2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

Type-2 Fuzzy Set Construction and Application for Adaptive Student Assessment System

Mei-Hui Wang, Chi-Shiang Wang, Chang-Shing Lee, Su-Wei Lin, and Pi-Hsia Hung

Abstract—**Student's performance is classified into four levels, including below basic, basic, proficient, and advanced levels. The descriptions of the performance standard make students understand their learning achievement via percentile rank (PR), a norm-referenced score, and *T* score (*T*). This paper develops an adaptive student assessment system and invites elementary-school students to do a test in mathematics. Additionally, one adaptive item selection strategy mechanism is developed to choose next item that meets the student's current estimated ability. After that, the response data are collected to execute the type-2 fuzzy set (T2FS) construction mechanism to build a personalized T2FS for each student's performance and a T2FS for all students with an identical level. Finally, the student evaluation mechanism is executed to show students and teachers some useful information to assist in their future teaching and guidance. The simulation results show the proposed approach is feasible to adaptively select items from the item bank and construct T2FS for students' ability. In the future, we plan to use the technologies of optimization and computational intelligence to infer each student's ability in the test based on the constructed T2FSs.**

I. Introduction

TODAY'S educational goal not only hopes to provide students with completed education and learning environment to enhance students' overall competitiveness, but also hopes to understand their learning performance and condition to offer them corresponding guidance and teaching. Generally speaking, performance of students is classified into four levels, including below basic, basic, proficient, and advanced levels. The descriptions of the performance standard make students understand their learning achievement via a norm-referenced score, percentile rank (PR). However, evaluation of students' learning achievement is an important topic for schools, teachers, parents, and students. But, how to objectively evaluate to understand student's learning achievement and literacy is still an effort that domain experts try to make in the world.

Psychometrics is a science that combines psychological testing and assessment and it is a quantitative method that applies to assessment, scaling, and evaluation [1]. Testing theory is classified into classical test theory (CTT) and modern test theory. Item response theory (IRT) is the structure of modern test theory [1, 2]. IRT considers that there exists a certain kind of relationship between the performance of students with a different ability and an item. This relationship is called an item characteristic curve (ICC) to represent the possible relationship between students' ability (or a certain latent trait) and probability that students correctly answer this item [1, 2]. According to 3 parameter logistic (3PL) model of the dichotomous scoring, one item has three parameters, namely a, b, and c, to represent this item's discrimination, difficulty, and guessing, respectively [1, 2]. One student's ability can be estimated based on his/her response to all the selected items, real-time change of his/her ability, and parameters of the selected items.

Recently, with the rapid growth of computer and computational intelligence technologies, integrating computer-based test with computational intelligence is still a challenge. Even though, incorporated with the technologies of the computational intelligence, there has been considerable research on education: Todai robot project team members, executed by National Institute of Informatics, Japan, have been developing a computer program by integrating multiple artificial intelligence technologies and their aim is to achieve a high score on the National Center Test for University Admissions by 2016 and to pass the University of Tokyo entrance exam in 2021 [3, 4]. Huang et al. [5] developed an adaptive testing system to efficiently conduct an adaptive test to reliably estimate students' ability. Lee [6] presented a computational method to efficiently estimate the ability of students in a Web-based learning environment by capturing their problem solving processes. Badaracco and Martinez [7] proposed a multi-criteria decision model (MCDM)-based item selection algorithm to enhance the accuracy of diagnosis and the adaptation of computerized adaptive tests (CAT) to students' competence level. Hwang and Chang [8] proposed a formative assessment-based approach for improving the learning achievements of students in a mobile learning environment.

Real life is full of uncertainty, but two important kinds of uncertainties are linguistic and random [9]. The former is associated with words, and the latter is associated with unpredictability [9]. The membership degree is a crisp for type-1 fuzzy set (T1FS), where a type-2 fuzzy set (T2FS) has fuzzy grades of membership that are bounded in [0, 1]. The concept of T2FS was first proposed by Zadeh in 1975; however, not so many persons had extended a type-1 fuzzy logic system (T1FLS) to a type-2 FLS (T2FLS) until the Karnik and Mendel's work in 1998 [10, 11]. This is because characterizing a T2FS is not as easy as characterizing a T1FS. In the last decade, T2FS has been applied to many research topics and shown to acquire a good performance. For

The authors would like to thank the financial support by the National Science Council of Taiwan under the grant NSC 102-2221-E-024-005.

Mei-Hui Wang, Chi-Shiang Wang, and Chang-Shing Lee are with the Dept. of Computer Science and Information Engineering, National University of Tainan, Taiwan. (corresponding author to provide phone: 886-6-2606123 ext. 7709; fax: 886-6-2606125; e-mail: leecs@mail.nutn.edu.tw).

Su-Wei Lin and Pi-Hsia Hung are with Dept. of Education, National University of Tainan, Taiwan.

example, Bernardo et al. [12] presented a genetic T2FLS for modeling the financial applications. Lee et al. [13] combined fuzzy markup language (FML), which was proposed by Acampora and Loia [14, 15], with T2FS to evaluate the diet healthy level. And, they also combined T2FS with fuzzy ontology to recommend a personal diabetic-diet menu [16]. Additionally, an architecture of the perceptual computer (Per-C) for making subjective judgments using Computing With Words (CWW) was proposed by Mendel in 2001 because words can mean different things to different people [17]. The paradigm of CWW has been gaining more attention [18] based on T2FS, for example, Bilgin et al. [18] employed the general type-2 fuzzy logic to dynamically model the human perceptions based on the human experience. Esposito and Pietro [19] proposed an interval type-2 fuzzy logic to automatically encode clinical practice guidelines (CPGs) by means of if-then rules. Liu and Mendel [20] proposed an interval approach (IA) to obtain interval T2FS models for words based on a type-2-fuzzistics methodology.

This paper develops an adaptive student assessment system (ASAS) to estimate the students' ability according to IRT. Based on the students' responses to the selected items, a personalized T2FS for each student's performance is constructed. Then, we category students' performance into four levels to construct the corresponding T2FS. In the future, combined particle swarm optimization (PSO) with fuzzy inference mechanism, the constructed T2FS will be the knowledge based of the system to infer each student's performance. The remainder of the paper is as follows. Section II introduces the proposed adaptive student assessment system. Section III introduces the type-2 fuzzy set construction mechanism for student performance evaluation. Simulation results are shown in Section IV. Finally, conclusion and future work are given in Section V.

II. Adaptive Student Assessment System

A. Item Characteristic Curve and Item Information

IRT is with the concepts of parameters invariance and information function. Therefore, the closer to examinee's ability the difficulty of the test, the smaller the estimated standard error and the more accurate the estimated examinee's ability [1, 21, 22, 23]. Fig. 1(a) shows the characteristic curves (ICC) for three items and it describes the relationship between the ability of individuals and the probability of their answering a test question correctly [23], where parameters a, b, and c represent this item's discrimination, difficulty, and guessing, respectively. Each item in the test has its own ICC. Fig. 1 indicates the following information: (a) The higher parameter a, the steeper the curve. In other words, Item 3 is with a better discrimination than Items 1 and 2. (b) The higher parameter b, the more difficult the lower-ability examinees answer correctly. (c) The probability of correct response to Item 2, $P(1|\theta)$, starts to increase when examinee's ability (θ) are over zero.

In this paper, 3PL model of the dichotomous scoring is adopted. The probability of correct response is calculated by Eq. (1), so the probability of response is expressed by Eq. (2).

$$P(U_i=1|\theta)=c_i+(1-c_i)\times\frac{e^{Da_i(\theta-b_i)}}{1+e^{Da_i(\theta-b_i)}}\text{, where } i = 1, 2, \ldots, N \quad (1)$$

where, (a) U_i denotes the response pattern of the i^{th} item which is shown in Eq. (3), (b) a_i, b_i, and c_i are discrimination, difficulty, and guessing of the i^{th} item, respectively, (c) θ is the ability of examinee, and (d) D equals 1.7.

The response pattern (**U**) composed of N items' response is expressed by Eq. (4). The examinee's joint probability for **U**, $P(U_1, U_2, \ldots, U_N|\theta)$, and the item information function for the i^{th} item, $I_i(\theta)$, are computed by Eqs. (5) and (6), respectively. Fig. 1(b) shows a three-item information function that demonstrates the following results: (a) Item 3 is more difficult than Items 1 and 2; hence, the item information functions for Items 1 and 2 are centered at a lower ability level than the one for Item 3. (b) Because Items 1 and 2 are less discriminating than Item 3, their corresponding item information functions are lower than Item 3 [1, 21, 22, 23].

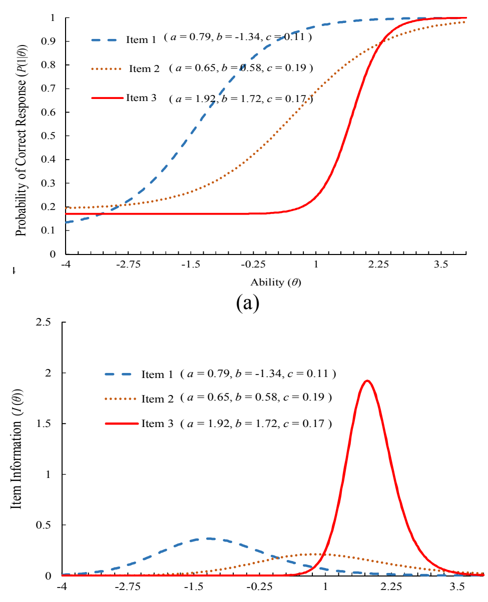
$$\begin{aligned}
&P(U_i|\theta)\\
&= P(U_i=1|\theta)^{U_i}P(U_i=0|\theta)^{1-U_i}\\
&= P_i^{U_i}\, Q_i^{1-U_i} \qquad (2)
\end{aligned}$$

$$U_i=\begin{cases}1, \text{make a right response to the } i^{\text{th}} \text{ item}\\ 0, \text{make an wrong response to the } i^{\text{th}} \text{ item}\end{cases} \qquad (3)$$

$$\mathbf{U}=(U_1, U_2, \ldots, U_N) \qquad (4)$$

$$\begin{aligned}
&P(U_1, U_2, \ldots, U_N|\theta)\\
&= P(U_1|\theta)\times P(U_2|\theta)\times P(U_3|\theta)\ldots\times P(U_N|\theta)\\
&= \prod_{i=1}^{n} P(U_i\mid\theta)=\prod_{i=1}^{n} P_i^{U_i}Q_i^{1-U_i} \qquad (5)
\end{aligned}$$

$$I_i(\theta)=D^2a_i^2\left(\frac{Q_i(\theta)}{P_i(\theta)}\right)\left(\frac{P_i(\theta)-c_i}{1-c_i}\right)^2 \qquad (6)$$



(b)

Fig. 1. (a) Item characteristic curve and (b) item information.

B. Adaptive Student Assessment System Structure

The proposed adaptive student assessment system structure is shown in Fig. 2 and is described as follows:

- Item bank is first established by domain experts. Then, students surf on the Internet to do a test via the provided interface.
- The item selection strategy mechanism adaptively chooses next item from the established item bank based on real-time students' ability. The details of item selection algorithm is described in Section II.C. After finishing the test, the students' response to the selected items are stored into the response data.
- The T2FS construction mechanism builds T2FSs for the performance of each student and the students with an identical level.
- The student performance evaluation mechanism demonstrates the diagnosis reports for students and teachers to understand whether the students have achieved the performance that they were supposed to have for this test.
- Teachers are able to retrieve response data to further understand students' learning situation.

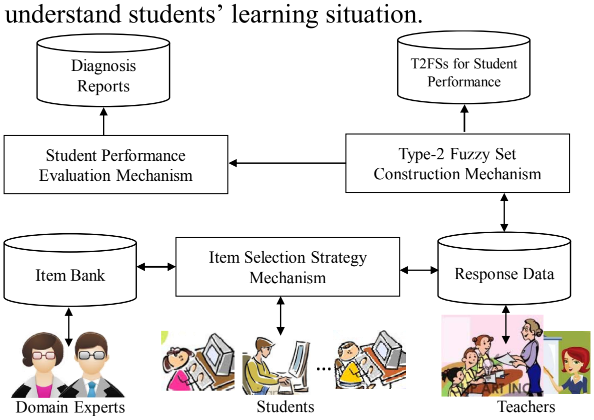


Fig. 2. Adaptive student assessment system structure.

C. *Item Selection Strategy Mechanism*

This subsection describes the item selection strategy mechanism. Fig. 3 shows the flowchart of the item selection strategy mechanism and its descriptions are as follows:

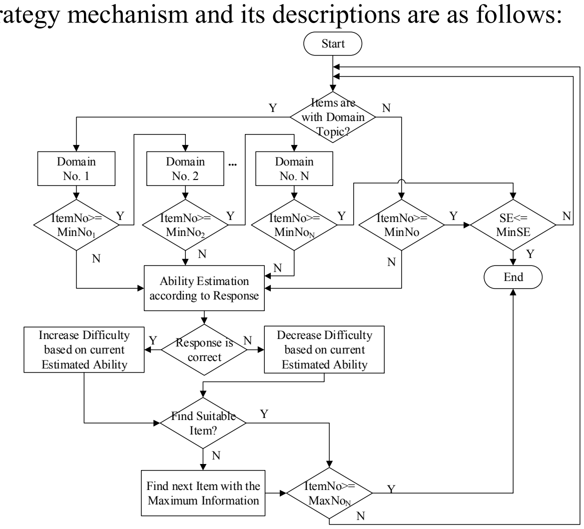


Fig. 3. Item selection strategy mechanism flowchart.

- Check if this subject's items are grouped into N domain topics, for example, Domain No. 1, No. 2, …, No. N in the test.
- If Yes, then the adopted mechanism selects the enough items from each domain topic until the minimum number of items for each domain ($MinNo_1$, $MinNo_2$, …, $MinNo_N$) is reached. If No, then the adopted mechanism selects next item from all possible ones until the minimum number of items ($MinNo$) is reached. No matter whether this subject's items are grouped into different domain topics, the adopted mechanism selects next item according to this student's current estimated ability and his/her response to last item.
- When the number of tested items ($ItemNo$) reaches the minimum number of items and standard error of estimation (SE) is less than or equal to the minimum SE ($MinSE$), the adopted mechanism ends the test. On the contrary, the adopted mechanism checks whether $ItemNo$ reaches the maximum number of items ($MaxNo_N$) or not. If Yes, then the adopted mechanism ends the test. If No, then the adopted mechanism selects a suitable item for this student to continue the test. Table I shows the detailed algorithm of the item selection strategy mechanism.

TABLE I. ITEM SELECTION STRATEGY MECHANISM ALGORITHM.

Input:

1. θ: Real-time estimated ability for this student and its range is between +4 and -4
2. U_i: Student's response to the i^{th} item and its value is 1 (answer correctly) or 0 (answer incorrectly) for dichotomous scoring
3. Items: $Item_1$, $Item_2$, …, $Item_N$ and each item is with three parameters, including a_i, b_i, and c_i /*N denotes the number of items and i=1, 2, …, N*/

Output:

$NextItem$ /*Selected $(i+1)^{th}$ item*/

Method:

Step 1:

Step 1.1: $Range\_gap \leftarrow DV_{Range\_gap}$ /*DV_{Range_gap} denotes the default value of $Range\_gap$*/

Step 1.2: $Offset\_gap \leftarrow DV_{RangeOffset\_gap}$ /*DV_{Offset_gap} denotes the default value of $Offset\_gap$*/

Step 1.3: $Offset \leftarrow 0$

Step 1.4: $NO_{SelectItem} \leftarrow 0$

Step 1.5: $MaxNO_{SelectItem} \leftarrow DV_{MaxNOSelectItem}$ /*$DV_{MaxNOSelectItem}$ denotes the default value of $MaxNO_{SelectItem}$*/

Step 1.6: $NextItem \leftarrow$ Null

Step 2: Do Until ($NO_{SelectItem} >= MaxNO_{SelectItem}$)

/*$NO_{SelectItem}$ denotes the number of times that the selection strategy mechanism has tried to select the suitable item from the item bank to meet the current student's estimated ability*/

/*$MaxNO_{SelectItem}$ denotes the maximum number of times that the selection strategy mechanism is allowed to select the suitable item from the item bank to meet the current student's estimated ability*/

Step 2: If U_i equals 1 /*Answer correctly, so select next item that is more difficult than this student's current estimated ability*/

Step 2.1: Set the range of next item's difficulty to the interval $DifficultyRange \leftarrow [\theta, \theta + Range\_gap + Offset]$

Step 3: If U_i equals 0 /*Answer incorrectly, so select next item that is easier than this student's current estimated ability*/

Step 3.1: Set the range of the next item's difficulty to the interval $DifficultyRange \leftarrow [\theta - Range\_gap - Offset, \theta]$

Step 4: If there is an item whose difficulty is in the interval $DifficultyRange$

Step 4.1: $NextItem \leftarrow Item_{Selected}$ /* $Item_{Selected}$ denotes the selected item from the item bank*/

Step 4.2: $NO_{SelectItem} \leftarrow MaxNO_{SelectItem}$

Step 5: If there is no any items whose difficulty are in the interval $DifficultyRange$

Step 5.1: $Offset \leftarrow Offset + Offset\_gap$

Step 5.2: $NO_{SelectItem} \leftarrow NO_{SelectItem} + 1$

Step 6: End Do Until

Step 7: If *NextItem* equals Null
Step 7.1: Sort the information of items that have not been tested by this student in an ascending order
Step 7.2: Select the item that provides maximum information and is also most close to this student's estimated ability
Step 7.3: Set this item to $Item_{Selected}$, that is, $NextItem \leftarrow Item_{Selected}$
Step 8: End

III. Type-2 Fuzzy Set Construction Mechanism for Student Performance Evaluation

A. Introduction to Type-2 Fuzzy Set

Students' ability exists so many uncertainties that it is hard to express the knowledge of the students' ability only by a crisp value. In this sub-section, suppose the variable of interest is "*Student Performance* (*SP*)," where a proficient student is used as an example. A T1FS *SP* is constructed and shown in Fig. 4(a). In this paper, a trapezoidal function can be expressed as the parameter set [*BS*, *BC*, *EC*, *ES*]. For instance, the membership function *SP*, shown in Fig. 4(a), can be denoted as [0.6, 0.9, 0.9, 1.25]. Hence, if the input of *SP* is SP', the membership value of a T1FS will be certain and its value is 1. On the other hand, if we want to reflect all the experts' opinions, a T2FS for *SP* is constructed and shown in Fig. 4(b), where *SP* is a primary variable and *u* is a secondary variable. Below is the brief descriptions of a T2FS [9, 10, 11]:

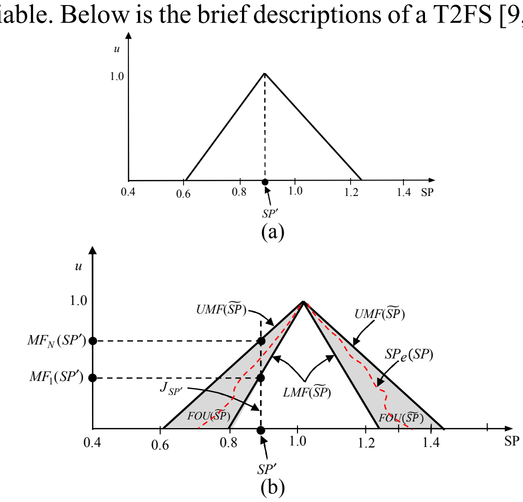


Fig. 4. (a) T1FS and (b) T2FS for SP.

- A T2FS is characterized by a fuzzy membership function (MF), and its membership value for each element of this set is a fuzzy set in [0, 1] [9]. Meanwhile, the MFs of T2FSs are three-dimensional and consist of a bounded region that we call the footprint of uncertainty (FOU), like the gray-shaded area in Fig. 4(b).
- The trapezoidal function for a T2FS can be expressed as a parameter set {[BS_L, BC_L, EC_L, ES_L], [BS_U, BC_U, EC_U, ES_U]}. For instance, the membership function $\widetilde{SP}$, shown in Fig. 4(b), can be denoted as {[0.8, 1.0, 1.0, 1.25], [0.6, 1.0, 1.0, 1.42]}.
- The upper membership function (UMF) and lower membership function (LMF) of $\widetilde{SP}$ are two type-1 MFs that bound the FOU. The UMF is associated with the upper bound of FOU($\widetilde{SP}$) and is denoted $UMF(\widetilde{SP})$, and the LMF is associated with the lower bound of FOU($\widetilde{SP}$) and is denoted $LMF(\widetilde{SP})$ [9].
- Suppose there are exactly *N* MFs in Fig. 4(b), then at each value of *SP*, there can be as much as *N* membership values. For example, the membership values of input SP' are $MF_1(SP')$,…, and $MF_N(SP')$. As a result, the primary membership values lie in the interval [$MF_1(SP')$, $MF_N(SP')$] when *SP* is SP', and $J_{SP'}$ is its primary membership.
- The dashed wavy curve, shown in Fig. 4(b), is an embedded T1FS, $SP_e(SP)$.

B. Type-2 Fuzzy Set Construction Mechanism

There are two methods to construct student's personalized performance and they are described as follows: (a) Method 1 (See Steps. 1-5 in Table II): Divide all tested items into some groups, for example, each group is composed of five items. Find minimum-item-difficulty and maximum-item-difficulty values for each group. Then, average all-group minimum-item-difficulty and maximum-item-difficulty values to generate begin support (BS) and end support (ES) of T1FS *Item Difficulty* (*ID*), respectively. After that, average BS and ES to get begin core (BC) and end core (EC) of T1FS *ID*. (b) Method 2 (See Steps. 6-12 in Table II): Find the minimum-item-difficulty and maximum-item-difficulty values for all tested items. Then, set minimum-item-difficulty and maximum-item-difficulty values to BS and ES of T1FS *ID*. After that, average BS and ES to get BC and EC of T1FS *ID*. Repeat the similar procedures to construct T1FS *Estimated Ability* (*EA*). Finally, construct T2FS Student Performance ($\widetilde{SP}$) according to the parameters of T1FSs *ID* and *EA* (See Steps. 5 and 12 in Table II). Table II shows the algorithm to construct each-individual T2FS $\widetilde{SP}$.

TABLE II. Algorithm to Construct Each-Individual T2FS $\widetilde{SP}$.

Input:
1. Item Difficulty (B) = {b_1, b_2, b_3, …, b_N} /*N denotes the number of the tested items*/ /*b_1, b_2, b_3, …, and b_N represent the 1st , 2nd, 3rd, …, and Nth item's difficulty, respectively.*/
2. Estimated Ability (θ) = {θ_1, θ_2, θ_3, …, θ_N} /*N denotes the number of test items*/ /*θ_1, θ_2, θ_3, …, and θ_N represent this student's estimated ability when he/she finishes making a response to the 1st , 2nd, 3rd, …, and Nth item, respectively.*/

Output:
1. $\widetilde{SP}_1$: T2FS for this student's performance by method 1
2. ID_1: T1FS for items' difficulty by method 1
3. EA_1: T1FS for this student's estimated ability by method 1
4. $\widetilde{SP}_2$: T2FS for this student's performance by method 2
5. ID_2: T1FS for items' difficulty by method 2
6. EA_2: T1FS for this student's estimated ability by method 2

Method:
/*Method 1 to construct T2FS $\widetilde{SP}_1$*/
Step 1: Initialize the parameters
Step 1.1: $Quotient \leftarrow$floor (N / M) /*floor() returns the largest integer less than or equal to the specified number and M denotes a divisor*/
Step 1.2: $Remainder \leftarrow$mod (N / M) /*mod() returns the remainder*/
Step 1.3: If *Remainder* is not zero
Step 1.3.1: $Quotient \leftarrow Quotient + 1$
Step 1.4: $MinBSet \leftarrow \emptyset$, $MaxBSet \leftarrow \emptyset$, and $AvgBSet \leftarrow \emptyset$
Step 1.5: $MinThetaSet \leftarrow \emptyset$, $MaxThetaSet \leftarrow \emptyset$, and $AvgThetaSet \leftarrow \emptyset$
Step 2: For $i \leftarrow 1$ to *Quotient*
Step 2.1: $BSet \leftarrow \emptyset$ /*Store the item difficulty of items*/
Step 2.2: $ThetaSet \leftarrow \emptyset$ /*Store the estimated ability after making a response to the tested item*/
Step 2.3: $Count \leftarrow 0$
Step 2.4: For $j \leftarrow 1$ to M

Step 2.4.1: index ← $(i - 1) \times M + j$
Step 2.4.1.1: If $index > N$
Step 2.4.1.1.1: Break
Step 2.4.1.2: $Count$←$Count$ + 1
Step 2.4.1.3: Add b_{index} to $BSet$
Step 2.4.1.4: Add θ_{index} to $ThetaSet$
Step 2.5: Sort $BSet$ in an ascending order
Step 2.6: Add $BSet_1$ to $MinBSet$ and add $BSet_{Count}$ to $MaxBSet$
Step 2.7: Sort $ThetaSet$ in an ascending order
Step 2.8: Add $ThetaSet_1$ to $MinThetaSet$ and add $ThetaSet_{Count}$ to $MaxThetaSet$
Step 3: Construct T1FS ID_1
Step 3.1: $BS_{ID} \leftarrow \frac{\sum_{k=0}^{Quotient} MinBSet_k}{Quotient}$
Step 3.2: $ES_{ID} \leftarrow \frac{\sum_{k=0}^{Quotient} MaxBSet_k}{Quotient}$
Step 3.3: $BC_{ID} \leftarrow (BS_{ID} + ES_{ID}) / 2$
Step 3.4: $EC_{ID} \leftarrow BC_{ID}$
Step 3.5: $ID_1 = [BS_{ID}, BC_{ID}, EC_{ID}, ES_{ID}]$
Step 4: Construct T1FS EA_1
Step 4.1: $BS_{EA} \leftarrow \frac{\sum_{k=0}^{Quotient} MinThetaSet_k}{Quotient}$
Step 4.2: $ES_{EA} \leftarrow \frac{\sum_{k=0}^{Quotient} MaxThetaSet_k}{Quotient}$
Step 4.3: $BC_{EA} \leftarrow (BS_{EA} + ES_{EA}) / 2$
Step 4.4: $EC_{EA} \leftarrow BC_{EA}$
Step 4.5: $EA_1 = [BS_{EA}, BC_{EA}, EC_{EA}, ES_{EA}]$
Step 5: Construct T2FS $\widetilde{SP}_1$ =
{[MAX(BS_{ID}, BS_{EA}), AVG(BC_{ID}, BC_{EA}), AVG(EC_{ID}, EC_{EA}), MIN(ES_{ID}, ES_{EA})],
[[MIN(BS_{ID}, BS_{EA}), AVG(BC_{ID}, BC_{EA}), AVG(EC_{ID}, EC_{EA}), MAX(ES_{ID}, ES_{EA})]}
={[BS_{SPL}, BC_{SPL}, EC_{SPL}, ES_{SPL}], [BS_{SPU}, BC_{SPU}, EC_{SPU}, ES_{SPU}]}

/*Method 2 to construct T2FS $\widetilde{SP}_2$*/
Step 6: $BSet$← ∅ /*Store the item difficulty of items*/
Step 7: $ThetaSet$← ∅ /*Store the estimated ability after testing items*/
Step 8: Sort item difficulty (B) = $\{b_1, b_2, b_3, \ldots, b_N\}$ in an ascending order and then $BSet$←B
Step 9: Sort estimated ability (θ) = $\{\theta_1, \theta_2, \theta_3, \ldots, \theta_N\}$ in an ascending order and then $ThetaSet$← θ
Step 10: Construct T1FS ID_2
Step 10.1: $BS_{ID} \leftarrow BSet_1$
Step 10.2: $ES_{ID} \leftarrow BSet_N$
Step 10.3: $BC_{ID} \leftarrow (BS_{ID} + ES_{ID}) / 2$
Step 10.4: $EC_{ID} \leftarrow BC_{ID}$
Step 10.5: $ID_2 = [BS_{ID}, BC_{ID}, EC_{ID}, ES_{ID}]$
Step 11: Construct T1FS EA_2
Step 11.1: $BS_{EA} \leftarrow ThetaSet_1$
Step 11.2: $ES_{EA} \leftarrow ThetaSet_N$
Step 11.3: $BC_{EA} \leftarrow (BS_{EA} + ES_{EA}) / 2$
Step 11.4: $EC_{EA} \leftarrow BC_{EA}$
Step 11.5: $EA_2 = [BS_{EA}, BC_{EA}, EC_{EA}, ES_{EA}]$
Step 12: Construct T2FS $\widetilde{SP}_2$ =
{[MAX(BS_{ID}, BS_{EA}), AVG(BC_{ID}, BC_{EA}), AVG(EC_{ID}, EC_{EA}), MIN(ES_{ID}, ES_{EA})],
[[MIN(BS_{ID}, BS_{EA}), AVG(BC_{ID}, BC_{EA}), AVG(EC_{ID}, EC_{EA}), MAX(ES_{ID}, ES_{EA})]}
={[BS_{SPL}, BC_{SPL}, EC_{SPL}, ES_{SPL}], [BS_{SPU}, BC_{SPU}, EC_{SPU}, ES_{SPU}]}
Step 13: End

Next, this sub-section also introduces the method to construct T2FS $\widetilde{SP}$ for all students with the same achievement performance. The procedures are similar to Table II; however, the difference is as follows: (a) Group students into four levels according to their T score. (b) For each-group students, average their first-item difficulty to generate a new item difficulty for the first item. Repeat this step for other items. (c) Execute Methods 1 and 2 in Table II to generate a T2FS for each-group students' performance. Fig. 5 shows its flowchart to construct different-level students' T2FS $\widetilde{SP}$.

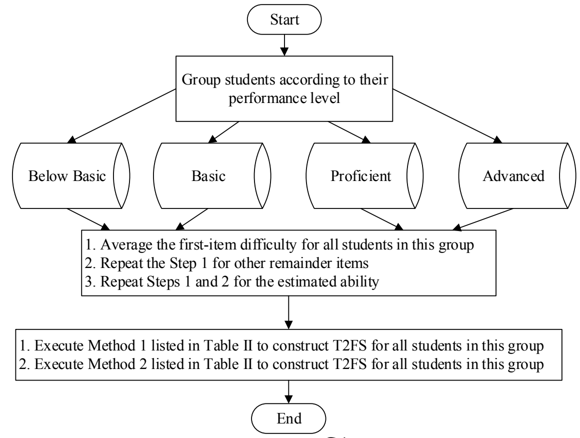


Fig. 5. Flowchart to construct T2FS $\widetilde{SP}$ of all students with the same performance level.

C. Student Evaluation Mechanism

This sub-section introduces the student evaluation mechanism. After testing, student's estimated ability (θ) will be transformed into T score (T) and percentile rank (PR). T score is transformed by Eq. (7), where mean value and standard error are 50 and 10, respectively. Additionally, student's performance is divided into four levels: Below Basic ($T < 40$), Basic ($54 > T >= 40$), Proficient ($65 > T > 54$), and Advanced ($T >= 65$). After testing, the developed system provides a diagnosis report to allow students and teachers to understand each student's individual achievement performance.

$$T \text{ score} = (\theta \times 10) + 50 \qquad (7)$$

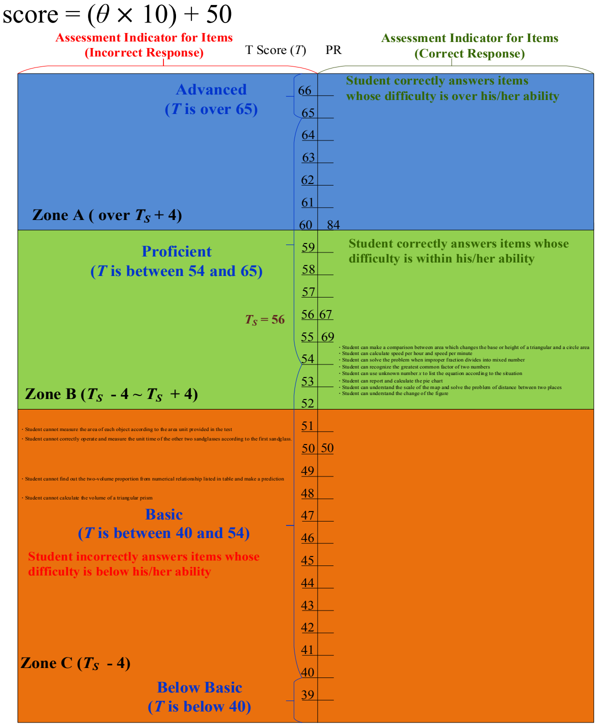


Fig. 6. Diagnosis report for one student whose T_S is 56.

Fig. 6 shows an example of a diagnosis report that this student's T score (T_S) is 56 and PR is 74. Table III lists the descriptions of the Zones A, B, and C. So, Fig. 6 indicates that (a) Zone A is the area when T is over T_S plus 4, (b) Zone B is the area when T is between T_S plus 4 and T_S minus 4, and (c) Zone C is the area when T is below T_S minus 4.

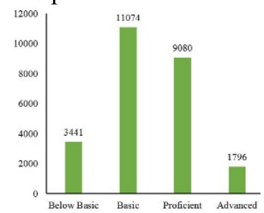
TABLE III. DESCRIPTIONS OF ZONES A, B, AND C.

Zone A	• In Zone A, item's difficulty is over this student's ability. As a result, this student has a potential to correctly answer the items in this zone. • If this student makes a right response to these items in Zone A, it means that the student can be advanced more if teachers provide further guidance for this student.
Zone B	• In Zone B, item's difficulty is with this student's ability. Therefore, this student challenges these items in Zone B. • If this student can correctly answer these items, it means that this student has achieved the achievement performance that he was supposed to have for this test.
Zone C	• In Zone C, items' difficulty is below this student's ability. • If this student cannot correctly answer these items, it means that this student has not yet achieved the achievement performance that he was supposed to have for this test. • The assessment indicators of the items that this student makes an incorrect response in Zone C are the main remedial director that teachers can provide for the student.

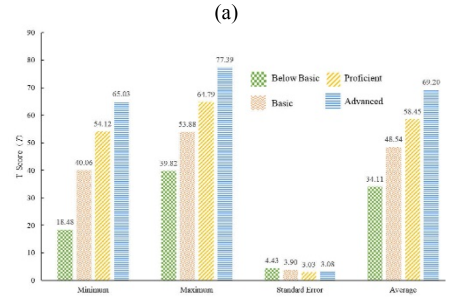
IV. SIMULATION RESULTS

This section shows the simulation results. In this paper, one adaptively student assessment system is developed and some four-grade students were invited to do a test for their mathematic ability. Fig. 7(a) shows four-level number of involved students, which indicates that most involved students' achievement level belongs to basic level. Fig. 7(b) shows the values of minimum T score, maximum T score, standard error, and average T score for all below-basic, basic, proficient, and advanced students, which indicates the higher the level, the bigger the values. Fig. 8 shows the item-difficulty variance for all tested items from the viewpoints of different achievement performances and it indicates the following situations: (a) Below Basic: The item difficulty keeps reducing during the test, (b) Basic: The variance in item difficulty keeps a stable situation, (c) Proficient: The item difficulty has a trend to gradually increase, and (d) Advanced: The item difficulty has a distinct increase than proficient-level students. However, the established item bank has no enough difficult items for advanced students after Item 16, so the item is selected according to item's information instead of item's difficulty at this time. This causes the item difficulty gradually to decrease after Item 16 till the end of the test.

Fig. 9 shows the variance in estimated ability, item difficulty, and item discrimination when one student does a test. After a test, this student's T score is 74.9, PR is 99, and his/her estimated ability is 2.49. There is a strong relationship between the estimated ability curve and the item difficulty curve. When this student's estimated ability increases, the item difficulty of the selected item is also increased. However, after Item 20, the item difficulty has a sharp decrease. This situation represents that the established item bank has no enough difficult items provided for this high-level student. Therefore, the item information that has a strong relationship with the item discrimination is adopted to be the main factor to select next item (Item 21) for this student. As Fig. 9 shows, the item discrimination of Item 21 is higher than all the other items. Figs. 10 and 11 show the constructed T2FSs for one below-basic-level and one advanced-level students, respectively. Table IV lists parameters of constructed FSs for a basic student and a proficient student.



(a)



(b)

Fig. 7. Bar charts for (a) number of students and (b) minimum, maximum, standard error, and average for below-basic, basic, proficient, and advanced levels.

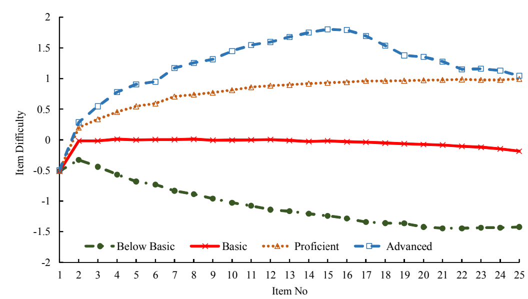


Fig. 8. Curves of variance in item difficulty for different levels.

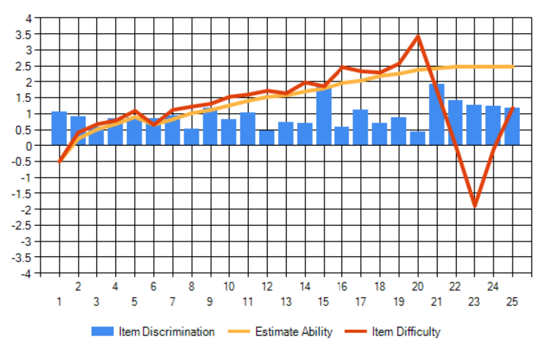
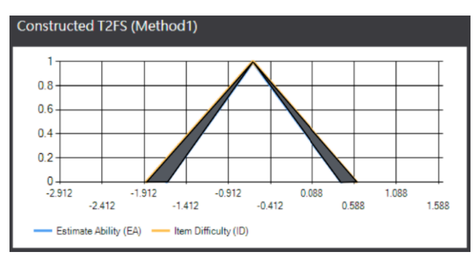
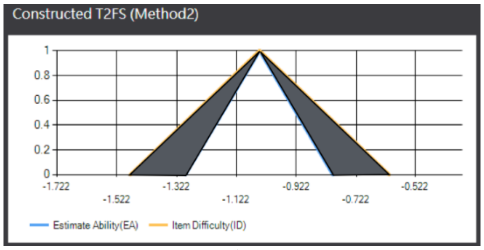


Fig. 9. Curves of variances in selected item difficulty, estimated ability, and item discrimination during a test.

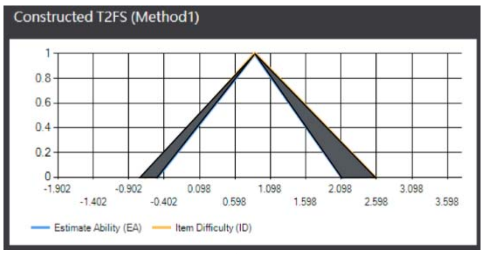


(a)

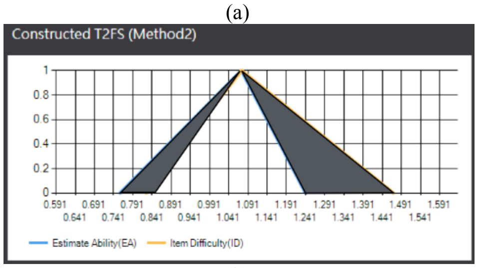


(b)

Fig. 10. Constructed T2FSs for a below-basic-level student whose T score is 34.24, PR is 6, and θ is -1.54 by (a) Method 1 and (b) Method 2.



(a)



(b)

Fig. 11. Constructed T2FSs for an advanced-level student whose T score is 71.09, PR is 98, and θ is 2.109 by (a) Method 1 and (b) Method 2.

TABLE IV. PARAMETERS OF CONSTRUCTED FSs FOR A BASIC AND A PROFICIENT STUDENTS.

Basic Student (T score is 47.3, PR is 39, and θ is -0.27)		
Constructed FS	Method 1	Method 2
T1FS *ID*	[-0.75, -0.01, -0.01, 0.73]	[-0.288, 0.001, 0.001, 0.29]
T1FS *EA*	[-0.5, -0.032, -0.032, 0.436]	[-0.148, 0,.03 0.03, 0.208]
T2FS $\widetilde{SP}$	{[-0.5, -0.021, -0.021, 0.436], [-0.75, -0.021, -0.021, 0.73]}	{[-0.148,0.015, 0.015, 0.208], [-0.288, 0.015, 0.015, 0.29]}
Proficient Student (T score is 62.85, PR is 90, and θ is 1.285)		
Constructed FS	Method 1	Method 2
T1FS *ID*	[-0.41, 0.595, 0.595, 1.6]	[0.68, 0.981, 0.981, 1.282]
T1FS *EA*	[-0.5, 0.378, 0.378, 1.357]	[0.68, 0.868, 0.868, 1.056]
T2FS $\widetilde{SP}$	{[-0.41, 0.486, 0.486, 1.357], [-0.5, 0.486, 0.486, 1.6]}	{[0.68, 0.924, 0.924, 1.056], [0.68, 0.924, 0.924, 1.282]}

V. CONCLUSION AND FUTURE WORK

This paper aims to develop an adaptive student assessment and construct T2FSs for an individual student and students with an identical level according to students' response data. The simulation results show that the proposed approach is feasible to construct T2FSs for students. In the future, incorporated the technologies of optimization and computational intelligence, we will motivate the use of two different methods to construct the student's personalized performance.

ACKNOWLEDGMENT

The authors would like to thank the involved students for their kind tests.

REFERENCES

[1] M. N. Yu, Item Response Theory (IRT) and its Application. Psychological Publishing Co., Ltd, Taiwan, Sept. 2009. (in Chinese)

[2] S. E. Embretson and S. P. Reise, Item Response Theory. Taylor & Francis, 2000.

[3] National Institute of Informatics, "Todai Robt Project," 2013 [Online]. Available: http://21robot.org/.

[4] National Institute of Informatics, "Todai Robot Project," *National Institute of Informatics News*, no. 46, Jul. 2013 [Online]. Available: http://www.nii.ac.jp/userdata/results/pr_data/NII_Today/60_en/all.pdf .

[5] Y. M. Huang, Y. T. Lin, and S. C. Cheng, "An adaptive testing system for supporting versatile educational assessment," *Computer & Education*, vol. 52, no. 1. pp. 53-67, Jan. 2009.

[6] Y. J. Lee, "Developing an efficient computational method that estimates the ability of students in a Web-based learning environment," *Computers & Education*, vol. 58, no. 1, pp. 579-589, Jan. 2012.

[7] M. Badaracco and L. Martinez, "A fuzzy linguistic algorithm for adaptive test in intelligent tutoring system based on competences," *Expert Systems with Applications*, vol. 40, no. 8, pp. 3073-3086, Jun. 2013.

[8] G. J. Hwang and H. F. Chang, "A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students," *Computers & Education*, vol. 56, no. 4, pp. 1023-1031, May 2011.

[9] J. M. Mendel, "Type-2 fuzzy sets and systems: an overview," *IEEE Computational Intelligence Magazine*, vol. 2, no. 2, pp. 20-29, Feb. 2007.

[10] N. N. Karnik and J. M. Mendal, "Introduction to type-2 fuzzy logic systems," in *Proceeding of 1998 World Congress Computational Intelligence* (*IEEE WCCI 1998*), Anchorage, Alaska, USA, May 4-9, 1998, pp. 915-920.

[11] J. M. Mendel, Uncertain Ruled-Based Fuzzy Logic Systems: Introduction and New Directions. Prentice Hall, USA, 2001.

[12] D. Bernardo, H. Hagras, and E. Tsang, "A genetic type-2 fuzzy logic based system for the generation of summarised linguistic predictive models for financial applications," *Soft Computing*, vol. 17, no. 12, pp. 2185-2201, Dec. 2013.

[13] C. S. Lee, M. H. Wang, G. Acampora, C. Y. Hsu, and H. Hagras, "Diet assessment based on type-2 fuzzy ontology and fuzzy markup language," *International Journal of Intelligent System*, vol. 25, no. 12, pp. 1187-1216, Dec. 2010.

[14] G. Acampora and V. Loia, "Fuzzy control interoperability and scalability for adaptive domotic framework," *IEEE Transactions on Industrial Informatics*, vol. 1, no. 2, pp. 97-111, May 2005.

[15] G. Acampora, C. S. Lee, M. H. Wang, and V. Loia, On the Power of Fuzzy Markup Language. Springer-Verlag, Germany, Jan. 2013.

[16] C. S. Lee, M. H. Wang, and H. Hagras, "A type-2 fuzzy ontology and its application to personal diabetic-diet recommendation," *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 2, pp. 374-395, Apr. 2010.

[17] J. M. Mendel and D. Wu, "Perceptual reasoning for perceptual computing," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 6, pp. 1550-1564, Dec. 2008.

[18] A. Bilgin, H. Hagras, A. Malibari, M. J. Alhaddad, and D. Alghazzawi, "Towards a linear general type-2 fuzzy logic based approach for computing with words," *Soft Computing*, vol. 17, no. 12, pp. 2203-2222, Dec. 2013.

[19] M. Esposito and G. D. Pietro, "Interval type-2 fuzzy logic for encoding clinical practice guidelines," *Knowledge-Based Systems*, vol. 54, pp. 329-341, Dec. 2013.

[20] F. Liu and J. M. Mendel, "Encoding words into interval type-2 fuzzy sets using an interval approach," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 6, pp. 1503-1521, Dec. 2008.

[21] R. K. Hambleton and R. W. Jones, "An NCME instructional module on comparison of classical test theory and item response theory and their applications to test development," *Educational Measurement: Issues and Practice*, pp. 38-47, 1993.

[22] M. Galdin and L. Laurencelle, "Assessing parameter invariance in item response theory's logistic two item parameter model: a Monte Carlo investigation," *Tutorials in Quantitative Methods for Psychology*, vol. 6, no. 2, pp. 39-51, 2010.

[23] M. J. Ree, "Estimating item characteristic curves," *Applied Psychological Measurement*, vol. 3, no. pp. 371-385, 1979.