ENHANCING FUZZY INFERENCE SYSTEM BASED CRITERION-REFERENCED ASSESSMENT WITH AN APPLICATION

Kai Meng Tay
Electronic Engineering Department
Faculty of Engineering,
University Malaysia Sarawak
kmtay@feng.unimas.my

Chee Peng Lim
School of Electrical & Electronic Engineering
University of Science Malaysia, Malaysia
cplim@eng.usm.com

Tze Ling Jee
Electronic Engineering Department
Faculty of Engineering,
University Malaysia Sarawak
jessie_jtl@yahoo.co.uk

KEYWORDS

Fuzzy inference system, education assessment, criterion-referenced assessment, monotonicity property.

ABSTRACT

An important and difficult issue in designing a Fuzzy Inference System (FIS) is the specification of fuzzy sets, and fuzzy rules. The aim of this paper is to demonstrate how an additional qualitative information, i.e., monotonicity property, can be exploited and extended to be part of an FIS designing procedure (i.e., fuzzy sets and fuzzy rules design). In this paper, the FIS is employed as an alternative to the use of addition in aggregating the scores from test items/tasks in a Criterion-Referenced Assessment (CRA) model. In order to preserve the monotonicity property, the *sufficient conditions* of the FIS is proposed. Our proposed FIS based CRA procedure can be viewed as an enhancement for the FIS based CRA procedure, where monotonicity property is preserved. We demonstrate the applicability of the proposed approach with a case study related to a laboratory project assessment task at a university, and the results indicate the usefulness of the proposed approach in the CRA domain.

INTRODUCTION

Educational assessment is a process of forming judgment about quality and extent of students' achievement or performance, and, therefore, by inference a judgment about the learning that has taken place. Judgment usually is based on information obtained by requiring students to attempt some specified tasks, and to submit their work for an appraisal of its quality. Scoring refers to the process of representing students' achievement by numbers or symbols.

With respect to criterion-referenced assessment (CRA), ideally, students' grade should be determined by comparing their achievements with a set of clearly stated criteria for learning outcomes and standards for some particular levels of performance. The aim of CRA is to report students' achievement with reference to a set of objective reference points. It can be a simple pass-fail grading schema, or a single grade or percentage (Sadler, 2005). From the literature, the use of CRA in essay writing (While, 2002) clinical performance (Nicholson *et al*., 2009) have been reported.

Scoring usually refers to test items/tasks rather than to the overall achievement (Sadler, 2005, White, 2002, Nicholson *et al*., 2009, Joughin, 2008). To ease the assessment process, in common practice, a score is given to each item or task, with the use of rubric. Scores are then aggregated to produce a final score. Scores from different test items/tasks are usually added together and then projected (Sadler, 2005). A score can be weighted before being added to reflect the relative importance of each task (Sadler, 2005).

The use of fuzzy set related techniques in education assessment models is not new. Biswas (1995) presented a fuzzy set related method to evaluate students' answer scripts. This work was further enhanced by Chen and Lee (1999). Ma and Zhou (2000) presented a fuzzy set related method to assess student-centred learning. Saliu (2005) suggested the use of the FIS in CRA, as a Constrained Qualitative Assessment (CQA) method, with a case study.

In this paper, an FIS-based CRA model is explained. The model can be viewed as an alternative to the use of addition in aggregating the scores from all test items/tasks, and to produce a final score. The idea of replacing simple or weighted addition with a more complicated algorithm is not new (Sadler, 2005). It is pointed out that aggregation of scores can be done by some dedicated algorithm or mathematical equation. The FIS is used owing to several reasons. First, the criteria in rubric can be qualitative rather than quantitative (Sadler, 2005). As an example, a score of 4 in a rubric does not mean two times better than that of a score of 2. The FIS acts as a solution to qualitative assessment, and keeps qualitative assessment accountable. Second, the relative importance of each task can be different. The importance of each task

Proceedings 24th European Conference on Modelling and Simulation ©ECMS Andrzej Bargiela, Sayed Azam Ali David Crowley, Eugène J.H. Kerckhoffs (Editors)
ISBN: 978-0-9564944-0-5 / ISBN: 978-0-9564944-1-2 (CD)

depends on the learning outcomes. Third, an FIS can be used as an alternative approach to model or to customize the relationship between the score of each task and the aggregated score.

With respect to the FIS, it can be viewed as a method to construct a multi-input, non-linear model in an easy manner (Jang, *et al.*, 1997). In this work, our investigation focuses on the monotonicity property of an FIS-based CRA model. The importance of the monotonicity property in FIS-based CRA has been pointed out by Saliu (2005). It was suggested that the failure of an FIS-based CRA model to fulfil monotonicity property is an anomaly, and effort should be put to overcome this problem. However, there are relatively few articles addressing the problem of designing monotonic FIS (Kouikoglou and Phillis, 2009). Noted that the importance of the monotonicity property in other assessment and selection problems has been highlighted in Kouikoglou and Phillis (2009), Broekhoven and Baets (2008, 2009).

In this paper, the monotonicity property of an FIS and the *sufficient conditions* for the FIS to be of monotonicity, as pointed in Kouikoglou and Phillis (2009) as well as Tay and Lim (2008a, 2008b), are reviewed. An FIS-based CRA model is then presented. The monotonicity property for the FIS-based CRA model is defined. The *sufficient conditions* for the FIS to be of monotonicity is applied to CRA. The applicability of our proposed approach is demonstrated with a case study related to laboratory project at Universiti Malaysia Sarawak, Malaysia.

THE PROPOSED FIS-BASED CRA MODEL

Figure 1 depicts a flow chart of our proposed FIS-based CRA model. Learning usually starts with definition of learning objectives and learning outcomes. From the learning objectives and learning outcomes, test items/tasks and their assessment criteria are designed. Consider a laboratory project with three test items/tasks, i.e., *electronic circuitry design*, *electronic circuitry development*, and *presentation*. Table 1 shows the scoring rubric for *electronic circuitry design*. Each partition of the rubric can be represented by a fuzzy set. Figure 2 depicts the membership functions of *electronic circuitry design*. Each membership function is assigned a *linguistic term*. For example, a score of 3 to 5 is assigned to *Satisfactory*, which refers to *"The circuit is simple (3~4 necessary ICs). Some unnecessary components are included. Able to apply moderately the learned knowledge. Simulate only parts of circuit and briefly explain the circuit operation*". The same is applied to the rubrics of *electronic circuitry development*, and *presentation*.

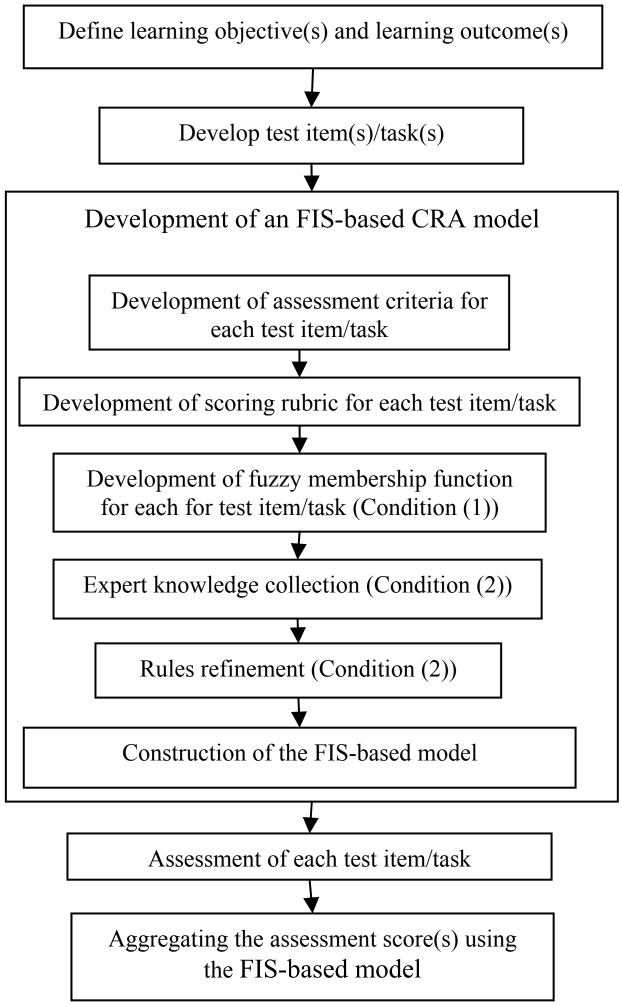

Figure 1**.** The proposed CRA procedure using the fuzzy inference system

In this paper, membership functions of *electronic circuitry design*, *electronic circuitry development*, and *presentation* are labeled as μ^{a}_{Dg}, μ^{b}_{Dv}, and $\mu^{c}_{\Pr}$, respectively. For example, for the test item/task *electronic circuitry design*, a score from 3 to 5 is represented by μ^{2}_{Dg}. The *final score* varies from 1 to 100, and is represented by seven fuzzy membership functions, i.e., "*Excellent*", "*Very good*", "*Good*", "*Fair*", "*Weak*", "*Very weak*" and "*Unsatisfactory*", respectively. The corresponding b scores are assumed to be the point where membership value of B is 1.

Table 1. Scoring Rubric of *Electronic Circuitry Design*

Score	Linguistic Terms	Criteria
10	Excellent	The circuit is complex (≥ 10 necessary ICs). Able to apply knowledge in circuit design. Able to simulate and clearly explain the operation of designed circuit.
9~8	Very good	The circuit is moderate (7~9 necessary ICs). Able to apply most of the learned knowledge. Able to simulate and clearly explain the operation of the circuit.
7~6	Good	The circuit is moderate (5~6 necessary ICs/Components). Some unnecessary components are included. Able to apply most of the learned knowledge. Able to simulate the circuit and briefly explain circuit operation.
5~3	Satisfactory	The circuit is simple (3~4 necessary ICs). Some unnecessary components are included. Able to apply moderately the learned knowledge. Simulate only parts of circuit and briefly explain the circuit operation.
2~1	Unsatisfactory	The circuit is simple (1~2 necessary ICs). Some components are not included and unnecessary components are added. Only able to apply some of the learned knowledge. Unable to simulate and explain the operation of designed circuit.

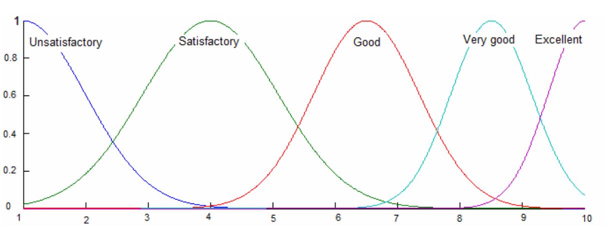

Figure 2. Membership function for one of the test item, i.e. *electronic circuitry design*

A fuzzy rule base is a collection of knowledge in the *If-Then* format from experts. It describes the relationship between *electronic circuitry design*, *electronic circuitry development*, and *presentation* and the *final score*. As an example, Figure 3 shows two rules collected from lecturers who are responsible for the assessment.

Rule 1
If *electronic circuitry design* is **Good** and *electronic circuitry development* is **Good** and *presentation* is **Unsatisfactory** then *Final Score* is **Weak**

Rule 2
If *electronic circuitry design* is **Very good** and *electronic circuitry development* is **Very good** and *presentation* is **Good** then *Final Score* is **Good**

Figure 3 An example of two fuzzy production rules

In this paper, a simplified Mamdani FIS is used to evaluate the *final score*, as shown in Equation (1), which is a zero-order Sugeno FIS model.

$$Final\ score = \frac{\sum_{a=1}^{M_{Dg}}\sum_{b=1}^{M_{Dv}}\sum_{c=1}^{M_{Pr}} \mu_{Dg}^{a}\times\mu_{Dv}^{b}\times\mu_{Pr}^{c}\times b^{a,b,c}}{\sum_{a=1}^{M_{Dg}}\sum_{b=1}^{M_{Dv}}\sum_{c=1}^{M_{Pr}} \mu_{Dg}^{a}\times\mu_{Dv}^{b}\times\mu_{Pr}^{c}} \quad (1)$$

A REVIEW ON THE *SUFFICIENT CONDTIONS*

If for all x^a and x^b such that $x^a < x^b$, then for a function f to be monotonically increasing or decreasing, the condition $f(x^a) \le f(x^b)$ or $f(x^a) \ge f(x^b)$ must be fulfilled, respectively. From the literature, there are a lot of investigations on the monotonicity property of FIS models. One attempt is to differentiate the output of an FIS with respect to its input(s). Won *et al.* (2002) derived the *sufficient conditions* for the first-order Sugeno fuzzy model with this approach. The *sufficient conditions* for a zero-order Sugeno FIS model to be monotonicity has also been reported (Kouikoglou and Phillis, 2009, Tay and Lim, 2008a, 2008b).

For an FIS to be monotone, the *sufficient conditions* state that two conditions are needed. *Condition (1)* is related to how a membership function should be tuned in order to ensure that the FIS satisfies the monotonicity property. Assume both μ^p and μ^q are differentiate-able. For $\mu^p \prec \mu^q$, Equation (2) has to be fulfilled

$$\frac{\mu^p{}'(x)}{\mu^p(x)} \le \frac{\mu^q{}'(x)}{\mu^q(x)} \quad (2)$$

Assume that the Gaussian membership function, $G(x) = e^{-[x-c]^2/2\sigma^2}$, is used in the FIS-based CRA model. The derivative of $G(x)$ is $G'(x) = -\left((x-c))/\sigma^2\right)G(x)$. Using Equation (2), the ratio of the Gaussian membership function returns a linear function, i.e., $E(x)=G'(x)/G(x)=-\left(1/\sigma^2\right)x+\left(c/\sigma^2\right)$. *Condition (2)* highlights the importance of having a monotonic rule base in the FIS model.

AN FIS-BASED CRA MODEL WITH THE *SUFFICIENT CONDITIONS*

The monotonicity property is important to the FIS-based CRA model to allow valid and meaningful comparisons among students' performance to be made. It describes the relationship between a single test item/ task with the aggregated final score. Generally, it is possible to explain the importance of the monotonicity property in CRA with the theoretical properties of a *length function* e.g. monotonicity and sub-additivity (Inder, 2005). For example, if a student obtains a higher score in *electronic circuitry design*, he/she should have a higher *final score*. For two students who are awarded the same scores in *electronic circuitry development* and *presentation*, the student with a higher score in *electronic circuitry design* should not have a lower score than that of the other.

To preserve this property, the *sufficient conditions* is applied to the FIS-based CRA model. *Condition (1)* is used to generate the membership function for *electronic circuitry design*, *electronic circuitry development*, and *presentation*, as illustrated in Figure 1. Figure 2 depicts the membership functions of *electronic circuitry design* that obey the condition. The membership functions of *electronic circuitry design* (as illustarted in Figure 2) can be projected, and $E(x)=G'(x)/G(x)$ allows the membership functions of *electronic circuitry design* to be visualized, as in Figure 4. For example, the membership function of *Excellent* is projected, and its linear line is greater than that of *Very Good* over the universe of discourse. Since $E_{Excellent}(x)>E_{VeryGood}(x)>E_{Good}(x)>E_{Satisfactory}(x)>E_{Unsatisfactory}(x)$, *Condition (1)* is fulfilled.

Condition (2) is used to check the validity of the corrected rule base. If the rule base collected does not fulfill *Condition (2)* , a feedback is sent to lecturer incharge so that rule set that fulfill *Condition (2)* is provided.

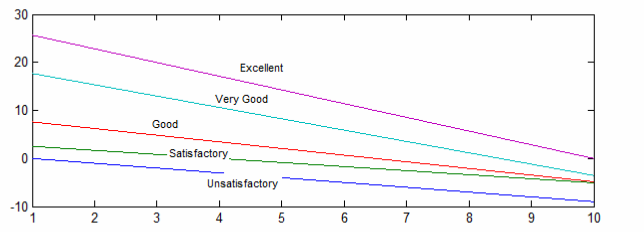

Figure 4 Visualization of the membership functions of *electronic circuit design*

A CASE STUDY

A case study was conducted to evaluate the proposed FIS-based CRA model. Asseemsent of a laboratory project by second year undergraduate students at Universiti Malaysia Sarawak was performed. The students were required to perform three test items/tasks: (1) to design a digital electronic system based on the knowledge learned at their digital system subject, as well as their creativity and technical skills; (2) to develop the system either using a *printed circuit board* or a *breadboard*; (3) to present and demonstrate their work(s).

Table 2 summarizes the assessment results with the FIS-based CRA model. Column "*No.*" shows the label of each student's project. Columns "*Dg*", "*Dv*", and "*Pr*" list the score of each test item/task, respectively. Column "*Final score*" shows the results from the FIS-based CRA model. Column "*Expert's knowledge*" shows the linguistic term associated with each project.

Figure 5 depicts one of the completed projects. The project was given a score of 7 for *electronic circuit design* because it consisted of about five components, and the student was able to explain the operations of the designed system. The student was given a score of 6 for *electronic circuit development* as the system worked well, and all electronic components were installed on the *breadboard* correctly. However, the electronic system was messy. The student was awarded a score of 7 for *project presentation*. The *final score* obtained by the student was 50.0102 (from the FIS-based CRA model).

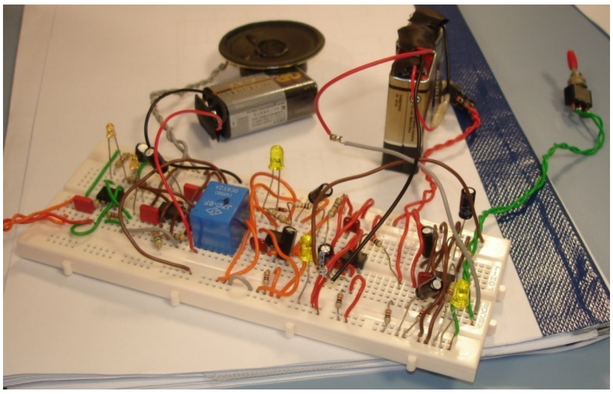

Figure 5 A digital system built by student #8

Table 2 Assessment with FIS based CRA

No.	Score of each task			Assessment with FIS based CRA	
	Dg	*Dv*	*Pr*	*Final score* (%)	Expert's knowledge
					Linguistic term
1	4	4	6	39.9691	Weak
2	5	4	6	40.2218	Weak
3	5	4	7	40.4292	Weak
4	7	4	6	40.4292	Weak
5	5	5	7	42.1384	Weak
6	6	8	5	48.8705	Fair
7	5	7	7	49.2797	Fair
8	7	6	7	50.0102	Fair
9	7	7	6	50.8660	Fair
10	8	6	6	51.1921	Fair
11	7	7	8	63.4142	Good
12	7	9	8	75.3164	Very good
13	8	8	10	78.1755	Very good
14	10	8	8	93.2697	Excellent
15	10	9	8	94.1681	Excellent

From the experiment, the FIS-based CRA model is able to produce an aggregated final score in accordance with expert's knowledge. This can be observed as the *final score* is in agreement with *Expert's knowledge*.

The importance of the monotonicity property can be explained by comparing the performance of students labeled "1" and "2", with scores of 4 4 6 and 5 4 6 ("*Dg*", "*Dv*" and "*Pr*" respectively). Both the students obtained the same score for *Dv*" and "*Pr*. However, student "1" was awarded a lower score (*Dg*=4) than that of student "2" (*Dg*=5) in the design task. The monotonicity property suggests that the final score of student "2" should not be lower than that of student "1", in order to allow a valid comparison of their performance.

From the observation, the FIS-based CRA model is able to fulfill the monotonicity property. There are no illogical predictions found in this case study. Figure 6 depicts a surface plot of the *total score* versus *electronic circuit design* and *project presentation* when *electronic circuit development=5*. An monotonic curve is obtained. In summary, as long as *Condition (1) and Condition (2)* are fulfilled, the monotonicity property can be ensured.

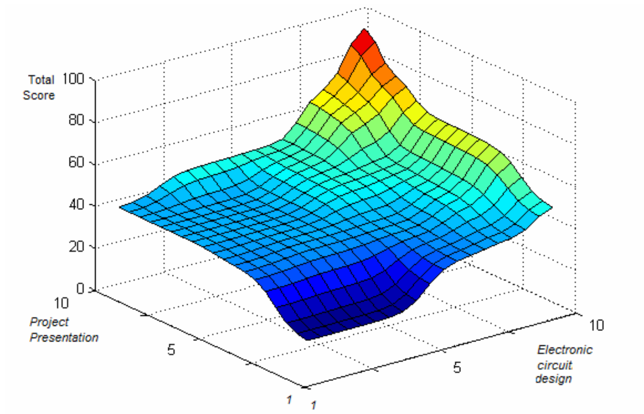

Figure 6 A surface plot of the *total score* versus *electronic circuit design* and *project presentation* when *electronic circuit development=5*

SUMMARY

In this paper, we have proposed an approach to construct an FIS-based CRA model. It is argued that the FIS-based CRA model should possess some theoretical properties of a length function. This is important to ensure the validity of the model and to allow valid and meaningful comparisons among students' performance to be made. The *sufficient conditions* is incoperated totp an FIS to ensure that the monotonicity property is fulfiled.

A case study has been conducted to evaluate the proposed approach. The experiment was conducted with data and information obtained from a laboratory project assessment problem at a university in Malaysia. The proposed FIS-based CRA model has been demonstrate to produce results that are in line with expert's knowledge.

For further work, we plan to examine how the FIS-based CRA model is able to fulfil another properties of a length function, i.e., sub-additivity. Further studies are also needed in order to vindicate the usefulness of the proposed FIS-based CRA model in the education assessment domain.

REFERENCES

Biswas. R. (1995) An application of fuzzy sets in students' evaluation. *Fuzzy Sets and Systems*, 74, 187-194.

Broekhoven, E.V. and Baets, B.D. (2008), Monotone Mamdani–Assilian models under mean of maxima defuzzification, *Fuzzy Sets and Systems*, 159, 2819 – 2844

Broekhoven, E.V. and Baets, B.D. (2009), Only Smooth Rule Bases Can Generate Monotone Mamdani–Assilian Models Under Center-of-Gravity Defuzzification, *IEEE Trans. on Fuzzy Systems*, 17, 1157-1174.

Chen, S.M. and Lee, C.H. (1999) New methods for students' evaluation using fuzzy sets. *Fuzzy Sets and Systems*, 104, 209-218.

Inder K. R. (2005), *An introduction to Measure and integration*, Alpha Science International.

Jang, J.S.R., Sun, C.T. and Mizutani, E. (1997) *Neural-Fuzzy and soft Computing*. Prentice-Hall.

Joughin, G., (2008), *Assessment, Learning and Judgement in Higher Education*. London: Springer Netherlands.

Kouikoglou, V.S. and Phillis, Y.A. (2009) On the monotonicity of hierarchical sum-product fuzzy systems, *Fuzzy sets and systems*, 160(24), 3530-3538

Ma, J. and Zhou. D. (2000) Fuzzy Set Approach to the Assessment of Student-Centered Learning. *IEEE Transactions on Education*, 43(2), 237-241.

Nicholson, P. Gillis, S., and Dunning, A., (2009) The use of scoring rubrics to determine clinical performance in the operating suite. *Nurse Education Today*, 29(1), 73-82.

Sadler, D.R (2005) Interpretations of criteria-based assessment and grading in higher education. *Assessment and Evaluation in Higher Education*, 30 (2), 175–194.

Saliu, S. (2005) Constrained Subjective Assessment of Student Learning, *Journal of Science Education and Technology,* 14(3), 271-284.

Tay, K.M., and Lim, C.P. (2008a) On the Use of Fuzzy Inference Techniques in Assessment Models: Part I: Theoretical Properties, *Fuzzy Optimization and Decision Making*, 7(3), 269-281.

Tay, K.M. and Lim, C.P. (2008b) On the Use of Fuzzy Inference Techniques in Assessment Models: Part II: Industrial Applications, *Fuzzy Optimization and Decision Making*, 7(3), 283.-302.

White, D.C. (2002) Criterion Based Assessment using the Support of a Computer**.** *Proceedings of the 35th Hawaii International Conference on System Sciences.*

Won, J.M., Park, S.Y. and Lee, J.S. (2002) Parameter conditions for monotonic Takagi-Sugeno-Kang fuzzy system, *Fuzzy Sets and Systems*, 132, 135-146.