Indefinite Integral Automatic Analysis Mechanism Study — Using Basic Indefinite Integrals Involving Quadratic Square Roots as an Example

M. Y. Ting

Abstract—**These researchers prepared a university calculus problem is structural model of knowledge "on the square root of the product of the second relates to the indefinite integral trigonometric substitution" effective testing program. After calculus courses and tests, in this study, Bayesian network analysis was used to find the best combination of recognition rate of Bayesian network, and easily identified these error types so that students might learn "trigonometric substitution" unit. Then, using the Bayesian probability analysis, the establishment of a "trigonometric substitution" unit modeled adaptive Bayesian network diagnostic test systems for example, the establishment of such structure and how it was used as a basis for developing remedial teaching statistical information to analyze.**

Index Terms—**Indefinite integral, constructed-response item, error type, Bayesian network.**

I. INTRODUCTION

Unfortunately, for the overwhelming majority of students, the Calculus is not a body of knowledge, but a repertoire of imitative behavior patterns [1]. Although many students successfully perform tasks in calculus program, they lack of understanding of the basic concepts of Calculus [2], [3]. Therefore, this study would apply determine the conceptual structure, and then using Bayesian network analysis to develop the integration diagnostic test system, and to participate in the follow-up remedial teaching.

II. PROCEDURE FOR PAPER SUBMISSION

A. Purposes

The purposes of this study were: 1). to determine the conceptual structure of the article, and 2). establishment of "indefinite integral" on-line testing system.

B. Content Quiz

However, Calculus forms a bridge in the form of basic math and science. The basic Calculus, including infinite series, limits, derivative function, integral, composes the majority of modern university mathematics education.

C. Error Types

In the process of mathematical calculations, mistakes at crucial steps produced the wrong final answer, according to

the disposition of key species, known as the error type [4]. Few literatures involved on the integration problem of Quadratic Square Root. After integration, the student might not correctly convert back to the original variable. In this study, the first year of university students were located in the computer classroom, any incomplete reaction was considered to be an invalid replay. 176 first-year students answered valid samples.

TABLE I: THE PROPOSITION CARD

Item 1. Evaluate the integral	$\int \dfrac{x+1}{\sqrt{1-x^2}}dx$
Item 2. Evaluate the integral	$\int \dfrac{1}{\sqrt{1+x^2}}dx$
Item 3. Evaluate the integral	$\int \dfrac{\sqrt{x^2-1}}{x}dx$

1) B1: Mix formulas:

For example, in $\int \dfrac{1}{\sqrt{x^2+1}}dx$, the integral of a function's n-th power, ($\int f^n(x)dx = \dfrac{f^{n+1}}{n+1}+c$) mixed with Derivative, students mistakenly regarded as $\dfrac{\sqrt{x^2+1}}{x}+c$. This should be $\ln|\sqrt{x^2+1}+x|+c$.

2) B2: There is an error in the ant derivative:

For example, in $\int \dfrac{1}{\sqrt{x^2+1}}dx$, students mistakenly regarded $\int \sin\theta d\theta$ as $\cos\theta + c$. This should be $-\cos\theta + c$.

3) B3: After integration, the student did not exchange the answer for the original variables:

For example, in $\int \dfrac{1}{\sqrt{x^2+1}}dx$, students mistakenly regarded $\ln|\sqrt{1+x^2}+x|+c$ as $\ln|\sec\theta+\tan\theta|+c$.

4) B4: Make wrong about trigonometric identities:

For example, in $\int \dfrac{\sqrt{x^2-1}}{x}dx$, students mistakenly regarded $\int \tan^2 x dx$ as $\int(1-\sec^2 x)dx$ This should be $\int(\sec^2 x-1)dx$

5) B5: After integration, the students wrong substituted the answer back to the original variable:

For example, in $\int \dfrac{x+1}{\sqrt{1-x^2}}dx$, students mistakenly regarded $-\sqrt{1-x^2}+\sin^{-1}x+c$ as $-\sqrt{1+x^2}+\sin x+c$.

D. Test Presentation

Students used the correct expressions to build the answers, as shown in Table II.

TABLE II: THE CODE REPRESSIONS

Code	Expression
「int_」	the integral
「∧」	a power of superscript symbol
「frac {a} {b}」	a / b

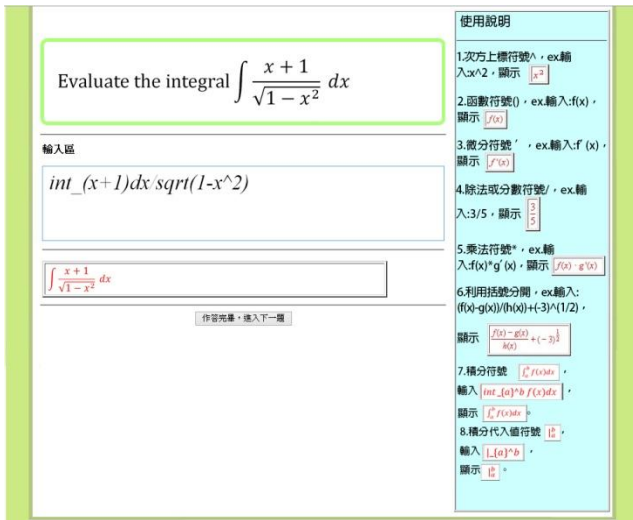Expressions presented in the lower display area, as shown in Fig. 3:

Fig.3. The test form

TABLE III: THE PROPOSITION CARD

( )1. $\int \dfrac{x+1}{\sqrt{1-x^2}} dx = ?$

(1) $-\sqrt{1-x^2} + \sin^{-1} x + c$ (2) $\sqrt{1-x^2} + \sin^{-1} x + c$

(3) $-\cos\theta + \theta + c$ (4) $-\sqrt{1+x^2} + \sin^{-1} x + c$

( )2. $\int \dfrac{1}{\sqrt{1+x^2}} dx = ?$

(1) $\dfrac{\sqrt{1+x^2}}{x} + c$ (2) $\sqrt{1+x^2} + c$

(3) $\ln | \sqrt{1+x^2} + x | + c$ (4) $\ln|\sec\theta + \tan\theta| + c$

( )3. $\int \dfrac{\sqrt{x^2-1}}{x} dx = ?$

(1) $\sec^{-1} x - \sqrt{x^2-1} + c$ (2) $\sqrt{x^2-1} + \sec^{-1} x + c$

(3) $\sqrt{x^2-1} - \sec^{-1} x + c$ (4) $\dfrac{(x^2-1)^{3/2}}{x^2} + c$

In this study, the valid sample of 186 freshmen, analyses were found out the reason leading to the wrong answers.

These multiple-choice items changed to the Constructed-Response items (CR), i.e. CR1~CR3.

E. Analysis Process

The analysis process of CR as follows:
Decision steps
- whether a student gave an answer

If the answer was that blank, then the student was considered "no answer." Results were encoded <99>. Otherwise, the student was considered "answer it" to enter the step 2.
- Decide whether the response was "right", "Just key the item" or "Known bugs".

If the answer was not wrong, then the student was considered to be "right." As a result, he was coding <key>. If the answer and the item were the same, it was considered "just key the item." Results were encoded <98>. Otherwise, the student was seen as "known error" to enter step 3.
- A judgment of error, to determine its type, if it was not known error, and then re-classified as a "new error."
- Judge the error to determine its new error type.

CR1 has seven error types, more than the original default error types. In this study, for example, the analysis of CR1 was regarded as a reasoning tool for diagnostic tests in the future to help teachers diagnose and analyze the type of error, as the basis for remedial teaching.

CR2~CR3 are so analysis. The additional error types of students to the Constructed-Response items are discovered as follows:

6) B6: No plus the indefinite constant:

For instance, $\int \dfrac{1}{\sqrt{1+x^2}} dx = \ln | \sqrt{1+x^2} + x |$.

This should be $\int \dfrac{1}{\sqrt{1+x^2}} dx = \ln | \sqrt{1+x^2} + x | + c$.

7) B7: No regularities to answer the question:

For instance, $\int \dfrac{1}{\sqrt{1+x^2}} dx = \dfrac{1}{x^2}(1+x^2)^{1/2} + c$.

Teachers understood the students' error types, to help remedial teaching.

F. Sub-skill

Integrands involving $\sqrt{x^2 \pm 1}, \sqrt{1-x^2}$, required the use of the following sub-skills. S1: Triangle hypothesis, S2: After the triangle variable is assumed that students convert the original variable to integration of a new variable. S3: Find the related antiderivative. S4: Transform these results into the original variable formula.

G. Bayesian Network

Currently Bayesian network (BN) takes domain knowledge to conclude a meaningful symbol, and then combines with data symbols so that a set of variables, the possibility of application of a causal relationship between variables, as well as the interaction between them, and the conditional probability table (CPT) to represent the quantitative knowledge.

For some partition $\{b_1, b_2, ..., b_n\}$ of the event space, the b_i is endowed $P(b_i)$ and $P(a|b_i)$. Using Bayes' Theorem to calculate the probability of b_i:

$$P(b_i | a) = \frac{P(a|b_i)P(b_i)}{P(a)} = \frac{P(a|b_i)P(b_i)}{\sum P(a|b_i)P(b_i)},$$

where $P(b_i)$: the prior probability ; $P(a|b_i)$: the sample probability; $P(b_i | a)$: the posterior probability.

We want to calculate, assuming our new evidence is true, what is the probability that b_i is true. Combination of the prior

probability and the sample probability will, according to Bayes' theorem, estimate the probability, then, can effectively draw inferences. The application of Bayesian inference means that a BN can be readily updated, when new knowledge becomes available [5]. BNs are therefore valuable tools for addressing uncertainty in data and combining observations, model simulation and expert knowledge [6].
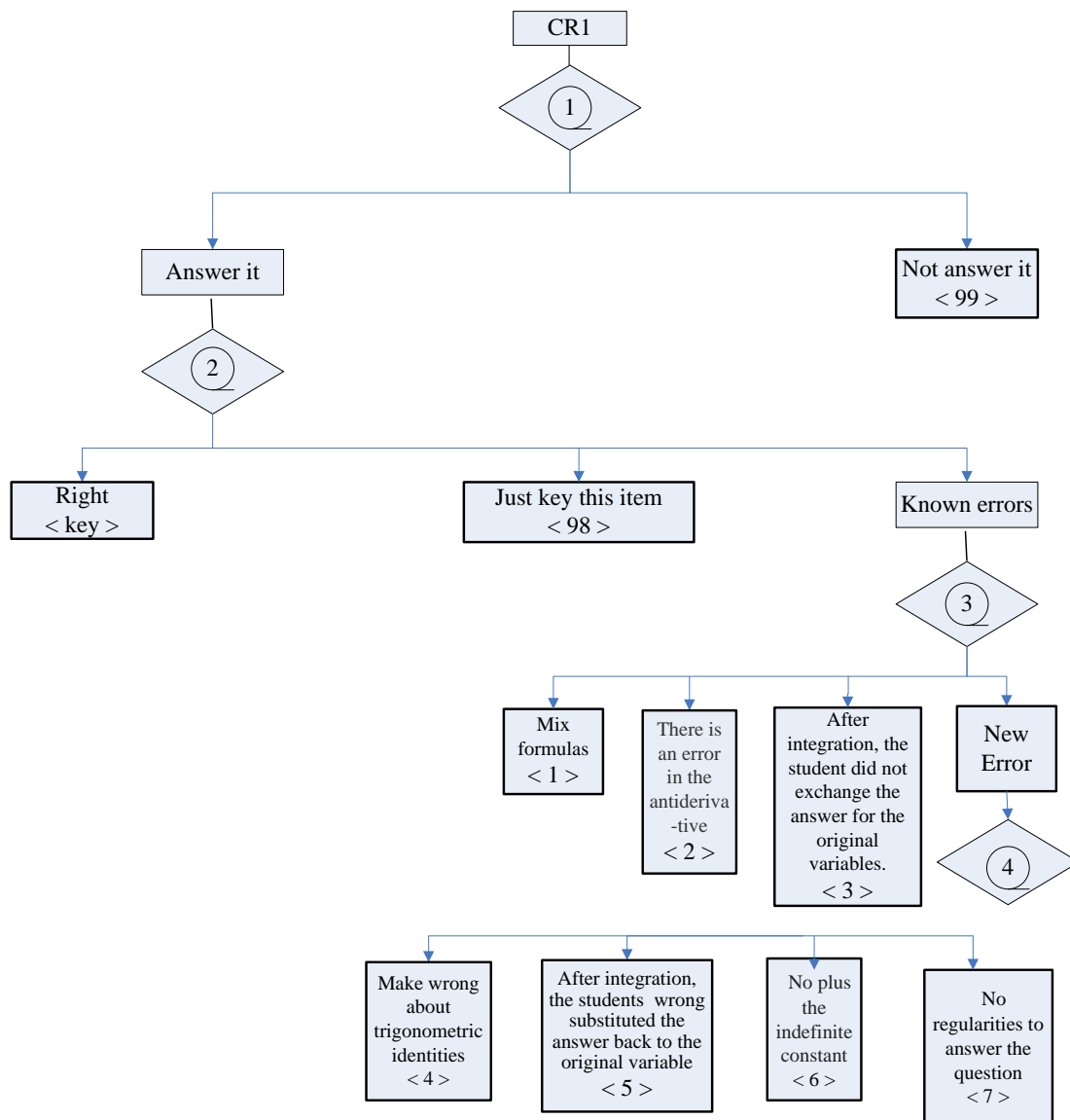

Fig. 5. The analysis of process for CR1.

Some obvious advantages, related to the Bayesian network, described as below.
1) BNs give a useful causal relationship between two
2) variables.
3) BNs take advantage of various input data.
4) BNs supply a visual decision support instrument.
5) BNs aid structural and parameter learning.
6) BNs allow to add the new evidence in.

Jayneel, Shahram and Thomas (2012) used causal BNs to represent data center capacity planning decision process.

The method used in the sensitivity of their analysis, decision-making process in order to reduce the influence of the most also provide an opportunity, if there are also factors that reduce unused data center capacity. Their method used sensitivity analysis to reduce the factors that influence the decision process the most while supplying an opportunity, if one exists, to also reduce unused data center capacity [7].

Network with discrete data results in better classification performance (an increase of up to 11.7% of detected cancers), easier to interpret and is more suitable compared to the data, based on Bayesian network expert Gaussian probability parameters [8]. Minoo, David, Amina, Kane, Scott, Nalin, and Kristin (2014) progressed a novel approach for incorporating expert rules into BNs for classification of Mycobacterium tuberculos is complex clades [9]. Their results showed that incorporating rules into problems can drastically increase classification accuracy if data alone are insufficient [10]. Eva, Luis, Gladys, Paula and Sandra (2012) investigated the integration and evaluation of existing general Bayesian student model existing computerized test system to Aveiro University Mathematics Education Project [9].

Their findings suggested that a high degree of consensus among the scores given by the experts and also among the diagnosis provided by the BSM in the written exam and expert's average [11]. Mostafa, Mahmoud and Bahram (2014) inspected a network with fuzzy parameters versus the network

with crisp parameters, accuracy result of predictions is provided which shows improvements in the predictions [12].

Mojtaba, Hossein, and Haghifam (2015) offered BN method which was applied to the IEEE reliability test system and the results demonstrate that the offered method is effective and is flexible in applications [11]. The IEEE Reliability Test System (RTS) was evolved in order to supply a common test system which could be used for comparing the results obtained by different methods [13]. BN had been a popular predictive modeling formalism, effective method from the above studies. The method used BN in this study to analyze the items, these types of error in the test and skills (Fig. 4 to 5). In these networks, the probability distributions were a variety of reaction database using a variety of candidates, including expert opinions drawn.

TABLE IV: THE BN DIFFERENCES BETWEEN MODEL 1 AND 2

Increasing items	CR1~CR3	
Increasing connected lines between items and error types	Error Types	Items
	B1	CR1,CR3
	B2	CR1,CR2
	B3	CR1,CR2
	B4	CR3
	B5	CR2,CR3
	B6	CR1, CR2,CR3
	B7	CR1, CR2,CR3

TABLE V: THE RECOGNITION RATE OF CONSTRUCTED-RESPONSE ITEMS

		Estimates	
		True	False
True values	True	n_{11}	n_{10}
	False	n_{01}	n_{00}

Note: $n_{ij} = 1$, If the measured value is consistent with the true value; otherwise $n_{ij} = 0.$

BN, from left to right, is divided into three levels: item (from the item 1to the item 3, and from CR1 to CR3, error types (from error type 1 to error type 5 that is, B1 ~B5) and skills (from skill 1 to skill 4, that S1~S4), respectively. Further, in this figure, the solid arrow line represents a number of lines of correlation between the left and right. In the BN figure, the output node is adapted to have, zero and one, two states. Setting this code into MATLAB7.0, and then supplying a different classification decision, will cause a corresponding recognition rate. Then supplying other classification decision values can search the best one of recognition rates. According to the students' responses, the study would the compute the joint probability through BN to evaluate whether students master these skills, there were still certain types of errors, to understand the abilities and errors of students. In this study, BN was regarded as an inference instrument for diagnostic tests to help teachers analyze and diagnose the types of errors and whether students have the skills or not, as the basis for the future remedial teaching. This study researched and analyses of the effectiveness of the above-mentioned multiple-choice (I1~I3), Constructed response items (CR1~CR3), in order to compare BN modes. As follows:

1) Model 1

The model consisted of three multiple-choice items, from left to right nodes are related topics: from item1 is Item3

(namely I1~I3); error types: from error type1 to error type 5 (namely B1~B5). and skills: from skill1 to skill 4 (namely S1~S4).

2) Model 2

The model consists of three multiple-choice questions of model1 in addition to three Constructed-Response items (CR1~CR3).
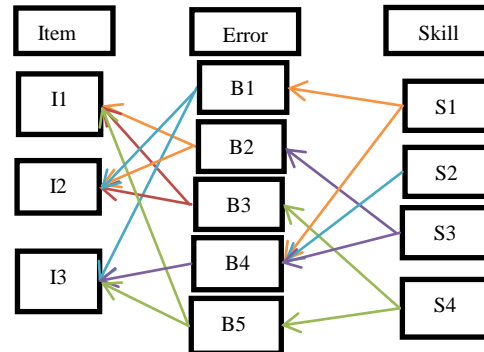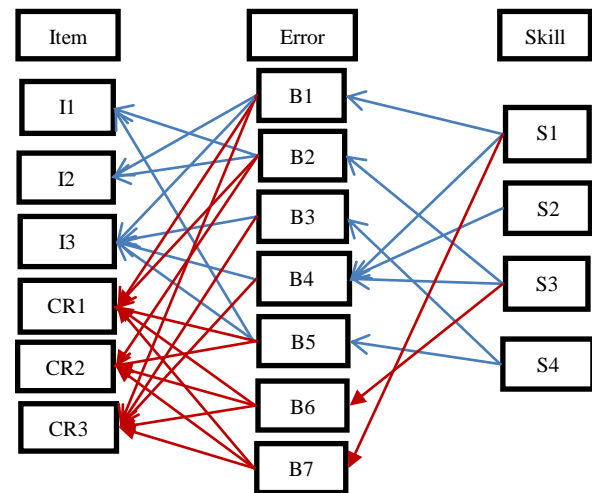

Fig. 4. The BN figure of model 1.

(where the blue lines represent the same connected lines like model 1, and the red lines represent the new connected lines)
Fig. 5. The BN figure of model 2.

According to these two modes, the present study was conducted to assess the effectiveness. The average skills' recognition rates of from the model 1 to 2, were 0.6387, 0.7501, respectively. The average errors' recognition rates, from model 1 to 2, were 0.7882 and 0.7942, respectively. Average recognition rates of Model 1 to 2, were 0.7135 and 0.7722, respectively. Results are expressed that their overall recognition rates of the BNs, which increased the CR items (model 2) was better than one, which had only the multiple - choice items (model 1).

III. RESULTS AND DISCUSSION

In recent years, Bayesian network has become an important method for modeling real-world applications in the decision-making problems [14]. The study used the "indefinite integral"as a university research, to establish a computerized diagnostic test, based on the ideas of the problem, combined with the BN, including multiple-choice and constructed- response items. The automatic analysis system was developed with the desire that, through this system, teachers can apply the model as a reference to supply

remedial teaching. This chapter described the results, based on the research purposes.

According to the research purposes and the results, the results are expressed as follows:

1) A computerized diagnostic system, which added Constructed Response items to the multiple-choice items, was built. It had the advantage of computer data storage that can store complex responses of Constructed-Response items.

2) The average of the recognition rates was 0.7722. If the mechanism was suitable for automatic analysis of assessment test, the burden of teachers would significant be decrease.

3) In this study, the recognition rate (0.7722) of model 2 was better than one, which had only the multiple- choice items (model 1).

BN had been a common predictive modeling and an efficient method from the researches done by predecessors. The study also had the same result.

According to the purposes and the results of this research, the discussion was presented as follows:

1) The Constructed-Response test interface of the study is unique. In addition, it was implied prior to applicable testing that the suitable instruction for use be reviewed.

2) The automatic analysis mechanism of Constructed-Response items acts well in diagnostic tests and can provide students with the error types.

3) This study concentrated on the diagnostic test and the automated analysis mechanism. Future research can be worked on the remedial teaching.

4) This study used settle the Constructed-Response items of the indefinite integral, and this interface can also be used to learn Calculus other units.

IV. CONCLUSIONS

Conclusions of this research were presented as follows:

1) to build an "indefinite integral about integrands involving quadratic square root" online test system.

2) to diagnosis and analyze the effectiveness of indefinite integral's computerized responses of Constructed-Response items. In the study, the recognition rate of model 2 was 0.7722. However, recognition rate of model 1, only had the multiple-choice items, was 0.7135, and.

3) The automatic analysis mechanism was suitable for computerized test, the burden of teachers would be decrease.

REFERENCES

[1] D. Tall, "Functions & calculus," *International Handbook of Mathematics Education*, Springer Netherlands, 1996, pp. 289-325.

[2] B. MCornu, "Limits," *Advanced Mathematical Thinking,* Dordrecht: Kluwer Academic Publishers, 1991, pp. 153–166.

[3] T. A. Jensen, "A study of the relationship between introductory calculus students' understanding of function and their understanding of limit," Montana state University, 2009, p. 229.

[4] T. T. Kathleen, "Error reduction strategies for whole number operations in grade four," presented at the Doctoral Dissertation, Brigham Young University, p. 120, 1987.

[5] J. L. Ticehurst, R. A. Letcher, and D. Rissik, "Integration modeling and decision support: The coastal lake assessment and management (CLAM) tool," *Mathematics and Computers in Simulation*, vol. 78, pp. 435-449, July 2008.

[6] L. Uusitalo, "Advantages and challenges of Bayesian networks inenvironmental modeling," *Ecological Modelling*, vol. 203, pp. 312-318, 2007.

[7] P. Jayneel, S. Shahram, and A. M. Thomas, "Knowledge based data center capacity reduction using sensitivity analysis on causal Bayesian belief network. *Information Knowledge Systems Management*," vol. 12, pp. 125-148, 2012.

[8] V. Marina, J. F. L. Peter, S. Maurice, and K. Nico, "On the interplay of machine learning and background knowledge in image interpretation by Bayesian networks," *Artificial Intelligence in Medicine,* vol. 57, pp. 73-86, 2012.

[9] M. Eva, D Luis, C. Gladys, O. Paula, and D. Sandra, "Using Bayesian networks to improve knowledge assessment," *Computers & Education*, vol. 60, pp. 436-447, 2012.

[10] M. Aminian, D. Couvin, A. Shabbeer, K. Hadley, S. Vandenberg, N. Rastogi, and K. P. Bennett, "Predicting mycobacterium tuberculosis complex clades using knowledge-based Bayesian networks," *BioMed Research International*, vol. 2014, pp. 1-12, 2014.

[11] E. Mojtaba, S. Hossein, and M. R. Haghifam, "Incorporation of protection system failures into bulk power system reliability assessment by Bayesian networks," *Generation, Transmission & Distribution, IET*, vol. 9, no. 11, pp. 1226 -1234, 2015.

[12] G. A. Mostafa, N. Mahmoud, and S. G.Bahram, "Learning parameters of fuzzy Bayesian network based on imprecise observations," *International Journal of Knowledge-based and Intelligent Engineering Systems*, vol. 18, pp. 167-180., 2014.

[13] R. Billinton and R. N. Allan, "Reliability assessment of large electric power systems," Kluwer Academic Publisher, Massachusetts, United States, 1988.

[14] G. A. Mostafa, N. Mahmoud, and S. G. Bahram, "Learning parameters of fuzzy Bayesian Network based on imprecise observations," *International Journal of Knowledge-Based and Intelligent Engineering Systems*, vol. 18, pp. 167-180, 2014.

Mu-Yu Ting is currently an associate professor of Center for General Education at National Formosa University, Taiwan. She received her Ph.D. degree in Graduate Institute of Educational Information and Measurement from National Taichung University of Education, Taiwan. Her research interests include knowledge structure analysis, calculus teaching and dynamic assessment.