Building and Applying the Bayesian networks based adaptive test System－Using Rounding & estimating with decimals Unit in the Fifth Grade as A Example

Shu-Chuan Shih [a], Bor-Chen Kuo [b], Yu-Lung Liu [c]

[a] *Department of Mathematics Education, National Taichung University, Taiwan*

[b] *Graduate School of Educational Measurement and Statistics, National Taichung University, Taiwan*

[c] *Department of Computer Science and Information Engineering , Asia University ,Taiwan*

[a] *ssc@mail.ntcu.edu.tw*

Abstract: This study aims at combining Bayesian networks with item ordering theory to build the Bayesian networks based adaptive test System and explores the efficiency of using this computerized adaptive diagnostic test system in practical instruction. The domain content chosen is the rounding and estimating with decimals unit. The results show that the proposed BNAT in this study is robust and has the good performances in testing speed, prediction accuracy and diagnostic accuracy of bugs and sub-skills.

Keywords: Computerized Adaptive Diagnostic Test (CADT), Bayesian networks, item ordering Theory, Bayesian networks based adaptive test(BNAT)

Introduction

In recent years, the fast increase in power and availability of information technologies has made computerized adaptive testing (CAT) an economically viable alternative to paper-and-pencil testing [12]. Although the CAT is developed to obtain an efficient estimate of an examinee's ability based on the assumed item response theory (IRT) model and perform a shorter test than the traditional paper-based tests, but it still can't provide diagnostic capability to diagnose students' cognitive skills at a more specific level, which could explain, for instance, individual student's problem-solving strategies, bugs, and knowledge-states in the sub-skills of the domain.

Knowledge structure based adaptive test (KSAT) provides some solution to the problem given above[7]. However, KSAT only can diagnose students' knowledge-states in the sub-skills of the domain. It lacks capability to diagnose students' bugs. So, this paper is focused on building the computerized adaptive diagnostic test (CADT) system can afford the diagnosis of students' bugs and sub-skills simultaneously.

However, it has been known that manifestations of the bugs are often "unstable" within an individual student's performance [3]. This instability contributes significantly to the overall complexity and unpredictability in bug diagnosis [8]. Several techniques have been used to implement cognitive models that incorporate uncertainty. One of the best-known uncertainty modeling techniques is Bayesian network, which incorporate uncertainty by using a probability-based approach [6,11]. Hence, this study will combine

Bayesian networks with item ordering theory to extend the KSAT in order to build a more sophisticated CADT which is called Bayesian networks based adaptive test (BNAT). After that, this system is used to predict the occurrences of students' bugs and sub-skills in the domain of rounding and estimating with decimals, and to evaluate the effectiveness of the proposed BNAT.

1. Bayesian Networks

Bayesian networks have become one of the most widely used tools for managing and assessing uncertainty in a number of fields [6,11]. A Bayesian network is a pair (D, P) which combine graph theory with conditional probabilities, where D is a directed acyclic graph (DAG), consisting of a set of nodes and directed arcs. The nodes represent random variables, and each node can take on a set of possible values. The arcs signify direct dependence between the connected nodes.

In addition to the graphical structure, $\mathbf{P} = \{p(x_1|\pi_1), \cdots, p(x_n|\pi_n)\}$ is a set of conditional probability distributions (CPDs) associated with each node in the network, π_i is the set of parents of node X_i in D. By computing the CPDs of all nodes in the network, we can represent the joint distribution of all variables in the network economically and can compute any desired probabilistic information with a given Bayesian network.

In the last few years, researchers of educational assessment had also studied the applications of Bayesian networks in education [2,8,9,10]. One line of such efforts was involved it's application to cognitive diagnosis [8,9]. According to the literatures [9,10], Mislevy suggested a Bayesian network framework for proceeding probability-based inference in cognitive diagnosis. Three key points were described in this framework: building Bayesian networks for a student model (SM-BN); constructing tasks that students can be provided an opportunity to reveal their performance in targeted knowledge and skill; and creating Bayesian networks for evidence models (EM-BNs) that describes how to extract the key items of evidence from what a student does in the context of a task, and pointers to parent student-model variables. This is where the inference network and local computation come into play.

In this study, we use the Bayesian network framework suggested by Mislevy, and combine this framework with item ordering structures estimated by item ordering theory to develop new adaptive testing algorithms.

2. Knowledge Structure Based Adaptive Test and Item Ordering Theory

The Adaptive Testing Algorithm of Knowledge Structure Based Adaptive Test

The adaptive testing algorithm of KSAT is based on the students' knowledge structure[7]. As shown in Fig.1, if the subject gets a top skill (item C) correct then it is inferred that he or she also has its prerequisites (items F, G, H, I). This algorithm can not only estimate students' knowledge structures quickly but predict students' profiles using fewer items than in original paper-based test. The number of links has an impact on both testing speed and prediction accuracy rate. As the number of links increases, the testing speed may get up but the prediction accuracy rate may decrease.

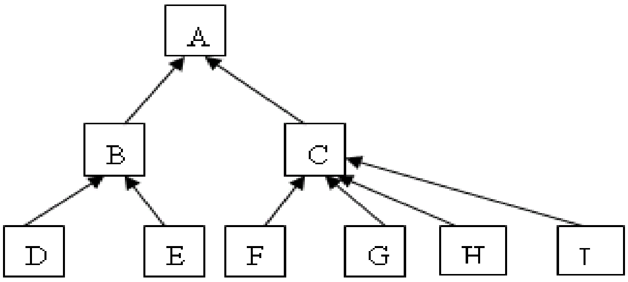

Fig. 1. The students' knowledge structure

Item Ordering Theory

In KSAT, ordering theory (OT)[1] is used for estimating item ordering structures and the students' knowledge structure. The theory of OT is described briefly in the following:

Let $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ denote a vector containing n binary item scores variables. Each individual taking n-item test produces a vector $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ containing 1(correct) and 0(wrong). Then the joint and marginal probabilities of items j and k can be represented in Table 1.

Table 1 The joint and marginal probabilities of item j and k

		Item k		
		$X_k = 1$	$X_k = 0$	Total
Item j	$X_j = 1$	$P(X_j = 1, X_k = 1)$	$P(X_j = 1, X_k = 0)$	$P(X_j = 1)$
	$X_j = 0$	$P(X_j = 0, X_k = 1)$	$P(X_j = 0, X_k = 0)$	$P(X_j = 0)$
	Total	$P(X_k = 1)$	$P(X_k = 0)$	1

For OT, let $\varepsilon^*_{jk} = P(X_j = 0, X_k = 1) < \varepsilon$, usually $0.02 \le \varepsilon \le 0.04$ [1], then item j could be linked forwards to item k. The relation is denoted as $X_j \to X_k$ means that X_j is a prerequisite to X_k. If $X_j \to X_k$ and $X_k \to X_j$, then the relation is denoted as $X_j \leftrightarrow X_k$ and it means item j and k are equivalent.

In this study, we integrate the students' knowledge structure estimated by OT into SM-BN to develop the innovative adaptive test algorithms of BNAT.

3. The Domain Area

Before constructing Bayesian networks, knowledge structure and preparing the items according to them, we have to do an analysis of the domain of rounding and estimating with decimals. A number of studies had been carried out on this area [4, 5]. According to the literatures and the competence indicator 5-n-10 of grade 1-9 mathematics curriculum guidelines, 29 sub-skills and 9 bugs chosen as key points in developing diagnostic tests. Descriptions of these sub-skills and selected bugs are shown as Table 2.

Table 2 Sub-skills and selected bugs of the Bayesian network in this study

5-n-10	The students can understand the general rule for rounding to a number of decimal places is to look at the digit after the one they wants to round to, and round up or down depending on whether this digit is '5 or more' or '4 or less', then use this rule to estimate calculations of decimals.
bugs	
B1	Have very little idea about decimal place value.

B2	Can't find the right "rounding digit".
B3	Confuse rounding up with rounding down.
B4	Depend on the number of digits they wants to round to round up or down.
B5	The procedure in round down is wrong, use the digit of the given place to subtraction one.
B6	All numbers to the right of the required place value are replaced by zeros.
B7	Can't infer the range of the actual number based on a round number.
B8	Don't use the general rule for rounding to estimate calculations of decimals.
B9	Use the general rule for rounding to get an approximate answer *after* accurate calculation
Sub-skills	
S01	Can use the general rule for rounding to round decimals to tenths, hundredths, thousandths, and so on.
S02	Can find the interval of the actual number based on a given round number (the number is rounded to the nearest unit).
S03	In math word problems, can find the interval of the actual number based on a round number (the number is rounded to the nearest unit).
S04	Can find the interval of the actual number based on a given round number (the number is rounded to the nearest tenth).
S05	In math word problems, can find the interval of the actual number based on a round number (the number is rounded to the nearest tenth).
S06	Can find the interval of the actual number based on a given round number (the number is rounded to the nearest hundredth).
S07	In math word problems, can find the interval of the actual number based on a round number (the number is rounded to the nearest hundredth).
S08	Can round decimals to the nearest unit, and then add the rounded decimals to solve addition estimation problems of decimals.
S09	Can round decimals to the nearest unit, and then add the rounded decimals to solve addition estimation word problems of decimals.
S10	Can round decimals to the nearest tenth, and then add the rounded decimals to solve addition estimation problems of decimals.
S11	Can round decimals to the nearest tenth, and then add the rounded decimals to solve addition estimation word problems of decimals.
S12	Can round decimals to the nearest hundredth, and then add the rounded decimals to solve addition estimation problems of decimals.
S13	Can round decimals to the nearest hundredth, and then add the rounded decimals to solve addition estimation word problems of decimals.
S14	Can round decimals to the nearest unit, and then subtract the rounded decimals to solve subtraction estimation problems of decimals.
S15	Can round decimals to the nearest unit, and then subtract the rounded decimals to solve subtraction estimation word problems of decimals.
S16	Can round decimals to the nearest tenth, and then subtract the rounded decimals to solve subtraction estimation problems of decimals.
S17	Can round decimals to the nearest tenth, and then subtract the rounded decimals to solve subtraction estimation word problems of decimals.
S18	Can round decimals to the nearest hundredth, and then subtract the rounded decimals to solve subtraction estimation problems of decimals.
S19	Can round decimals to the nearest hundredth, and then subtract the rounded decimals to solve subtraction estimation word problems of decimals.
S20	Can round numbers to the specified place value, and then multiply the rounded numbers to solve multiplication estimation problems of integers.
S21	Can round numbers to the specified place value, and then multiply the rounded numbers to solve multiplication estimation word problems of integers.
S22	Can round decimals to the nearest unit, and then multiply the rounded decimals to solve multiplication estimation problems of decimals.
S23	Can round decimals to the nearest unit, and then multiply the rounded decimals to solve multiplication estimation word problems of decimals.
S24	Can round decimals to the nearest tenth, and then multiply the rounded decimals to solve multiplication estimation problems of decimals.

S25	Can round decimals to the nearest tenth, and then multiply the rounded decimals to solve multiplication estimation word problems of decimals.
S26	Can round decimals to the nearest hundredth, and then multiply the rounded decimals to solve multiplication estimation problems of decimals.
S27	Can round decimals to the nearest hundredth, and then multiply the rounded decimals to solve multiplication estimation word problems of decimals.
S28	Can round numbers to the specified place value, and then divide the rounded numbers to solve division estimation problems of integers.
S29	Can round numbers to the specified place value, and then divide the rounded numbers to solve division estimation word problems of integers.

4. Building the Bayesian Networks Based Adaptive Test

The Test Development Process of Bayesian Networks Based Adaptive Test

Gathering bugs in the domain of rounding and estimating with decimals.
Building the experts' knowledge structure and design items according to each sub-skill node of this structure.
Constructing the original paper-based tests that the items are designed to elicit bug symptoms. There are three parallel tests in the domain of this study and each test has 35 items, which are carefully constructed so that the bugs and sub-skills can appear in various types of tasks. All items are 4 options (A, B, C, D) multiple-choice format and no response item is coded N.
Trying out the original paper-based test and proceeding with item and reliability analysis. The sample size of tryout is 301 fifth grade students. The results are all values of the item difficulties and discriminations fall within an acceptable range and the test has a Cronbach's α reliability coefficient of 0.94.
Using OT to estimate item ordering structures and the students' knowledge structure.
As Fig. 3 shown below, Bayesian network is proposed and constructed by integrating the students' knowledge structure into sub-skills level. In BNAT system, the students' knowledge structure is used as the basis of the adaptive test algorithm. Next, we train the network using the data in tryouts.
Convert the original paper-based tests to BNAT. Testing interface as shown in Fig.2.

Fig. 2 Testing interface of BNAT

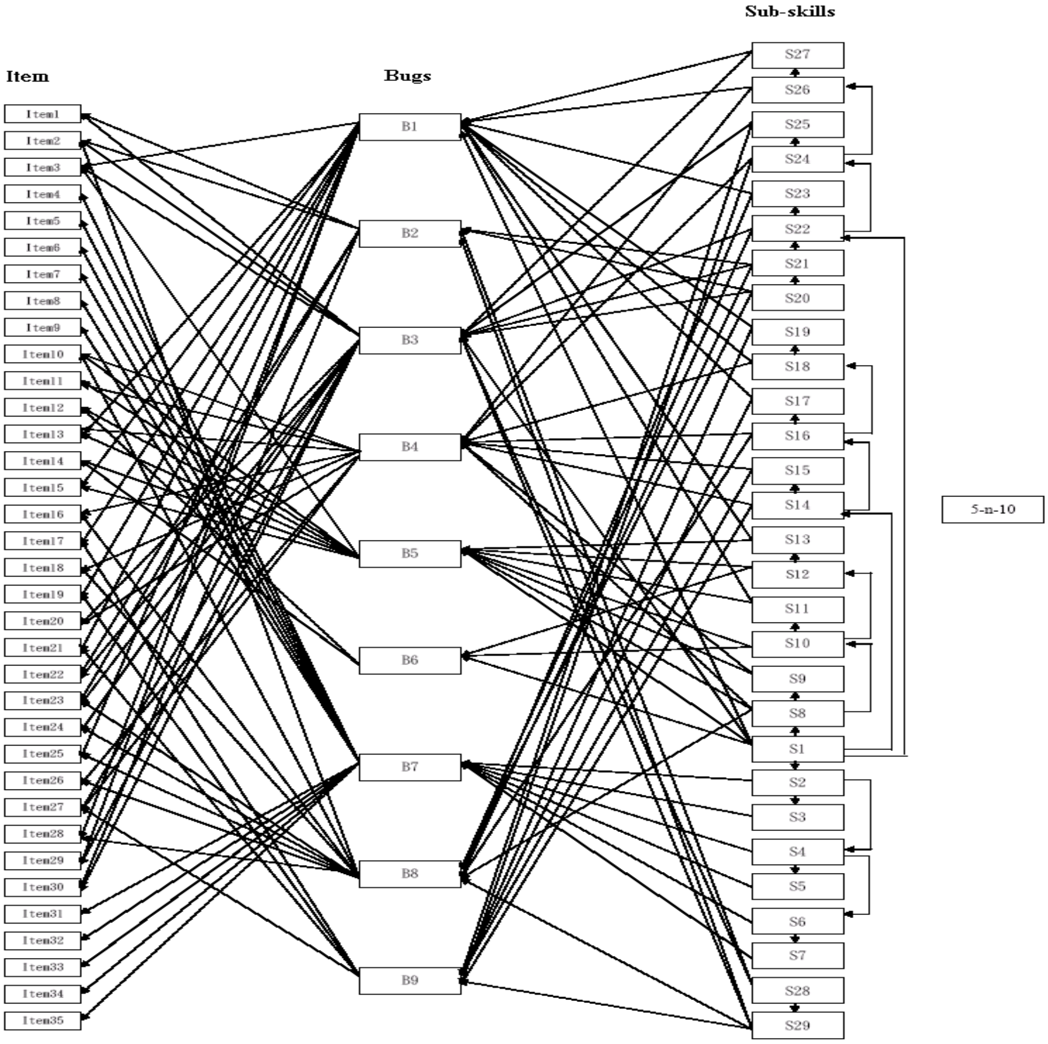

Fig. 3. Bayesian network in the unit of rounding and estimating with decimals

The Method of Evaluation

In order to assess how well the BNAT would work, the BNAT is administered to 64 fifth grade students twice: pretest (before remedial instruction) and posttest (after remedial instruction). Computing method of Evaluation are as follows:

Computing testing speed of BNAT system: $\frac{\text{full - length of test - average test length}}{\text{full - length of test}}$

Computing prediction accuracy rate of the adaptive test algorithm used in BNAT system:

$$\frac{\text{the number of itms that are predicted accurately by } BNAT}{\text{full - length of test}}$$

Computing classification accuracy rate in predicting the existence of bugs and sub-skills of BNAT system:

$$\frac{\text{the number of sample that bug(or sub - skill) are classified accurately by } BNAT}{\text{the number of testing sample}}$$

5. Evaluations and Discussions

As Table 3 and Table 4 shown below, overall testing speed results for each test are reported.

Regardless of pretest, posttest or average, we can find the numbers of items administrated by BNAT are all less than 25.33. That is, more than 9.67(35-25.33) items are omitted and prediction accuracy can exceed 97.6%.

Table 3 testing speed of BNAT

	Full-length of test	average test length	average test time	testing speed
Pretest	35	25.33	27.5 min	27.6%
Posttest	35	25.19	22.9 min	28%
average	35	25.26	25.2 min	27.8%

Results of the Table 4 showed that all average classification accuracies rate of bugs and sub-skills also are greater than 96.2%.

Table 4 Prediction accuracy rate and classification accuracy rate of BNAT

	Prediction accuracy rate	Average classification accuracy rate of bugs and sub-skills
Pretest	97.6%	96.7%
Posttest	97.7%	96.2%
average	97.7%	96.5%

6. Conclusions

The results show that the proposed BNAT in this study is robust and the performances in both testing speed and prediction accuracy are well.Using the proposed BNAT in this study to diagnose the existence of bugs and sub-skills in individual students also can get good performance.

References

[1] Airasian, P. W., & Bart, W. M. (1973). Ordering theory: A new and useful measurement model. *Educational Technology*, *5*, 56-60.

[2] Almond, R. G., & Mislevy, R. J. (1999). Graphical models and computerized adaptive testing. *Applied Psychological Measurement, 23*, 223-237.

[3] Brown, J. S., & Burton, R. R. (1978). Diagnostic Models for Procedural Bugs in Basic Mathematical Skill. *Cognitive Science*, *2*(2), 155-192.

[4] Forreste, M. A., & Pike, C. D. (1998). Learning to estimate in the mathematics classroom: A conversation-analytic approach. *Journal for Research in Mathematics Education*, *29*(3), 334-357.

[5] Hanson, S. A., & Hogan, T. P. (2000). Computational estimation skill of college students. *Journal for Research in Mathematics Education*, *31*(4), 483-500.

[6] Jensen, F. V. (2001). *Bayesian networks and decision graphs*. New York: Springer.

[7] Kuo B. C., Hsieh Y. C., Chung C. H.,& Tsai K. Y.(2005). *A study of knowledge structure based adaptive test and remedial instruction system.* TaiWan E-Learning Forum, National Taiwan Normal University, Taiwan.

[8] Lee, Jihyun. (2003). *Diagnosis of bugs in multi-column subtraction using Bayesian networks*. Unpublished ph.D, Columbia University

[9] Mislevy, R.J. (1995). Probability-based Inference in Cognitive Diagnosis. In P.Nichols, S.Chipman & R. Brennan. (Eds.) *Cognitively Diagnostic Assessment.* Hillsdale, NJ: Lawrence Erlbaum Associates.

[10] Mislevy, R.J., Almond, R.G., Yan, D., & Steinberg, L.S. (1999). Bayes nets in educational assessment: Where do the numbers come from? In K. B. Laskey & H. Prade (Eds.) *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence* (pp.437-446). San Francisco: Morgan Kaufmann Publishers, Inc.

[11] Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann.

[12] Zara, A. R., & Kingsbury, G. G. (1989). Procedure for selecting items for computerized adaptive tests. *Applied Measurement in Education*, *2*(4), 359-375.