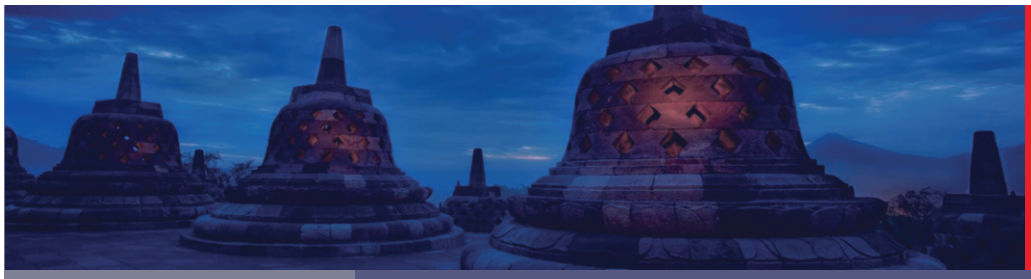

ISSN: 2407-1501

PROCEEDING

INTERNATIONAL CONFERENCE ON EDUCATIONAL RESEARCH AND EVALUATION (ICERE)

"Assessment for Improving Students' Performance"

May 29 – 31 2016
Rectorate Hall and Graduate School
Yogyakarta State University
Indonesia

Organized by:
Study Program of Educational Research and Evaluation
Graduate School, Yogyakarta State University
in Cooperation with Indonesian Educational Evaluation Association (HEPI),
and Center for Educational Assessment (PUSPENDIK) Ministry of Education and Culture

Proceeding
International Conference on Educational Research and Evaluation (ICERE) 2016

Publishing Institute
Yogyakarta State University

Director of Publication
Prof. Djemari Mardapi, Ph.D.

Address
Yogyakarta State University
ISSN: 2407-1501

Foreword of the Chairman

Assalamualaikum wr. wb.

Good morning ladies and gentlemen.

Praise be to Allah who has given abundant blessings so that we can hold this international conference.

This conference is aimed at improving the quality of assessment implemented in schools and other institutions. The quality of assessment determines students' ways of learning, so that it is hoped that the quality of education improves. Besides, this conference is a means of information exchanges in the forms of seminars dealing with results of research in educational assessment and evaluation. The expectation is that there is always improvement in educational assessment and evaluation methods, including in it is the instrument – both cognitive and noncognitive instruments.

The participants of this conference are the lecturers and teachers who teach educational assessment and evaluation, practitioners of assessment and evaluation, and researchers of assessment and evaluation. This conference can be held in cooperation with the Graduate School, Yogyakarta State University, Association of Educational Evaluation of Indonesia (HEPI), and Centre for Educational Research, Ministry of Education and Culture of Indonesia, supported by the Australian Council for Educational Research (ACER), Intel, Intan Pariwara Publisher, and many other institutions. For this reason, on behalf of the Organizing Committee, I would like to thank the Rector of Yogyakarta State University, Prof. Dr. Rochmat Wahab, M.Pd., M.A., and the Director of Graduate School, Yogyakarta State University, Prof. Dr. Zuhdan Kun Prasetyo, M.Ed., and all other institutions for their assistance and contribution that have made this conference possible. I would like to thank HEPI's Local Coordination Unit and all sponsors for supporting this conference and also all the audience for participating in this conference.

To the committee members, both in Jakarta and Yogyakarta, I would like to thank them for the hard work they have performed and for the togetherness so that this conference can be held.

Last but not least, we apologize for all the incoveniences you might encounter during this conference. Please enjoy the conference.

Wassalamu'alaikum wr. wb.

Prof. Djemari Mardapi, Ph.D.

Foreword of the Chairman of Himpunan Evaluasi Pendidikan Indonesia (HEPI)

Assalamu'alaikum Wr. Wb.

Indonesian Association for Educational Evaluation (HEPI) is a professional organization in education holding in the high esteem the principles of professionalism and knowledge development in the field of educational and psychological measurement, assessment, and evaluation. HEPI was established in November 19, 2000 in Yogyakarta, with a vision to become a professional organization that excels in the field of evaluation and measurement in education and psychology in Indonesia. Its mission is to develop up-to-date methodologies of evaluation, assessment, measurement, and data analysis in education and psychology, as well as studies of policies and technical implementation of the field for improving Indonesian education quality.

As a professional organization, HEPI brings together experts, practitioners and interested persons in the field of evaluation, assessment, and measurement of education, psychology and other social sciences. HEPI is open to anyone who has the interest the field with no restriction in terms of educational background and working experiences. Hopefully, through HEPI, members of the association can sustainably develop themselves as professionals. The existence of HEPI is also expected to contribute to the improvement of the quality of national education through research, consultancy, seminar, conference, publication, and training for members of the organization and for public audiences.

HEPI organizes annual workshop and conference in cooperation with the Regional Chapter of HEPI and universities. In 2016, for the first time HEPI organized **International Conference on Educational Research and Evaluation: Assessment for Improving Student's Performance** in May 29-30 2016 in Yogyakarta. This conference is jointly organized by HEPI and Yogyakarta State University and supported by the Center for Educational Assessment the Ministry of Education and Culture, Australian Council for Educational Research (ACER), INTEL Indonesia, and Intan Pariwara Publisher.

It is important to note that the choice of the HEPI 2016 conference theme is driven by the fact that the quality of our national education is still under expectation as shown by the results from School National Exam and international surveys conducted by some international agencies. HEPI believes that a number of factors contribute to the low quality of national education, including low teacher's knowledge and skills in classroom and school assessment. Therefore, improving the competence of teachers in classroom and school assessment is urgently required. In this context HEPI as a professional organization and individual members of the organization have to play an active role in improving teachers' competence in quality learning assessment.

In line with 2016 conference theme, HEPI invited two respected guest speakers , namely, Professor Geofferey Masters, Ph.D., Director of the Australian Council for Educational Research (ACER), who presented a paper on Assessment to Improve Student Competency and Professor Frederick Leung, Ph.D., from the University of Hong Kong, who delivered a paper on the International Assessment for Improving Classroom Assessment.

As a tradition, in 2016 conference HEPI organized two pre-conference workshops. The first workshop is on the conceptual introduction of Rasch model by Jahja Umar, Ph.D., senior lecturer at the Faculty of Psychology, State Islamic University Jakarta and the second workshop was delivered by Heru Widiatmo, Ph.D., researcher at American College Testing (ACT) Iowa, United States on Measuring Higher Order Thinking Skills (HOTS).

On behalf of HEPI, I would like to express my heartfelt gratitude to Rector of the Yogyakarta State University, invited speakers, resource persons, HEPI regional chapters, sponsors, speakers, participants, invited guests, and organizing committee who have worked hard in making this international coneference a success. Thank you very much for your participation and support and we are looking forward to seeing you in the next conference.

Last but not least, we hope that all of us get much benefit from this conference for enhancing Indonesian quality education through quality assessment.

Wassalamualaikum wr. wb.

Chairman,
BAHRUL HAYAT, Ph.D.

Table of Contents

USING FUZZY LOGIC TO SELECT ITEM TEST IN COMPUTERIZED BASE TESTING

Haryanto
haryanto.ftuny@gmail.com

ABSTRACT --- The development of the computerized base testing (CBT) with fuzzy logic algorithm aimed to investigate: the performance of fuzzy inference system in making decision to select the appropriate test items for the students. This research and development (R&D) study consists of two parts: (1) developing CBT program with *fuzzy* logic, (2) the CBT program testing was conducted to the samples from the students at Senior High School of Yogyakarta. Data were collected through observation, documentation, questionnaire and test. The data were analyzed using quantitative descriptive technique. The result of research showed that: (1) the CBT program is able to work properly according to the function and the type of users, and (2) the CBT program can managing (a) administer the test *bank*, (b) organize test items automatically, (c) organize the test items based on students' competence, (d) randomize the answer in the options, and (e) record the result of the test simultaneously and individually. Based on the data analysis the result of applying the CBT program for testing students' competence shows that: (1) the inference system of the program could appropriately provide adapted test items to the students based on the correct-incorrect responses they gave, (2) the collection of test items provided by the program to each participant can appropriately describe the competence of each students. In conclusion, the CBT program is using *fuzzy* logic algorithm can perform its function well to select the suitable test items and to measure the result of students' achievement in the learning process.

***Keywords**: Fuzzy logic, Inference systems, Item tes, Adaptive testing*

I. INTRODUCTION

The development and advancement of Science, Technology and the Arts (Science and Technology), particularly the field of Information Technology (IT), showed a significant increase for the benefit of education. It can be seen, among others: the use of IT for learning the subjects in the curriculum, the learning process with media-based IT, process test computer-assisted (PP No: 22, 2005; Baumgartner & Jackson, 1995; Bates & Poole, 2003; Chee & Wong, 2003).

Target of increasing the quality of education, through the measurement of IT-based, has sought to revive (Chee & Wong, 2003). Use of IT for testing purposes intended for effectiveness and efficiency of the implementation and operation of the test (Chee and Wong 2003 and Towndrow & Vallence 2004). IT era through Computerized Base Test (CBT), the computer is programmed to process the order to see the grains exam questions and receive answers from the participant's response (Chee and Wong, 2003; Towndrow & Vallence. 2004).

Based on observations of the programs used by institutions / agencies organizing computer-based testing, there are some things that need to be studied and researched further. The process of making random test item bank of questions in computer-based testing, raises the problem of mismatch between the level of difficulty about the ability of the test taker. On the other hand, retrieval of test items sequentially question bank that is widely used, weakens the validity of the test results.

Recent developments, along with the bank about the business of making and improving the quality of tests, computer technology is very helpful for this purpose (Hambleton, Swaminathan, & Rogers, 1991). Improved quality conformance test item is intended to test the ability of the test taker or often referred to as "Computerized Adaptive Test / CAT)". Called computerized (computerized) for the implementation of testing using the computer. Adaptive, that the level of difficulty of test items that appear based on the results of the analysis algorithms in a program, to match the abilities of participants. Suitability items are based on the responses of right and wrong answers to the test taker.

Various theories of evolution of computing to support human needs are constantly sought. One of these developments for the benefit of the test is based computational algorithms "Artificial Intelligence (AI)" or artificial intelligence (Luger, 2005). This theory describes an attempt to make the engine computer can "learn" so that it can work with the "intelligence" as humans. Problems associated with testing the learning outcomes, the computer is programmed with the provision of knowledge to determine the ability of test takers through intelligent analysis of the responses to participants' answers to the test, to choose the appropriate test items.

AI approach in this study is fuzzy logic. Fuzzy logic inference systems are used for the purposes of (decision) to a problem that is both qualitative and quantitative (Yan, Ryan, & Power, 1994). Application of the test is to select test items are appropriate and in accordance with the ability of the test participants, in order to obtain the results in the form of a valid measurement.

Based on the description above the expected research goals are to generate:

1. Model CBT program inference system with fuzzy logic algorithm to select test items right.
2. CBT program that is capable of providing reliable test results informai.

II. CBT MODEL WITH FUZZY LOGIC

Hambleton, Swaminathan, & Rogers (1991) in his book "Fundamentals of Item Response Theory" says that what is meant by testing adaptive computerized constitute "would be to give every examinee a test that is" tailored "or adapted, to the examinee's ability level" , It also said that tests with long items have been fixed, it is not efficient if given to all participants of the test. This was due to the ability of diverse participants should reasonably be given only to a few questions that correspond to the level. He said that the test can be shortened without losing precision test, if the test items prepared (administered) in accordance with the capabilities of each participant tests. For that, of course, requires complete information about the ability of the test participants, so that each participant will receive a set of test questions according to his ability (administered a unique set of items).

Masters & Keeves (1999) mentions that in the adaptive test, test items were selected from a question bank, based on the rule (rule) the selection of test items previously defined. Programmable computer with the ability to have a high convenience for use in the testing process is adaptive (adaptive testing).

Baeck, Fogel, & Michalewicz, (1997), the outline suggests that the evolution of computing in the field of measuring instruments and measurement starts from a conventional classical, to modern intelligent. According to Goldberg (1989) and Mitchell (1997), the desired goal improvement (evolution) computing is to make the engine computer not only as a tool to count, but also optimizing the workings of the machine that has the ability and intelligence of an expert (teaching, scoring, testing, and interpreting) humane (humane = able to learn the character of testee's ability to be able to judge correctly). Understanding adaptive in this research is the ability to customize the item difficulty with the test participants' ability test (the examinee's ability) based on the response of right and wrong answers. The ability to computer hardware programmed with fuzzy logic.

Model testing is modern with IRT differentiated by the number of parameters of test items, namely the model of the parameter (Rasch model), two parameters, and the three parameters (Hambleton, Swaminathan, & Rogers, 1991). Van der Linden & Hambleton (1997), mentions these parameters are difficult items, different power items, and guesses. IRT models for dichotomous two-parameter test items (item difficulty, item different power) is as follows (van der Linden and Hambleton (1997); Hambleton, Swaminathan, & Rogers (1991); Lord (1980)):

$$P_i(\theta) = \int_{-\infty}^{a_i(\theta - b_i)} \frac{1}{\sqrt{2\pi}} e^{-z^2/z} dz \quad \text{atau} \quad P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$$ i = test item 1, 2, 3, … , n

$P_i(\theta)$: probabilitas peserta dengan kemampuan θ untuk menjawab butir ke-i dengan benar.
θ : tingkat kemampuan peserta uji
n : banyaknya butir tes
e : nilai transcendental yang besarnya 2,718
b_i : tingkat kesulitan butir
a_i : daya beda butir

Figure 1 shows the characteristic curve and distribution capabilities of the two groups, namely the poor performance and high capability. It is seen that the group with high ability to have a greater probability of correctly answered the item correctly than the group with low ability. Thus the characteristics of the item does not depend on the group.

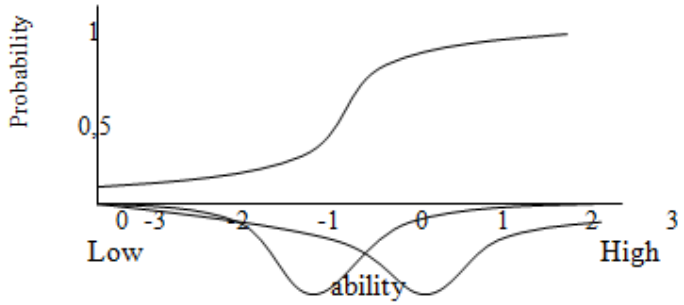

Figure 1. Item Characteristic Curves and Distribution Capabilities
(Adopted from Hambleton, Swaminathan and Rogers (1991))

Rasch model assumptions are based on the specific requirements of objectivity (Hambleton, Swaminathan, & Rogers, 1991). Ke- assumptions: (1) their unique ordering (unidimensionality) according to people's ability and item difficulty of the test. (2) the implications of the assumptions to-1, the local independence, the value of the likelihood of people answering correctly a test item does not depend on the answers to other test items. (3) are also implications of the assumptions to-1, namely equality of discrimination, interference ratio which is represented by the slope (slope) maximum grain characteristic curve assumed to be similar for all test items. (4) uni dimensionality set of test items in a test device measures only one dimension. (5) conduct random guessing (random guessing behavior). Rasch Model requires that for each test item, the value of the possible answers correct asymptotik tend to zero in line with the declining level of ability. (6) when the ability of people increases, the value of the possibility of correct answers to a test approaching one (as opposed to assuming all four).

Fuzzy logic system is one branch of science that studies on artificial intelligence (artificial intelligence / AI). The basic principles of intelligent systems (Luger, (2005); Nilsson, (1980)) is to make the computer through certain programming techniques to be able to think, take the right decisions and act in ways as humans do.

The ability of fuzzy logic in describing the student's ability to use reasoning monotonous. Through the monotonous reasoning, will obtain a degree of certainty regarding the ability of students is based on responses provided during the testing process. The certainty factor is a unity of fuzzy logic in order to describe the ability of students. If there are two fuzzy areas related by simple implications, namely:

IF x is A, THEN y is B.

The transfer function is expressed by the notation y = f ((x, A), B).

Mathematically described as follows:

$$\mu A[x] = (x-a)/(b-a) = q \qquad \text{and} \qquad \mu B[y] = 1-2[(d-y)/(d-c)]^2 = q$$

In this case the known value of x to find the value of y through the value of q as reasoning monotomous (Figure 2 shows a graph of the process).

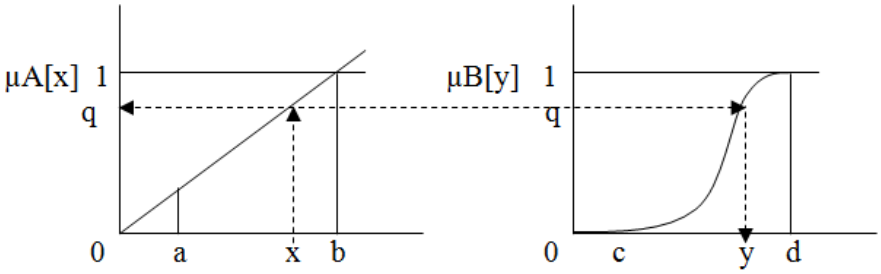

Figure 2. Reasoning Monotone

The value of the output of fuzzy inference system in the form of item difficulty is further given to test participants. Test items taken from the banks about the difficulty level according to the results of fuzzy logic inference. During the exam, the test participants' ability calculated estimated using maximum likelihood. Ability obtained in the form of a score of -3 to +3. In its application, the result of the ability to be converted into a value of 0 to 10 or 0 to 100.

An assessment of the ability of students, first performed by calculating the value of pi (θ) and qi (θ) of each item test. In this study, using two parameters, namely: the power difference (ai) and the index of difficulty (bi), thereby calculating the value of pi (θ) (chance-takers with the ability of [θ] answered grain test all i really) and qi (θ) (odds-takers with the ability [θ] responded to the test item-i wrong) is:

$$p_i(\theta) = \frac{1}{1 + e^{-Da_i(\theta - b_i)}} \qquad p_i(\theta) + q_i(\theta) = 1$$

Furthermore, with the known value of pi (θ), pi (θ) and θ can be calculated likelihood values L (U | θ) with the formula:

$$L(U \mid \theta) = \prod_{i=1}^{n} p_i^{u} q_i^{1-u}$$

where:
n: a lot of test items *u: the students' answers on the test item*

To determine the ability of test takers, calculated beforehand using Maximum Likelihood Estimation L (θ | U) with the formula:

$$L(\theta \mid U) = \frac{L(U \mid \theta)}{\sum L(U \mid \theta)}$$

Based on the results of L (θ | U) to the value of θ from -3.0 to 3.0, the estimate of the student's ability test participants is the value of θ from the L (θ | U) is the highest (maximum). So estimation capability test participants determined by the formula:

$$\text{Estimate}\,\theta = \text{Maximum}\,[L(\theta \mid U)]$$

Such estimates implies that the participants opportunities to test the ability of [θ] answer test items max L (θ | U) x 100% correct. On the other hand, means the chances of the test participants with the ability of [θ] answer test items [1 - max L (θ | U)] x 100% wrong.

III. RESEARCH METHODE

This study uses research approaches Research and Development. There are two stages in the implementation process, namely: the development phase of CBT with fuzzy logic algorithm and the implementation phase of the product. In the product development stage, the process is done is referring to the model Pressman (1997) and Rolston (1988). The second stage is to apply the product by following the steps proposed by Borg and Gall (1983).

The study was conducted in: (a) Laboratory of Computer Education Electrical Engineering FT UNY, for process development / product development CBT program. (B) SMA Negeri 6 Yogyakarta Special Region of Yogyakarta, as a place for the implementation / research products CBT program.

Subject of the study include: (a) Teachers of Mathematics, to the process of extracting information and identification requirements necessary systems in CBT program. (B) Students to test the functionality and performance of products CBT program.

Data required in this study include: (a) The qualitative data in the form of the required functions of the program, potential users of the program, and minimum standard hardware used program, and the results of internal testing of the functionality of the program. (B) The quantitative data in the form of teacher responses to product CBT program and the results of performance testing program in the process of student learning outcomes.

Data collection techniques include: (a) For the development of the product, using: (1) observation, concerning: the accuracy of instruction, verification and validation of products (alpha-beta testing). (2) The questionnaire and interview, regarding: the completeness and accuracy of the program's functionality. (3) Documentation, regarding: materials, shapes and models tests. (B) For the implementation of the product, using: (1) observation, regarding: truth, accuracy, functionality, and product capabilities in analyzing the ability of students. (2) Documentation, concerning: the data value student achievement test results.

Data analysis technique used is quantitative descriptive and evaluative. Quantitative descriptive, carried out to assess the fuzzy inference system in selecting test items corresponding to the level of ability of the test taker. Evaluative descriptive analysis technique conducted to determine the feasibility, capability and effectiveness of CBT in its function of measuring the ability of the test taker.

IV. RESULT RESEARCH

Ratings response to CBT program, conducted through questionnaires to teachers of mathematics studies. Data is collected after the CBT program teachers use for various activities in accordance with their authority.

Scores of teacher responses to the CBT program at each aspect in the range of> 3 to 4. The teacher's response to CBT programs on aspects of performance (1) Operational / use program, (2) the program display (3) the relevance of the program with the test material (4) the benefit of the program in support of teachers' duties, obtaining very good value. It shows that the response of teachers to CBT program with fuzzy logic as a whole is considered very good, or worth using.

Beta testing is used to determine the ability of CBT program work in measuring the ability of students through the test items done. Grain test question bank which is equipped with a difficulty index parameters and different power. Based on the test results can be described a number of the group's ability as follows: (1) There were 17 students expressed less with categories ability -3 to <1. (2) There are 7 students expressed both by category ability ≥1 to <2. (3) There were 14 students expressed very well by category ability ≥2 to 3.

Description accuracy CAT programs work on the selection of test items given to students illustrated by the analysis of the average correlation values and standard deviations in theory with the work of fuzzy inference CBT program. Results of correlation analysis of the level of difficulty of test items highest correlation values (r) of 0.93 and the lowest (r) of 0.09 and an average value of correlation (r) of 0.72. It shows that CBT program with fuzzy logic inference systems managed by either selecting test items that match the students' abilities. That is, that the test items were selected from a question bank between the theoretical analysis with the work of the CBT program has a high relationship (mean r = 0.72).

The results of the analysis of standard deviations (sb) is theoretically about the difficulty level of test items obtained the highest score of 0.12 and the lowest sb = sb = 0.03, and the average value sb = 0.09. Sb analysis results are fuzzy about the level of difficulty of test items obtained the highest score of 0.15 and the lowest sb = sb = 0.05, and the average value sb = 0.11. It also shows that CBT program with fuzzy logic managed by either selecting test items that match the ability of the test taker.

MLE analysis results were also obtained Likelihood function graphs concerning the description of the participants' ability tests. The ability to graph functions obtained from the

analysis of the level of difficulty grains taken from a question bank. Values illustrated by the ability of the curve with a limit of -3 <θ <+3 it shows that the ability of students in normal distribution. Based on the analysis of data as a whole, shows that the fuzzy logic inference systems managed by either selecting test items are appropriate and in accordance with the ability of the test taker. In that case, the test taker with a high capacity majority receiving test items with a high difficulty level. Participants test-skilled receiving test items with a moderate level of difficulty. Participants test with low ability majority receiving test items with a low degree of difficulty. On the other hand, a lot of grain received test takers with one another are also diverse (not the same). Participants test with high ability and low ability or receive many grains of less than capable of being. It shows that CBT program with fuzzy logic has also been successful with a nice set many items are given according to student ability (adaptive).

CBT program developed can also display packaging test items to the location of randomized answer options. This meant that the inter-takers can not cooperate with each other in answering the test. The uniqueness and at the same time more value from product CBT program is the use of artificial intelligence techniques (artificial intelligence) for the development of fuzzy logic inference systems. CBT program product is also capable of: packing the test items automatically, carry out the administration of test items, save the test results, and do the scoring.

The ability of the CAT program with the fuzzy logic algorithm as described above, has been as expected by item response theory. That each student receives the difficulty level test items that match the characteristics of the item information. On the other hand, each student will receive a number of test items that vary according to their abilities. This is in accordance with the demands of the nature of adaptability tests of CAT

V. CONCLUSIONS

The conclusions that can be drawn based on the data analysis and discussion are as follows: (1) Model CAT with fuzzy logic algorithm is able to select test items with the appropriate level of difficulty based on right and wrong answers students. In addition, it is also able to arrange a lot of test items each student according to their ability level. (2) CAT models inference system with fuzzy logic algorithm is able to set the level of difficulty of test items for each student according to their ability level. (3) The accuracy of the system inference in selecting test items that match the students' answers in response average value error / standard deviations are 0.093 and 0.112 in fuzzy theory. Based on the graphical analysis and correlation of test items between theoretical analysis with fuzzy results with the average value of r = 0.72 per student. (4) The ability of the model CAT with fuzzy logic algorithm to estimate the students' abilities based test items received is obtained: (a) 34% of students are very good capacity (b) 21% of students the ability to both, and (c) 45% of students less ability good. (5) The findings of this study, are: (a) CAT models inference system with fuzzy logic algorithm is able to conduct elections in an adaptive items. (b) System inference CAT models with fuzzy logic algorithm: Able to manage database question bank for the computerized exam. Being able to choose adaptively test items in accordance with the students' ability in computerized exam.

REFFERENCE

[1] Baeck, T., Fogel, D.B., & Michalewicz, Z. (1997). “Handbook on evolutionary computation”. New York: IOP Press.

[2] Baumgartner, T.A., & Jackson, A.S. (1995). “Measurement for evaluation in physical education and exercise science” (5th ed.). New York: WCB Brown & Benchmark Publishers.

[3] Bostock, S.J. (1997). “Designing web-based instruction for active learning”. Englewood Cliffs: Educational Technology Publications.

[4] Borg, W.R., & Gall, M.D. (1983). “Educational research: An introduction” (4th ed.). New York: Longman Inc

[5] Brennan, R.L, (2006). “Educational measurement” (4th ed.). New York: American Council on Education.

[6] Chee, T.S., & Wong, A.F.L. (2003). “Teaching and learning with technology”. Singapore: Prentice Hall.

[7] Echols, J.M., & Shadily, H. (1986). “Kamus Inggris Indonesia”. Jakarta: PT. Gramedia.

[8] Eiben, A.E., & Michalewicz, Z. (1999). “Evolutionary computation”. New York: IOP Press.

[9] Goldberg, D.E. (1989). “Genetic algorithms in search: Optimization & machine learning”. New York: Addison-Wesley.

[10] Hambleton, R.K., Swaminathan, H., & Rogers, H. (1991). “Fundamentals of item response theory”. New Delhi: Sage Pub. Inc.

[11] Hornby, AS. (1984). “Oxford advanced learner’s dictionary of current English”. London: Oxford University Press.

[12] Lin, R.L. (1989). “Educational measurement” (3rd ed.). New York: American Council on Education, Macmillan Publishing Company.

[13] Lord, F.M. (1980). “Applications of item response theory to practical testing problems”. Englewood Cliffs: Lawrence Erlbaum Associates, Publishers.

[14] Luger, G.F. (2005). “Artificial intelligence, structure and strategies for complex problem solving” (5th ed). New York: Addison Wesley.

[15] Masters, G.G., & Keeves, J.P. (1999). “Advances in measurement in educational research and assessment”. New York: Pergamon Press.

[16] Mitchell, T.M. (1997). “Machine learning”. New York: McGraw-Hill Companies, Inc.

[17] Nilsson, N.J. (1980). “Principles of artificial intelligence”. NewYork: Tioga Publishing Company.

[18] Pressman, R.S. (1997). “Software engineering: A practitioner’s approach”. New York: Mc Graw hill Book, Inc.

[19] Rolston, D.W. (1988). “Principles of artificial intelligence and expert systems development”. Singapore: McGraw Hill Book, Co.

[20] Terano, T., Asai, K., & Sugeno, M. (1992). “Fuzzy systems theory and its applications”. New York: Academic Press, Inc.

[21] Tim Prima Pena (2007). “Kamus besar bahasa Indonesia”. Jakarta: Gita Media Press.

[22] Towndrow, P.A., & Vallence, M. (2004). “Using IT in the language classroom: A guide for teachers and students in Asia” (3rd ed.). Singapore: Longman Pearson Education South Asia Pte. Ltd.

[23] van der Linden, W.J., & Hambleton, R.K. (1997). “Handbook of modern item response theory”. New York: Springer Verlag.

[24] Wang, L.X. (1994). “Adaptive fuzzy systems and control: Design and stability analysis”. Englewood Cliffs: Prentice Hall Inc.

[25] Widrow, B., & Stearns, S.D. (1985). “Adaptive signal processing”. Englewood Cliffs: Prentice Hall Inc.

[26] Yan J., Ryan M., & Power J. (1994). “Using fuzzy logic”. Englewood Cliffs: Prentice Hall Inc.

[27] Zimmermann, H.J. (1991). “Fuzzy set theory: And its applications” (2nd ed.). London: Kluwer Academic Publishing.