Fuzzy Students' Knowledge Modelling System through Revised Bloom's Taxonomy

W. T. Ng, C. S. Teh

Faculty of Cognitive Sciences and Human Development, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia. csteh@unimas.my

Abstract-The conveniences of web-based educational systems have attracted a large heterogeneous group of learners with various knowledge levels, learning goals, and others learning characteristics, to study online. To enhance the effectiveness of the web-based educational system in delivery knowledge, a system should be capable to identify the learners' learning characteristics, and adapt the instructional process accordingly. Hence, this paper presented a students' knowledge modelling system that is capable of infer and updating the students' knowledge level in accordance to the cognitive processes dimension in the Revised Bloom's Taxonomy. However, the students' knowledge modeling process consists of tasks and factors that are vague and unmeasured, thus Fuzzy Logic is integrated into the students' knowledge modeling system to deal with such uncertainties. The proposed fuzzy students' knowledge modeling system uses fuzzy sets to represent students' knowledge level and other influencing factors, and uses Mamdani type inference technique to determine and update knowledge levels.

Index Terms—Cognitive Processes Dimension; Fuzzy Logic; Knowledge Modelling System; Web-Based Educational System.

I. INTRODUCTION

Concurrent with the advances of computer and web technologies, the number of learners using web-based educational systems has increased. The main attractive of web-based educational systems is that the learners can gain knowledge through electronic information and communication technologies although they are separated with the instructor in space and time [1, 2]. In fact, the learners have different knowledge levels, cognitive and metacognitive abilities, learning needs, and others learning characteristics. Therefore, it is ineffective to deliver same learning materials to all learners through same instructional conditions.

To effectively deliver knowledge to the heterogeneous group of learners, the web-based educational systems should be capable of analysis the learning characteristics of the learners and their learning outcomes, and adapt the instructional process accordingly, like the teaching process of real classrooms education [3]. A system with such capabilities of collecting, reasoning and maintaining learners' learning characteristics is known as user modelling system. For an educational system, the important learners' learning characteristic is the learners' knowledge level [4]. Hence, this paper proposed to model knowledge level in accordance to the cognitive processes dimension in the Revised Bloom's Taxonomy, which described six major categories of intellectual knowledge development.

However, the learners' knowledge level is ambiguous in

description and subject to change. Therefore, the user modelling system should be capable of dealing with such vagueness in reasoning and updating the knowledge level of the learners and the corresponding changes occurred during the learning process throughout the learners' interactions with the web-based educational system [5].

Therefore, this paper presents a system that uses Fuzzy Set Theory to model learners' knowledge level in accordance to cognitive processes dimension in the Revised Bloom's Taxonomy [6]. An overview of cognitive process dimension and user modelling system in web-based educational system is provided first. Following is the description of the proposed fuzzy students' knowledge modelling system. The implementations of the proposed system are presented and discussed in the end.

A. Cognitive Processes Dimension

Bloom's Taxonomy, named after Benjamin Bloom, was originally published in 1956 [7]. This original Taxonomy is a framework for creating and classifying learning goals and measuring learning outcomes across subject matter and grade levels. The original Taxonomy consisted of three domains - cognitive domain, affective domain and psychomotor domain. This paper focuses on the cognitive domain as it describes the intellectual knowledge development. The cognitive domain of the original Taxonomy has six major categories - *Knowledge, Comprehension, Application, Analysis, Synthesis*, and *Evaluation* [8]. Each category represents difference cognitive skills and learning goals.

Among several revisions proposed to the original Taxonomy, a revision published in 2001 [9], referred to as the revised Taxonomy, extended the original Bloom's Taxonomy to two dimensions - knowledge dimension and cognitive processes dimension. The cognitive processes dimension has six major categories like the original Bloom's Taxonomy, but changed the order of the *Synthesis* category and the *Evaluation* category, and renamed them to *Remember*, *Understand*, *Apply*, *Analyse*, *Evaluate*, and *Create* [6]. The changes are shown in Figure 1.

These six categories of the cognitive processes dimension in the revised Taxonomy are used to categorize learners' knowledge level in the proposed fuzzy students' knowledge modeling system.



Figure 1: Revised Taxonomy in comparison of the original Taxonomy

B. User Modelling Systems

User modelling system in web-based educational systems could be described as a system with capabilities of collecting and inference learners' learning characteristics, and maintaining these information in form of user models [10]. User modelling technique was originally proposed in the field of Intelligent Tutoring Systems (ITS), namely Student Modelling technique [11]. User models in ITS is known as student models, thus user modelling system is also known as Student Modelling System (SMS).

Brusilovsky & Millán [10] described three aspects that is related to user model - the nature of the user's information, the structure of the information is represented, and the way of constructing and maintaining the represented information in the user model. In term of education, the nature of the represented information can be described as the learners' learning characteristics, such as knowledge level, misconceptions and learning styles.

Brusilovskiy [12] provided three possible forms of user model in reflecting the learners' learning characteristics stereotype model, overlay model, and error model. Take example of modelling a learner's knowledge level, the stereotype model assigns the learner into certain stereotypes based on the state of the learner's knowledge. Example of such system is in [13] that modelled students into five knowledge stereotypes - novice, beginner, intermediate, advanced and expert. For the form of overlay model, the learner's knowledge level is reflected as subset of the model of expert-level knowledge of the domain, namely domain knowledge model. Ways of using overlay model to represent both domain and student knowledge are demonstrated in [3, 5]. An extension of overlay model is error model representing both errors and misconceptions performed by the user during the learning process. Such model is demonstrated in [14] to facilitate the learning of conceptual database design.

Meanwhile, the construction of user models indicated the user modelling approaches, for example, [5, 15] used Fuzzy Logic to model students' knowledge levels, [16] used Bayesian Network to classify learners' cognitive states, and [3] used Fuzzy Cognitive Map to illustrate the learners' knowledge levels and the prerequisite relationships between the domain concepts.

II. FUZZY STUDENTS' KNOWLEDGE MODELLING SYSTEM

A. Overview of Fuzzy Set Theory

Prof. Lotfi A. Zadeh [17] formalized Fuzzy Set Theory in 1965, to mathematically capture uncertainty and lexical imprecision in representing information. Contrasted to classical set theory that classifies a characteristic or an element whether belong to a class or not belong, fuzzy set theory describes the characteristic or element belong to a class in certain extent, namely degree of membership.

Let *X* be the universe of discourse with generic elements marked as x_i , and set *A* is a class in *X*. In classical set theory, set *A* is defined as crisp set using characteristic function $f_A(x_i)$ as shown in Equation (1). x_i belongs to set *A* if and only if $f_A(x_i)=1$, else $f_A(x_i)=0$.

$$f_A(x_i) = \begin{cases} 1, \ x_i \in A \\ 0, x_i \notin A \end{cases}, \ x_i \in X \\ f_A(x_i) \to \{0, 1\} \end{cases}$$
(1)

Fuzzy set theory extends the truth value of x_i belong to a set A from {0, 1} to the range of [0, 1]. The truth value, also known as the degree of membership is determined through membership function $\mu_A(x_i)$ as shown in Equation (2).

$$A \subseteq X$$

$$A = \left\{ \frac{\mu_A(x_i)}{x_i} \right\}, \ x_i \in X$$

$$\mu_A(x_i) \rightarrow [0, 1]$$
(2)

Whereby, if x_i is totally in set A, $\mu_A(x_i)=1$, or if x_i is not belong to set A, $\mu_A(x_i)=0$. Else, if x_i is partially belong to set A, $\mu_A(x_i)$ takes values in the interval [0, 1]. Note that the line of fraction in membership function symbolizes the association of membership degree $\mu_A(x_i)$ with a particular element x_i .

In term of Fuzzy Set Theory, the universe of discourse, X is a linguistic variable or fuzzy variable with linguistic values which are the elements x_i . Each linguistic value is defined as a fuzzy set in order to represent imprecise concepts [18]. For example, the proposed system has a linguistic variable which is used to represent the difficulty level of the quiz questions, is defined by taking values of '*easy'*, '*moderate'*, and '*difficult'*. These linguistic values are imprecise in nature, but defined precisely using Fuzzy Set Theory [18].

B. Fuzzy Rules and Fuzzy Inference

Fuzzy rules take fuzzy variables to connect antecedent(s) with consequent in form of IF-THEN rules, where the IF part represents the rule's antecedent, and the THEN part represents the rule's consequent [4]. The rule's antecedent defines the condition to activate the rule, while the rule's consequent assigns certain fuzzy sets from the fuzzy output variable as conclusion to the given input combination.

Based on the Fuzzy Set Theory, fuzzy inference works as reasoning mechanism which consists of a set of fuzzy rules that deals with vagueness and imprecision in information to generate decisions [4]. The proposed system used Mamdani model for fuzzy inference. Fuzzy inference process of the Mamdani model is performed in four steps - fuzzification of the input, rule evaluation, aggregation of the rule outputs, and defuzzification of the aggregation output [4].

Fuzzification is a process of fuzzifying crisp inputs into linguistic values which is related to a fuzzy variable, through the membership functions. Rule evaluation is a process of matching the fuzzy input variables to the fuzzy output variables based on the fuzzy rules, and assigning certain degree of membership to the given fuzzy output variable. For fuzzy rules that have multiple antecedents, fuzzy operators -AND operator and OR operator, are used to compile the result of the antecedent evaluation into single degree of membership to be assigned to the fuzzy output variable. Aggregation is a process of compiling one or multiple degrees of membership value assigned to all possible fuzzy output variables through rule evaluation. The last step is defuzzification, a process of transforming the results of aggregation into a crisp output value.

One of the defuzzification techniques is Centre of Gravity (CoG) method, that takes a sample of fuzzy output values, *x*, and their degrees of membership, $\mu_A(x_i)$ to the related fuzzy output variable, *A*. The obtained COG is crisp value.

$$COG = \frac{\sum_{x=a}^{b} \mu_A(x) \cdot x}{\sum_{x=a}^{b} \mu_A(x)}$$
(3)

where a and b are counters.

C. Brief Intro to the Proposed System

The proposed fuzzy students' knowledge modelling system structured the students' knowledge level in form of overlay model. The domain knowledge for a course is provided by subject-matter expert(s), and further categorized into a set of domain concepts C, and numbered based on the sub-topics of every chapter in the course. An example of domain concepts is shown in Table 1.

Table 1 Example of Domain Concepts

Chapter	Title / Domain Concept	Representation
1	Introduction to Fuzzy Logic	c ₁₀
1.1	What is Fuzzy Logic?	c ₁₁
1.2	Crisp sets and Fuzzy sets	c ₁₂
1.3	Basic terms of fuzzy sets	c ₁₃

As the student model is an overlay over the domain knowledge, the set of domain concepts is also used in modelling students' knowledge level with addition of a set of labels attached on each concept (as shown in Figure 2). For every domain concepts in the student model, it has a set of labels representing six categories of the students' knowledge level which derived from the cognitive processes dimension - *Remember*, *Understand*, *Apply*, *Analyse*, *Evaluate* and *Create*.



Figure 2: Graph of Domain Model and Students' knowledge model

The students' knowledge level is described using linguistic variable, *performance level* with seven linguistic values formed by three terms (unknown, known, learned) and three quantifiers (slightly, partially, and completely) - *completely unknown, slightly known, partially known, completely known, slightly learned, partially learned,* and *completely learned.* Each of them is associated with a fuzzy set and corresponding membership function as shown in Figure 3.

D. Determination and Updating Knowledge Level

A way to collect data about students and their knowledge level is through assessments, such as a set of quizzes provided by subject-matter expert(s) [15]. For every set of quizzes, the subject-matter expert(s) compiled certain number of domain concepts and educational objectives that are intended to evaluate through the quiz questions. The educational objectives can be represented through the six categories of cognitive processes dimension - *Remember*, *Understand*, *Apply*, *Analyse*, *Evaluate* and *Create*. The quiz question examines the students' knowledge about a domain concept on a particular category in the cognitive processes dimension.

Moreover, each quiz question has its mark allocation determined by the subject-matter expect(s). The proposed system evaluates the students' performance in accordance to each quiz question, which means the students' knowledge level is inferred in accordance to the domain concepts.

For every quiz questions, the scores obtained by the students are normalized based on the mark allocated for the particular question. Take an example of a quiz question assigned with 4 marks, if a student scored 4 marks in that question, his/ her normalized score for that question is 1 mark, or if he/ she scored 2 marks, the normalized score will be 0.5 marks.

Besides rating the scores achieved by the students in a quiz, other influencing factors including the difficulty level of a quiz question can be taken into measurement of the students' knowledge level. Similar with the setting of cognitive processes categories for each question, the difficulty levels for each question are determined by the subject-matter expert(s). Considering the vagueness and imprecision in categorizing the score level and difficulty level, and describing the knowledge level of the students, the proposed system used several fuzzy variables and rules to infer the students' knowledge level.

The proposed system takes two crisp inputs, which are the normalized scores and the difficulty levels for each question in a quiz. These two crisp inputs are fuzzified by mapping them over the membership functions of the fuzzy variables, *score level* and *difficulty level* as shown in Figure 3(A) and Figure 3(B).

The fuzzy variable of 'score level' has range from 0 ('low') to 1 ('high') with five fuzzy sets - low, low average, average, high average, and high, describing the normalized score achieved by the students (as shown in Figure 3(A)). Meanwhile, the 'difficulty level' variable shown in Figure 3(B) describes the difficulty of the quiz question through five categories - very easy, easy, moderate, difficult, and very difficult, ranged from 1 ('very easy) to 10 ('very difficult'). After fuzzifying these two crisp inputs, the proposed system matches the fuzzy input variables to the fuzzy output variable, which is the students' performance level based on the fuzzy rules. The fuzzy output variable of students' performance level describes the students' knowledge level for a particular concept through seven fuzzy sets - completely unknown, slightly known, partially known, completely known, slightly learned, partially learned, and completely learned as shown in Figure 3(C).

Next, the proposed system performs aggregation and defuzzification process, whereby the output of the system is a set of crisp values that describes the students' performance level for corresponding domain concepts in accordance to the categories in the cognitive processes dimension. Such crisp value is recorded into corresponding label, which is attached on the domain concepts in the student model.



Figure 3: Fuzzy variables - (A) Score Level, (B) Difficulty Level, (C) Performance Level

The proposed system inferred students' knowledge level through a set of twenty-five fuzzy rules. Each fuzzy rule takes fuzzy input variables - *difficulty level* and *score level* as the rule's antecedent to activate the particular rule, and assigns one of the fuzzy sets from the variable of *performance level* as the rule's conclusion to the given input combination (as shown in Figure 4). As each fuzzy rule has multiple antecedents, fuzzy operator 'AND' is applied to evaluate the conjunction of the degree of membership values obtained through the membership functions of the rule's antecedent. The result of the antecedent evaluation is applied to the membership function of the rules' conclusion.

(1) IF difficulty level is 'very easy'	AND	score level is 'low'
THEN performance level is 'col	mpletely un	known'
(2) IF difficulty level is 'easy'	AND	score level is 'low'
THEN performance level is 'sli	ightly unkno	own'

Figure 4: Example of fuzzy rule used to infer knowledge level

III. SYSTEM IMPLEMENTATION AND DISCUSSION

Cooperating with a subject-matter expert, several assessments were conducted over a real-world course with a class of 174 students for 8 weeks. This course involves an introductory module about Fuzzy Logic, whereas all students are new to the given module. The data collection is focused on the data regarding the students' knowledge level on that module.

All teaching materials were provided by the subject-matter expert. The Fuzzy Logic module consists of seven sub-topics, and each sub-topic represents a domain concept. The domain concepts are numbered based on the sub-topics, as shown in Table 1.

To collect data regarding the students' knowledge level, assessments were given in the class as paper and pencil quizzes. First quiz was conducted after three weeks of lectures to evaluate the knowledge level of the students for the first four sub-topics. After three weeks conducted the first quiz, second quiz was conducted to evaluate the students' knowledge level for another three sub-topics. On the 8th week, mid semester examination was conducted to evaluate the students the students' knowledge level for most of the sub-topics in the module.

Every set of the quizzes contains the details of the targeted sub-topic and the categories in the cognitive processes dimension to be tested, and the difficulty levels and mark allocation for each question, which are provided by the subject-matter expert. Take example of a question in Quiz 1, where the question is "Give TWO differences of the Boolean Logic and Fuzzy Logic" with total marks of 4. This question is set to test the students' knowledge level on sub-topic 1 in the cognitive processes category of "Understand", and its difficulty level is rated as 3 out of 10. The information about the questions of the quizzes is represented in form of matrixes.

Take the first quiz as example, which includes five questions that evaluate students' knowledge level for the first four sub-topics, with equivalent of 4 marks. Denoted by Q_i , where i = 1, 2, 3, 4, 5, representing the questions in the first quiz. Q_1 and Q_2 tested the students on their knowledge about sub-topic 1, c_{11} in the cognitive processes category of "*Understand*" and "*Apply*" respectively, while Q_3 is about sub-topic 2, c_{12} in category of "*Remember*", Q_4 and Q_5 is

about sub-topic 3 and 4, c_{13} and c_{14} in category of "Apply" respectively.

The subject-matter expert marked the students' answer after conducting the assessment. The scores obtained by students for each question in a quiz were recorded in table, as shown in Table 2. The obtained scores were pre-processed, including checking for missing values, normalizing the data range and transforming the data representation. This process is to ease the students' knowledge modelling process and to enhance the system performance.'

Table 2 Results of the first quiz

Students	Q_1	Q_2	Q3	Q_4	Q5
1	2	4	1	4	4
2	3	1	3	2	1
174	2	1	0	3	1

Next, the processed data is input to the proposed system which is implemented using MATLAB version 2013a. The students' knowledge model in the proposed system is initialized using default value, which is "completely unknown" for each cognitive processes category in every domain concept. Once the processed data is given, the system updated the value representing the knowledge level for a particular cognitive processes category in a domain concept. Next, the system generated tables and charts to report the performance level of the students for difference domain concepts according to the cognitive processes categories throughout the assessments.

As shown in Table 3, the system could generate a table that shows the average performance level of the students in any assessment, classified by different domain concepts in accordance to the cognitive processes categories. It helps the subject-matter expert to know the students' average performance in an assessment, while review the students' learning progress in different concepts in accordance to various cognitive processes categories. For example, in the first quiz, the average performance level indicates that the students demonstrated better performance in the category of "Understand", compared to the category of "Apply" for the sub-topic 1.

 Table 3

 Average Performance Level of the Students in the First Quiz

Concept	Cognitive Processes Category	Average Performance Level (%)
	Understand	31%
C ₁₁	Apply	28%
C ₁₂	Remember	18%
c ₁₃	Apply	35%
c ₁₄	Apply	25%

In addition, the subject-matter expert could know the distribution of the students by the results of any assessment for a sub-topic in one category of cognitive processes dimension. Figure 5 showed a histogram illustrating the students' performance level in the first quiz, in specific is the students' performance level on the sub-topic 1 in the cognitive processes category of "Understand".



Figure 5: Histogram of the students' performance level of the first quiz for sub-topic 1 in category "Understand"

Moreover, the subject-matter expert could have an overview on the learning progress of the students throughout different assessments. Take the first quiz and mid-semester examination as example (as shown in Figure 6), both includes questions that evaluate the students' knowledge level on the sub-topic 1 in the category of "Apply", and on the sub-topic 2 in the category of "Understand". The corresponding average performance level in the first quiz and mid-semester for sub-topic 1 is 28% and 63%, while for sub-topic 2 is 18% and 20%.



Figure 6: Average performance level of the students in the first quiz and mid-semester examination.

Through the illustration of tables and charts, the proposed system is capable of displaying the performance level of various students achieved in different assessments through a way that is meaningful and easy to understand.

IV. CONCLUSION

This paper presented a way of integrating Fuzzy Set Theory and Revised Bloom's Taxonomy into a student knowledge modelling system to confront the uncertainty and human subjectivity in modelling students' knowledge level. The Fuzzy Students' Knowledge Modelling System can work together with an adaptation model which uses the information stored in the student model to analyse and decide the presentation of teaching materials, leading adaptation effects into the web-based educational system to meet the needs of heterogeneous group of individuals.

Through the student model, the web-based educational system can distinguish between different learners and adapt

the instructional conditions in order to facilitate the learning and teaching process between the learners and the system.

The Fuzzy Students' Knowledge Modelling System can be used to measure the effectiveness of the teaching contents or strategies to the students. However, this work requires detail analysis on the relationships between various sub-topics or chapters in a course, whereby the knowledge acquisition of a topic may affect the knowledge acquisition of another topic. In addition, the system evaluation should be conducted using simulated students before conducting with real students. This is to ensure the validity of the evaluation method and system performance in inferring students' knowledge level as indicated in [16].

ACKNOWLEDGMENT

Special thanks to Universiti Malaysia Sarawak for the financial supports provided to this study through UNIMAS Postgraduate Scholarship.

REFERENCES

- [1] "Distance learning," Encyclopædia Britannica. 2016.
- [2] H. Henderson, *Encyclopedia of Computer Science and Technology*. New York: Infobase Publishing, 2009.
- [3] K. Chrysafiadi and M. Virvou, Advances in Personalized Web-Based Education, vol. 78, 2015.
- [4] A. Kavčič, A. Navia-Vázquez, and R. Pedraza-Jiménez, "Student modelling based on fuzzy inference mechanisms," in *Computer as a Tool. The IEEE Region 8 EUROCON 2003.*, 2003, vol. 2, pp. 379–383.
- [5] A. Kavcic, "Fuzzy User Modeling for adaptation in Educational Hypermedia," *IEEE Trans. Syst. Man Cybern. Part C (Applications Rev.*, vol. 34, no. 4, pp. 439–449, Nov. 2004.

- [6] [6] D. R. Krathwohl, "A Revision of Bloom's Taxonomy: An Overview," *Theory Pract.*, vol. 41, pp. 212–218, 2002.
- [7] B. S. Bloom, M. D. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl, Eds., Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain. New York: David McKay.
- [8] C. Munzenmaier and N. Rubin, "Bloom's Taxonomy: What's Old Is New Again," *Perspectives (Montclair).*, pp. 1–47, 2013.
- [9] L. W. Anderson, D. R. Krathwohl, P. W. Airasian, K. A. Cruikshank, R. E. Mayer, P. R. Pintrich, J. Raths, and M. C. Wittrock, Eds., A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives. New York: Longman, 2001.
- [10] P. Brusilovsky and E. Millán, "User Models for Adaptive Hypermedia and Adaptive Educational Systems," in *The Adaptive Web*, P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 3–53.
- [11] M. C. Polson and J. J. Richardson, Eds., Foundations of intelligent tutoring systems. Hillsdale: Lawrence Erlbaum Associates, 1988.
- [12] P. Brusilovsky, "The Construction and Application of Student Models in Intelligent Tutoring Systems," *J. Comput. Syst. Sci. Int.*, vol. 32, no. 1, pp. 70–89, 1994.
- [13] A. Grubišić, S. Stankov, and B. Žitko, "Stereotype Student Model for an Adaptive e-Learning System," Int. J. Comput. Electr. Autom. Control Inf. Eng., vol. 7, no. 4, pp. 440–447, 2013.
- [14] A. Weerasinghe and A. Mitrovic, "Facilitating Adaptive Tutorial Dialogues in EER-Tutor," in *Artificial Intelligence in Education*, G. Biswas, S. Bull, J. Kay, and A. Mitrovic, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 630–631.
- [15] Z. Jeremić, J. Jovanović, and D. Gašević, "Student modeling and assessment in intelligent tutoring of software patterns," *Expert Syst. Appl.*, vol. 39, no. 1, pp. 210–222, 2012.
 [16] E. Millán and J. L. Pérez-de-la-Cruz, "A Bayesian Diagnostic
- [16] E. Millán and J. L. Pérez-de-la-Cruz, "A Bayesian Diagnostic Algorithm for student modeling and its evaluation," *User Model. Useradapt. Interact.*, vol. 12, no. 2/3, pp. 281–330, 2002.
- [17] L. A. Zadeh, "Fuzzy sets," Inf. Control, vol. 8, pp. 338-353, 1965.
- [18] C. Sammut and G. I. Webb, Eds., *Encyclopedia of Machine Learning*. Boston, MA: Springer US, 2010.