International Journal of Engineering & Technology, 7 (4.31) (2018) 87-94

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Design and Development of Topic-based Students' Knowledge Modelling System using Fuzzy Set Theory and Visual Analytics

Wen Thing Ng[1], Chee Siong Teh[2]*, Chwen Jen Chen[3]

[1]*Faculty of Cognitive Sciences and Human Development, Universiti Malaysia Sarawak, Sarawak, Malaysia*
**Corresponding author E-mail: csteh@unimas.my*

Abstract

To the field of education, advances in technologies raise opportunities and challenges in delivering personalised learning to heterogeneous group of learners while monitoring and analysing their learning progress and learning outcomes. This paper addresses the challenges of gaining a big picture about learners' knowledge that is essential for providing personalised learning. Next, a system that models students' knowledge is proposed in which topic-based approach, fuzzy set theory, and visual analytics were applied to design and develop the proposed system. The proposed system was designed with the capabilities to model, analyse and report individual learner's knowledge, whereas its applicability was presented and discussed based on a real-world case study.

Keywords: *Fuzzy set theory; Knowledge modelling system; Topic-based student modelling system; Visual analytics.*

1. Introduction

Throughout the past decades, the technologies of supporting learning and instructions are changing and evolving. Among the emerging technologies, recent research shows that personalised learning has great potential in transforming learning and instruction [1]. In this paper, personalised learning is referred to any approach that can be used to guide individual learners to the most appropriate educational contents in accounts of any relevant learning characteristics of the learners [2]. The learning characteristics of the learners may vary in terms of their knowledge level, abilities and disabilities, learning needs, learning goals, preferences, and so forth [3], [4].

Compared to the traditional one-method-fits-all approach, personalised learning takes the diverse needs of learners into consideration in order to guide the learners to the most appropriate educational contents [5], resulting in better learning performance for group of heterogeneous learners [6], [7]. In other words, knowing whom to adapt [8] and which decisions have to be made to guide the learners to the most appropriate educational contents [2] are the important aspects in providing personalised learning.

This study employs the student modelling techniques to gain the big picture of learners' knowledge. Such techniques were introduced in the field of Intelligent Tutoring Systems (ITS) for collecting, analysing and maintaining the learning characteristics of learners to form student models [9]. Although student modelling techniques were proposed in the field of ITS, such techniques were applied in different educational applications that adopt personalised learning [10]–[12]. Hence, in this paper, a topic-based, structural student modelling system that is capable of collecting, reasoning, and keeping track of learners' knowledge is proposed.

The fuzzy sets theory [13] was also utilised for reasoning the knowledge of the learners. Often, to assess the knowledge of the learners, a set of assessments is generated by instructors and to be solved by the learners [11]. However, the learners' knowledge may not be reflected accurately based on the assessments results as the reasoning of the knowledge can be influenced by other factors, such as the difficulty level of an assessment, time spent to solve the assessment and so forth. [11]. Indeed, these factors as well as learners' knowledge are ambiguous and vague in nature. The process of determining the knowledge level tends to be uncertain [11]. Therefore, this study employed the fuzzy set theory, which is capable of handling such vagueness and uncertainties mathematically, for reasoning learners' knowledge by taking into consideration the various imprecise factors.

A visual analytics component was also incorporated into the proposed system to provide detailed insight into learners' knowledge. With the aid of this visual analytics component, the burden of analysing whom to adapt and what learners have learned can also be reduced. As the result, the feasibility of adopting personalised learning among the heterogeneous group of learners is enhanced [14].

The subsequent sections of this paper are organized as follows: Section 2 presents the background and relevant works; Section 3 and 4 elaborate on the design and development of the students' knowledge modelling system, which employed the topic-based approach, fuzzy set theory and visual analytics; Section 5 discusses the implementation of the proposed system; and Section 6 provides the conclusion and recommendations of future work.

2. Background and Related Works

This section focuses on the background and related past studies on student modelling system, fuzzy set theory and visual analytics.

2.1. Student Modelling System

Student modelling system was originally proposed in the field of Intelligent Tutoring Systems (ITS) [9]. It is referred as a system that is capable to collect and infer the particular learning characteristics of a learner and maintain these characteristics in the form of student models [9], [15]. A student model can represent the knowledge level in various forms with the overlay student model as the most preferred form [16]. The overlay student model [17] structures the knowledge level of an individual learner in relation to a domain model, which consists of the expert-level knowledge for a particular domain [15], [16]. Using the overlay student model to represent the knowledge level of a learner offers great flexibility in modelling the knowledge level and in representing various domains [18].

The complexity of structuring knowledge level using the overlay student model depends on the granularity of the domain model, and the estimation of the knowledge level [18], [19]. Past studies on student modelling mostly focused on constructing the domain model by fine-graining the domain knowledge into dozens to hundreds of concepts in relation with the educational content element [10], [19]. The concept-based approach requires the instructors to decompose the domain knowledge into fine-grained concept level, link the concepts to each other, and index them with educational contents [19]. Although fine-grained domain models provide great precision in delivering personalised learning, the granularity of the domain model increases the complexities in authoring the domain model, and increases the difficulty for the instructors to comprehend and master the way of authoring domain model [19], [20].

In concern of such issue of complexity, this study applied the topic based approach for constructing the domain model in the proposed student modelling system. Topic-based approach was introduced by Sosnovsky and Brusilovsky [21], namely *topic-based knowledge modelling*. The topic-based approach structures the domain model based on the way an instructor structures the body of the domain knowledge into course curricula [19]. In Sosnovsky and Brusilovsky [21], the topic-based domain model of the system only consisted of 22 topics for C programming languages, while the concept-level domain model had several hundred concepts. Such approach is much simpler to be understood and applied by instructors. The significant difference between the topic-based approach and the concept-based approach is that each educational activity contributes to only one topic in topic-based approach, while it is linked to multiple concepts in the concept approach [21]. Inspired by the works of QuizGuide, the following section presents how this study incorporated the topic-based approach into the proposed students' knowledge modelling system for modelling the knowledge level of an individual learner.

2.2. Fuzzy Set Theory

Fuzzy set theory is formalized by Prof. Lotfi A. Zadeh in 1965 [13], to mathematically represent uncertainty and lexical imprecision of the linguistic variables. Let X be a linguistic variable representing the universe of discourse, with elements namely linguistics values, x_i. Each element of x_i represents particular imprecise concepts as fuzzy sets. Take example of a linguistic variable namely difficulty level, it could be defined with the linguistic values of 'easy', 'moderate', and 'difficult'. These values are vague in nature, but could be defined precisely using fuzzy set theory [22]. To define whether the particular elements of X belong to a set A, degree of membership with range of [0, 1] was formulated using membership function $\mu_A(x_i)$ as shown in Equation (1).

$$\mu_A(x_i) = \begin{cases} 1,\ x_i \text{ is totally belong to set A} \\ 0 < \mu_A(x_i) < 1,\ x_i \text{ is partly in set A} \\ 0,\ x_i \text{ is totally not belong to set A} \end{cases} \quad (1)$$

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

Various past studies reported the use of the fuzzy set theory to model the knowledge level of an individual learner. For example, FuzKSD [23] used fuzzy cognitive map to represent the knowledge of the learners for various concepts in a particular domain. Meanwhile, DEPTH [11] and TRAM [24] used a fuzzy inference system to inference knowledge of the learners.

2.3. Visual Analytics

According to the definition introduced in the call for papers of the 1st International Conference on Learning Analytics and Knowledge 2011 (https://tekri.athabascau.ca/ analytics/), learning analytics involve the measurement, collection, analysis and reporting of data about learners and the corresponding contexts, in order to understand and improve learning and the corresponding learning environments. According to the study by Ruipérez-Valiente et al. [14], there are two approaches of applying learning analytics into educational systems. One of the approaches is through visual analytics that presents the data about learners and relevant learning contexts using various types of visualizations and reports, such as reports for students, activities and performance in a given period and so forth. [25].

ALAS-KA [14] is one of the various past studies that followed the visual analytics paradigm. It is a visual analytics module that extends the supports of learning analytics in the Khan Academy platform, which reports visualization and recommendations for the entire class as well as individual student. These include the visualization of the affective states for the class group, total time distribution for each user in different activities and so forth. Other works include GLASS [26] that presents the most active events and learners in the system and CAMera [27], a tool for monitoring and reporting learners' learning behaviour to aid self-regulated learning.

3. Topic-Based Student Modelling System

This section presents the works of designing and developing a students' knowledge modelling system based on the instructor-oriented topic-based approach and with visual analytics applied for modelling and reporting the knowledge of an individual learner. The proposed students' knowledge modelling system was designed in structural design of a topic-based system as illustrated in Fig. 1. There are five important components, which are the domain model, goal model, content model, overlay student model, and visualization analytics model.

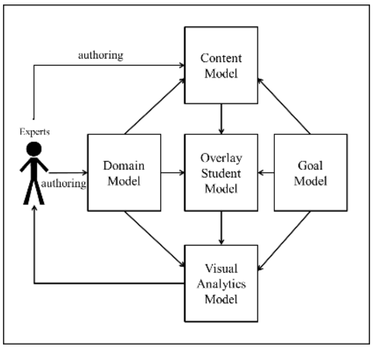


Fig. 1: Structural design of topic-based students' knowledge modelling system.

3.1. Topic-Based Domain Model

This study designed the domain model of the proposed system using the topic-based approach, namely topic-based domain model. As illustrated in Fig. 2, the domain model represented the knowledge of a particular domain by structuring the elementary elements, namely *Topics* in the domain knowledge, in the form of nodes. The nodes were organized based on the curricula of a particular course domain, which was provided by the instructor. For each module structured in a course domain, it is denoted as m, while each of the topics structured under the module is denoted as n.

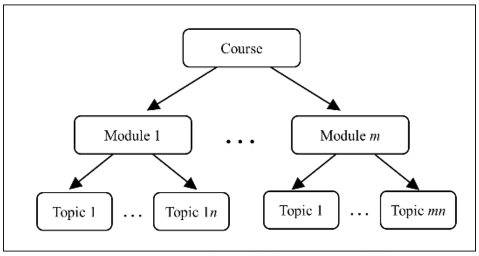


Fig. 2: Topic-based domain model

3.2. Topic-Based Goal Model

The purpose of having a topic-based goal model in the proposed system is to represent the instructional objectives dedicated by the instructors. The ultimate goal of instructional objective is to eliminate the gap between the knowledge level of the expert of a domain and the individual learner [16]. Consequently, to gain better insight of the knowledge level of the individual learners, we utilized the Revised Bloom's Taxonomy as metrics of instructional objectives to indicate the dedicated learning progress, learning outcomes, and learning objectives. There are six categories of cognitive processes dimension in the revised taxonomy, which are the categories of *Remember*, *Understand*, *Apply*, *Analyse*, *Evaluate* and *Create*. Each category represents different cognitive skills that require different degree of cognitive processing in learning.

The topic-based overlay student model of the proposed system associated the domain model with the goal model. In accordance to every *topic* in the domain model, the student model further structured the knowledge level of individual learner based on the six categories of cognitive processes dimension. Considering the potential mass scale of a domain model, for example, dozens of topics were listed in particular curricula of a course domain, the representation of the cognitive processes dimension was simplified to only five levels, namely *Bloom Levels* in order to reduce the storage consumption of the goal model (as shown in **Fig. 3**). These five levels were named as *Level 0*, *Level 1*, *Level 2*, *Level 3*, and *Level 4*, which corresponds to the fuzzy set of '*Bloom Level*'. As shown in Table 1, these levels represent different extents of the cognitive processes dimension in term of the six categories in the revised Taxonomy. By default, the goal model is initialized by assigning default value of zero to each Bloom Level.

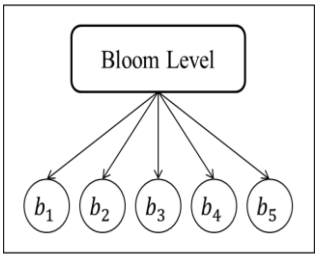


Fig. 3: Topic-based goal model

Table 1: Different extents of *bloom levels* in topic-based goal model

Bloom levels	Extensions of *bloom levels*
Level 0	This level represents the cognitive processes dimension in which is likely in the category of 'Remember'
Level 1	This level represents the cognitive processes dimension in which is mostly in the category of 'Remember' and likely in the category of ''Understand'
Level 2	This level represents the cognitive processes dimension in which is mostly in the category of 'Understand' and likely in the category of 'Apply'
Level 3	This level represents the cognitive processes dimension in which is mostly in the category of 'Apply' and likely in the categories of 'Analyse, Evaluate, and Create'
Level 4	This level represents the cognitive processes dimension in which mostly in the categories of 'Analyse, Evaluate, and Create'

3.3. Topic-Based Content Model

A set of assessments was used for assessing the data about the knowledge level of an individual learner, and thus the topic-based content model served as the connection between the configurations of assessments, the *topics* enlisted in the domain model, and the dedicated instructional objectives, *Bloom Level*. Take example of a sequence of *Topics* that is structured based on particular curriculum of a course. For assessing the knowledge level of individual learners, several assessments were conducted throughout the course, whereby each assessment was denoted as *Quiz*, and each *Quiz* consisted of a sequence of individual *Questions*. For every *Questions* in each *Quiz*, they are varied in terms of difficulty level in answering them, and the maximum marks that the individual students could score. In addition, following the topic-based content organisation, each *Question* was associated with one of the course *Topics*, and with one of the *Bloom Levels*. The associations with both Topics and Bloom Level are indicated using dash line in Fig. 4.

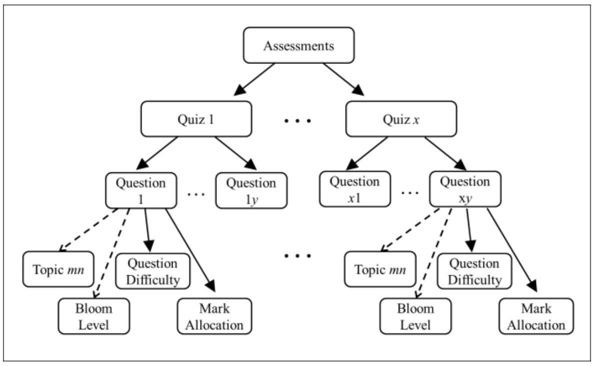


Fig. 4: Topic-based content model

3.4. Topic-Based Overlay Student Model

The topic-based overlay student model of the proposed system represented the knowledge level of individual learners in a particular domain. The student model was designed to be associated with the domain model and goal model (as shown in **Fig. 5**). In other words, the knowledge levels of individual learners were structured based on the set of *topics* in the domain model, and further classified into corresponding *bloom level*.

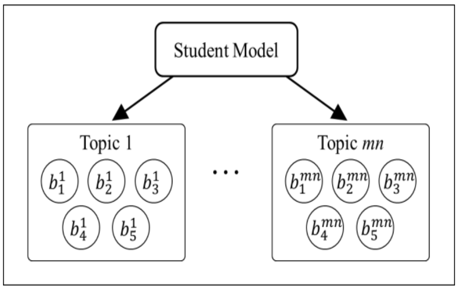


Fig. 5: Topic-based overlay student model

Qualitative measures were applied to indicate the degree of which the individual learners know about the particular *topics* in particular *bloom levels*. The values of the qualitative measures were corresponded to the fuzzy set of '*performance level*' that represented the output of the fuzzy inference system used to estimate the knowledge level of individual learners. The values were formed by three terms (unknown, known, learned) and three quantifiers (slightly, partially, and completely). Together there are seven terms to describe the knowledge level that individual learners acquired - *completely unknown*, *slightly known*, *partially known*, *completely known*, *slightly learned*, *partially learned*, and *completely learned*.

3.5. Visual Analytics Model

To ease the interpretations and comprehensibility of the student model generated by the proposed system, the visual analytics component was incorporated into the proposed system. The visual analytics generated relevant visualizations and reports based on the data modelled and analysed by the proposed system The visualizations and reports make the reading of student model be more meaningful to the instructors.

This study applied several design rules into the visual analytics component, which was in reference to the work by Ruipérez-Valiente et al. [14]. The rules are listed as follows:

- Applying colours meaningfully. Different tones of black colour were applied to illustrate different aspects of the visualization contexts. For example, from performance level of Completely Unknown to Completely Learned, we applied the colour of light grey to black to indicate them.
- Divide the visualizations into comprehensive modules according to their semantics, whereby we divided the visualization analytics into three parts representing different contexts of the presented visual information.
- Use the same standard for all visualizations.

There are two parts in the visual analytics model of the proposed system. The first component, namely *Content Visualization,* presented visual information about the topic-based domain model and content model, such as table list of the *topics* in the course, histogram of the *Topics* that is related with the *Questions*, etc. These visualizations and reports served as a tool for instructors to review the curricula focus of the course and evaluate the needs of shifting focus in accordance to the performance of the students in the course

The second component presented the learning outcomes and learning progress according to the Revised Bloom's Taxonomy in term of a group of individual learners, namely *Class Visualization*, and in term of an individual learner, namely *Individual Visualization*. The visualizations and reports in *Class Visualization* presented the knowledge level of a group of individual learners in average and served as tool for the instructors to review the instructional strategies of the course from a wider view of perspective. In addition, the average performance presented in *Class Visualization* could be used as a mean to compare the performance of an individual learner, which was presented in *Individual Visualization*. Meanwhile, *Individual Visualization* displayed the strength and weakness of an individual learner in learning particular *topics* of the course, and mastering certain cognitive skills described through *bloom levels*.

4. Fuzzy Knowledge Inference System

In this study, the fuzzy set theory was utilized to design and develop a Mamdani fuzzy inference system for estimating the knowledge level of an individual learner. The proposed fuzzy inference system follows Mamdani type inference processes including fuzzifying crisp inputs into fuzzy input variables, matching the fuzzy input variables to the fuzzy output variables based on the fuzzy rules, assigning degree of membership to the given fuzzy output variable, and transforming the fuzzy output variable into crisp output using defuzzification technique. The defuzzification technique used in the proposed fuzzy inference system is Centre of Gravity (CoG) method (as shown in Equation (2)), whereby x as the values of fuzzy output variables, $\mu_A(x)$ as their degrees of membership in relation to the fuzzy output variable, A, and a and b are counters.

$$COG = \frac{\sum_{x=a}^{b} \mu_A(x) \cdot x}{\sum_{x=a}^{b} \mu_A(x)} \quad (2)$$

The proposed fuzzy inference system consists of two rule bases (as illustrated in Fig. 6). The top rule base is named as Revised Bloom's Taxonomy Evaluation, which is to inference the bloom level that a Question is related to. The bottom rule base, namely Students' Performance Level Evaluation, is for inferring the knowledge level of individual learners in accordance to the Bloom Level, the difficulty level of answering, and the normalized scores that a student obtained for that particular Question. The inputs required to activate the proposed fuzzy inference system are recorded in the topic-based content model.

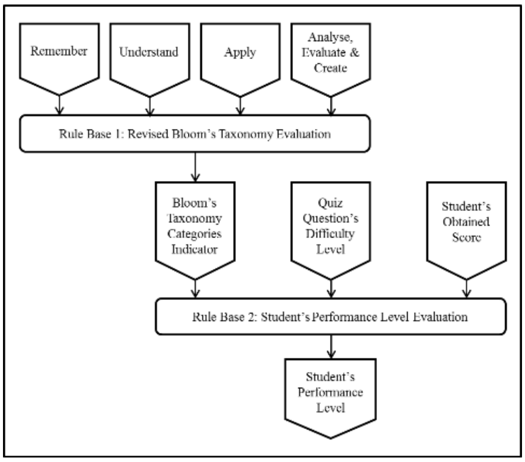


Fig. 6: Fuzzy inference system to inference knowledge level

The top rule base is designed based on past research works that used the original Taxonomy to design a stereotype student model [12], whereby we modified it to be using Revised Bloom's Taxonomy. It receives four crisp inputs that describe the degree to which different categories of cognitive processes dimension, 'Remember', 'Understand', 'Apply', and 'Analyse, Evaluates, and Create' were involved in a particular Question. Each input is ranged with values between 0 and 1, whereby 0 is referred as 'none' and 1 is referred as 'mostly' (as shown in Fig. 7(A)).

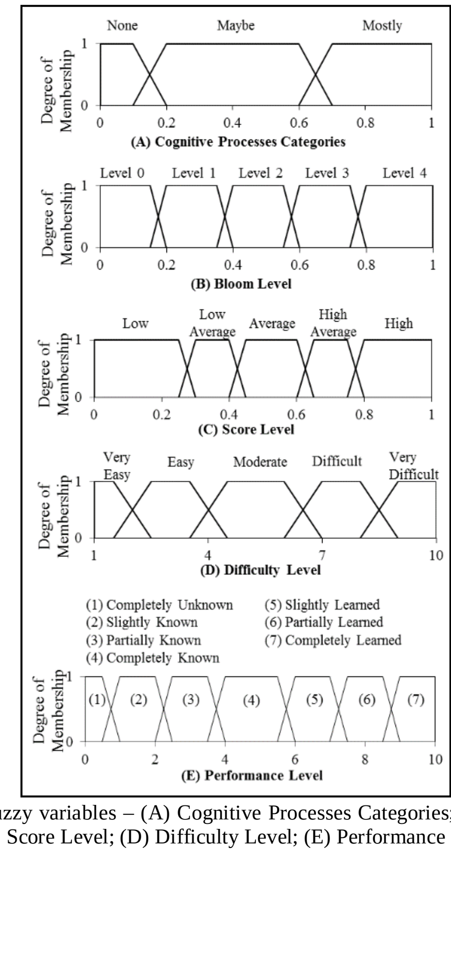


Fig. 7: Fuzzy variables – (A) Cognitive Processes Categories; (B) Bloom Level; (C) Score Level; (D) Difficulty Level; (E) Performance Level.

The fuzzy output variable of the top rule base is named as Bloom Level with fuzzy five sets – Level 0, Level 1, Level 2, Level 3, and Level 4 (as shown in Fig. 7(B)). The variable of Bloom Level is ranged with values between 0 and 1 to indicate different extents to which the cognitive processes dimension that a particular Question was set to test the knowledge level of individual learners (as described in Table 1). To inference the bloom levels, the top rule base is based on nine fuzzy rules (as shown in Fig. 8), which we deducted in reference of the past research work [12].

(1) IF ***Remember*** is '*none*' AND ***Understand*** is '*none*'
AND ***Apply*** is '*none*' AND ***Analyse, Evaluate and Create*** is '*none*'
THEN *Bloom level* is '*level 0*'
(2) IF ***Remember*** is '*maybe*' AND ***Understand*** is '*none*'
AND ***Apply*** is '*none*' AND ***Analyse, Evaluate and Create*** is '*none*'
THEN *Bloom level* is '*level 0*'
.
.
.
(9) IF ***Analyse, Evaluate and Create*** is '*Mostly*'
THEN *Bloom level* is '*level 4*'

Fig. 8: Example of fuzzy rules used to inference *bloom levels*

Taking the fuzzy output variable of Bloom Level as one of the inputs to the bottom rule base, the bottom rule base inference the knowledge level of individual. Other inputs received are the normalized scores that individual learners obtained for particular questions (as shown in Fig. 7(C)) and the difficulty of answering the question (as shown in Fig. 7(D)). The output of the bottom rule base is the knowledge level of individual learners, namely performance level with fuzzy sets - completely unknown, slightly known, partially known, completely known, slightly learned, partially learned, and completely learned (as shown in Fig. 7(E)). Details about these input variables of score level and difficulty level, and the output variable of performance level could be seen in previous work [3].
For the bottom rule base, as we extended our previous work with addition fuzzy input variable of bloom level, the bottom rule base inferred students' knowledge level through a set of 125 fuzzy rules (as shown in Fig. 9). The output of the bottom rule base, which represented the knowledge level of individual learner, is recorded into the topic-based overlay student model.

(1) IF ***Bloom level*** is '*level 0*' AND ***difficulty level*** is '*very easy*'
AND ***score level*** is '*low*'
THEN *performance level* is '*completely unknown*'
(2) IF ***Bloom level*** is '*level 0*' AND ***difficulty level*** is '*very easy*'
AND ***score level*** is '*low average*'
THEN *performance level* is '*slightly unknown*'
.
.
.
(125) IF ***Bloom level*** is '*level 4*' AND ***difficulty level*** is '*very difficult*'
AND ***score level*** is '*high*'
THEN *performance level* is '*completely learned*'

Fig. 9: Example of fuzzy rules used to inference knowledge level

5. System Implementation and Discussion

To demonstrate the implementation of the proposed topic-based students' knowledge modelling system, a case study was conducted with the cooperation of an instructor for an undergraduate course. The undergraduate course had 174 enrolled students and all were new to the course. For this case study, only one of the modules in the course with seven topics in total was taken into consideration. Throughout the lectures of the module, the instructor conducted three assessments with to evaluate students' learning progress and learning outcomes of the class in accordance to the dedicated instructional objectives. All assessments were given in the form of paper and pencil. The total duration of the lectures of the module was 8 weeks. First assessment was conducted after three weeks of lectures to evaluate the knowledge level of the class for the first four topics in the module, while the second assessment was conducted after three weeks of conducting the first assessment to evaluate the knowledge level of the class on another three topics. On the 8th week, a third assessment was conducted as mid semester examination to evaluate the knowledge level of the class for most of the topics in the module.
The purpose of conducting assessments was to evaluate the knowledge level of various topics in a module. In other words, each question in the assessments was set to test the knowledge level of the class on one topic of the module. Moreover, each question in an assessment was configured with different maximum marks that the students could score. The maximum marks were used to normalize the marks the students obtained for each question, in order to ensure the consistency of the input of the scores achieved by the students. The instructor could also indicate the difficulty level in answering each question using a scale with range of 1, which meant "very easy", to 10, which meant "very difficult". In addition, the instructor could rate the degree to which that particular question was set to test the knowledge level on particular category of cognitive processes dimension, with values ranged from 0, which meant "not related", to 1, which meant "mostly related".
Take example of a question in the first assessment that has total marks of 4 marks, the instructor justified that question was to test students' knowledge level for the Topic 1 of the module in the cognitive processes category of 'Understand'. For this question, the instructor selected value of 1 for the category of 'Understand' to indicate this question was mostly involved the category of 'Understand', while other categories were rated as 0, as other categories were not related. Meanwhile, the instructor also rated the difficulty level of the question was 3 out of 10. These configuration regarding the assessments and assessment questions were collected and recorded in the topic-based content model.
After the instructors marked the assessment answers of the class for each assessment, the marks were also collected and recorded in tabular form. The table of marks was pre-processed, including checking for missing values, normalizing the data range and transforming the data representation. The purpose of pre-processing was to ease the process of knowledge inference and to enhance the system performance.
Meanwhile, no intervention on teaching was conducted during the lectures based the results of assessments. Subsequently, the instructor intended to visualize the learning progress and learning outcome of all students throughout all assessments.
For initialization, the proposed system assigned default value, which is *completely unknown,* for each *bloom level* in every *topic* structured in the topic-based overlay model. Subsequently, the system initialized the topic-based goal model by default. Next, we inputted the relevant data into the proposed system for constructing the topic-based domain model and topic-based content model. With all relevant data entered, the proposed system activated the fuzzy inference system to infer the knowledge level of each *Student A a*nd updated the record in the topic-based overlay student model. At the end of the system execution, the proposed system generated the required visualizations and reports using the component of visual analytics, to ease the interpretations and comprehensibility of the student model outputted by the proposed system. The following sub-sections presented several examples of visualizations and reports generated by the proposed system..

5.1. Content Visualisation

The visualizations and reports generated by the *Content Visualization* presented the information regarding the curriculum of the course, and the configuration of the assessments and assessment questions. Take example of the configuration of the assessments in the case study, Fig. 10 showed the histogram of the topics that are related to the assessment questions. The histogram shows that the instructor evenly evaluated the knowledge level of class on every topic in the module, except for the Topic 5 and Topic 7. Based on this diagram, the instructor could reflect on the tendency of testing the knowledge level of the class on topics of the module, and thus gaining more insight about configuring assessments.

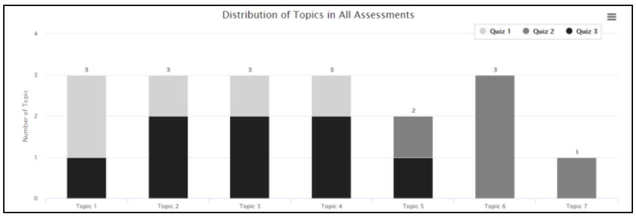


Fig. 10: Histogram of topics that are related to the assessment questions

Another similar diagram is showed in Fig. 11, but the focus of this histogram was on the Bloom Levels that the instructors set on the assessment questions for evaluating the knowledge level of the class. The histogram shows that most of the questions are related to Bloom Level 3, which was tend to test the students' knowledge level in the category of "Apply" and likely in the categories of "Analyse", "Evaluate", and "Create. On the other hand, Bloom Level 0 and Bloom Level 1 are not tested throughout all assessments conducted during the lectures of module. Although the histogram shows imbalance in setting up the questions for testing the knowledge level in particular cognitive process categories, however, if the instructional objectives set by the instructor were mainly focused on evaluating the category of 'Apply' or even higher categories, such tendencies matched with the dedicated instructional objectives. Hence, through this histogram, the instructor could review whether the configuration of assessments matched with the dedicated instructional objectives.

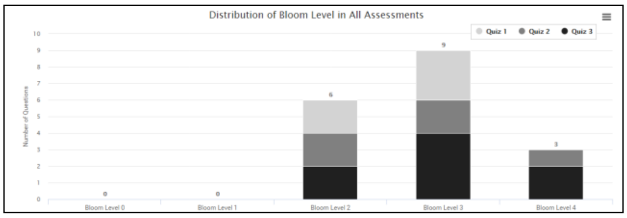


Fig. 11: Histogram of Bloom Levels that are related to the assessment questions.

The example diagrams of *Content Visualization* showed that it served as a tool for instructors to review the focus of the course, the configurations of assessments, and the instructional strategies of the course. Furthermore, the instructor could evaluate and analyse the needs of shifting the focuses or the instructional strategies.

5.2. Class Visualisation

Through *Class Visualization*, various visualization and reports were generated to present the knowledge level of the class in average and the average performance level is displayed in percentage value. Fig. 12 illustrates the overall performance level of all students for all topics and all assessments conducted in the case study. The chart shows that 44% of the students performed in the category of 'Completely Unknown', while only 23% of the students performed in the category of 'Completely Learned'. The instructor could gain a wider view of the knowledge level of the class through this chart.

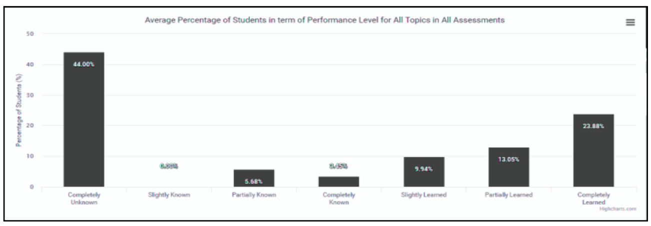


Fig. 12: Average performance level of the class in percentage for all topics throughout all assessments.

Furthermore, the *Class Visualization* could present visual information about the knowledge level of the class in accordance to the *Bloom Levels* through pie charts. Take example of the performance of the class in both Bloom Level 2 and Bloom Level 3 (as illustrated in Fig. 13), the percentage of average performance level of 'Completely Unknown' in Bloom Level 3 (46.16%) is slightly lower than the Bloom Level 2 (49.22%). Meanwhile, the percentage of 'Completely Learned' and 'Partially Learned' in Bloom Level 3 (27%, 17.70%) is higher than the Bloom Level 2 (16.08%, 7.47%). The pie charts aided the instructors in comparing the knowledge level of the class in between of various cognitive process categories, and further evaluate the effectiveness of the instructional strategies.

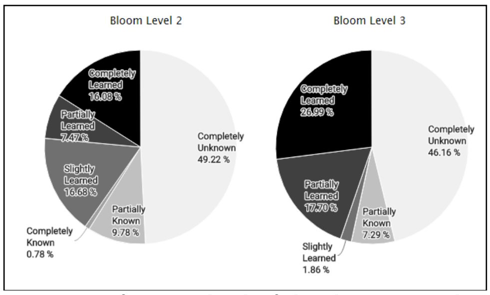
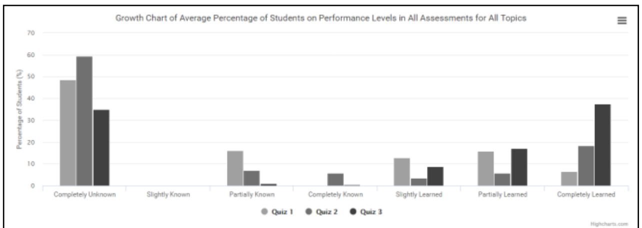


Fig. 13: Average performance level of the class in accordance to the Bloom Levels.

In addition, the instructor could review the growth in knowledge level of the class throughout all assessments. As shown in Fig. 14, the column bar presented the growth in average performance level of the class for all topics throughout all assessments. In this case study, first assessment covered first four topics of the module, the second assessment covered another three topics, and the third assessment covered most of the topics. This chart illustrated that the knowledge level of the class in terms of Completely Unknown is 48.51% and 59.54% in first two assessments, and at the third assessment, the percentage dropped to 35.06%. Meanwhile, in terms of Completely Learn, a growing trend of the knowledge level is shown throughout three assessments, whereby from 6.55% to 37.64%. This chart illustrated the gaining of knowledge level of the class for all topics in the module. For detailed insight of the learning of each topic, the Class Visualization could present similar growth chart in terms of single topic. With the visualization of the growth in the knowledge level of the class, the instructor could review the effectiveness of particular instructional strategies and the learning materials provided to students, and lead to improving the quality of the course, and the gaining of knowledge among the students.



	Quiz 1 (%)	Quiz 2 (%)	Quiz 3 (%)
Completely Unknown	48.51	59.54	35.06
Slightly Known	0.00	0.00	0.00
Partially Known	16.21	7.13	0.93
Completely Known	0.00	5.75	0.57
Slightly Learned	12.76	3.45	8.69
Partially Learned	15.98	5.86	17.10
Completely Learned	6.55	18.28	37.64

Fig. 14: Growth in average performance level of the class in percentage for all topics throughout all assessments.

The example diagrams of *Class Visualization* presented the visualizations and reports that the instructors could utilize for reviewing the instructional strategies of the course, the learning outcomes and learning progress of the class from a wider view of perspective.

5.3. Individual Visualisation

The purpose of having Individual Visualization is for the instructor to review and evaluate the performance of a particular individual learner, in comparison to the average performance of the class. The visualizations and reports aid the instructor to gain better insight about the learning status of the class by detailing the performance of the class into individual level. The visualizations and reports displayed for Class Visualization, similarly presented in Individual Visualization with context shifted into particular individual learner.

Take a student, Student A in the case study as example, several visualizations and reports of Individual Visualization were presented in the following. An example of the scores and performance level achieved by Student A for one of the assessments is illustrated in Fig. 15. The scores and performance level are displayed in tabular form for ease of reading, and with additional details on the configuration of each assessment question, the instructor could gain detailed information about the performance of Student A in that particular assessment.

Quiz Scores in Quizzes

Student's Matric No: [redacted]

Quiz 1

	Mark Allocation	Scores	Related Topics	Bloom Levels	Difficulty Levels	Performance Level
Question 1	4	3	Topic 1	Level 2	3	Partially Learned
Question 2	4	3	Topic 1	Level 3	7	Partially Learned
Question 3	4	3	Topic 2	Level 2	3	Partially Learned
Question 4	4	3	Topic 3	Level 3	8	Partially Learned
Question 5	4	4	Topic 4	Level 3	6	Completely Learned
Total Scores	20	16				

Fig. 15: Individual scores and performance level in one of the assessments

Besides the table of scores and performance achieved by Student A, Individual Visualization also presented how Student A performed for all topics throughout all assessments in average and in comparison of the average knowledge level of the class (as shown in Fig. 16). The average knowledge level of Student A is presented as column chart, while the average knowledge level of the class is presented as line chart. The diagram shows that Student A performed majority in the level of Completely Learned (55.56%). In comparison with the average knowledge level of the class, the performance of Student A is better, as the average knowledge level of Student A in term of Completely Unknown (11%) is lower than the average knowledge level of the class (44%), while Completely Learned (55.56%) is higher than the average (23.88%). Through this diagram, the instructor could know that Student A is well performed throughout the lectures of module.

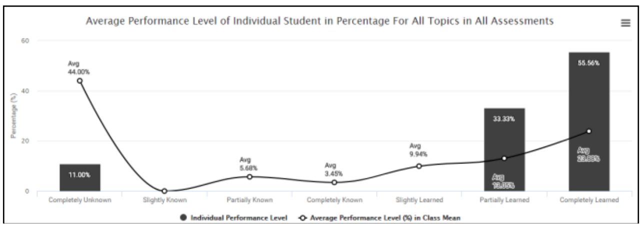


Fig. 16: Individual average performance level in percentage for all topics throughout all assessments.

In case of the instructor intends to evaluate the growth in knowledge level of Student A, Individual Visualization could present the growth chart of the knowledge level of Student A for all topics throughout all assessments (as shown in Fig. 17). As shown in the diagram, Student A performed in terms of Completely Unknown (40%) in second assessment, however, at the third assessment, the Student A showed the gaining of knowledge level, and decreased the corresponding percentage to zero, while increasing the performance level in term of Completely Learned from 20% to 88%. Through the diagram of growth chart, the instructor could gain insight of the learning progress of Student A throughout the lectures of module.

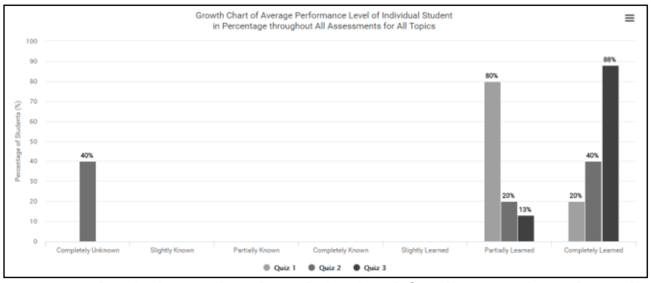


Fig. 17: Individual growth in knowledge level for all topics throughout all assessments.

The instructor could also review the average performance level of Student A in accordance to the Bloom Levels through pie charts. Take example of the performance of Student A in the Bloom Levels of Level 2 and Level 3 (as illustrated in Fig. 18), the percentage of average performance level of Completely Learned in the Bloom Level 2 (66.67%) is higher than the Bloom Level 3 (44.44%). This showed that Student A gained more knowledge in terms of Bloom Level 2 than in Bloom Level 3. Meanwhile, the average performance level of Student A in terms of Completely Unknown and Partially Learned in Bloom Level 3 (11.11% and 44.44%) is slightly higher than the Bloom Level 2 (0% and 33.33%). From the pie charts, the instructors could compare the average knowledge level of Student A in various cognitive process categories, and further evaluate the effectiveness of the instructional strategies to support the learning of Student A in various cognitive process categories.

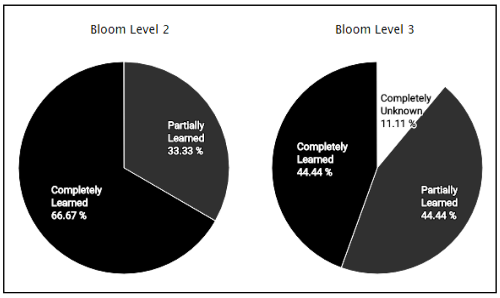


Fig. 18: Individual average performance level in accordance to Bloom Level 2 and Bloom Level 3.

The example diagrams of Individual Visualization presented that through visualizations and reports, the instructors could evaluate the strength and weakness of an individual learner in learning particular topics of the course, and certain cognitive skills as described through bloom levels. Such visualizations and reports could aid the instructors in customizing the learning materials and instructional strategies to deliver personalised learning. Consequently, the quality of the lectures of modules, and the effectiveness of lecturing could be improved together with the increase of gaining knowledge among the students.

In short, the implementation of the proposed topic-based students' knowledge modelling system was presented through a real-world case study. Take the case study as example, various relevant visualizations and reports were presented to illustrate the way the instructor could utilize them for evaluating the learning status, learning outcomes, and learning progress of a group of individual learners, or a particular individual learner. Consequently, the instructor could gain better insight about the lectures of module, the assessment configurations, and the effectiveness of the lectures in supporting the students in learning

6. Conclusion

This paper presents the design and development of a topic-based students' knowledge modelling system that employed the fuzzy set theory and visual analytics. The purpose of proposing such a system is to contribute towards the efforts of adopting personalised learning in practical education. The proposed system consists of a topic-based domain model that is designed using the topic-based approach, which is a friendlier approach for instructors. The domain model consists of a set of *topics* that is structured based on the curricula of a particular course domain. Moreover, a topic-based overlay student model is used to structure and model the knowledge level of individual learners in accordance to the set of topics and the instructional objectives indicated by the instructors. To infer the knowledge level of the individual learner, this proposed system utilizes the fuzzy set theory by constructing a fuzzy inference system. The proposed fuzzy inference system considers various vague factors that may influence the inference of knowledge levels, which include the assessment results. To illustrate the modelled knowledge level in the proposed system, a visual analytics component was included to provide a wider range of possibilities for interpretations and reading of the analytics outcomes. By utilizing the visualizations and reports generated by the visual analytics component, the instructors can monitor and evaluate the knowledge level of each individual learner, and improve or tailor the instructional strategies during teaching to improve the learning of each individual learner.

The topic-based design of the proposed system is flexible in representing varies domain, and is easy to learn and comprehend by instructors. Future studies may work on applying the proposed system to model knowledge level of individual learners in other domains. As the topic-based design of the proposed system is simple and friendly to the instructors, future studies may design and develop relevant authoring tools to aid the instructors in authoring the domain model, configuring the assessments and assessment questions.

The design of the proposed system separates the components for modelling, reasoning, and reporting the knowledge level of the students. Such design offers flexibility in applying various knowledge modelling techniques, knowledge inference techniques, such as Bayesian Networks and so forth as well as visual analytics methods. Furthermore, such design is open for embedding or be embedded into other components or systems. For example, the proposed system may be incorporated into an adaptive learning system, to provide personalised learning for individual learners based on the knowledge level modelled by the proposed system.

The visual analytics components can also be expanded by including more possible visualizations and reports by considering other possible factors that might influence the inference of the knowledge level. Possible factors include the duration taken to answer a particular assessment question, the errors and misconceptions of the individual learners towards the set of topics and so forth. With additional factors considered to infer the knowledge level, the instructors may explore and evaluate the knowledge level of individual learners in other aspects as well.

Acknowledgement

The authors acknowledge the financial support rendered by Universiti Malaysia Sarawak through UNIMAS Special Grant, grant no. F04/SpFRGS/1536/2017.

References

[1] Spector J (2018), Future trends of designing learning in the global context. *Authentic Learning Through Advances in Technologies*, Chang TW, Huang R, and Kinshuk, Eds. Springer, Singapore, pp. 205–216.

[2] Brusilovsky P (2016), Domain modeling for personalized guidance. *Design Recommendations for Intelligent Tutoring Systems: Volume 4 - Domain Modeling*, Sottilare R, Graesser A, Hu X, Olney A, Nye B, and Sinatra A, Eds. Orlando, FL: U.S. Army Research Laboratory, pp. 165–183.

[3] Ng WT & Teh CS (2017), Fuzzy students' knowledge modelling system through Revised Bloom' s Taxonomy. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, Vol. 9, No. 2–9, pp. 169–174.

[4] Shute VJ & Zapata-Rivera D (2012), Adaptive educational systems. *Adaptive technologies for training and education*, Durlach PJ & Lesgold A, Eds. Cambridge: Cambridge University Press, pp. 7–27.

[5] Bates S (2014), Personalised learning: Implications for curricula, staff and students. Paper presented at the Universitas 21 (U21) Educational Innovation Conference, The University of New South Wales (UNSW), Australia.

[6] Shute VJ (2006), Tensions, trends, tools, and technologies: Time for an educational sea change. *ETS Research Report Series*, Vol. 2006, No. 1, pp. i-49.

[7] Woolf BP, Ed. (2010), *A roadmap for education technology [National Science Foundation Grant #0637190]*. Washington: Computing Community Consortium.

[8] Volarić T, Vasić D & Brajković E (2017), Adaptive tool for teaching programming using conceptual maps. *Advanced technologies, systems, and applications*, Hadžikadić M & Avdaković S, Eds. Switzerland: Springer International Publishing, pp. 335–347.

[9] Polson MC & Richardson JJ (1988), *Foundations of intelligent tutoring systems*. New Jersey: Lawrence Erlbaum Associates.

[10] Chrysafiadi K & Virvou M (2015), *Advances in personalized web-based education*, Vol. 78. Cham: Springer International Publishing.

[11] Jeremić Z, Jovanović J & Gašević D (2012), Student modeling and assessment in intelligent tutoring of software patterns. *Expert Systems with Applications*, Vol. 39, No. 1, pp. 210–222.

[12] Grubišić A, Stankov S & Žitko B (2013), Stereotype student model for an adaptive e-learning system. *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, Vol. 7, No. 4, pp. 440–447.

[13] Zadeh LA (1965), Fuzzy sets. *Information and Control*, Vol. 8, pp. 338–353.

[14] Ruipérez-Valiente JA, Muñoz-Merino PJ, Leony D & Delgado Kloos C (2015), ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Computers in Human Behavior*, Vol. 47, pp. 139–148.

[15] Brusilovsky P & Millán E (2007), User models for adaptive hypermedia and adaptive educational systems. *The Adaptive Web*, Brusilovsky P, Kobsa A & Nejdl W, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 3–53.

[16] Chrysafiadi K & Virvou M (2013), Student modeling approaches: A literature review for the last decade. *Expert Systems with Applications*, Vol. 40, No. 11, pp. 4715–4729.

[17] Stansfield JL, Carr BP & Goldstein IP (1976), *Wumpus advisor I: A first implementation of a program that tutors logical and probabilistic reasoning skills*. Cambridge, Massachusetts: Massachusetts Institute of Technology.

[18] Martins AC, Faria L, De Carvalho CV & Carrapatoso E (2008), User modeling in adaptive hypermedia educational systems. *Journal of Educational Technology & Society*, Vol. 11, No. 1, pp. 194–207.

[19] Sosnovsky S & Brusilovsky P (2015), Evaluation of topic-based adaptation and student modeling in QuizGuide. *User Modeling and User-Adapted Interaction*, Vol. 25, No. 4, pp. 371–424.

[20] Johnson L, Adams Becker S, Cummins M, Estrada V, Freeman A & Hall C (2016), NMC horizon report: 2016 higher education edition. *The New Media Consortium*, Austin, Texas.

[21] Sosnovsky S & Brusilovsky P (2005), Layered evaluation of topic-based adaptation to student knowledge. *User Modeling 2005*, Ardissono L, Brna P & Mitrović A, Eds. Berlin, Heidelberg: Springer, pp. 47–56.

[22] Sammut C & Webb GI (2010), *Encyclopedia of Machine Learning*. Boston, MA: Springer US.

[23] Chrysafiadi K & Virvou M (2015), Fuzzy Logic for adaptive instruction in an E-learning environment for Computer Programming. *IEEE Transactions on Fuzzy Systems*, Vol. 23, No. 1, pp. 164–177.

[24] Voskoglou MG & Subbotin IY (2015), A Fuzzy Model for Measuring the Student Learning Based on Bloom's Taxonomy. *Egyptian Computer Science Journal*, Vol. 39, No. 2, pp. 43–55.

[25] Mehdi Khosrow-Pour DBA (2018), *Encyclopedia of information science and technology (4th ed.)*. Hershey: IGI Global.

[26] Leony D, Pardo A, de la Fuente Valentín L, de Castro DS & Kloos CD (2012), GLASS: A learning analytics visualization tool. *LAK '12 Proceedings of the 2nd international conference on learning analytic and knowledge*, pp. 162–163.

[27] Schmitz HC, Scheffel M, Friedrich M, Jahn M, Niemann K & Wolpers M (2009), CAMera for PLE. *Learning in the synergy of multiple disciplines.* Cress U, Dimitrova V & Specht M, Eds. Berlin, Heidelberg: Springer, pp. 507–520.