$ is answered, the correlative evaluation rule will be activated and revise the vector $\overline{SCOP}$ After doing the test, a result vector $\overline{SCOP}=(S_1, S_2, S_3 \ldots S_n)$ is obtained. A normalized vector is then taken as

$$T=[\frac{S1}{E1}\times 100\%, \frac{S2}{E2}\times 100\%, \frac{S3}{E3}\times 100\%, \ldots, \frac{Sn}{En}\times 100\%]$$

$$=(t_1, t_2, t_3 \ldots t_n),$$

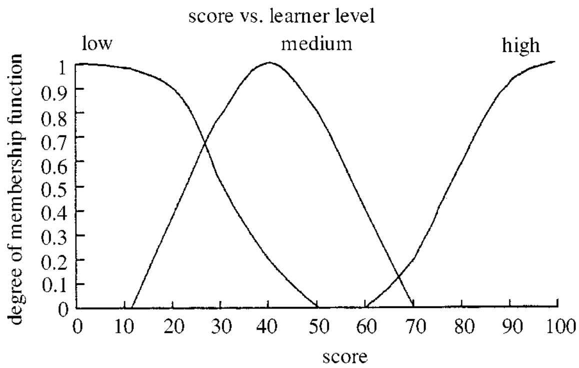

Fig. 4. The membership function of score vs. learner's level.

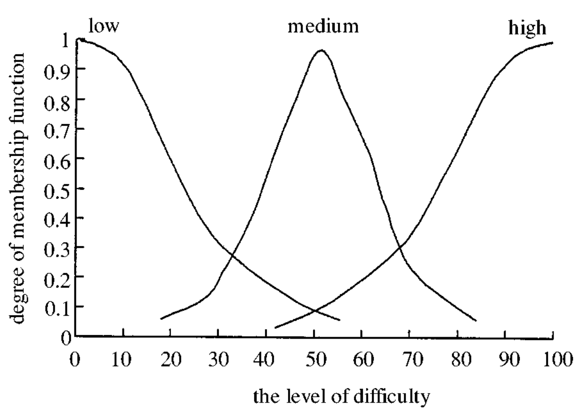

Fig. 5. The membership funciton of the level of difficulty.

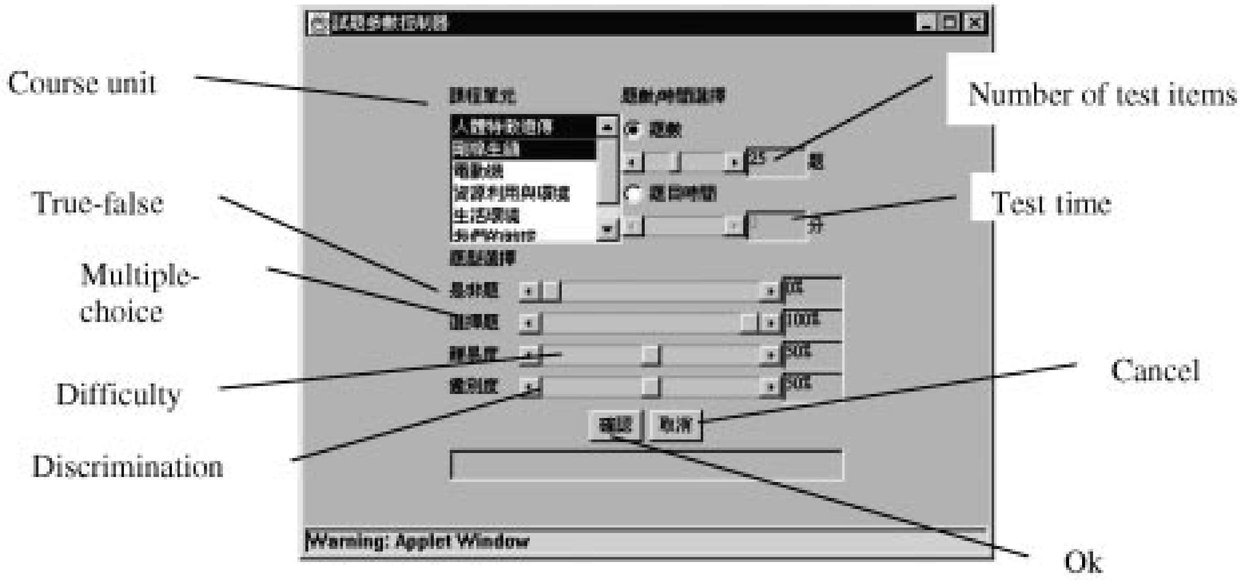

Fig. 6. Parameter controller page.

where t_i stands for the degree of students' acquaintance with KC-i.

If a student wants to take a test for self-assessment, the system must decide the level of difficulty according to the student's learning status. In such a personalized test, the level of discrimination will be ignored; instead, two indices of the level of the student, that is, IQ and the test history, will be taken into consideration. In the followings, we shall introduce how fuzzy concepts are used to represent the relationships among the indices and the learner's level.

First, we define the fuzzy membership function of IQ vs. learner level as shown in Fig. 3. The corresponding fuzzy rules are defined as follows:

IF IQ is high THEN learner's level is high
IF IQ is medium THEN learner's level is medium
IF IQ is low THEN learner's level is low

Second, we define the membership function of score vs. learner's level as shown in Fig. 4.

The corresponding fuzzy rules are defined as follows:

IF score is high THEN learner's level is high
IF score is medium THEN learner's level is medium
IF score is low THEN learner's level is low

The learner's level can be determined according to the membership function and fuzzy rules. Finally, we define the fuzzy rules of learner's level vs. difficulty as follows:

IF learner level is high THEN the level of difficulty is high
IF learner level is medium THEN the level of difficulty is medium
IF learner level is low THEN the level of difficulty is low

The membership function of learner's level vs. difficulty is shown in Fig. 5.

After determining the learner's level, we can obtain the level of difficulty by means of the above fuzzy rules and membership function. The system will then allocate test items for the student based on the level of difficulty and the default level of discrimination (50).

When a student fails to learn a unit well, the diagnostic rule offers an effective method for diagnosing what reinforcement work needs to be done. The well-learned key concepts and the poor-learned key concepts will be indicated and some suggestions will be given for further learning. Based on the learning log files and analysis of test results, the tutoring system can also be evaluated. If the tutoring script for some key concepts does not perform well for most students, the tutoring system can adjust its tutoring strategies in the next learning stage (Hwang, 1998).

V. Administration and Evaluation

We develop the Intelligent Testing and Diagnostic System server in the Windows NT environment. The test sheet constructing rules are represented by

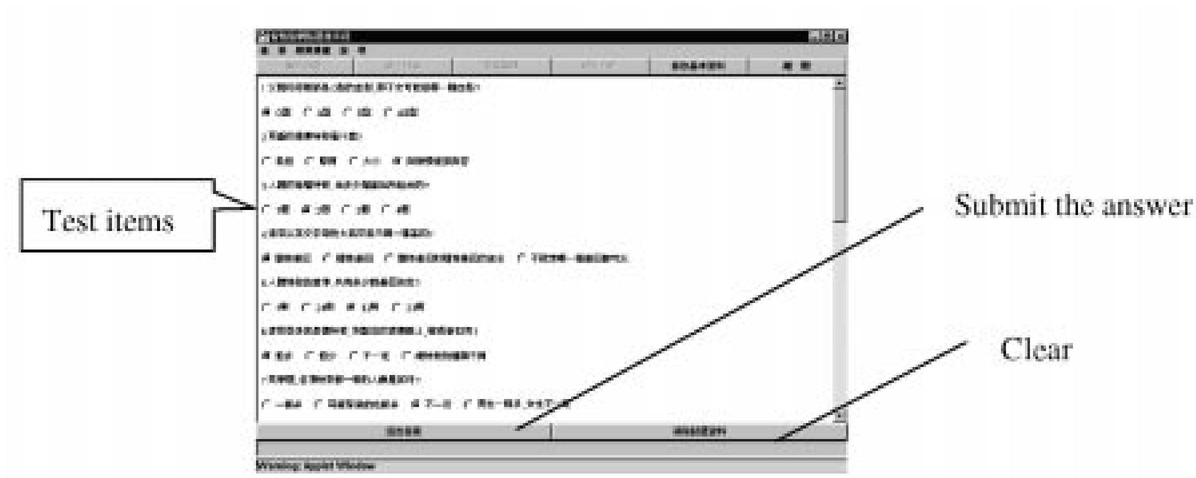

Fig. 7. A page depicting test items.

Fig. 8. On-line group discussion window.

the format of CLIPS, a well-known expert system shell developed by NASA (Giarratano & Riley, 1989), and are depicted in the appendix. Users can access this system through a WWW browser. To construct a test sheet, the user is asked to give the parameters concerning the test as shown in Fig. 6.

After setting the parameters, a test sheet is generated as shown in Fig. 7.

For a student who wishes to take an examination as self-assessment, an on-line discussion window is provided. The student can either initiate a new discussion group, or choose an existing group. Students are then allowed to communicate with each other and answer the same test items jointly as shown in Fig. 8.

After students submit the answers to the test items, the server gathers the answers from clients and store them as five text files: question-id file, user's answers file, correct solutions file, the related concepts file, and the weight of each concept in the questions. These five text files are sent to the fuzzy expert system and two text files are generated after performing the inference process, which generates a well-learned concepts file and a poor-learned concepts file. Based on these files, the server will present the learning guidance to the users as shown in Fig. 9.

To evaluate the performance of the testing and diagnostic system, a small group test has been done with 30 students (17 boys and 13 girls) from the Ai-Lan primary school in Nantou. Some experimental results are given as follows:

A. How do you feel by using computer to do the test?

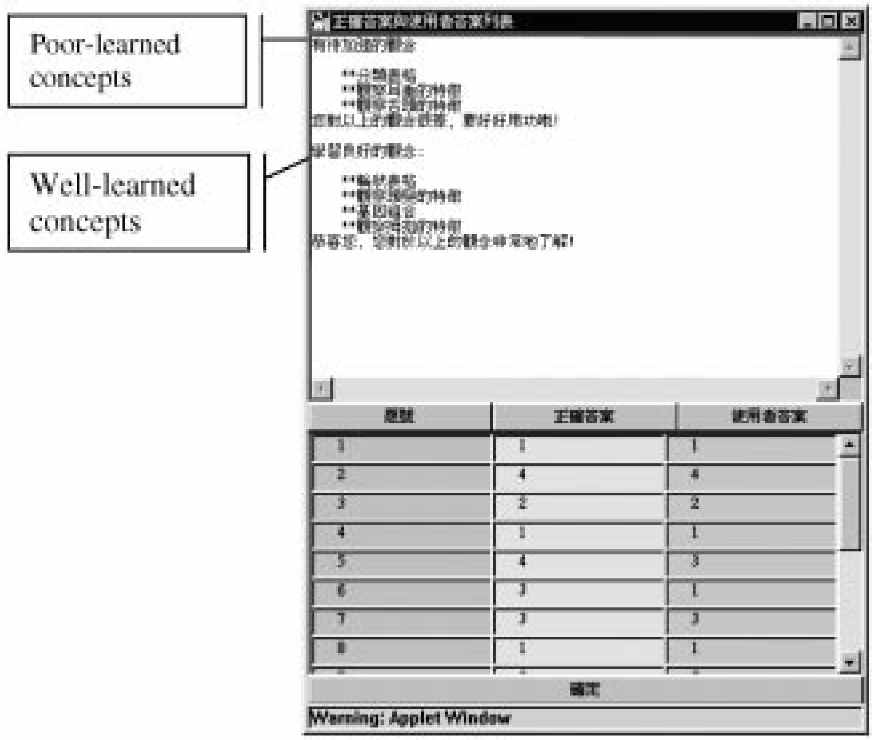

Fig. 9. Learning guidance presented by the system.

Choice		Number of students	Percentage
Very exciting	Boy	13	76.47%
	Girl	8	61.54%
No special feeling	Boy	4	23.52%
	Girl	5	38.46%
	Sum	30	

B. What is the bset feature of this system?

Choice	Number of students	Percentage
Network-based feature	5	16.7%
Learning guidence	24	80%
User interface	1	3.3%
Other feature	0	0%
Sum	30	100%

C. Will you be willing to take other tests with this system in the future?

Choice	Number of students	Percentage
Yes	25	83.3%
No	3	10%
(No answer)	2	6.7%
Sum	30	100%

According to the testing results, we find many students are willing to use the system, and most of them prefer the learning guidance function. Some students have suggested us to represent the contents of the learning guidance with some natural language-like sentences, which is one of the improvements to be made in the near future.

V. Conclusions

In this paper, we present a network-based intelligent testing and diagnostic system. The system has been implemented with JAVA language and can be accessed through a WWW browser to achieve the goal of cooperative testing and learning. The system is integrated with a network-based tutoring system, which keeps user log files concerning the learning behavior of each student of make more precise decisions for both testing and tutoring. According to a small group test with 30 primary school students, it can be seen that our approach is worth further studying. We are currently trying to enlarge the system with better user and management interfaces so that it can be applied to a wide variety of training courses.

Appendix

The test sheet constructing rules with CLIPS format

```
(defrule calculate
        ?phase <- (phase calculate)
        ?rel<-(relation ?relation)
        ?sol<-(solution-sum ?solsum)
        ?ans<-(answer-sum ?anssum)
        (value ?value)
    =>
        (retract ?phase )
        (retract ?rel ?sol )
        (assert (solution-sum =(+ ?solsum ?value)))
        (assert (phase read-solution-answer))
        (if (eq ?relation equal)
                then
                (retract ?ans)
                (assert (answer-sum =(+ ?anssum ?value)))))

(defrule suggest
        ?phase <- (phase suggest)
        ?con <-(concept ?concept)
        ?answer-sum <-(answer-sum ?anssum)
        ?solution-sum <-(solution-sum ?solsum)
    =>
        (retract ?phase ?con ?answer-sum ?solution-sum)
        (assert (phase value-initial))
        (if (= ?solsum 0)
                then
                (printout t "solution-sum=0")
                else
                (if (>(/ ?anssum ?solsum) 0.8)
                        then
                        (printout high ?concept )
                        (printout high " ")
                        else
                        (if (<(/ ?anssum ?solsum) 0.6)
                                then
                                (printout low ?concept )
                                (printout low " ")
                                else
                                (printout middle ?concept)
                                (printout middle " ")))))
```

Reference

Fan, J. P., Tina, K. M., & Shue, L. Y. (1996). Development of knowledge-based computer-assisted instruction system. *1996 International Conference Software Engineering: Education and Practice*, Dunedin, New Zealand.

Giarratano, J. R., & Riley, G. (1989). *Expert Systems: principles and programming*, PWS-KENT Publishing.

Hsaio, H. W., Tseng, S. S., & Hwang, G.J. (1993). An evaluation model for the development of intelligent CAI systems. *International Conference on Computer in Education*.

Ho, R. G. (1997). A Networking environment of item bank, testing and practicing system, *1997 Project Report of National Science Council on Computer-Assisted Learning and Distance Instruction of*, pp. 144-162.

Hopper, S. (1992). Cooperative learning and computer-based instruction. *Educational Technology Rresearch & Development*, **40**(3), 21-38.

Hsu, C. S., Tu, S. F., Yeh, S. Y., Chu, C. Y., & Hwang G. J. (1997). Development of an intelligent testing and evaluation system on computer networks. *1997 National Computer Symposium, Taiwan*, pp. D129-134.

Hwang, G. J. (1998). A Tutoring Strategy Supporting System for Distance Learning on Computer Networks, *IEEE Transactions on Education*, Nov., 1998.

Hwang, G. J., Chen, F. R., & Tseng, S. S. (1995). On the development of an intelligent CAI for chemistry courses. *Proceedings of the National Science Council of R.O.C.*, Part D, **5**(1), 9-18.

Sun, C. T., & Chou, C. (1996). Experiencing CORAL: design and implementation of distance cooperative learning. *IEEE Transactions on Education*, **39**(3), 357-366.

Wainer, H. (1990). *Computerized adaptive testing: A primer*. Hillsdale, NJ: Lawrence Erlbaum Associates.

遠距學習環境中智慧型測驗及診斷系統之研製

黃國禎

國立暨南國際大學資訊管理研究所

摘　要

將人工智慧與電腦輔助教學結合，使電腦輔助教學更具智慧，是人工智慧研究者和電腦輔助教學研究者所致力的目標。而如何擬出一份適當而有效率，且又能配合學生特質的試題，更是教育界所關注的焦點。以往的教學及測驗系統大多針對單一使用者設計，學生遇到問題往往求助無門。在本論文中，我們嘗試透過網路建一套智慧型測驗及診斷系統，可以根據所要求的難易度、鑑別度等參數，動態進行測驗及學習障礙之診斷。本系統同時可允許來自各地的學生共同完成一份測驗，使學生從討論中得知學習的盲點，並提高學生的學習興趣。此外，運用模糊專家系統的推理機制，本系統會依據測驗結果給予學生進行學習障礙診斷並給予學習指導，以真正達到改進學習的效果。