74129

DEVELOPMENT OF KNOWLEDGE MANAGEMENT FOR ADAPTIVE HYPERMEDIA LEARNING SYSTEM

SHAHLIZA ABD HALIM
NOR BAHIAH HJ AHMAD
PM DR NAOMIE HJ SALIM
PM DR SITI MARYAM HJ SAMSUDIN

FAKULTI SAINS KOMPUTER DAN SISTEM MAKLUMAT
UNIVERSITI TEKNOLOGI MALAYSIA

2006

Abstract

Adaptive hypermedia learning systems (AHLS) aims in adaptively accommodating learning materials based on individual differences in user. This kind of system faces problems in cognitive overload where too many information are given to students resulting to some of them become lost in hyperspace. In order to reduce the overload, user or leaner should be given materials well suited with their learning ability and also in their style of learning. Our project concentrates in individualizing the learning material and navigational paths to adapt with different users learning styles and knowledge acquisition. Learning ability is associated by their knowledge acquisition in using the system captured via user model. Knowledge acquisition per se is not enough as research on learning shows students learns effectively when they were thought with methods that suits their learning style. Pedagogical framework for this project comprise of Myers Briggs Type Indicator (MBTI) personality factor as learning strategy and Howard Gardner and Honey & Mumford theory for learning technique and method. Our contribution is on the learning materials which are structured based on the pedagogical framework. Another contribution is mainly on the theoretical and practical mechanism for the integration of computational intelligence techniques mainly to personalize the user both at the presentation and navigation level. Theoritical and practical aspects of this project are discussed; e.g., instructional material suitable for learning style chosen, experiments on Kohonen neural network and rough sets for classification of knowledge acqusitions, fuzzy logic architecture for the adaptation in learning style, architecture of the prototype developed and also its evaluation. The contribution of this project is based on these seven papers describe herewith:

Paper I discusses the problems in personalizing instructional material that matches the differences of students according to student's fuzzy membership to certain learning styles.

Paper II focuses on the use of computational intelligence techniques such as Kohonen self-organizing maps in the classification of student models and Fuzzy logic in the adaptation of learning material and navigation path.

Paper III discusses the structure of the learning material which are; 1) theory, 2) example, 3) exercise and 4) activities.

In paper IV, a generation of rough set rules is implemented in identifying the status of students' knowledge acquisition in hypermedia learning.

Paper V describes the use of back propagation neural network in classifying the student.

Paper VI presents the architecture and design for the development of an adaptive hypermedia learning system for teaching and learning Data Structure.

Paper VII discusses the evaluation of an adaptive hypermedia learning system developed. We use summative evaluation where we address the educational impact of this system on students and its practical acceptability in terms of usability.

// Abstrak

Sistem Pembelajaran Hipermedia Adaptif (AHLS) mempunyai matlamat untuk menyesuaikan material pembelajaran secara adaptif berdasarkan keunikan yang terdapat pada setiap individu. Sistem seumpama ini menghadapi masalah bebanan kognitif lebihan dimana terlalu banyak maklumat diberikan kepada pelajar menyebabkan sesetengah daripada pelajar “sesat dalam ruang hiper”. Untuk mengurangkan bebanan tersebut, pengguna atau pelajar perlu diberikan bahan yang bersesuaian dengan kemampuan pembelajarannya dan juga gaya pembelajarannya. Projek ini menumpukan kepada persembahan bahan pembelajaran dan laluan penyusuran yang unik bagi setiap indivdu bagi membolehkan adaptasi dilakukan terhadap gaya pembelajaran dan perolehan pengetahuan yang berbeza dikalangan pengguna atau pelajar. Kemampuan pembelajaran individu dilihat berdasarkan perolehan pengetahuan semasa menggunakan sistem dimana ianya disimpan di dalam model pengguna. Perolehan pengetahuan semata-mata tidak mencukupi dimana penyelidikan dalam bidang pembelajaran menunjukkan pelajar dapat mempelajari dengan lebih efektif apabila mereka diajar dengan kaedah yang bersesuaian dengan gaya pembelajaran mereka. Rangkakerja pedagogi bagi projek ini adalah terdiri daripada faktor personaliti Myers Briggs Type Indicator (MBTI) sebagai strategi pembelajaran dan teori Howard Gardner serta Honey&Mumford sebagai kaedah dan teknik pembelajaran. Sumbangan projek ini adalah dari sudut material pembelajaran yang telah distrukturkan berdasarkan kepada rangkakerja pedagogi. Sumbangan utama adalah mekanisme secara teori dan praktikal bagi integrasi teknik kepintaran komputan bagi memberikan material yang khusus kepada pengguna pada aras persembahan dan navigasi. Aspek teori dan praktikal bagi projek ini telah dibincangkan contohnya material pembelajaran yang bersesuaian dengan gaya pembelajaran yang dipilih, eksperimen terhadap rangkaian neural kohonen dan juga set kasar bagi pengklasifikasian perolehan pengetahuan, senibina bagi logik kabur untuk pengadaptasian gaya pembelajaran, senibina bagi prototaip yang telah dibina dan juga penilaian terhadap prototaip. Sumbangan projek ini adalah berdasarkan kepada tujuh kertas kerja yang diterangkan di bawah:-

Kertas I membincangkan masalah-masalah yang wujud dalam adaptasi bahan pengajaran bagi persembahan bahan pembelajaran yang sepadan dengan keunikan pelajar berpandukan hubungan kabur pelajar dengan gaya pembelajaran pelajar terbabit.

Kertas II memfokuskan kepada aplikasi teknik-teknik pengkomputeran pintar seperti pemetaan sendiri Kohonen dalam pengelasan model pelajar dan logik kabur dalam adaptasi bahan pembelajaran dan laluan penyusuran.

Kertas III pula membincangkan bahan-bahan pembelajaran yang terdiri daripada ; 1) teori, 2) contoh, 3) latihan dan 4) aktiviti.

Dalam Kertas IV penjanaan peraturan bagi set kasar telah diimplimentasikan dalam mengenalpasti status perolehan pengetahuan pelajar dalam pembelajaran hipermedia.

Kertas V menerangkan aplikasi rangkaian neural rambatan balik dalam mengelaskan pelajar.

Kertas VI membentangkan senibina dan reka bentuk pembangunan sistem pembelajaran adaptif hipermedia bagi pengajaran dan pembelajaran Struktur Data.

Kertas VII membincangkan penilaian yang dilakukan bagi sistem pembelajaran hipermedia adaptif yang telah dibina. Kami menggunakan Penilaian Penyimpulan dimana penilaian adalah terhadap impak pendidikan sistem ini dan penerimaan praktikal terhadap kebolehgunaannya.

TABLE OF CONTENT

INTRODUCTION

Introduction to Knowledge Management and Knowledge Discovery

Knowledge management is a process which refers to a range of practices used by organizations to identify, create, represent and distribute knowledge for reuse, awareness and learning across the organization. This process deals with methods, models and strategies to capture, reuse and maintain knowledge. Knowledge discovery on the other hand is the process of identifying valid, novel, potentially useful and ultimately understandable patterns in data. It involves a new generation of techniques and tools to intelligently and automatically assist humans in analyzing the mountains of data for useful knowledge. In Adaptive Hypermedia Learning System (AHLS), knowledge management is crucial as it will organize knowledge base to store the learning sources and store user profile and at the same time restrict the amount of knowledge represented to the user. Knowledge discovery on the other hand can be use to discover knowledge about student behavior during their learning time.

Introduction to Adaptive Hypermedia Learning System (AHLS)

An AHLS is an educational system that incorporates the features of hypermedia and intelligent tutoring system. Such system can create a learning environment that gives students freedom to explore and learn but implicitly guided by the system. Currently a prototype of an AHLS, called SPAtH, has been developed. This system has three main components student model, domain-expert model, and adaptive engine. Student model contains the student knowledge including individual information, learning history and level of knowledge acquisition; the domain-expert model stores the teaching materials consisting of notes, practices examples, check-point questions, help/hints and solutions to each questions and the teaching strategies that match the learner's knowledge acquisition status namely poor, average and good. The adaptive engine contains all the functions required as the interface between the user and the system, as well as the system adaptiveness. Concern arises over how to deal with exponential increases in the amount of available knowledge and increasingly complex processes.

Managing knowledge represents the primary opportunity for achieving substantial savings, significant improvements in human performance, and competitive advantage. Repositories promote the preservation, sharing, retrieval and reuse of data. A knowledge repository organizes and stores data, information, knowledge, expertise, and experience for one domain.

Literature Review on Kohonen Network

Kohonen network has been widely used for classification task and the results are considered succeeded Cho (1997). Basic Kohonen network like vector quantization or k-means can be used for simple classification Sarle (1994).

The development of the self-organizing map (SOM) was introduced by Kohonen (1995). Such maps are the end result of analysis by what are known as Kohonen networks. The self-organizing Kohonen network is a type of neural network and is popular in areas that require visualization and dimension reduction of large, high dimensional data sets. Kohonen network are a vector quantization method, which can preserve the topological relationships between input vectors when projected to a lower dimensional display space. The idea was to create a neural network to represent the input space using the topological structure of a grid to store neighborhood relations. In contrast to most neural network methods that use the desired result to compute the weights of the network, Kohonen network need no reference output (unsupervised learning).

The network is trained by finding the weight vector, which is most similar to an input vector. This weight vector, and its neighbors are then updated to make them more similar to the input vector. The output of training phase is a weight vectors. After training, the new testing data set is feeded into the SOM. During testing, the weights gained during training are used to classify the data set into a number of groups based on the similarity they shared.

The groups are not easy to interpret manually especially to recognize the similarity features shared among the group members. It makes the process of evaluating which group represents which kind of data become difficult. Many researchers developed self-organizing techniques to automatically grouping a set of data to groups and using different techniques to interpret and labeling those groups (Kaski *et al*.,1998; Rauber, 1999; Drobics et al., 2000).

Literature Review on Fuzzy Logic

Fuzzy logic method, proposed by Zadeh [10], has proved to be very effective in handling the vagueness and uncertainty intrinsically existing in the knowledge possessed by people. It is also computationally undemanding and is most suitable for processing imprecise input data, as it supports natural description of knowledge and reasoning in the form of imprecise concepts, operators and rules [Negnevitsky 2002].

In ITS, fuzzy logic techniques have been used due to their ability to handle imprecise information, such as student's actions, and to provide human descriptions of knowledge and of student's cognitive abilities [Stathacopoulou et al. 1999]. The fuzzy logic system consists of three main stages: Fuzzification, Rule Evaluation and Defuzzification. Fuzzy rules based in fuzzy logic provide a qualitative description of the input-output relationship of a system using fuzzy linguistic term.

Research Problem

We define three major problems that lead us to the proposed solution in this project:

i. Management of large and complex knowledge in AHLS

Even though there has been many research conducted recently, researchers still face a problem to manage a large and complex knowledge in hypermedia learning system. This is because of too many information are given to the students that makes some of them become *lost in hyperspace* (Brusilovsky, 1996). The system had difficulty in providing the most suitable and useful information and learning material to the students. We provide an organized database to store the learning souces and student data and allows the access to these data efficiently.

ii. Identifying method for managing student model intelligently and method for individualizing the learning material

With SPAtH, we apply computational intelligence approach to help managing the learning material and students' data. We do this by using fuzzy logic to evaluate student's preferable learning material based on student personality factor and individualizing it to suit the individual preferences. Some of the techniques implemented in this project to solve the above questions are neural network such as back propagation and Kohonen's self-organizing maps, fuzzy logic and rough set.

iii. Integrating the learning style and knowledge acquisition for adaptivity in hypermedia learning system

It is an important part of the research which is to integrate the learning style and knowledge acquisition to make it possible for the adaptivity in SPAtH. Our research has developed the design of adaptation based on the learning style information and level of knowledge acquisition. To determine the level of knowledge acquisition, we have selected the attributes of adaptation based on the work done by Papanikolaou et al.(2003) and Paridah et. al. (2001). The attributes selected are the learning time, number of backtracking, number of getting help function and the score earned while doing the exercise. Meanwhile, the pedagogical and learning style refer to student's personality factor based on (MBTI) as explained in Norreen & Naomie (2005). We also have identified the structure of the learning material that the system should offer to learners with different styles and characteristics.

Research Objective

To provide the solution to the above problem statement, we define three objectives:

i. To identify suitable computational intelligence techniques for managing knowledge stored in student model particularly in its classification based on knowledge acquisition and domain model specifically in adapting learning materials based on the mix traits of different learning styles for AHLS.

ii. To test and enhance the methods for personalizing the learning style using Fuzzy Logic and classification of user model based on knowledge acquisition using Kohonen neural network and Rough Sets.

iii. To develop prototype software that integrates the mechanism of adaptivity in learning style and knowledge acquisition for AHLS.

iv. To evaluate the prototype software based on its impact in student learning particularly in its usability.

Reasearch Scope

This research is conducted to achieve the above objectives within the following scope:

i) Study soft computing techniques suitable for adaptation in Hypermedia Learning System such as Rough Sets, Kohonen Network and Fuzzy Logic.
ii) Learning style is based on Myers Brigss Type Indicator (MBTI).
iii) The prototype of an AHLS will be used for teaching and learning Data Structure for university students.

Report Overview

Report on the research is explained by the following papers:

i. PAPER I, "Fuzzy Logic Approach to Evaluate Student's Preferable Learning Material Based on Student Personality Factor", discusses the problems in personalizing instructional material for representing learning material that matches the differences of students according to student's fuzzy membership to certain learning styles.

ii. PAPER II, "Individualizing the Learning Material and Navigation Path in an Adaptive Hypermedia Learning System", focus on the use of computational intelligence technique such as Kohonen self-organizing maps in the classification of student models and Fuzzy logic in the adaptation of learning material and navigation path.

iii. PAPER III, “Individualizing Learning Material Of Adaptive Hypermedia Learning System Based On Personality Factor (MBTI) Using Fuzzy Logic Techniques”, propose the solution to the inflexible linking provided in conventional hypermedia learning system that can cause teaching and learning to be less effective.

iv. PAPER IV, “Rough Set Generation for Identifying Status of Student’s Knowledge Acquisition”, implementation of a generation of rough set rules in identifying the status of students knowledge acquisition in hypermedia learning.

v. PAPER V, “Student Classification Using Neural Network in Adaptive Hypermedia Learning System: Preliminary Work”, describe the use of back propagation neural network in classifying the student that is needed for the system to provide suitable learning module to each individual student by taking consideration of students’ knowledge level and their performances as they go through the system.

vi. PAPER VI, “Development of an Adaptive Hypermedia Learning System”, presents the architecture and design for the development of an AHLS for teaching and learning Data Structure.

vii. PAPER VII, “Summative Evaluation for the Usability of Intelligent Tutoring System (SPATH)” discusses the evaluation of an adaptive hypermedia learning system called SPATH. We use summative evaluation where we address the educational impact of this system on students and its practical acceptability in terms of usability.

PAPER I:

FUZZY LOGIC APPROACH TO EVALUATE STUDENT'S PREFERABLE LEARNING MATERIAL BASED ON STUDENT PERSONALITY FACTOR

Norreen Binti Haron
Software Engineering Lab
Department of Software Engineering
Faculty of Computer Science and Information System
Universiti Teknologi Malaysia
norreen.haron@gmail.com

Naomie Binti Salim
Faculty of Computer Science and Information System
Universiti Teknologi Malaysia
naomie@fsksm.utm.my

Abstract

AHLSs provide adaptation to user's learning styles. However, most of the AHLSs incorporate learning style based on the notion that each student has only one learning style which is not necessary true in the real life. Incorporating several learning styles of students into the system to allow for better matching of material to students cannot be done using crisp algorithms due to the fuzzy nature of the learning styles in each individual. In this research, a system based on fuzzy logic has been developed in which a student can have mixed traits of different styles, each with a certain percentage of membership, rather solely having one particular learning style. We also address problems in personalizing instructional material for representing learning material that matches the differences of students according to student's fuzzy membership to certain learning styles. There are four input and four output linguistic variables considered in this paper, where the inference rule of fuzzy reasoning consists of four antecedents and four consequents. The antecedents are represent based on the student's personality factor (Myers-Briggs Type Indicator (MBTI)); extrovert score, introvert score, sensor score and intuition score, and the consequents represented the student's preferable learning material; (theory, example, exercise and activities). Based on fuzzy set and fuzzy rule theory, the vagueness and uncertainty intrinsically existing in the knowledge possessed by expert is computed whilst providing qualitative description of the input-output relationship using fuzzy linguistic terms. Triangle fuzzy set, Mamdani inference and center of gravity (COG) defuzzification techniques are used in the system. Comparison of the defuzzified values of the fuzzy rule base system and the value from conventional system shows that the fuzzy rule base system has better performance in structuring the learning material.

Keywords

Fuzzy logic, Fuzzy sets, Fuzzy ruled base, Learning style, AHLS.

1. Introduction

Research on learning has shown that student learns differently, they process and represent knowledge in different ways and they learn more effectively when taught with preferred methods. Information about learning style can help system become more sensitive to the differences students using the system.

Although learning style theory is widely accepted amongst educational theorists in the context of e-learning environments [1,2,3,4,5], there is still no research on the adaptation to individualize student's learning material based on fact that students have more than one learning style in a certain degree. In particular the possibility of fluctuations in a learning style with changing tasks or content has not yet been addressed [6].

In this paper, a system based on fuzzy logic has been developed in which a student can have mixed traits of different styles, each with a certain percentage of membership, rather solely having one particular learning style. It aims to utilize the learning characteristics and provide a personalized learning environment, that exploit learning style and fuzzy logic techniques. We focus on using the fuzzy rule-based system that involved fuzzy sets and fuzzy logic. The learning style in this paper refer to the student's personality factor; Myers-Briggs Type Indicator (MBTI) [7, 8]. Based on the MBTI theories, the fuzzy logic techniques are then use to classify the student's preferable learning material.

In designing the fuzzy logic system, it is important to identify the main control variables and determine the term set which is at the right level of granularity for describing the values of each linguistic variable [9]. In this problem, the fuzzy system is represented by four input linguistic variables (or the antecedents), and four output linguistic variables (or the consequents). Also, each input and output may be represented by either a three-term-set or a five-term-set of linguistic values. After defining the fuzzy variables and its term sets, fuzzy rule base is then being constructed. The number of fuzzy rules being formed is directly related to the number of fuzzy term sets defined at the antecedents. At the end of the process, the crisps output shows the structure of the learning material, which is learning material that the student most preferable and which learning material that the student choose less attention.

2. Approach and Method

Fuzzy logic method, proposed by Zadeh [10], has proved to be very effective in handling the vagueness and uncertainty intrinsically existing in the knowledge possessed by people. Fuzzy rules based in fuzzy logic provide a qualitative description of the input-output relationship of a system using fuzzy linguistic term. Moreover, fuzzy linguistic rule appear close to human reasoning and in many real-world applications, and it is more adequate and flexible for knowledge representation than conventional crisp IF THEN rules. Thus, it is reasonable to use fuzzy logic system to classify students and determine the most suitable learning material for them.

Fuzzy logic techniques in this research are used to personalize the learning material where it denotes to structures of learning material. It calculates precisely the structures of learning material (theory, example, exercise and activities) that suit student's personality.

In this study, the fuzzy sets and fuzzy rules are constructed based on the standard of mastery or the criterion-referenced acquired from the human instructors' experience and knowledge about student's learning style.

The processes that implement in fuzzy logic system indicate 4 main stages; fuzzification, rule evaluation, aggregation and defuzzification. The fuzzification stage transforms crisp student personality, captured from pre-course questionnaires, into suitable linguistic values. The rule evaluation stage takes the fuzzified inputs, and applies them to the antecedents of the fuzzy rules. The aggregation stages use to combine all the fuzzy output derived from fuzzy rules, and last step, the defuzzification stage produces the output in crisp value. In this problem, the defuzzified value is derived based on the centre of gravity method. Figure 1 shows the flow of the fuzzy logic system.

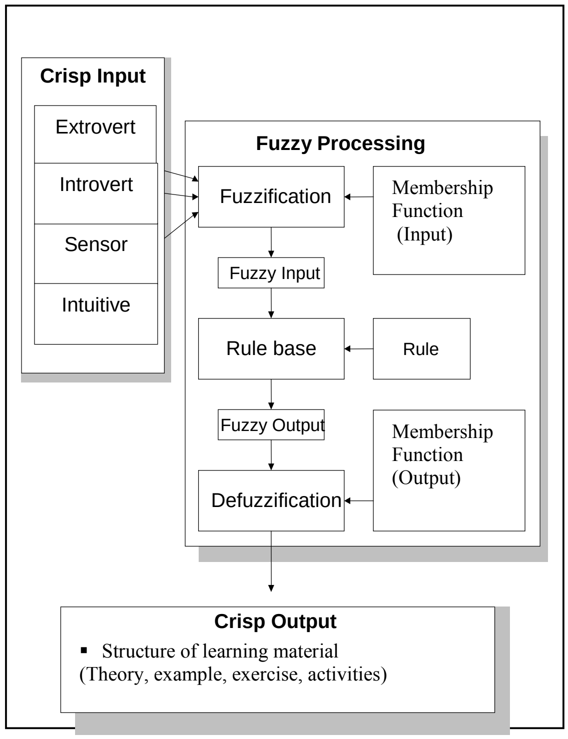


Figure 1: Flow of Fuzzy Logic System

In this paper, there are four inputs and four outputs of linguistic variables are being considered. The input linguistic variables are representing the student personality's scores; extrovert scores, introvert scores, sensor scores and intuitive scores.

The input expressed by:

$(x_a, x_b, x_c \text{ and } x_d)$. (1)

Whilst, the outputs are expressed by:

$(y_a, y_b, y_c \text{ and } y_d)$ (2)

that represents the student's acceptance level of learning material; theory, example, exercise and activities.

To identify student personality's scores as fuzzy numerical data in ranges (0.0, 1.0), all the score retrieve from pre-course questionnaire are gathered according to each personality and are calculated by dividing the total score of particular personality, s_i, with the total number of pre-course questionnaire answered for that particular personality, P, as follows:

$$X_a = \frac{\sum_{i=1}^{P} s_i}{P} \quad (3)$$

We also determine the membership function of fuzzy set in fuzzification stages. The membership functions of fuzzy set in this paper are based on Zimmerman [11] which is allows the element in set to have multiple degree of membership. For example:-

$\tilde{A} = \{(x, \tilde{A}(x)) \mid x \text{ e } R\})$ (4)

where x={100, 90, 80, 70, 60, …, n}

then $\tilde{A}$ ={(100, 1), (90, 0.9), (80, 0.8),

(70,0.7), (60, 0.6),…, (n, $\tilde{A}(x)$)}

-R is student,

-x is student personality's value,

-$\tilde{A}(x)$ is membership function x in $\tilde{A}$,

-($\tilde{A}$) is a fuzzy set for personality's value x with student's acceptance level of theory learning material.

The membership value 1, appoint to the student personalities x=100, where the student's acceptance level of theory learning material at the highest level. The student's acceptance level will decrease when number in personality x become lower.

In the second stages, rule evaluation, the inference rule of fuzzy reasoning consists of multiple antecedents and multiple consequents are expressed as below:

Ri :

IF X_a is N_1 and X_b is N_2

and X_b is N_3 and X_d is N_3

THEN

Y_1 is M_1 and Y_2 is M_2

and Y_3 is M_3 and Y_4 is M_4

where Ri, (i = 1, 2, … n) is the rule number, Ni, (i =1, 2 and 3) are the membership functions of the antecedent part, Mi, (i =1, 2, 3, 4 and 5) are the membership functions of the consequent part. The example of fuzzy rules applied in this problem as show in Table 1.

Input:-

{E-extrovert, I-introvert, S-sensor, N-intuitive}

Output:-

{1-theory, 2-example, 3-excersice, 4-activity}.

Table 1: Example of Fuzzy Rules

Input- (student personality)				**Output - learning Material**			
E	**I**	**S**	**N**	**1**	**2**	**3**	**4**
Low	High	Low	High	VHigh	High	Low	VLow
Low	High	Med	Med	High	Med	Low	Vlow
Low	High	High	Low	Med	VHigh	VLow	Low
:	:	:	:	:	:	:	:
High	Low	High	Low	VLow	Low	High	VHigh

In this paper, the fuzzy set is expressed by a triangular function as triangular function can provide an adequate representation of the expert knowledge [12] and at the same time significantly simplifies the process of computation. The fuzzy set is expressed by three parameters {a, b, c} as shown in figure 2.

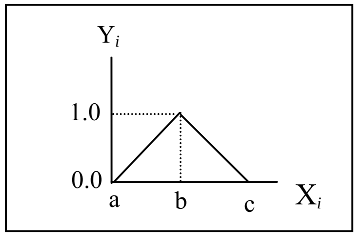


Figure 2- Triangular fuzzy set

The membership value is derived by the following formula as shown in Figure 3.

$$Y_i = \begin{cases} \dfrac{X_i - a}{b-a} & \text{if } a <= X_i < b \\ 1 & \text{if } X_i = b \\ 1 + \dfrac{b - X_i}{c-b} & \text{if } b < X_i <= c \\ 0, & \textit{otherwise} \end{cases}$$

Figure 3 - Triangular membership function

where a, b, and c are parameters that control the intervals of the triangle (as shown in figure 2), X_i is the input (i =1, 2, 3, 4), and Y_i is the output of fuzzification input i.

Overall process in second stages is expressed as following definition:-

$$\acute{L}i(z) := ((R\alpha i = [Ai(x_{a1}) \square Bi(x_{b1})]) \square Li(z))$$
for i= 1,2,3…n

(5)

Ĺi(z) is fuzzy output retrieved from selected rule (Rα) based on the rule evaluation , input data, antecedent (xa,xb,xc and xd) and output data, consequent, Li(z).

The maxima method in aggregation stage is based on the definition shown in equation (6):-

Ļ (z) = Ĺ1 (z) Ĺ2 (z)

= (Rα1 L1 (z)) (Rα2 L2 (z))   (6)

Ļ(z) are fuzzy set expressed from combination of all consequent values retrieved from selected rule, Ĺi(z), (i= 1,2,3…n), where Rαi is selected rule and Li(z), (i= 1,2,3…n), is fuzzy consequent value.

The defuzzified value of the fuzzy reasoning, are derived based on the Mamdani-style inference (centre of gravity-COG) as shown in equation (7) below:

$$COG = \frac{\sum_{x=a}^{c} Zb(x)\, x}{\sum_{x=a}^{c} Zb(x)} \qquad (7)$$

3. Experiment and Result

Table 2: The Comparison between Fuzzy rule base system and Conventional System

Data Input (student personality) %				**Fuzzy rule base system**				**Conventional System**			
E	I	S	N	Theory	Example	Exercise	Activity	Theory	Example	Exercise	Activity
30	70	60	40	0.554	0.683	0.317	0.446	0.000	1.000	0.000	0.000
46	54	40	50	0.526	0.470	0.500	0.428	1.000	0.000	0.000	0.000
50	30	80	20	0.187	0.569	0.25	0.750	0.000	0.000	0.000	1.000

Comparison of the defuzzified values of the fuzzy rule base system and the value from conventional system shows that the fuzzy rule base system present better performance in structuring the learning material. As shows in table 2, the conventional system can't structure the learning material and shows only one learning material to the user even though the user have mixed traits of different styles. Differ from fuzzy rule base system; the students can have different learning material according to their mixed traits of different styles, with a certain percentage of membership rather solely having one particular learning material to learn. Moreover, the result also shows that the learning material could be structure according to the defuzzified result.

4. Conclusion and Further Work

This paper has proposed a way to personalize the learning material for AHLS, which aims to provide learners with a customized learning environment. It emphasizes the combination of learning style theories and artificial intelligent techniques. Fuzzy logic techniques used to impart the learning content based on student's fuzzy personality data and instructional rules in order to support customisation that will allow learners to learn faster and understand the learning material much easier. Conversely, the challenge is to identify what those learning content (structure of learning material) for a given learner in online system based on student personality and which result is more precise to the learner's personality whether the traditional approach or using the fuzzy logic techniques.

For further work, the author suggested to test the evaluation and performance of this theory, the author will conduct two assessments; testing the performance of effectiveness and testing the accuracy of the system. In the first evaluation, the author divide the student into two groups where as both group have same knowledge level. The first group will use the system that proposes in this theory while other group will use a system without including this theory. Both groups will take a test before and after using both systems. The result from the test would show which of both group perform good result. This is base on the result from student's test after using the systems.

The second evaluation is to test the accuracy of the system. This test is use to test the precision of the system with the student choice of learning material. A prototype system will be build. In this prototype system, a sub topic of data structure subject will be show in different method whereas the method is base on the teaching style. Student must first answer a questionnaire before using this prototype system.

This questionnaire is use to identify the learning style of the student. In using the prototype system, student will be show several button that link to each of the teaching method. Based on the student questionnaire and student preferable of teaching method, the system will recognize the relationship between student learning style and the teaching style. The result will be use to compare with the result in fuzzy logic techniques. This comparison is use to find the precision of fuzzy logic techniques with student preferences.

5. References

[1] N. Stach, A. Cristea and P. De Bra. "Authoring of Learning Styles in Adaptive Hypermedia." in Procceding. of WWW Conference. 2004.

[2] J.E. Gilbert and C.Y. Han. "Adapting Instruction In Search Of a Significant Difference". Journal of Network and Computer applications. 1999. 22. 3.

[3] M. Kwok and C. Jones. "Catering For Different Learning Styles, Association for Learning Technology". -Association for Learning Technology Journal-(ALT-J). 1985. 3(1). 5-11.

[4] M. Grigoriadou, K. Papanikolaou, H. Kornilakis and G. Magoulas. " INSPIRE: An Intelligent System For Personalized Instruction In A Remote Environment". In Proceedings of 3rd Workshop on Adaptive Hypertext and Hypermedia . 2001. pp. 13-24.

[5] P. Paredes and P. Rodriguez. "Considering Learning Styles in Adaptive Web-based Education". The 6thWorld Multi Conference on Systemics, Cybernetics and Informatics. 2002. vol. II. pp. 481-485.

[6] C. Wolf. "iWeaver: Towards 'Learning Style'-based e-Learning in Computer Science Education. ACE2003". Conferences in Research and Practice in Information Technology. 2003.

[7] S. Carolyn. et. al. "Myers Briggs Type Preferences in Distance Learning Education". In International Journal of Educational Technology. 2001. 2(2).

[8] C. C. Bishop. et. al. "The Myers-Briggs Personality Type and Its Relationship to Computer Programming." Journal of Research on Computing in Education. 1994. 26(3). 358 -371.

[9] H. R. Berenji. “Fuzzy Logic Controllers In Yager R.R. & Zadeh L. A., An Introduction to Fuzzy Logic Applications in Intelligent System”. Kluwer Academic Publisher, 1992.

[10] L. A Zadeh. “Fuzzy Logic for Management of Uncertainty”. John Wiley & Sons, New York.1992.

[11] H.J. Zimmerman, “Fuzzy Sets Theory - and Its Applications”. Kluwer Publishing Company. 1985.

[12] M. Negnevitsky. “Artificial Intelligence. A Guide to Intelligent Systems”. First Edition, Pearson Education. 2002.

PAPER II:

INDIVIDUALIZING THE LEARNING MATERIAL AND NAVIGATION PATH IN AN ADAPTIVE HYPERMEDIA LEARNING SYSTEM

Bariah Yusob[1], Norreen Haron[2], Nor Bahiah Hj Ahmad[3], Shahliza Abd Halim[4]

[1] byvail@catcha.com , [2] n_reen@hotmail.com , [3] bahiah@fsksm.utm.my & [4] shahliza@fsksm.utm.my

Department of Software Engineering

Faculty of Computer Science and Information System

Abstract

This research aims to develop a prototype of an AHLS for teaching and learning Data Structure for university students. Our system integrates pedagogy of education and intelligence technique in the presentation of the learning material and the suggested navigation path. The system comprises of three main components, user profile model, domain model and adaptive engine. User profile model stores the learning activities, learning performances and interaction history of each student in the database. Adaptive navigation path will provide the annotated link based on the performance and the interaction history of each student. To reduce disorientation, each student will get different paths based on their level of knowledge acquisitions and learning style. Adaptive engine will determine the appropriate learning material and the navigation path based on the student's status that was retrieved from the user profile model.

The focus of this paper is on the use of computational intelligence technique in the classification of student models and in the adaptation of learning material and navigation path. Kohonen self-organizing maps is used to classify the student's status using simulated data. We found out that Kohonen was able to cluster students accurately based on the maps assigned. Meanwhile, the domain model focuses on the uses of fuzzy logic to dynamically adapt the choice of possible paths through the learning material based on the attributes captured in the student model. The material presented to the student is adapted based on the students learning style and

performance. By adapting the user both at presentation and navigation level, we hope that this study can solve disorientation and lost in hyperspace problem that usually occur in conventional hypermedia learning system.

Keywords: Adaptive hypermedia learning system, personalization, user profile, self-organizing map, learning style.

1.0 Introduction

The Adaptive Hypermedia Learning Systems HLS (AHLS) is a computer based learning system in which interactive and dynamic learning module is customized to each student. Research on learning has shown that each individual student learns differently and processes and represents knowledge in different ways. Therefore, it is important to diagnose the learning style because some students learn more effectively when taught with preferred methods. Information about the learning style can help system become more sensitive to the differences students using the system. Several systems adapting the learning style have been developed to date; however, it is not clear which aspects of learning characteristics are worth modelling, how the modelling can take place and what can be done differently for users with different learning style [Brusilovsky 2001]. There are serious consequences when student learning styles and teaching styles do not match. One of it, the students face on difficulties to understand what is being taught, this lead to decrease of student interest to continue study in the subject and the student need a long term to finish one lesson session [Hashim and Yaakub 2003].

This research aims to develop a prototype of an AHLS for teaching and learning Data Structure for university students. Our system integrates pedagogy of education and intelligence technique in the presentation of the learning material and the suggested navigation path. The system comprises of three main components, user profile model, domain model and adaptive engine. User profile model stores the learning activities, learning performances and interaction history of each student in the database. Adaptive navigation path will provide the annotated link based on the performance and the interaction history of each student. To reduce disorientation, each student will get different paths based on their level of knowledge acquisitions.

Adaptive engine will determine the appropriate learning material and the navigation path based on the student's status that was retrieved from the user profile model.

Domain model stores all the teaching materials including the learning objectives, lecture notes, examples, exercises and the answer for each question. To adapt to the user category, the flow of the learning material for each category of the student will be different. To assist the user in terms of navigation, an individualized navigation path is constructed for each student, suggesting the path or link that the student has already learned, forbidden (the prerequisite is not fulfill), ready to be learned and need revision. Each link for each node will be annotated in different colors suggesting the depth of knowledge that the user already acquired. This way, the user can choose which node has greater priority to be learned and thus enabling him to optimize the path he plan to explore while studying. By adapting the user both at presentation and navigation level, we hope that this study can solve disorientation and lost in hyperspace problem that usually occur in hypermedia learning system.

In this paper, a framework for learning path personalization in adaptive learning system is introduced. It aims to utilize the learning characteristics and to provide a personalized learning environment, that exploit pedagogical model and fuzzy logic techniques. The pedagogical model and learning style are referring to the student's personality factor; Myers-Briggs Type Indicator (MBTI) [Carolyn et al. 2001; Bishop et al. 1994]. Based on the MBTI theory, the fuzzy logic techniques are then use to classify learning material (structure of learning material, type of learning material and additional link).

Fuzzy set theory, proposed by Zadeh [Zadeh 1992], has proved to be very effective in handling the vagueness and uncertainty intrinsically existing in the knowledge possessed by people or implied in numerical data. Rules based on fuzzy logic provide a qualitative description of the input-output relationship of a system using fuzzy linguistic terms. Fuzzy linguistic rule appear close to human reasoning and in many real-world applications, and thus it is more adequate and flexible for knowledge representation than conventional crisp IF THEN rules. This is the major reason why fuzzy classification rules are adopted in this paper.

2.0 Approaches And Methods To Implement Adaptivity

In the construction of an AHLS, the first issue that needs to be considered is how to identify the user features and to develop the content that reflects the personality's principles. The proposed architecture in this paper is based on this question. The architecture involved three main phases, as can be seen in Figure 1.

Based on Figure 1, the user profile model stores the information about learners in the learning system. The profiles were extracted from both explicit and implicit user profile. The explicit information is the information that the learner gave willingly or directly and he/she is aware that the information is kept in the database. The implicit information is the information the system collects without the learner acknowledgement. It records the learner's activity and behavior as he/she navigates through the system.

In this work, we test the learner's knowledge by giving them some exercises to be completed after finishing a concept. We keep the score that represent the explicit data because learner has to finish the exercises to gain scores. The implicit data used are the learning time, number of backtracking and number of getting help. From these data, we use Kohonen network to classify the learners' into three categories as shown in Figure 1.The process of identifying the learner's learning features is difficult [Brusilovsky 2001]. Moreover, it is not clearly defined which aspect in learning features that really useful for learner's modelling. Besides, the process of developing and identifying the learner's attributes in the learner's model will take a very long time. Therefore, we use a simulated data that represent the actual learner's data.

In the second phase, the pedagogical framework is used as a guide in presenting a good learning strategy. A good learning strategy is influences by the combination of learning approach, method and techniques. This pedagogical framework is derived from pedagogical expert model. In this paper, the learning strategy is base on the MBTI personality factor, whilst the method and techniques are base on Howard Gardner theory [Dara-Abrams 2002; Gardner and Korth 2001] and Honey & Mumford theory [Schroeder 1993]. The MBTI personality factors that use in this paper indicate of four types; Extrovert (E); Introvert (I); Sensing (S) and Intuition (N). Extrovert and introvert are illustrating the student preferable condition in focusing attention while the sensing and intuition type, illustrating the student preferable way in taking information.

The extrovert students prefer and focus on the outer activity and are energize by interaction with others. They prefer to talk, participate and interaction with people. While the introvert students prefer and focus on the inner activity. They prefer reading, listening to others and writing. Sensing student prefer concrete information, facts and procedures. They are good in memorization and like to go systematically, they also learn best with instruction that allow them to use their sense. Intuition student prefer discovering possibilities and relationship. They also like courses that involve a lot of experimentation and experiences.

The pedagogical framework comprise the steps on how content is develop to reflect those personality principles. Table 1 shows the relationship between learning strategy and learning method.

Table 1: The Relationship between Learning Style and Method.

Learning Style (MBTI)	**Method (Howard Gardner and Honey & Mumford)**
Extrovert	1. Visual 2. Kinaesthetic 3. Interpersonal
Introvert	1. Verbal/Linguistic 2. Intrapersonal
Sensing	1. Verbal/Linguistic 2. Intrapersonal
Intuitive	1. Logical-mathematic 2. Kinaesthetic
Extrovert-Sensor	- Experiment
Extrovert-Intuitive	- Exercise
Introvert-Sensor	- Example
Introvert-Intuitive	- Theory

3.0 User Classification Based On Kohonen Network

Kohonen network has been widely used for the classification purposes and it produced an excellent results [Cho 1997; Vendlinski and Stevens 2000]. Basic Kohonen algorithm such as Vector Quantization or k-means clustering can be used as a simple classifier [Sarle 1994].

The Kohonen's self-organizing map (SOM) was introduced by Professor Teuvo Kohonen at the University of Helsinki in 1982. The idea was to create a neural network to represent the input space using the topological structure of a grid to store neighborhood relations. In contrast to most neural network methods that use the desired result to compute the weights of the network, SOMs need no reference output themselves (unsupervised learning).

A SOM defines a mapping of an n-dimensional input space *R* to an m-dimensional output space *C* (we use m=2). The output space consists of N_c neurons. They are embedded in the topological structure of *C*, which may be an m-dimensional grid or any other graph structure. To each neuron of the output space, a parametric weight vector in the input space $w_s = [\mu_{s1}, \mu_{s2}, \ldots, \mu_{sn}]^T \in R$ is associated.

Eq. 6 define the mapping $\varnothing$ from the input space *R* to the topological structure *C*:

$$\varnothing_w : R \rightarrow C\,,$$

$$x \rightarrow \varnothing_w(x)\,, \qquad (1)$$

where,

$\varnothing_w(x)$ is defined as,

$$\varnothing_w(x) = \arg\min_i \{\| x - w_i \|\}\,. \qquad (2)$$

Every input sample is mapped to the neuron of the output layer whose weight vector is closest to the input.

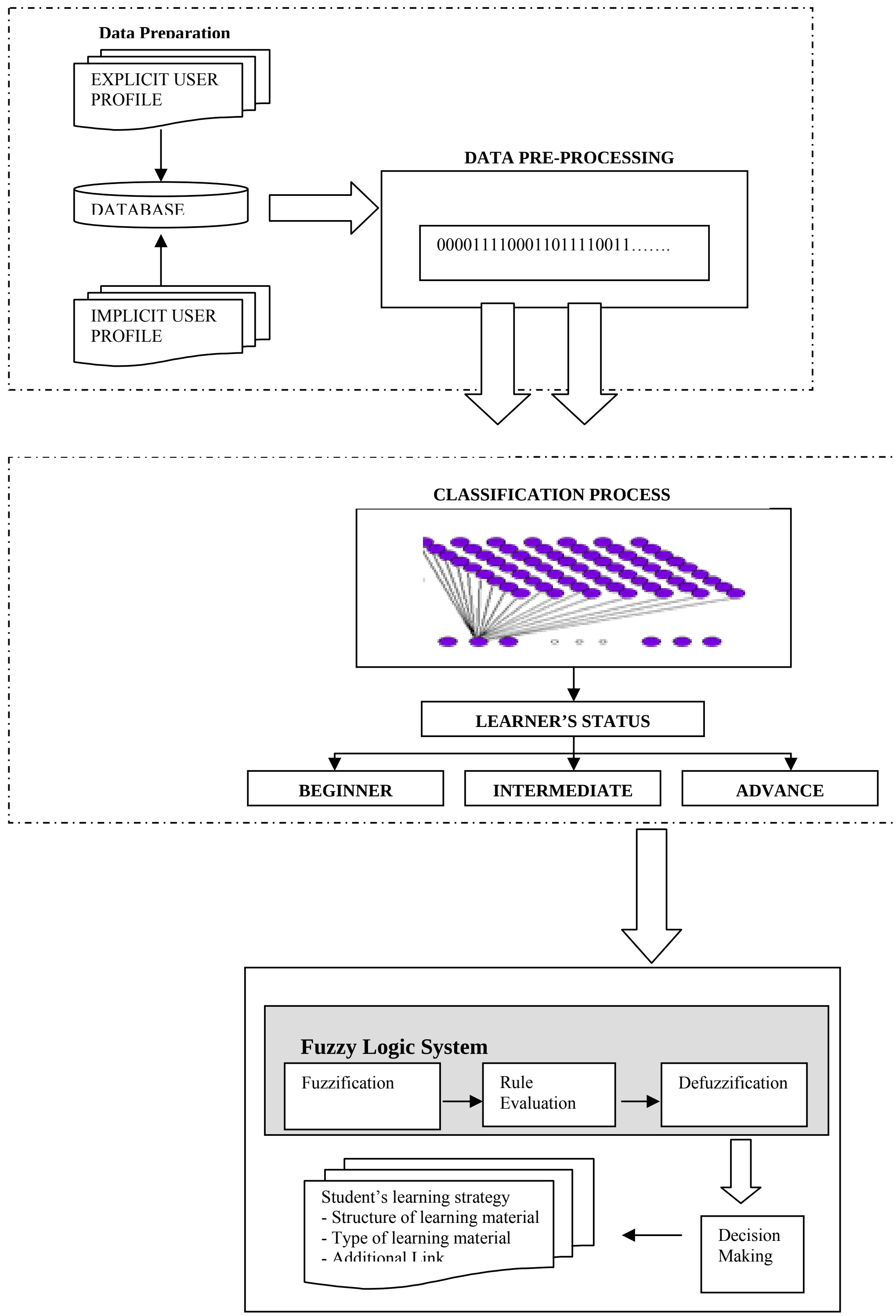


Figure 1: The Architectures of the AHLS.

The goal of Kohonen learning process is to reduce the *distortion error:*

$$d(x) = \sum_{x \in I} \sum_{i \in C} h_{ci} \| x - w_i \| \frac{1}{|I|}, \quad (3)$$

where,

I denotes the set of input samples,

h_{ci} denotes the neighbourhood relation in C between neuron i and the best matching one c.

3.1 Experiment with Self-Organising Maps (Kohonen)

Table 2 shows the attributes and the values defined as simulated data. To develop a simulated data, we use the criteria shown in Table 3. When the learner log into the learning system, the system will count the time he/she spent on learning a concept. The system suggests the time to be spent on each concept. Learning time is calculated based on the percentage the learner takes to finish learning from the suggested time. The learner's learning time is calculated as follows:

Suggested time for n concept = 1200 seconds

Total time spent by learner a = 900 seconds

Percentage = $\frac{900}{1200} \times 100$

t = 75.00.

Table 2: Simulated data

Attribute	Value
Learning time, *t*	0.00 – 100.00 %
Number of backtracking, *b*	0 – 5
Number of getting help function, *h*	0 – 5
Score, *s*	0.00 – 100.00 %

Table 3: Criteria for learner's classification

Attribute	Beginner	Intermediate	Advanced
Learning Time, t	$t > 80\%$	$30\% \le t \le 80\%$	$t < 30\%$
Numb. Of Backtracking, b	$b > 4$	$2 \le b \le 4$	$b < 2$
Numb. Of Using Help, h	$h > 4$	$2 \le h \le 4$	$h < 2$
Score, s	$s < 30\%$	$30\% \le s \le 80\%$	$s > 80\%$

The number of backtracking shows that the learner is not fully mastering the concept, lose direction or change his/her learning goal. The number of backtracking is defined by counting how many times the learner reopen any pages in particular concept. In this research, help function is a list of definition and explanation on terms used in the notes given. This attribute shows that the more help the learner gets, the more he/she is having a difficulty in understanding a concept. The number of getting help is defined by counting how many time the learner click on the help button in particular concept. To test the learner's level of mastering, the system provides an exercise at the end of the learning period. The score is calculated by the percentage of correct answers given. The SOM structure is defined as shown in Table 4.

All the data must be transformed into a standard format to get a valid and accurate classification. The transformation of the data is included in the pre-processing phase using a normalization method. We used a normalization method that was defined by [Rao 1995] as follows:

$$x_n = \frac{1}{\sqrt{\sum (x_n)^2}} \times x_n \tag{4}$$

where, X_n is the input data for n.

Table 4: Parameter settings for the SOM training

Parameter	**Value**
Size	10x10
Dimensionality	2
Shape	Sheet
Map lattice	Rectangular
Neighbourhood	Gaussian
Learning rate	0.5
Iteration	5000
Size of training sample	1050
Size of testing sample	450

In the training phase, input data is given to the Kohonen network. The weights are captured after completing the training phase. The size of training sample is 1050. In the testing phase, there is no target data is provided. We used 450 dataset to the network. The network classifies the data based on the weights and outputs are obtained. When the testing results were obtained, the percentage of the classification accuracy was calculated.

3.2 Result of the Experiment

The training process does not consist of the class of learner. The map shown in Figure 2 is the mapping of the weights produced from the network learning through the data sample given. During the training, the network learns the data and generates the weights by calculating the nearest distance to the real data presented. The number 0, 1 and 2 are the representation of the classes defined during data simulation whereby 0 represents class for beginner, 1 for intermediate and 2 for advance.

```
2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 1
2 2 2 2 2 2 2 2 1 1
0 0 2 1 2 2 2 1 1 1
0 0 0 0 2 0 1 1 1 1
0 0 0 0 0 1 1 1 1 1
0 0 0 0 0 1 1 1 1 1
0 0 0 0 1 1 1 1 1 1
0 0 0 0 1 1 1 1 1 1
0 0 0 0 1 1 1 1 1 1
```

Figure 2: Result Map From Network Training

Table 5 shows the accuracy percentage of network classification in testing process. From 450 data presented, the networks are able to classify 445 data correctly. From the result shown, we conclude that Kohonen network is capable of classifying the learners' data into categories. It gives more than 90% accuracy in both training and testing phase. The Kohonen's SOM is definitely a good tool to classify data into a number of groups without supervision. It will be very useful in this study because it can deal with more complex and bigger sample of data when it is applied to the real learners' data in the learning system's database.

Table 5: Result from network testing

Numb. Of Correct Classification	Numb. Of testing data	Accuracy
445	450	98.89 %

4.0 Fuzzy Approach

In the second phase, fuzzy techniques are used to personalize the learning path where it denote to 3 output, structures of learning material, type of learning material and additional link. The first output is to calculate precisely the structures of learning material. The learning materials are structured in the form of theory, example, exercise and experiment that suit student's personality. Next stage is the calculation of the type of learning material that suit to the students whether the student prefer more visual or linguistic learning material. The last output is the additional link that link to e-mail and forum link. Student need to answer questionnaires before start learning process. This process is the same as the traditional process. The differences are in terms of the techniques in personalizing the student learning material. The traditional techniques totally use the result from the questionnaires to personalize the learning material. The highest result from the student's questionnaires will be considered as the student's personality and the learning method is referring to the personality. Meanwhile, the fuzzy system will calculate the student's personality and propose the suitable learning method to the student. The problem that occurs in traditional techniques is producing the suitable learning method once the students have equal result in their personality.

Fuzzy logic is computationally undemanding and is most suitable for processing imprecise input data, as it supports natural description of knowledge and reasoning in the form of imprecise concepts, operators and rules [Negnevitsky 2002]. In ITS, fuzzy logic techniques have been used due to their ability to handle imprecise information, such as student's actions, and to provide human descriptions of knowledge and of student's cognitive abilities [Stathacopoulou et al. 1999]. The fuzzy logic system consists of three main stages: Fuzzification, Rule Evaluation and Defuzzification. The fuzzification stage transforms crisp student's personality data, captured in the student database, into suitable linguistic values.

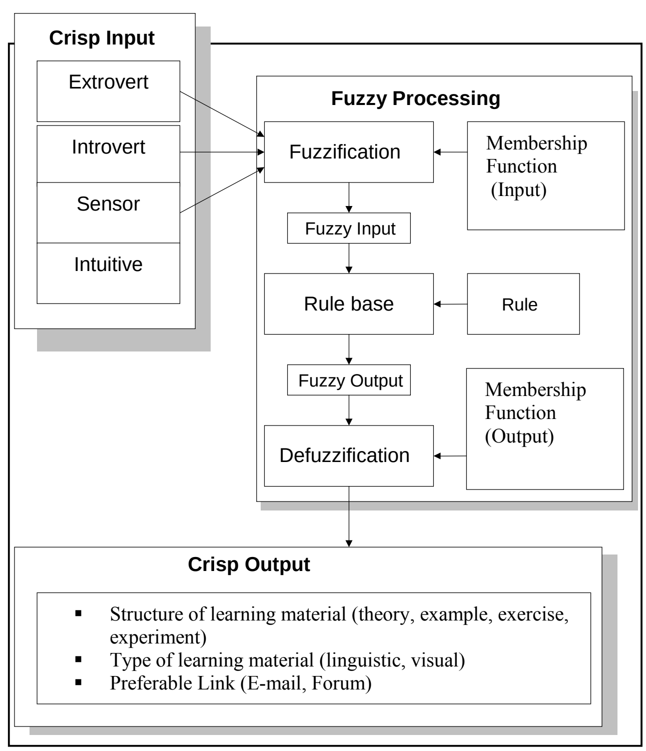


Figure 3. Architecture of fuzzy system

In this paper, the linguistic variables are based on the MBTI where it provides 4 linguistic variables; (Extrovert, introvert, sensor and linguistic). Whilst the fuzzy set that use for this paper is triangular fuzzy set where it hold 3 parameter {a, b, c} and can be seen in table 3. The rule evaluation stage takes the fuzzified inputs, and applies them to the antecedents of the fuzzy rules. The defuzzification stage produces the output in crisp value. In this problem, the defuzzified value is derived based on the maxima and sum method in Mamdani-style inference. The suitable defuzzified will be choose appropriately based on the result.

The triangular membership function can be specified by three parameters {*a, b, c*} as shown in figure 4: In figure 5, *a*, *b*, and *c* are parameters that control the intervals of the triangle, x_i is the input, and y_i is the output of fuzzification input *i*.

$$y_i = \begin{cases} \dfrac{x_i - a}{b - a} & \text{if } a <= x_i <= b \\ 1, & \text{if } x_i = b \\ 1 + \dfrac{b - x_i}{c - b} & \text{if } b < x_i <= c \\ 0, & \textit{otherwise} \end{cases}$$

Figure 4. Triangular membership function

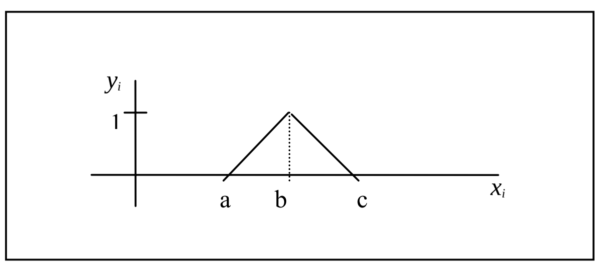


Figure 5. Triangular fuzzy set

The next process is the process to generate the fuzzy rule-base. This process classified the student learning material according to the student personality. The fuzzy rule based system is consisting of IF/THEN statement and combine with AND/OR operation as the example in figure 6.

First Rules:	IF (a is A1 AND (b is B3) AND (c is C1) AND (d is D3) THEN (e is E2), (f is F5), (g is G4), (h is H2), (i is I1), (j is J1) and (k is K3)

Figure 6: Fuzzy Rules Evaluation

In this figure a, b, c and d each of it represent the Student's Personality (extrovert, introvert, sensor and intuitive. Whereas (A1,A2,A3), (B1,B2,B3), (C1,C2,C3), (E1,E2,E3), (J1,J2,J3) and (K1,K2,K3) represent the input /output {High, Medium, Low}, and (F1,F2,F3,F4,F5), (G1,G2,G3,G4,G5), (H1,H2, H3, H4,H5) and (I1,I2,I3,I4,I5) represent the output. Table 6 illustrates more examples of the rules.

Table 6: Example of fuzzy rules

Input				**Output**						
Extrovert	Introvert	Sensor	Intuitive	Inter **Personal**	Theory	Example	Exercise	Experiment	Visual	Linguistic
Low	High	Low	High	Low	VHigh	High	Low	VLow	Low	High
Low	High	Med	Med	Low	High	Med	Low	Vlow	Med	High
Low	High	High	Low	Low	Med	VHigh	VLow	Low	Med	High
:	:	:	:	:	:	:	:	:	:	;
High	Low	High	Low	High	VLow	Low	High	VHigh	High	Low

The defuzzification process used in this paper is Mamdani inference-style where it involves difference operation of defuzzification. The best defuzzification operation will be selected. The criterion of selecting the best defuzzification is base to the most similar result that fulfils the expert expectation. Based on the result, the defuzzifications Centre of Area/Gravity (COG) have the most similarity. For example, see table 7 below.

5.0 Conclusion And Further Work

This paper has proposed a way to personalize the course content for AHLS, Which aims to provide learners with a customized learning environment. It emphasizes the combination of pedagogical theories and artificial intelligent techniques. It is important to note that for a given dataset and defined SOM properties, the SOM training process is dependent on the learning parameter settings.

Table 7: Comparison of Mamdani defuzzification operation with aggregation – max
(input introvert =0.7, extrovert =0.3, sensor = 0.6, intuitive = 0.4)

	Exercise	Example	Theory	Experiment	Inter Personal
Mamdani Defuzzification (aggregation max)					
COA/G	0.317	0.683	0.446	0.554	0.354
Bisector	0.29	0.71	0.39	0.6	0.29
MOM	0.245	0.75	0.245	0.75	0.13
LOM	0.33	0.84	0.33	0.84	0.26
SOM	0.16	0.66	0.16	0.66	0

Further research is required to identify the most suitable parameter setting for real learners' data. Our future work will seek to apply different types of network structure such as mapping topology and lattice, and restructure the neighbourhood radius formulation to improve the Kohonen network.

In particular, for adapting the MBTI theories in AHLS, a specific pedagogical model must be prescribe. In this paper, we outline the pedagogical framework containing the approach, method and techniques that suit for AHLS. This first stage are use to describe how content is reflect to the MBTI personality in online system. Fuzzy logic techniques are then used to impart the learning content based on student's fuzzy personality data and instructional rules in order to support customisation that will allow learners to learn faster and understand the learning material much easier. Conversely, the challenge is to identify what those learning content (structure and type of learning material) for a given learner in online system based on student personality and which result is more precise to the learner's personality whether the traditional approach or using the fuzzy logic techniques. Fuzzy logic model provides an efficient way to reason the student's learning method based on the student's personality. For further research, the fuzzy logic model may need to hybrid with genetic algorithm for tuning the membership function and the scaling function for fuzzy input and output that result to better fuzzy logic techniques.

Acknowledgement

This research is conducted under IRPA grant, project vote 74129: The Development of Knowledge Management for AHLS.

6.0 References

BISHOP-CLARK, C. ET AL, 1994. The Myers-Briggs Personality Type and Its Relationship to Computer Programming. J*ournal of Research on Computing in Education* 26, 3, 358-371.

BRUSILOVSKY, P. 2001. Adaptive Hypermedia. User Modeling and User-Adapted Instruction, *In 2001 Kluwer academic Publishers,* 11, 1 – 2, 87-110.

CAROLYN, S. ET AL, 2001. Myers Briggs Type Preferences in Distance Learning Education. *In International Journal of Educational Technology* 2, 2.

CHO, S.B. 1997. Self-organizing Map with Dynamical Node-splitting: Application to Handwitten Digit Recognition. In *Neural Computation, 9*, 1345-1355.

DARA-ABRAMS, B. P. 2002. Applying Multi-Intelligent Adaptive Hypermedia to Online Learning. *In Proceedings of E-Learn2002, Montreal, Canada*.

GARDNER, B.S. AND KORTH, S. J., 2001. Using Psychologist Type Theory to Maximize Student Learning. *26th International Conference Learner-Centered Universities for the New Millennium.* Rand Afrikaas University, Johannesburg, South Africa.

HASHIM. S AND YAAKUB. R. 2003. *Terbitan Pertama Psikologi Pembelajaran dan Personaliti*. PTS Publications & Distributors Sdn. Bhd (Pelita).

NEGNEVITSKY, M. 2002. *Artificial Intelligence. A Guide to Intelligent Systems*, First Edition, Pearson Education.

RAO, V.B. 1995. *C++ Neural Networks and Fuzzy Logic*. MTBooks, IDG Books Worldwide.

SARLE, W.S. 1994. Neural Networks and Statistical Models. In *Proceedings of the Nineteenth Annual SAS Users Group Internationanl Conference*. Cary, NC: SAS Institute, 1538-1550.

SCHROEDER, C.C. 1993. *New-Students-New Learning Style.* 25(5): 21-26.

STATHACOPOULOU R., MOGOULAS G. D, GRIGORIADOU M. 1999, Neural Network-based fuzzy Modeling for the student in Intelligent Tutoring Systems, *Proceedings of the International Joint Conference on Neural Network, IJCNN'99*, 3517-3521.

VENDLINSKI, T. AND STEVENS, R. 2000. The Use of Artificial Neural Nets (ANN) to Help Evaluate Student Problem Solving Strategies In: *B. Fishman & S. O'Connor-Divelbiss (Eds.): Fourth International Conference of the Learning Science*. 108-114.

ZADEH, L. A., 1992. *Fuzzy Logic for Management of Uncertainty*. John Wiley & Sons, New York.

PAPER III:

INDIVIDUALIZING LEARNING MATERIAL OF ADAPTIVE HYPERMEDIA LEARNING SYSTEM BASED ON PERSONALITY FACTOR (MBTI) USING FUZZY LOGIC TECHNIQUES

Norreen Binti Haron

Faculty of Computer Science and Information System

Universiti Teknologi Malaysia

norreen.haron@gmail.com

Naomie Binti Salim

Faculty of Computer Science and Information System

Universiti Teknologi Malaysia

naomie@fsksm.utm.my

Abstract

The inflexible linking provided in conventional hypermedia learning system has some drawbacks that can cause teaching and learning to be less effective. Research on learning has shown that student learn differently since student process knowledge in different ways, with some students learning more effectively when taught with methods that suits their learning style. One solution to this problem is to develop an adaptive hypermedia learning system which basically incorporates intelligence and knowledge about the individual user learning style to assist learner to achieve learning objectives. Information about learning style can help system become more sensitive to the differences of students using the system. Domain modeling is also an important task in the development of an adaptive hypermedia learning system, as much semantic of the domain and support for the adaptive navigation have to be catered for and incorporated in the model. In this paper, a framework for individualizing the learning material structure in adaptive learning system is introduced. It aims to utilize the learning characteristics and provide a personalized learning environment that exploit pedagogical model and fuzzy logic techniques. The

learning material consists of 4 structures; 1) theory, 2) example, 3) exercise and 4) activities. The pedagogical model and learning characteristics are based on the student's personality factor (Myers-Briggs Type Indicator (MBTI)), whilst the fuzzy logic techniques are used to classify the structure of learning material which is based on student's personality factors. This paper focuses on the use of fuzzy logic techniques for adaptation of the content to the user, allowing a learning system to dynamically adapt the choice of possible learning structure through the learning material based on the user's personality factor, with the hope to provide an adaptive hypermedia learning system that is user-customized to support faster and more effective learning.

Keywords

Adaptive Hypermedia system, Pedagogical Framework, Personality Factors (MBTI), Learning Styles, Fuzzy Logic.

1.0 Introduction

The adaptive hypermedia learning system (AHLS) is a computer based learning system in which interactive and dynamic learning module is customized to each student. Research on learning has shown that student learn differently and process knowledge in different ways. Information about learning style can help system become more sensitive to the differences of students that use the system. Several systems adapting different learning styles have been developed to date. However, it is not clear which aspects of learning characteristics are worth modeling, how the modeling can take place and what can be done differently for users with different learning style [1]. These problems may lead to students facing difficulties to understand what is being taught, decrease of students' interest to continue their study in the subject time taken to finish a particular lesson session [2].

Currently, the adaptation of student's learning style to learning is totally based on the dominant student learning style, where the dominant result is mainly stated as one particular student's preferable learning material, ignoring other learning styles that a student may also posses. In reality, a student's learning style can be of mixed traits, each with a certain percentage of membership to the student's overall style.

This paper tends to model the fuzziness in student's learning style and the appropriate learning material method suitable for student's fuzzy learning styles membership.

A framework for learning path personalization in adaptive learning system is introduced. It aims to utilize the learning characteristics and provide a personalized learning environment, that exploit pedagogical model and fuzzy logic techniques as shown in Figure 1 below. The pedagogical model and learning style refer to student's personality factor based on the Myers-Briggs Type Indicator (MBTI) [3, 4]. Based on the MBTI theory, fuzzy logic techniques are then used to classify learning material (structure of learning material).

Fuzzy set theory, proposed by Zadeh [5], has proved to be very effective in handling the vagueness and uncertainty intrinsically existing in the knowledge possessed by people or implied in numerical data. Rules based on fuzzy logic provide a qualitative description of the input-output relationship of a system using fuzzy linguistic terms. Fuzzy linguistic rule is closer to human reasoning and in many real-world applications, and thus it is more adequate and flexible for knowledge representation than the conventional crisp IF THEN rules which is the major reason for its adaptation in this research.

As shown in Figure 1, the architecture for the learning strategy is based on the MBTI personality factor whilst the learning method and techniques is based on the Honey & Mumford theory. The four MBTI personality types used in this research are; Extrovert (E); Introvert (I); Sensing (S) and Intuition (N). Extrovert and introvert represent the student's preferable condition in focusing attention. Sensing and intuition, illustrate the student's preferable way in taking information.

Extrovert students prefer and focus on the outer activities and are energized by interaction with others. They prefer to talk, participate and interact with people. The introvert students prefer and focus on the inner activities such as reading, listening to others and writing. Sensing type of students prefer concrete information, facts and procedures. They are good in memorization and like to go systematically, they also learn best with instruction that allow them to use their senses. Intuition student prefer discovering possibilities and relationship. They also like courses that involve a lot of experimentation and experiences [6, 7].

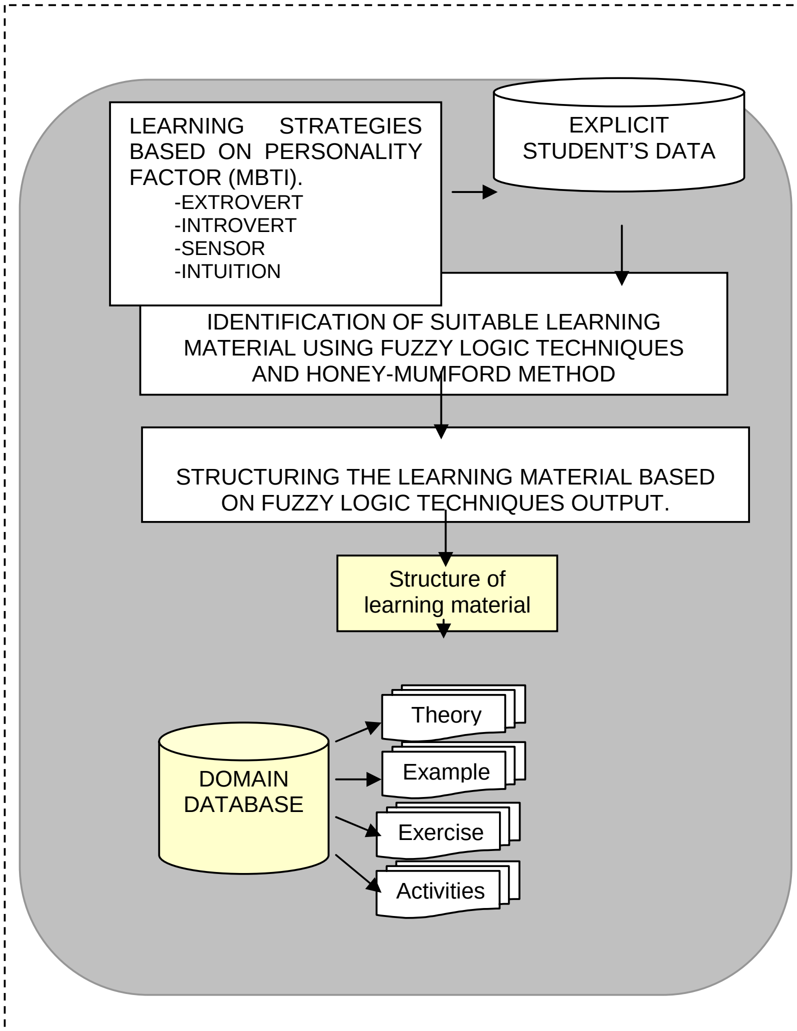


Figure 1: The Framework Of Fuzzy Logic Approach

2.0 Approach And Methods

The pedagogical framework consists of steps on how content is develop to reflect those personality principles. Table 1 below, shows the relationship between learning style and learning method.

Table 1: The Relationship Between Learning Style And Method.

Learning Style (MBTI)	Method (Honey & Mumford)
Extrovert-Sensor	Activity
Extrovert-Intuitive	Exercise
Introvert-Sensor	Example
Introvert-Intuitive	Theory

Fuzzy techniques are used to personalize the learning path where it has 4 outputs. The outputs indicate the structures of learning material (theory, example, exercise and activity) that suits student's personality, taking into account the most preferred learning material and the least preferred learning material.

Fuzzy logic is computationally undemanding and is most suitable for processing imprecise input data, as it supports natural description of knowledge and reasoning in the form of imprecise concepts, operators and rules [8]. In AHLS, fuzzy logic techniques have been used due to their ability to handle imprecise information, such as student's knowledge and their cognitive abilities [9]. Table 2 shows several examples of AHLS adapting intelligent techniques.

Table 2: Examples Of Previous Ahls Adapting Intelligent Techniques In User Modelling.

System	**Intelligent Techniques**	**Predict User Modeling**
KBS-Hyperbook [10]	Bayesian Network	User Knowledge Level
ALICE [11]	Fuzzy Logic Techniques	User Knowledge Level
iWeaver [12]	Bayesian Network	User Media Presentation
INSPIRE [13]	Neuro Fuzzy Techniques	User Knowledge Level

The processes implemented in fuzzy logic systems indicate 4 main stages; fuzzification, rule evaluation, aggregation and defuzzification. In this research, the fuzzification stage transforms crisp student personality, captured from pre-course questionnaires, into suitable linguistic values. The rule evaluation stage takes the fuzzified inputs, and applies them to the antecedents of the fuzzy rules. The aggregation stage combine all the fuzzy output derived from fuzzy rules, and the final stage, the defuzzification stage, produces the output in crisp value. In this problem, the defuzzified value is derived based on the maxima aggregation method in Mamdani-style inference. The suitable defuzzified value will be choosing appropriately based on the result. Figure 2 below shows the flow of the fuzzy logic system.

3.0 The Fuzzy Model

In this study, the fuzzy sets and fuzzy rules are constructed based on the standard of mastery or the criterion-referenced acquired from the human instructors' experience and knowledge about their students. In this problem, there are four input and four output linguistic variables being considered. The input linguistic variables represent the student's personality's scores: extrovert scores, introvert scores, sensor scores and intuitive scores. The fuzzy inputs are expressed by:

x_1, x_2, x_3 and x_4

Whilst, the outputs represents the student's acceptance level of learning material; theory, example, exercise and activities are expressed by:

y_a, y_b, y_c and y_d

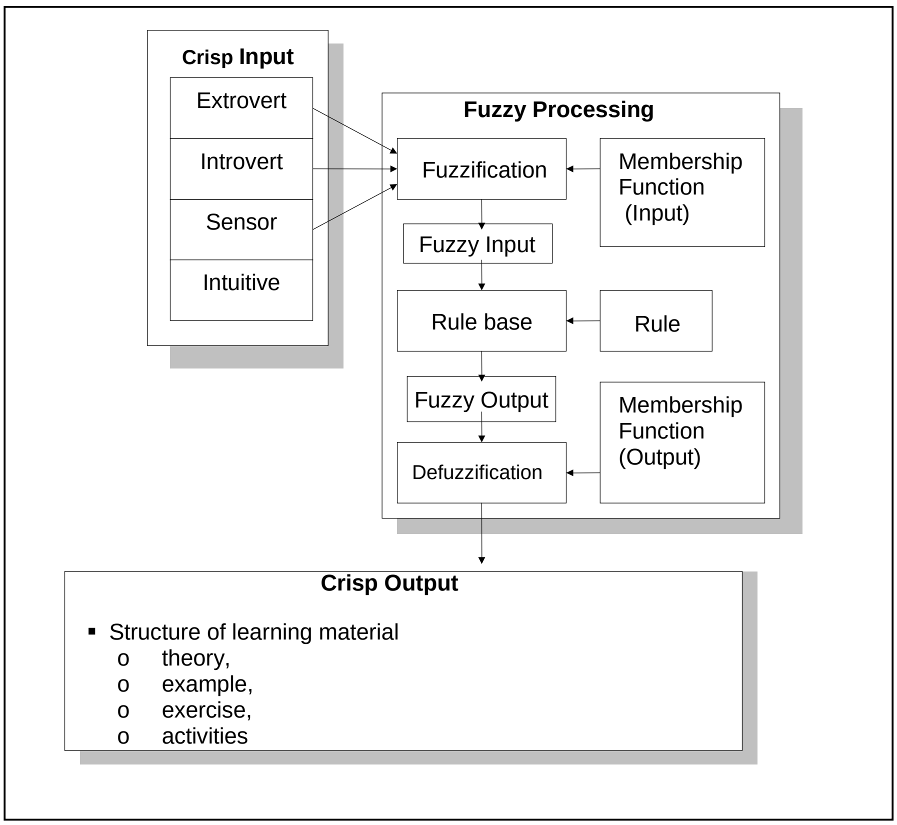


Figure 2: Flow Of Fuzzy System

In this paper, the fuzzy set is expressed by a triangular function as triangular function. This is chosen based on expert advises, interview session and survey session for the case at hand and at the same time significantly simplifies the process of computation. The fuzzy set is expressed by three parameters {a, b, c} as shown in figure 3 and figure 4 below:

$$y_i = \begin{cases} \dfrac{x_i - a}{b-a} & \text{if } a <= x_i <= b \\ 1, & \text{if } x_i = b \\ 1 + \dfrac{b-x_i}{c-b} & \text{if } b < x_i <= c \\ 0, & \textit{otherwise} \end{cases}$$

Figure 3: Triangular Membership Function

where a, b, and c are parameters that control the intervals of the triangle (as shown in figure 3 and figure 4), x_i is the input, and y_i is the output of fuzzification input i.

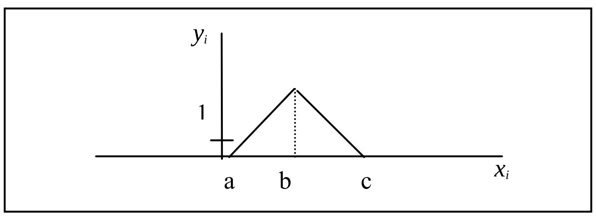
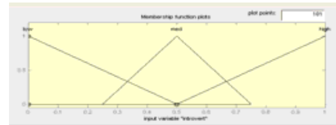
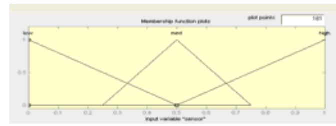
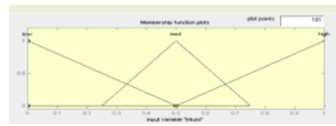
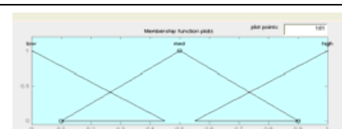
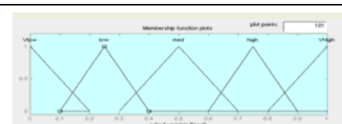
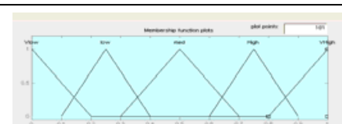
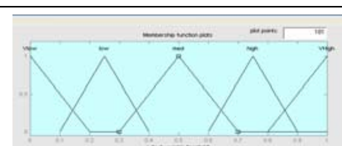


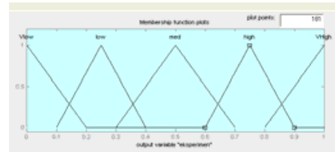
Figure 4: Triangular Fuzzy Set

Table 3 below show the fuzzification process that include the linguistic value, notation, numerical ranges and fuzzy set for the input and output of this study.

Table 3: Fuzzification Process

Linguistic Value	**Input / Output**	**Notation**	**Numerical Ranges (normalized) *Based on expert advise**	**Membership Function *Based on expert advise**
Extrovert	Input	Low Med High	[0, 0, 0.5] [0.25 0.5 0.75] [0.5 1 1]	

Introvert	Input	Low Med High	[0, 0, 0.5] [0.25 0.5 0.75] [0.5 1 1]	
Sensor	Input	Low Med High	[[0, 0, 0.5] [0.25 0.5 0.75] [0.5 1 1]	
Intuitive	Input	Low Med High	[0, 0, 0.5] [0.25 0.5 0.75] [0.5 1 1]	
Visual	Output	Low Med High	[0 0 0.45] [0.1 0.5 0.9] [0.55 1 1]	
Linguistic	Output	Low Med High	[0 0 0.45] [0.1 0.5 0.9] [0.55 1 1]	
Theory	Output	VLow Low Med High VHigh	[0.0, 0.0, 0.25] [0.2, 0.32, 0.45] [0.4, 0.5, 0.6] [0.55, 0.7, 0.85] [0.8, 1.0, 1.0]	
Excercise	Output	VLow Low Med High VHigh	[0.0, 0.0, 0.25] [0.2, 0.32, 0.45] [0.4, 0.5, 0.6] [0.55, 0.7, 0.85] [0.8, 1.0, 1.0]	
Example	Output	VLow Low Med High VHigh	[0.0, 0.0, 0.25] [0.2, 0.32, 0.45] [0.4, 0.5, 0.6] [0.55, 0.7, 0.85] [0.8, 1.0, 1.0]	

Activities	Output	VLow Low Med High VHigh	[0.0, 0.0, 0.25] [0.2, 0.32, 0.45] [0.4, 0.5, 0.6] [0.55, 0.7, 0.85] [0.8, 1.0, 1.0]	

The next stage is the process to generate the fuzzy rule-base. This process classifies the student's learning material according to the student's personality. The fuzzy rule based system consists of IF/THEN statement, multiple antecedents, multiple consequents and combined with AND/OR operation as expressed below:

R_i :

IF X_a is $N1$ AND X_b is N_2 AND X_c is N_3 AND X_d is N_3

THEN

Y_a is M_1 AND Y_b is M_2 AND Y_c is M_3 AND Y_d is M_4

where R_i, (i = 1, 2, … n) is the rule number, N_i, (i =1, 2 and 3) are the membership functions of the antecedent part, M_i, (i =1, 2, 3, 4 and 5) are the membership functions of the consequent part. The example of fuzzy rules applied in this problem as show in table 4.

Table 4: Example Of Fuzzy Rules

Input				**Output**					
Extrovert	**Introvert**	**Sensor**	**Intuitive**	**Theory**	**Example**	**Exercise**	**Activities**	**Visual**	**Linguistic**
Low	High	Low	High	VHigh	High	Low	VLow	Low	High
Low	High	Med	Med	High	Med	Low	Vlow	Med	High
Low	High	High	Low	Med	VHigh	VLow	Low	Med	High
:	:	:	:	:	:	:	:	:	;
High	Low	High	Low	VLow	Low	High	VHigh	High	Low

The defuzzification process used in this paper is Mamdani inference-style where it involves different operations of defuzzification. The best defuzzification operation will be selected. The criterion of selecting the best defuzzification is based on the most similar result that fulfils the expert expectation. Based on the result, the defuzzifications Centre of Area/Gravity (COG) have the most similarity.

4.0 Conclusion And Further Work

This paper has proposed a way to personalize the course content for adaptive hypermedia learning system, which aims to provide learners with a customized learning environment. It emphasizes on the combination of pedagogical theories and artificial intelligent techniques. In particular, for adapting the MBTI theories in adaptive hypermedia learning system, a specific pedagogical model must be prescribed. In this paper, we outlined the pedagogical framework containing the method and techniques suitable for adaptive hypermedia learning system. Fuzzy logic techniques are used to impart the learning content based on student's fuzzy personality data and instructional rules in order to support customization that will allow learners to learn faster and understand the learning material easier. The learning content (structure and type of learning material) for a given learner in online system based on student personality is identified. Fuzzy logic model provides an efficient way to reason the student's learning method based on the student's personality. We are in process of testing the fuzzy logic model by the students compared to crisp conventional method, based on two assessments; performance effectiveness and the accuracy of the system.

Acknowledgement
This research is conducted under IRPA grant, project vote 74129: The Development of Knowledge Management for Adaptive Hypermedia Learning System.

References

1. Brusilovsky, P. (2001). *Adaptive Hypermedia. User Modeling and User-Adapted Instruction.* Kluwer academic Publishers, Vol 11, No. 1 – 2, pp 87- 110.

2. Hashim, S, and Yaakub R. (2003). *Terbitan Pertama Psikologi Pembelajaran dan Personaliti*. PTS Publications & Distributors Sdn. Bhd (Pelita).
3. Carolyn, S. et. al. (2001). Myers Briggs Type Preferences in Distance Learning Education. *In International Journal of Educational Technology,* 2(2).
4. Bishop, C.C. et. al. (1994). The Myers-Briggs Personality Type and Its Relationship to Computer Programming, *Journal of Research on Computing in Education,* 26(3), pp. 358-371.
5. Zadeh, L. A. (1992). *Fuzzy Logic for Management of Uncertainty*. John Wiley & Sons, New York.
6. Gardner, B.S. and Korth, S. J. (2001). Using Psychologist Type Theory to Maximize Student Learning. *26th International Conference Learner-CenteredUniversities For the New Millennium.* Rand Afrikaas University, Johannesburg, South Africa.
7. Schroeder, C.C. (1993). New-Students-New Learning Style, 25(5), pp. 21-26.
8. Negnevitsky, M. (2002). *Artificial Intelligence. A Guide to Intelligent Systems*, First Edition, Pearson Education.
9. Stathacopoulou R., Mogoulas G. D, Grigoriadou M. (1999). Neural Network-based fuzzy modeling of the student in Intelligent Tutoring Systems, *In Proceedings of the International Joint Conference on Neural Network, IJCNN'99*, July 10-16, Washington DC, USA.
10. Henze. N., and Nejdl, W., (1999). Adaptivity in the KBS Hyperbook System, *In Proceedings of the 2nd Workshop on Adaptive Systems and User Modeling on the WWW*, May 11, Toronto.
11. Kavcic, A. (2001). Dealing with Uncertainty of User Knowledge: Fuzzy User Modelling in Educational Hypermedia, *Advances in Multimedia Communications, Information Processing and Education*, Invited Lecture, October, Leganés, Spain.
12. Wolf, C. (2002). iWeaver: towards an interactive web-based adaptive learning environment to address individual learning styles, online European Journal of Open and Distance Learning, (see www.eurodl.org). Retrieved December 17, 2003,
web site: http://www.eurodl.org/materials/contrib/2002/2HTML/iWeaver

13. Kyparisia, A., et al. (2003). Personalizing the Interaction in a Web-based Educational Hypermedia System: the case of INSPIRE, *Journal ACM, User Modeling and User-Adapted Interaction.* 13(3), pp. 213-267.

PAPER IV:

ROUGH SET GENERATION FOR IDENTIFYING STATUS OF STUDENT'S KNOWLEDGE ACQUISITION

Nor Bahiah Hj Ahmad, Siti Mariyam Hj. Shamsuddin & Siti Zaiton Mohd Hashim
Soft Computing Research Group,
Faculty of Computer Science and Information System,
Universiti Teknologi Malaysia, Malaysia.
bahiah@fsksm.utm.my, mariyam@fsksm.utm.my, szaiton@fsksm.utm.my

Abstract

In this study, a generation of rough set rules are implemented in identifying the status of students knowledge acquisition in hypermedia learning. Along with this concept is the process of subdivision of the universal set of all possible categories of identifying the status into a number of distinguishable categories called elementary sets. Four attributes, cpa, learning time, number of backtracking, and scores obtained by the student have been collected in the student model and the decision of the students knowledge acquisition whether as good, average or poor are transformed into the information table. In this study, each object represents students's performance. The decision is the students's knowledge acquisition status, identified by adopting expert advice as poor, average and good. They are represented as Class 1, Class 2 and Class 3 respectively. We apply the rules to 76 unseen objects and the result shows that the total percentage of correct classification is 93.4211%.

Keywords: Rough set, rule generations, classification, hypermedia learning, knowledge acquisition.

1. Introduction

More and more hypermedia learning system have been developed to aid teachers and students in teaching and learning. However hypermedia system has some drawbacks that can cause teaching and learning to be less effective. First, the link structures between the nodes are static and cannot adapt to the student's background, understanding and desires. Secondly, when the system has thousands of non-linear links, the students surfing the information can get lost and confused of where they are, what actually they are searching for and where to go next. To address this problem, we are developing an adaptive learning system with hypermedia technology. Adaptivity is incorporated by presenting suitable learning materials, guidance and aid during learning session according to different categories of learners based on their majoring background and knowledge acquisition status [1].

Rough set theory, introduced by Zdzislaw Pawlak in the early 1980's is a new mathematical tool to deal with vagueness and uncertainty. The methodology is concerned with the classificatory analysis of imprecise, uncertain or incomplete information or knowledge or knowledge expressed in terms of data acquired from experience. Here, objects are perceived through the information that is available about them through the values for a predetermined set of attribute. The main advantage of rough set is that it requires no additional information to the data represented in table. It does not need any preliminary or additional information about data, such as probability distribution in statistic, grade of membership or the value of possibility in fuzzy set theory [2].

Objects characterized by the same information are indiscernible in view of the available information about them. The similarity/indiscernible relation generated in this way is the mathematical basis of the rough set theory. Any set of all indiscernible objects is called elementary set and form the basic granule of knowledge about the universe. Any union of some elementary sets is referred to as crisp set, otherwise a set is rough (imprecise, vague). Vague concept is characterized by the lower and upper approximation. The lower approximation consists of all objects which definitely belong to the set (members of the set), and upper approximation contain all objects which possibly belong to the set. The difference between the upper and lower approximation constitute the boundary region of the set.

In rough sets the issue of knowledge can be viewed as partition or classification. The knowledge representation system can be perceived as a data table, columns which are labeled as attributes, and rows are labeled as objects. Each row represents a piece of information about corresponding objects. The data table can be obtained as a results of measurements, observations or represents knowledge of an agent or a group of agents. Thus in the study, we present the classification of student's knowledge acquisition status using rough set.

2. Basic ideas of Rough Set and implementation

Rough Set theory offers some important techniques in managing an information system (IS), and consists of several steps leading towards the final goal of generating rules from information/decision systems. The main steps of the rough set approach described by [3] are:

- ❖ The mapping of information from the original database into the decision system format (Information System Table)
- ❖ Completion of data
- ❖ Discretisation of data
- ❖ Computation of reducts from data
- ❖ Derivation of rules from reducts
- ❖ Filtering of rules

2.1 Mapping of information and knowledge representation

The first step, mapping of information into a decision system, depends on the format of the original formation. In some cases, the data set is already on a format ready to be imported into a decision table. In other cases, attributes (including the decision attribute) and object must be identified before data can be placed in the decision table. For some types of data, such as time series, additional mapping of the data is necessary. In rough sets the issue of knowledge representation can be viewed as partition (classification). The knowledge representation system can be perceived as a data table, columns which are labeled as attributes, and rows are labeled as objects. Each row represents a piece of information about corresponding objects. The data table can be obtained as a results of measurements, observations or represents knowledge of an agent or a group of agents. Representation of knowledge in tabular form has great advantage especially for its clarity. The data table may be perceived as a set of propositions about reality, and can be viewed as a model for special logic, called decision logic. Rough sets offers its normal form representation of formulas and the second employing the concept of *indiscernibility* to investigate whether some formulas are true or not. Indiscernibility leads to simple algorithms for data reduction and analysis.

In this study, the information table represents input data gathered from any domain in students's model. This information table describes students's performance, and the decisions are represented as status of knowledge acquisition. Each student is characterized by the CPA earned, duration of learning period, number of times doing back tracking, and finally scores obtained from exercises and tests [4]. The decision for this information table is classified by the experts as good, average or poor. Table 1 shows some of the informations with related decisions for the domain described above.

Table 1. Information table with decision

No	CPA	Learnin g Time	# Back tracking	Score	D
1	Good	Good	Good	Poor	Good
2	Avg	Good	Avg	Good	Good
3	Good	Good	Good	Poor	Good
4	Poor	Good	Poor	Good	Poor
5	Good	Good	Good	Poor	Good
6	Poor	Good	Poor	Good	Poor
7	Poor	Good	Poor	Good	Poor
8	Good	Poor	Poor	Good	Avg
9	Good	Poor	Poor	Good	Avg

2.2 Indiscernibility relation

The definition of discernibility is given to explain the indiscernibility relation. The basic definition of discernibility is given as input an attribute and two attribute values, and returns true if it is possible for the two values to be different. The standard rough set theory uses ordinary definition of inequality:

$$discerns(a, a(x_i), a(x_j)) = (a(x_i) \neq a(x_j)) \qquad (1)$$

From Table 1, we consider attributes {CPA, Learning Time, #Backtracking, Scores } and generate classes of indiscernibility objects for the selected attributes.

E1 = {1,3,5}

E2 = {2}

E3 = {4,6,7}

E4 ={8,9}.

E1 to E4 classes contain objects of the same condition or attribute values i.e., set of objects that are indiscernible by attributes described above. Table 2 shows the indiscernibility relations obtained from 100 cases.

Table 2. Indiscernibility relation of students knowledge of acquisition

Class	**CPA**	**Learning Time**	**# Back tracking**	**Score**	**D**
E1 (50x)	Good	Good	Good	Poor	Good
E2 (5x)	Avg	Good	Avg	Good	Good
E3 (30x)	Poor	Good	Poor	Good	Poor
E4 (15x)	Good	Poor	Poor	Good	Avg

For futher explanation, let us transform the data in Table 2 in simpler numerical representation as shown in Table 3.

a – CPA (1:poor, 2:average, 3:good)

b – Learning Time (1:poor, 2:average, 3:good)

c - #Backtracking (1:poor, 2:average, 3:good)

d - Score (1:poor, 2:average, 3:good)

Table 3. Simplified representation of students knowledge of acquisition

Class	**a**	**b**	**c**	**d**	**# Object**
E1	3	3	3	1	50x
E2	2	3	2	3	5x
E3	1	3	1	3	30x
E4	3	1	1	3	15x

2.3 Discernibility matrix

A *discernibility matrix* can be created by using the discerns predicate in Equation (1) Given an IS $\mathrm{A} = (U, A)$ and $B \subseteq A$, the *discernibility matrix* of A is M_B, where each entry $m_B(i, \ j)$ consists of the attribute set that discerns between objects x_i and x_j where $1 < i, j < n = |U / IND(B)|$. Table 4 shows the discernibility matrix of Table 3.

$m_B\,(i, j) = \{a \in B\text{: } discerns(a, a(\mathrm{x}_i), a(\mathrm{x}_j))\}$ (2)

Table 4. The discernibility matrix

	E1	**E2**	**E3**	**E4**	*f*
E1	x	acd	acd	bcd	c v d
E2	acd	x	ac	abc	a v c
E3	acd	ac	x	ab	a
E4	bcd	abc	ab	x	b

2.4 Completion

Completion of data is a preprocessing step that is used for decision tables with missing values. One way of making a data set complete is to simply remove objects which have missing values. Another way of dealing with the problem is to fill out missing values using a mean value computed from other objects in the decision table.

2.5 Discretisation

Discretization refers to the process of arranging the attribute values into groups of similar values. Discretization of real value attributes is an important task in data mining, particularly the classification problem. Empirical results are showing that the quality of classification methods depends on the discrimination algorithm used in preprocessing step. In general, discrimination is a process of searching for partition of attribute domains into intervals and unifying the values over each interval. Discretization involves searching for “cuts’ that determine intervals. All values that lie within each interval are mapped to the same value, in effect converting numerical attributes that can be treated as being symbolic. The search for cuts is performed on the internal integer representation of the input Decision System. In this study, we discretise the data using Boolean reasoning approach. Table 5 is some of the data of students level of knowledge acquisition and, Table 6 shows the result of this data after discretisation using Boolean reasoning technique accordingly.

Table 5. Data of students knowledge acquisition

CPA	TIME	SCORE	BACK TRACK	STATUS
3.5	82	70	1	3
2.8	56	88	2	3
3.7	84	50	1	3
2.4	61	89	5	1
3.7	75	55	1	2
2.6	55	90	4	1
1.8	61	92	5	1
3.7	120	91	7	2
3.8	102	88	6	2
3.8	66	87	0.3	3
2.8	86	73	2.2	2
2.5	102	59	3	1
2	112	54	3.6	1
2	114	53	3.8	1

Table 6. Data of students knowledge acquisition after discretization

1	0	0	0	1
2	0	0	0	1
2	0	0	0	1
3	0	1	0	2
2	0	1	0	3
3	0	1	0	2
3	0	1	0	2
2	0	1	0	3

3	0	0	0	3
3	1	1	0	1
2	1	1	0	3
3	0	1	0	2
3	1	1	1	2
3	1	1	1	2

2.6 Computation of reducts

A fundamental problem in information system/decision systems is whether the whole knowledge is always necessary to define some categories available in the knowledge considered. This problem arises in many practical applications and be referred as knowledge reduction. In reduction of knowledge the fundamental concept offered by RS is *reduct* and *core*. A reduct of knowledge is the essential part in which suffices to define all basic concepts occur in the considered knowledge, whereas the core is in a certain sense its most important part. Reducts are subsets of the original attributes with the original dependencies preserved. Rules can be generated using a minimal subset of attributes. The computation of dynamic reducts is similar to the computation of reducts. The difference is that dynamic reducts only preserve the original dependencies approximately, not accurately. The result is that attributes that have a minimal impact on the decision attributes are not used to generate rules. Thus, reducts with their cardinalities for students knowledge acquisition in this study are listed accordingly :

{backtrack}	1
{cpa, scored}	2
{cpa}	1
{time, backtrack}	2
{scored, backtrack}	2
{time, scored}	2
{cpa, backtrack}	2
{cpa, time}	2

2.7 Rule generation

Rules are generated from reducts. The rules may be of different types and on different formats, depending on the algorithms used. Measures of confidence and frequency usually accompany the rules. We listed some of the rules generated from our reducts on knowledge acquisition data below:

Rule 1:

backtrack({0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, 1}) => **status(3) –** this mean that the values given in the parentheses are the generated rules and imply that the status of students knowledge acquisition are good represented by numerical value 3, and the same for the others accordingly.

Rule 2:

cpa({2.8, 2.9, 3, 3.2, 3.4, 3.5, 3.8, 3.9}) AND **scored**({27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52}) => **status(3)**

Rule 3:

cpa({2.8, 2.9, 3, 3.2, 3.4, 3.5, 3.8, 3.9}) AND **scored**({100, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99}) => **status(3)**

Rule 4:

cpa({0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, 1, 1.1, 1.2, 1.4, 1.5, 1.6, 1.7, 1.9, 2, 2.4, 2.5}) => **status(1)**

Rule 5:

time({{12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 55, 56, 58, 60, 61, 66, 68, 70, 72, 74}, {75}}) AND **backtrack**({2.1, 2.2, 2.4, 2.5, 2.7, 2.9, 3, 3.6, 3.8, 3.9, 4.1, 4.2, 4.3, 4.4, 4.6, 4.7, 4.8, 4.9, 5.1, 5.2, 5.3, 5.4, 5.6, 5.7, 5.8, 5.9, 6.1, 6.2, 6.3, 7}) => **status(1)**

Rule 6:

scored({100, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99}) AND **backtrack**({2.1, 2.2, 2.4, 2.5, 2.7, 2.9, 3, 3.6, 3.8, 3.9, 4.1, 4.2, 4.3, 4.4, 4.6, 4.7, 4.8, 4.9, 5.1, 5.2, 5.3, 5.4, 5.6, 5.7, 5.8, 5.9, 6.1, 6.2, 6.3, 7}) => **status(1)**

Rule 7:

scored({53, 54, 55, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81}) AND **backtrack**({2.1, 2.2, 2.4, 2.5, 2.7, 2.9, 3, 3.6, 3.8, 3.9, 4.1, 4.2, 4.3, 4.4, 4.6, 4.7, 4.8, 4.9, 5.1, 5.2, 5.3, 5.4, 5.6, 5.7, 5.8, 5.9, 6.1, 6.2, 6.3, 7}) =>**status(2)**

Rule 8:

scored({27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52}) AND **backtrack**({2.1, 2.2, 2.4, 2.5, 2.7, 2.9, 3, 3.6, 3.8, 3.9, 4.1, 4.2, 4.3, 4.4, 4.6, 4.7, 4.8, 4.9, 5.1, 5.2, 5.3, 5.4, 5.6, 5.7, 5.8, 5.9, 6.1, 6.2, 6.3, 7}) =>**status(1)**

Rule 9:

time({{100, 102, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98}}) AND **scored**({53, 54, 55, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81}) => **status(2)**

Rule 10:

time({{12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 55, 56, 58, 60, 61, 66, 68, 70, 72, 74}, {75}}) AND **scored**({53, 54, 55, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81}) => **status(3**)

Rule 11:

cpa({2.8, 2.9, 3, 3.2, 3.4, 3.5, 3.8, 3.9}) AND **backtrack**({2.1, 2.2, 2.4, 2.5, 2.7, 2.9, 3, 3.6, 3.8, 3.9, 4.1, 4.2, 4.3, 4.4, 4.6, 4.7, 4.8, 4.9, 5.1, 5.2, 5.3, 5.4, 5.6, 5.7, 5.8, 5.9, 6.1, 6.2, 6.3, 7}) => **status(2)**

Rule 12:

cpa({2.8, 2.9, 3, 3.2, 3.4, 3.5, 3.8, 3.9}) AND **time**({{12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 55, 56, 58, 60, 61, 66, 68, 70, 72, 74}, {75}}) => **status(3)**

3. Classification

After a rule set has been generated, these rules should be tested on a data table. Preferrably, the rules should be tested on unseen objects, that is, objects that were not used in the rule generation process. After applying the rules on unseen objects, the predicted outcomes can be compared to the correct classification of the object.

In this study, each object represents students's performance as shown in Table 1. The decision is the students's knowledge acquisition status, identified by adopting expert advice as poor, average and good. They are represented as Class 1, Class 2 and Class 3 respectively. We apply the rules to 76 unseen objects. The result shows that 71 objects are correctly classified. The total percentage of correct classification is 93.4211%. The details are as shown in the Table 7.

Table 7. Result of students knowledge acquisition classification

		Predicted	**Class**	
Actual Class	1	2	3	
1	30	1	0	0.967742
2	0	20	1	0.952381
3	0		21	0.875
	1.0	0.833333	0.954545	0.934211

To see more closely on how the classifications are made, we present a few snapshots of the objects as shown in Table 8. For object 4, its value for attributes cpa, time, score and backtrack are 2.4, 61, 89 and 5 respectively (Refer Table 5 for original value of the object's attributes). When the rules are applied, it satisfies 3 of them, and this includes rule 4, rule 5 and rule 6 with classification identification as Class 1. Similarly for object 10 in which the attributes are represented as 3.8, 66, 87 and 0.3. It satisfies rule 1, rule 3 and rule 12 with 3 rules classify it as Class 3. Different rules may generate different classification results. When this is the case, the class predicted by most rules is chosen as illustrated by object 12. Its attribute values satisfy rule 1, rule 7 and rule 9 whereby rule 1 classifies it as Class 1, rule 7 and 9 classify it as Class 2. Hence the error exists. Overall results in this study shows that the error rate is less than 7%.

Table 8. Result of students knowledge acquisition classification

```
Object 4:   ok
            Actual=1  (1)
            Predicted=1  (1)
            Ranking=(1.0)  1  (1)
            3  rule(s)

Object 10: ok
            Actual=3  (3)
            Predicted=3  (3)
            Ranking=(1.0)  3  (3)
            3  rule(s)

Object 12: ERROR
            Actual=1  (1)
            Predicted=2  (2)
            Ranking=(0.727273)  2 (2)
            2 rule(s)
            Ranking=(0.272727)  1 (1)
            1 rule(s)
```

4. Conclusion

Four attributes, (cpa, learning time, number of backtracking, and scores obtained by the student), that have been collected in the student model and the decision of the students knowledge acquisition whether as good, average or poor are transformed into the information table. Based on the chosen attributes, classes of indiscernibility objects and the discernibility matrix are generated followed by discretisation using boolean reasoning approach. Rules generation is done using a minimal subset of attributes that are collected after the reducts computation. The rules are tested on a data table of unseen objects. The criterion of choosing the best classification model is based on the highest percentage of classification toward the new unseen data. The number of the classification percentage varies based on the number of reduct obtained in the process.

5. References

1. P. Brusilovsky, "Adaptive Hypermedia : An Attempt to Analyse and Generalize". *4th Intern. Conf. On User Modeling*, Massachusettes, USA 1994.

2. Z. Pawlak, "Rough Set : Theoretical Aspect of Reasoning About Data. " *Kluwer Publications*, 1991.

3. R. Greathe. "Knowledge Extraction From Process Data : A Rough Set Approach to Data Mining on Time Series", *MSc. Thesis*, 1999.

4. Siti Zaiton Mohd Hashim, Norazah Yusof, Paridah Samsuri and Nor Bahiah Ahmad, "Design and implementationof an adaptive web-based hypermedia learning system : a prototype", *International Conference on Organizational Development and Leadership in Education*, Kuala Lumpur, 2001.

PAPER V:

STUDENT CLASSIFICATION USING NEURAL NETWORK IN ADAPTIVE HYPERMEDIA LEARNING SYSTEM: PRELIMINARY WORK

Bariah Yusob, Nor Bahiah Hj Ahmad, Siti Mariyam Hj Shamsuddin, Shahliza Abd Halim, Norazah Yusof

Faculty of Computer Science and Information System

University of Technology Malaysia, 81310 Skudai

byvail@catcha.com

Abstract

Traditional Adaptive Hypermedia Learning System will pose disorientation and lost in hyperspace problem. Adaptive Hypermedia Learning System is the solution whereby the system will personalize the learning module presentation based on student classification: advanced, intermediate and beginner. Classification of student is needed for the system to provide suitable learning module to each individual student by taking consideration of students' knowledge level and their performances as they go through the system. Three classifiers determine student knowledge level. The first classifier determines the student initial status from data collected from explicit data extraction technique. Second classifier identifies student's status from implicit data extraction technique, and the third classifier will be executed if the student has finished doing exercises. Implicit extraction technique includes process of gathering and analysis of students' behavior provided by web log data while they navigate through the system. Explicit extraction technique on the other hand is a process of collecting students' basic information from user registration data. Finally, based on the AHLS architecture, this information will be integrated into user profile to perform classification using simple backpropagation neural network.

Keywords: student classification, web log analysis, neural network

1.0 Introduction

Adaptive Hypermedia Learning System (AHLS) is an approach to overcome the problems with traditional static hypermedia learning system (Brusilovsky, 1994). AHLS manage to present interactive and dynamic interface to provide suitable learning module to each student. Learning module presented is based on students' knowledge and skill level together with their preferences that can be seen through analysis of their behavior as they navigate along the system.

Learning module presentation that meets those features must be created through one very popular concept: personalization (Wang, 2000). This concept has been used as a trend in electronic application in World Wide Web including e-commerce, e-medicine and others. Through personalization, users are no longer treated equally. The AHLS will recognize users and provide services according to their personal needs. For example, a beginner student will not be presented with learning module that is too complex and have difficult paths. If the student receive unsuitable module that does not match his/her level, he/she will not be able to follow the learning process and finally the objective of the module is not achieved.

To apply personalization concept, student's information must be analyzed and extracted to get useful information about the student and to identify type of the student, whether he/she is advanced, intermediate or beginner. The aim of this project is to integrate two techniques of user data extraction in AHLS to create user profile. These techniques are explicit and implicit data extraction. Explicit technique includes data given by students when they register into the system, and score they gain as they perform exercises after finish one topic. Implicit user profile record students' behavior through their interaction with the system. At this stage, the students are not aware that their movements are recorded in web log file. Data collected are navigation path, clickstreams, together with date and time the node is requested.

This paper presents the preliminary work in order to classify student by manipulating data extracted explicitly and implicitly. We will use data stored in the system for the neural network training process to perform the classification task.

Specifically, the paper will first present a model of student classification using neural network. Second section contains the process of analyzing students' activity in web log data, including navigation, number of using help, number of backtracking and others. There are three types of classifier for determining the student's status, which will be explained in detail in the next section.

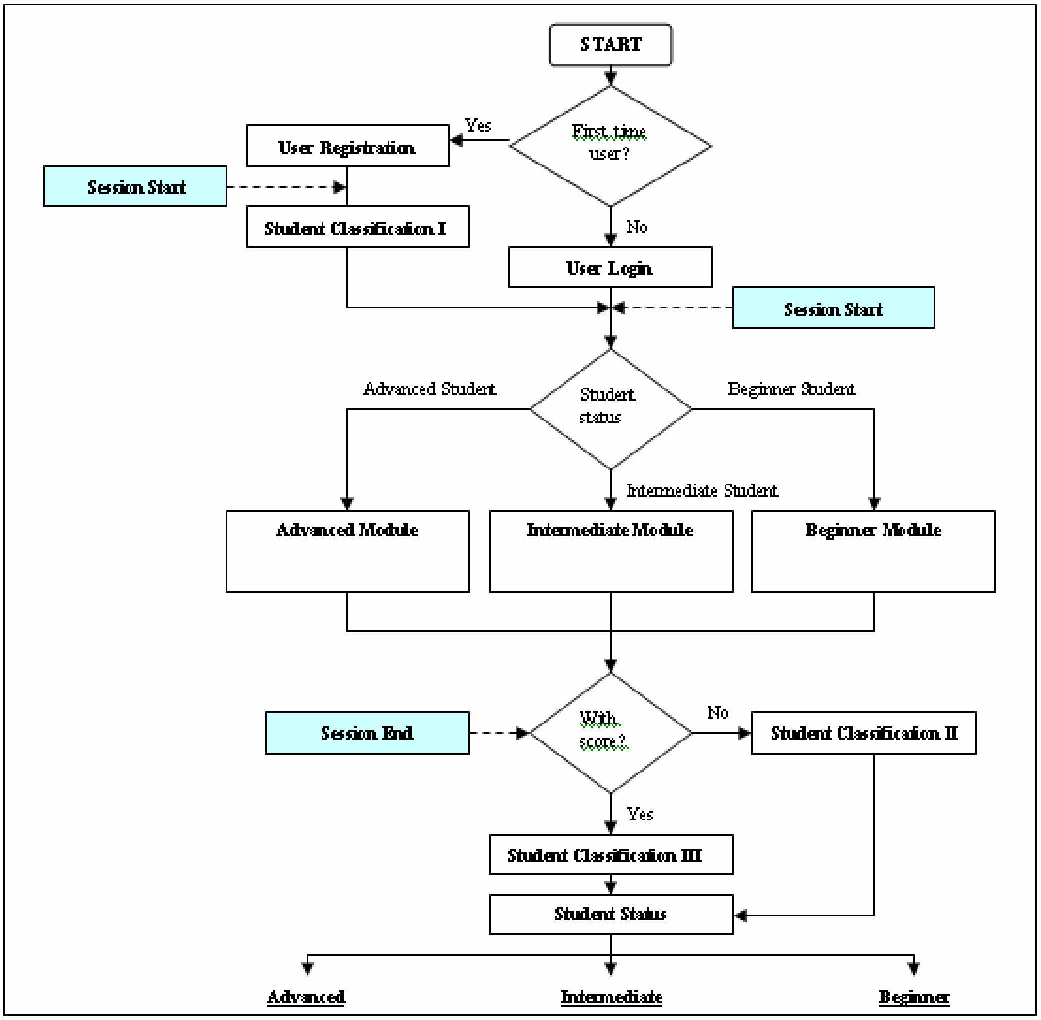


Figure 1: Workflow of Student Classification

2.0 Workflow Of Student Classification Using Neural Network

The system must first identify whether the student has register into the system before allowing them to start learning the module. For the classification purpose, there are situation need to be considered when student accessing the system. The situation includes first time user, student without score value and student with score value. In the following, we will give an overview of each situation:

i) Student who has not register into the system. It means that the student is a new user. The student must enter his/her information including name, metric number, ic number, cpa and programming knowledge. Once the student submit his/her registration data, the system will record the data into database and remember his/her metric number as a new user. Then, the system will execute first classifier to identify his/her initial status.

ii) The student has registered and will be asked to log into the system by entering his/her metric number so that the system could recognize him/her and will retrieve his/her previous status. Based on his/her last status, the suitable learning module will be presented. In this situation, there is possibility that the student simply finish learning the module without performing any exercises. Once he/she log out of the system, second classifier will identify his/her current status.

iii) When the user finished learning the module and has completed doing exercises, the score will be stored in system database. Third classifier will identify his/her current status using explicit and implicit data extraction.

Based on Figure 1, the system must first identify students' status from the previous learning session. Once their status is identified, the learning module will be presented according to their status. After they finish their session, then the last classifier will be applied to categorize them into their current status to be used in the next session when they come back for learning new module.

3.0 Explicit Data Extraction: Classifier I

For the first time user, she/he has to register into the system by filling his/her information into the student registration form. Explicit data is a data provided interactively and willingly by the student.

3.1 Student Registration

Data collected from student registration process are student's name, metric card number, identity card (ic) number, cpa and programming knowledge (prog). Figure 2 shows the student registration form. Registration data will be stored in system's database. Once student submit his registration, system will assign a new session that recognize the student by his metric number, noMatrik. Data collected from student registration will be used in student classification phase to identify the status.

Figure 2: Student Registration Form

3.2 Training Data Sample

Before data can be fed into neural network for student classification, one set of training data is needed. Training data is used to train the network to perform classification into desired groups. To obtain the training sample, one program is developed to calculate the desired output values. The program will be using input data from cpa and prog.

Table 1: Input Data

Input, x	Value	Weight, w
cpa, x_1	0.00 – 4.00	$w_1 = 0.7$
prog, x_2	1,2,3	$w_2 = 0.3$

Where,

$$\text{Output, } y = x_1 * w_1 + x_2 * w_2$$

Student's *cpa* will hold the value between 0.00 – 4.00.
prog will hold value whether 1 for beginner, 2 for intermediate and 3 for advanced. Weight is given based on the priority between these two data. cpa carries more priority than prog value.

Classification I:

If y > 2.00 then
    Status = beginner
Else if 2.00 ≤ y ≤ 3.00 then
    Status = Intermediate
Else if y > 3.00 then
    Status = advanced

Data Representation:

Beginner : 00
Intermediate : 01
Advanced : 11

4.0 Implicit Data Extraction: Classifier Ii

Implicit data extraction is a process of analyzing web log data, which contain students' activity through their interaction with the system.

4.1 Web Log Analysis

Web log analysis is a process of extracting useful information about user's behavior recorded in web log server file. In this project, we perform implicit technique by collecting web log data and analyze it to get students' path navigation through the system.

Figure 3 shows lines from a file using the following fields: date, time, client IP address and URI stem (requested node). The original data show sequences of requests by IP address and request time. In order to analyze the log data, we first need to define an individual session.

A session is defined by a unique IP address and a unique request time. It begins when user login or when system come across an IP address. The request time is defined as the beginning of a session. System then keep tracking that individual's requests continuously, and define the end of that particular session for that individual to be when a subsequent request does not appear within an hour. Figure 4 shows the log data extracted from the original web log file into one session.

```
#Software: Microsoft Internet Information Services 5.0
#Version: 1.0
#Date: 2003-09-19 16:03:54
#Fields: date time c-ip cs-uri-stem cs(Referer)
2003-09-19 16:03:54 127.0.0.1 /iishelp/iis/htm/core/iiauths.htm
http://127.0.0.1/iishelp/iis/htm/core/iiabasc.htm
#Software: Microsoft Internet Information Services 5.0
#Version: 1.0
#Date: 2003-09-19 16:21:21
#Fields: date time c-ip cs-uri-stem
2003-09-19 16:21:21 127.0.0.1 /spath/
2003-09-19 16:21:48 127.0.0.1 /spath/fT02.htm
2003-09-19 16:21:48 127.0.0.1 /spath/banner.htm
2003-09-19 16:21:48 127.0.0.1 /spath/sisiMD.htm
2003-09-19 16:21:48 127.0.0.1 /spath/EFRONT1.gif
2003-09-19 16:21:48 127.0.0.1 /spath/T02.htm
2003-09-19 16:21:51 127.0.0.1 /spath/T02F001.htm
2003-09-19 16:21:54 127.0.0.1 /spath/T02F002.htm
2003-09-19 16:21:54 127.0.0.1 /spath/T02F001R001.jpg
2003-09-19 16:22:02 127.0.0.1 /spath/T02F003.htm
2003-09-19 16:22:03 127.0.0.1 /spath/T02F003N001.htm
2003-09-19 16:22:15 127.0.0.1 /spath/T02F003N003.htm
2003-09-19 16:22:15 127.0.0.1 /spath/T02F002N003R001.gif
2003-09-19 16:23:30 127.0.0.1 /spath/
2003-09-19 16:23:41 127.0.0.1 /spath/T02.htm
2003-09-19 16:23:43 127.0.0.1 /spath/T02F001.htm
2003-09-19 16:23:48 127.0.0.1 /spath/T02F002.htm
2003-09-19 16:23:48 127.0.0.1 /spath/T02F001R001.jpg
2003-09-19 16:23:53 127.0.0.1 /spath/T02F003.htm
2003-09-19 16:24:09 127.0.0.1 /spath/T02F003N001.htm
2003-09-19 16:24:16 127.0.0.1 /spath/T02F003N003.htm
2003-09-19 16:24:16 127.0.0.1 /spath/T02F002N003R001.gif
```

Figure 3: Example of Original Web Log Data

Data collected from web log data include date of request, time of request, IP address/username and page request. In this phase, we also need a training data before using it with neural network to perform student classification task. One more program is developed to calculate the desired output. Input for this program will be time and request. Figure 5 is the example of output of this program. Recommended learning time is 60 minutes.

```
2003-09-19 16:21:21 127.0.0.1 /spath/
2003-09-19 16:21:48 127.0.0.1 /spath/fT02.htm
2003-09-19 16:21:48 127.0.0.1 /spath/banner.htm
2003-09-19 16:21:48 127.0.0.1 /spath/sisiMD.htm
2003-09-19 16:21:48 127.0.0.1 /spath/EFRONT1.gif
2003-09-19 16:21:48 127.0.0.1 /spath/T02.htm
2003-09-19 16:21:51 127.0.0.1 /spath/T02F001.htm
2003-09-19 16:21:54 127.0.0.1 /spath/T02F002.htm
2003-09-19 16:21:54 127.0.0.1 /spath/T02F001R001.jpg
2003-09-19 16:22:02 127.0.0.1 /spath/T02F003.htm
2003-09-19 16:22:03 127.0.0.1 /spath/T02F003N001.htm
2003-09-19 16:22:15 127.0.0.1 /spath/T02F003N003.htm
2003-09-19 16:22:15 127.0.0.1 /spath/T02F002N003R001.gif
2003-09-19 16:23:30 127.0.0.1 /spath/
2003-09-19 16:23:41 127.0.0.1 /spath/T02.htm
2003-09-19 16:23:43 127.0.0.1 /spath/T02F001.htm
2003-09-19 16:23:48 127.0.0.1 /spath/T02F002.htm
2003-09-19 16:23:48 127.0.0.1 /spath/T02F001R001.jpg
2003-09-19 16:23:53 127.0.0.1 /spath/T02F003.htm
2003-09-19 16:24:09 127.0.0.1 /spath/T02F003N001.htm
2003-09-19 16:24:16 127.0.0.1 /spath/T02F003N003.htm
2003-09-19 16:24:16 127.0.0.1 /spath/T02F002N003R001.gif
```

Figure 4: User Session 1

```
Date       Time      ip/host    request                              Visited  TimeTaken
2003-09-19 16:21:21 127.0.0.1 /spath/                                   0      00:27
2003-09-19 16:21:48 127.0.0.1 /spath/fT02.htm                           0      00:00
2003-09-19 16:21:48 127.0.0.1 /spath/banner.htm                         0      00:00
2003-09-19 16:21:48 127.0.0.1 /spath/sisiMD.htm                         0      00:00
2003-09-19 16:21:48 127.0.0.1 /spath/EFRONT1.gif                        0      00:00
2003-09-19 16:21:48 127.0.0.1 /spath/T02.htm                            0      00:03
2003-09-19 16:21:51 127.0.0.1 /spath/T02F001.htm                        0      00:03
2003-09-19 16:21:54 127.0.0.1 /spath/T02F002.htm                        0      00:00
2003-09-19 16:21:54 127.0.0.1 /spath/T02F001R001.jpg                    0      00:08
2003-09-19 16:22:02 127.0.0.1 /spath/T02F003.htm                        0      00:01
2003-09-19 16:22:03 127.0.0.1 /spath/T02F003N001.htm                    0      00:02
2003-09-19 16:22:15 127.0.0.1 /spath/T02F003N003.htm                    0      00:00
2003-09-19 16:22:15 127.0.0.1 /spath/T02F003N003R001.gif                0      00:01
2003-09-19 16:23:16 127.0.0.1 /spath/T02F003N003help.htm                0      00:14
2003-09-19 16:23:30 127.0.0.1 /spath/                                   1      00:03
2003-09-19 16:23:33 127.0.0.1 /spath/T02F003help.htm                    0      00:08
2003-09-19 16:23:41 127.0.0.1 /spath/T02.htm                            1      00:02
2003-09-19 16:23:43 127.0.0.1 /spath/T02F001.htm                        1      00:05
2003-09-19 16:23:48 127.0.0.1 /spath/T02F002.htm                        1      00:00
2003-09-19 16:23:48 127.0.0.1 /spath/T02F001R001.jpg                    1      00:05
2003-09-19 16:23:53 127.0.0.1 /spath/T02F003.htm                        1      00:16
2003-09-19 16:24:09 127.0.0.1 /spath/T02F003N001.htm                    1      00:07
2003-09-19 16:24:16 127.0.0.1 /spath/T02.htm                            1      00:00
2003-09-19 16:24:16 127.0.0.1 /spath/T02F003N003.htm                    1      00:

Count of backtracking = 9
Count of help         = 2
Learning time              = 105 s --> 1.75 min

Recommended learning time = 60 min

Percentage       = 2.9167%
```

Figure 5: The Example of Output of Program To Calculate Training Data For Classifier II

Based on Table 2 below, there are a few criteria identified to be an input data for student classification II. These include learning time, number of backtracking, and number of using help (Hashim et al., 2001).

We identify students' status by comparing the time taken by student with the time recommended by system. If the student took more than the recommended time, this indicates that he/she is a slow learner. On the other hand, if he/she took less time, he/she is a fast learner.

In this research, we assume that a beginner student will backtrack the relevant materials he/she has gone through earlier as shown in Table 2. The more the student use help, shows that the student is having problem understanding the module. So, we consider number of using help as one criteria to classify student's status.

Table 2: Criteria to Classification II

Criteria	Beginner, α	Intermediate, β	Advanced, γ
Learning Time, x_1	$x_1 > 100\%$	$80\% \leq x_1 \leq 100\%$	$x_1 < 80\%$
Freq. Of Backtracking, x_2	$x_2 > 5$	$3 \leq x_2 \leq 5$	$x_2 \leq 1$
Freq. Of Using Help, x_3	$x_3 > 5$	$3 \leq x_3 \leq 5$	$x_3 \leq 1$

Mathematically,

$$\alpha(x_i) \in \begin{Bmatrix} x_1 > 100\% \\ x_2 > 5 \\ x_3 > 5 \end{Bmatrix} \quad \beta(x_i) \in \begin{Bmatrix} 80\% \leq x_1 \leq 100\% \\ 3 \leq x_2 \leq 5 \\ 3 \leq x_2 \leq 5 \end{Bmatrix} \quad \gamma(x_i) \in \begin{Bmatrix} x_1 < 80\% \\ x_2 \leq 1 \\ x_3 \leq 1 \end{Bmatrix}$$

Data Representation:

Beginner	: 00
Intermediate	: 01
Advanced	: 11

5.0 Integration Of Explicit And Implicit Techniques: Classifier III

For the third classifier, we integrate the explicit and implicit data extraction to obtain input data for the student classification.

5.1 Training Data Sample

Students' behavior data are collected from analysis of web log data similar to the second classification process. The data considered are the requested pages and time of request. In the explicit user data extraction process, we collect students' score value from system's database as an additional input data into our program to calculate the desired output data for the training sample. The question and answer session is used to test the student's understanding of the material being learned. Table 3 shows the possible criteria for input data to calculate the desired output.

Table 3: Criteria To Classification III

Criteria	Beginner, α	Intermediate, β	Advanced, γ
Learning Time, x_1	$x_1 > 100\%$	$80\% \le x_1 \le 100\%$	$x_1 < 80\%$
Freq. Of Backtracking, x_2	$x_2 > 5$	$3 \le x_2 \le 5$	$x_2 \le 1$
Freq. Of Using Help, x_3	$x_3 > 5$	$3 \le x_3 \le 5$	$x_3 \le 1$
Score, x_4	$x_4 < 60\%$	$60\% \le x_4 \le 800\%$	$x_4 > 80\%$

Mathematically,

$$\alpha(x_i) \in \left\{ \begin{array}{l} x_1 > 100\% \\ x_2 > 5 \\ x_3 > 5 \\ x_4 < 60\% \end{array} \right\} \quad \beta(x_i) \in \left\{ \begin{array}{l} 80\% \le x_1 \le 100\% \\ 3 \le x_2 \le 5 \\ 3 \le x_2 \le 5 \\ 60\% \le x_4 \le 800\% \end{array} \right\} \quad \gamma(x_i) \in \left\{ \begin{array}{l} x_1 < 80\% \\ x_2 \le 1 \\ x_3 \le 1 \\ x_4 > 80\% \end{array} \right\}$$

Data Representation:

Beginner : 00

Intermediate : 01

Advanced : 11

6.0 Conclusion And Future Work

Data preprocessing is implemented through the process of extracting information from registration data (explicit technique) and web log analysis (implicit technique) to develop a complete the student profile that use the system. We defined three types of classifiers to be used in classifying the students' status based on input data collected from registration form, web log data or combination of the web log data and the students' score while doing the exercises.

The results from this phase is a set of training data including input data and desired output data that will be used in classification with neural network for future work. Interesting issue here concern the session identification using student's metric number. In this paper we consider IP address as a unique session identifier. For the future research, we will look at integration of metric number variable and IP address and request time to identify unique session.

Acknowledgements

This research is conducted under IRPA grant, vote project 74129: The Development of Knowledge Management For Adaptive Hypermedia Learning System, Faculty of Computer Science, UTM Skudai.

References

1. Bailey, D. And Thompson, D. (1990). *How To Develop Neural-Network Applications*. In AI Expert. June 1990. 38-45.

2. Brusilovsky, P. (1999). *Adaptive and Intelligent Technologies For Web-Based Education*. In C. Rollinger and C. Peylo (eds) Kunstliche Intelligenz, Special Issue on Intelligent Systems and Teleteaching. 4. 19-25.

3. Buser, D., Kauffman, J., Llibre, J.T., Francis, B., Sussman, D., Ullman C. And Duckett J. (2000). *Application, Sessions And Cookies*. Wrox Press Ltd. (2000). Beginning Active Server Pages 3.0. 289-330.

4. Catledge, L.D. and Pitkow, J.E. (1995). *Characterizing Browsing Strategies in The World Wide Web*. Georgia Institute of Technology, Atlanta, USA.

5. Cooley, R., Mobasher, B. And Srivastava, J. (1997). Web Mining: Information and Pattern Discovery on The World Wide Web. Departmanet of Computer Science, University of Minnesota, USA.

6. Cripps, A. (1996). *Using Artificial Neural Nets to Predict Academic Performances*. In Association of Computing Machinery, Inc.(ACM). 33-37.

7. Hashim, S.Z.M., Yusof, N., Samsuri, P. and Ahmad, N.B. (2001). *Design and Implementation of An Adaptive Web-Based Hypermedia Learning System: A Prototype*. In Proceedings of The International Conference on Organizational Development and Leadership in Education. Kuala Lumpur, Malaysia.

8. Jeon-Slaughter, H. (2002). *Report of Web Log Data Mining*. In E-Journal User Study: Recent Findings. Stanford University Libraries.
URL: http://ejust.stanford.edu/logdata.pdf

9. Moore, D.J., Hobbs, D.J., Mullier, D. And Bell, C. (1997). *Interaction Paradigms With Educational Hypermedia*. In In Proceedings of 23rd EUROMICRO Conference'97 New Frontiers of Information Technology. Budapest, Hungary.

10. Mullier, D.J. (2000). *The Application of Neural Network And Fuzzy Logic Techniques To Educational Hypermedia*. Ph.D Thesis. Leeds Metropolitan University, UK.

11. Mullier, D.J. (1995). *Using A Neural Network To Model Hypermedia Browsing – An Alternative To Traditional Intelligent Tutoring Methods*. Faculty of Information and Engineering Systems, Leeds Metropolitan University, UK.

12. Quentin-Baxter, M. (1993). *Hypermedia Learning Environments Limit Access To Information*. Faculty of Medicine Computing Centre, University of Newcastle Upon Tyne, UK.

URL: http://decweb.ethz.ch/WWW7/1941/com1941.htm

13. Wang, Y. (2000). *Web Mining and Knowledge Discovery of Usage Patterns*. In CS 748T Project Report (Part I).
URL: http://db.uwaterloo.ca/~tozsu/courses/cs748t/surveys/wang.pdf

14. Yusof, N., Salim, A., Samsuri, P., Hashim, S.Z.M. and Ahmad, N.B. (2000). *SPAtH: An Adaptive Web-Based Learning System*. In Proceedings of IASTED International Conference. Banff, Canada.

15. Zaiane, O.R. and Luo, J. (2001). *Towards Learners' Behaviour in Web-Based Distance Learning Environment*. Department of Computing Science, University of Alberta, Canada.

PAPER VI:

DEVELOPMENT OF AN ADAPTIVE HYPERMEDIA LEARNING SYSTEM

Nor Bahiah Hj Ahmad, [1]Shahliza Abd Halim, [2]Bariah Yusob, [3]Norreen Haron

Department of Software Engineering
Faculty of Science Computer and Information Systems

University of Technology Malaysia, 81310 Skudai

bahiah@fsksm.utm.my
[1]shahliza@fsksm.utm.my
[2]byvail@gmail.com
[3] norreen.haron@gmail.com

Abstract

This paper presents the architecture and design for the development of an adaptive hypermedia learning system for teaching and learning Data Structure. The learning system developed integrates pedagogy of education and intelligence technique in the presentation of the learning material and the suggested navigation path. The system comprises of three main components, user profile model, domain model and adaptive engine. The focus of this research is on the use of computational intelligence technique in the classification of student models and in the adaptation of learning material and navigation path. Both learning style and knowledge acquisition have been considered as feature of adaptation in the learning system. Supervised Kohonen network is used to classify the students into advance, intermediate and beginner based on their knowledge acquisition and performance. Meanwhile, fuzzy logic is used to dynamically adapt the choice of possible paths through the learning material based on the students learning style. The integration of the two intelligence technique is able to personalize the user both at the presentation and navigation level. This study is also expected to solve disorientation and lost in hyperspace problem that usually occur in conventional hypermedia learning system.

Introduction

The advancement in educational technology has transformed the processes of teaching and learning in higher education from traditional method into computer based approach. Web-based learning environment are now used extensively as components of course delivery in higher learning institutions. Academics are placing more course material on-line to supplement their lecture based delivery method. However, research showed that traditional web-based learning system is not always suitable for all learners. Not all learners is able to work independently in web-based learning systems. Student's individual differences such as student's background, learning styles, cognitive styles and prior knowledge possessed may need more attention and support from instructional designers. The flexibility of linking in hypermedia learning environment also impose problems related to cognitive overload and lost in hyperspace which caused confusions and frustrations among learners. Therefore, adaptive learning system is important in providing students the capability to tailor the learning material presented based on the information stored in the student profile.

The purpose of this research is to develop an adaptive hypermedia learning system for learning Data Structures at Faculty of Computer Science and Information System in Universiti Teknologi Malaysia. We have explored the computational intelligence technique that effectively able to adapt the learning material based on the student learning style. In this case, we combined the students personality factor (Myers-Briggs Type Indicator (MBTI)) and fuzzy logic techniques to produce a dynamic course adaptation which will present the appropriate structure of the learning material to the student [(Norreen & Naomie (2005)]. We also have experiment supervised Kohonen network in order to classify the student based on their level of knowledge acquisition [Bariah et al. (2004)]. Besides that, we follow the software engineering principle in developing the architecture design and the development of the system. Special attention also have been made in the design of the structure of the learning material in order to present the material that can imitate the way human teacher teach in the classroom.

This paper will describe how the adaptive hypermedia learning system has been developed. First, we discusses the adaptive hypermedia learning system concepts, followed by features of adaptation and system development. The architecture of the system and the design of the course material are explained in the system development phase. Supervised Kohonen that has been used for classifying students based on their knowledge level is discussed in the subsequent section.

Adaptive Hypermedia Learning System

Adaptive hypermedia learning system (AHLS) can be defined as the technology that allows personalization for each individual user in hypermedia learning environment. In this case, adaptive refers to the ability of the website to change its behaviour or responses in reaction to the way it is used. AHLS is the solution to the conventional hypermedia learning system which allows freedom to the user to navigate through the system according to their preferences and paces. AHLS on the other hand will present the learning material based on user preferences, goal, and knowledge acquisition and user characteristics. The adaptation of AHLS is implemented through a decision making and personalization engine which adapts the content according to a user model.

AHLS system mostly comprises of three main components, user profile model, domain model and adaptive engine [Brusilovsky (2001)]. User profile model stores the learning activities, learning performance and interaction history of each student in the database. The profiles were extracted from both explicit and implicit user profile. The explicit information is the information that the learner gave willingly or directly and he/she is aware that the information is kept in the database. The implicit information is the information the system collects without the learner acknowledgement. It records the learner's activity and behavior as he/she navigates through the system.

Adaptive navigation path provide the annotated link based on the interaction history of each student. To reduce disorientation, each student will get different paths based on their level of knowledge acquisitions. Adaptive engine will determine the appropriate learning material and the navigation path based on the student's status that was retrieved from the user profile model. Domain model stores all the teaching materials including the learning objectives, lecture notes, examples, exercises and the answer for each question.

Features Of Adaptation

Research activity in the e-learning domain that apply adaptive features in web-based learning has been very intense. Methods and techniques of adaptive hypermedia had been introduced by Brusilovsky. Information that usually used are user's goals/tasks, knowledge, background, hyperspace experience, and preferences [Brusilovsky (2001)]. At least two more items can be added to this list; the user's interests and individual traits [Brusilovsky,2003)]. User interests are not a new feature of the user to be modeled. User's individual traits is a group name for user features that together define a user as an individual. Examples are personality factors such as introvert or extravert, cognitive factors, and learning styles. Like user background, individual traits are stable features of a user that either cannot be changed at all, or can be changed only over a long period of time. [Kobsa et al. (1999)] suggested distinguishing adaptation to user data, usage data, and environment data. User data comprise various characteristics of the users, usage data comprise data about user interaction with the systems and environment data comprise all aspects of the user environment that are not related to the users themselves.

Among other variables that influence the success of learning is learning styles [Ford & Chen (2000)]. It is important to diagnose the students learning style because some students learn more effectively when taught with preferred methods. Information about the learning style can help system become more sensitive to the differences of students using the system. Understanding learning styles can improve the planning, producing, and implementing of educational experiences, so they are more appropriately tailored to students' expectations, in order to enhance their learning, retention and retrieval [Zywno & Waalen (2002)]. Sadler-Smith (1997) identified four broad categories of learning style in an attempt to acknowledge and accommodate the range of aspects of individual differences referred in the educational psychology literature in an holistic way. Table 1 listed all the categories.

[Papanikolaou et al.(2003)] analyzed learners' studying behavior such as time spent and hits on resources and navigation traces by the different learning style categories to provide evidence about the way learners that belong to different learning style categories select and use educational.

Table 1 : Learning style categories [Sadler-Smith (1997)]

Learning Style Categories	**Method of learning**
cognitive personality	field dependence and field independence
Information-processing style	(converger, diverger, accommodator, assimilator) or activist, reflector, theorist, pragmatist
Instructional preferences	individual's propensity to choose or express a liking for a particular instructional technique or combination of techniques
Approaches to Studying	Deep, surface, strategic approach, lack of direction

Based on the literature done, our research has focused on the design of adaptation based on the learning style information and level of knowledge acquisition. To determine the level of knowledge acquisition, we have selected the attributes of adaptation based on the work done by [Papanikolaou et al.(2003)] and [Paridah et. al. (2001)]. The attributes selected are the learning time, number of backtracking, number of getting help function and the score earned while doing the exercise. Meanwhile, the pedagogical and learning style refer to student's personality factor based on (MBTI) as explained in [Norreen & Naomie (2005)]. We also have identified the structure of the learning material that the system should offer to learners with different styles and characteristics.

System Development

The adaptive hypermedia learning system developed in this research will be used by students taking Data Structure subject. This course is offered to second year computer science students in Faculty of Computer Science and Information System. The main objective of this course is to introduce the data structure concepts and to enable the students to apply the concepts in programming. Students are expected to master both the theory of the data structures and also the programming part. From observations, we found out that weak students usually have problems in understanding the theory and programming part. They need more explanation, more exercise and more practical in the computer laboratory compared to other students. Meanwhile, moderate students can understand the theory part easily but have difficulty in the programming part. Advance students don't have much problem in understanding the theory part and able to implement the program without many difficulties. In average, we conclude that some students need more explanation in the theory part, need to do a lot of exercises and need more practices in the programming part compared to others.

The overall architecture for the Adaptive Hypermedia System is shown in Figure 1. There are three main components in this system, adaptive engine, domain model and user profile model. Adaptive engine interface enable the user to interact with the system. User will be presented with the learning material and navigation path adapted based on the learning style and information stored in the user profile model. Adaptive engine generate features of adaptation, a tree structure to present the learning material structure.

User profile model stores the information about learners. The information are extracted from both explicit and implicit user profile. The explicit information is the information that the learner gave while fill in the registration form and the questionnaire while using the system for the first time. The implicit information is the information that the system collects while the user interact with the system. The user activity such as the URLs accessed and the behavior while he navigates through the system is recorded. The user behaviour being considered is the usage of help properties, such as referring to glossary, using back and previous button and the learning time taken while learning a concept. The information captured in the user profile will be normalized and given as input to supervised Kohonen self-organizing maps to identify the class of student, whether as beginner, intermediate and advanced.

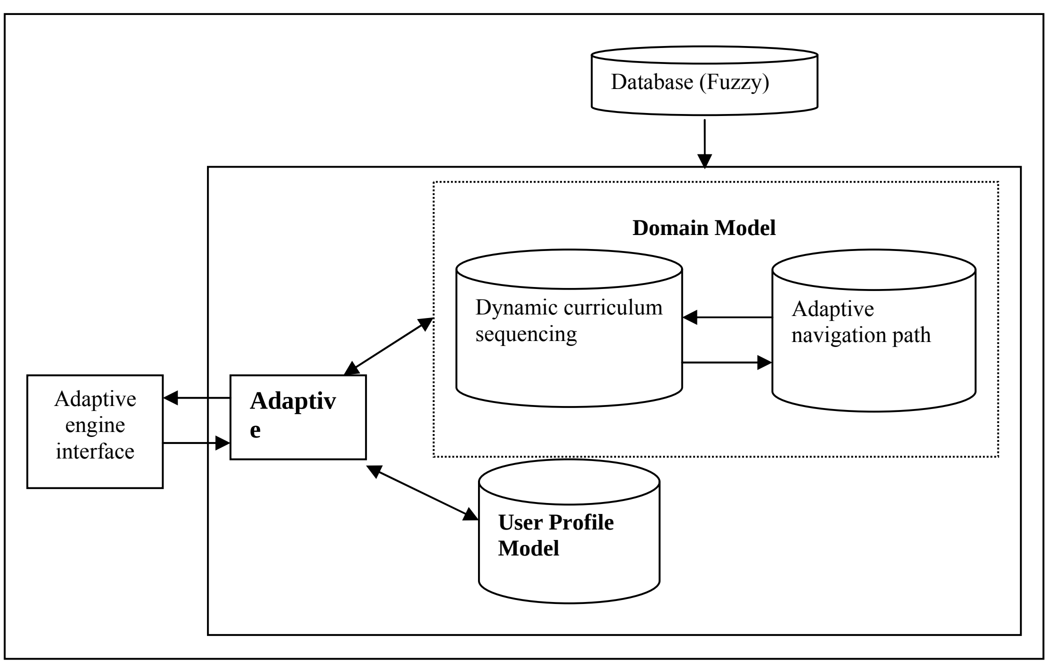


Figure1: Architecture of an adaptive hypermedia system

Domain model comprise two subcomponents, dynamic curriculum sequencing and adaptive navigation path. Dynamic curriculum sequencing will determine the student learning style and the sequence of learning material structure based on their learning style. In this research, we have identified four types of student learning style, Introvert-Sensor, Extrovert-Sensor, Introvert-Intuition and Extrovert-Intuition based on personality factor MBTI. Adaptive navigation path component further will assist the student while navigating through the system. The link showed to each student will be annotated with different colors such as yellow, red and green to signify the material that the student have already learned, forbidden because of prerequisite is not fulfilled and ready to be learned. The link annotation can help prevent the student from disorientation while navigates.

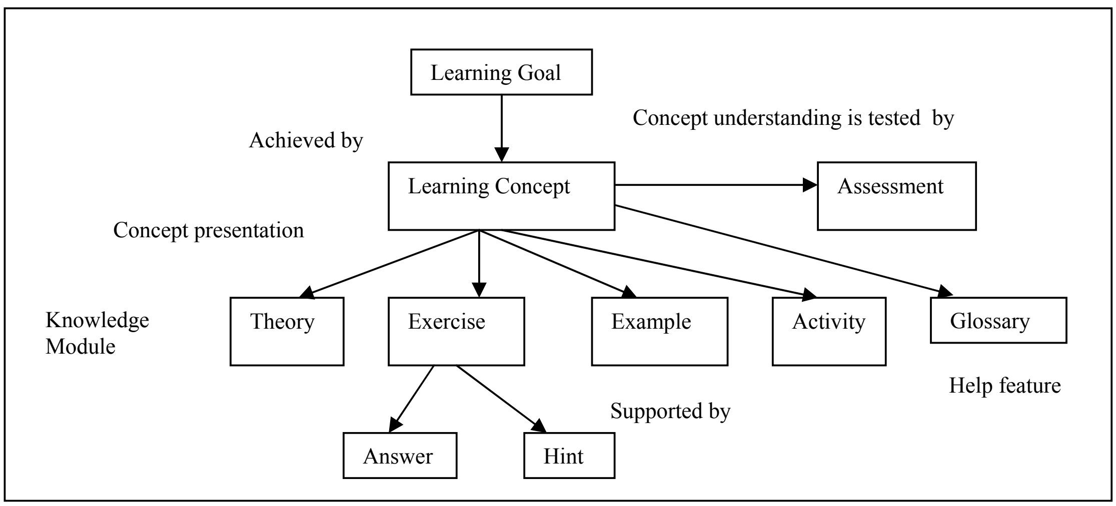


Figure 2: The structure of the learning material

Table 2 : The coding scheme for the URL for each category

Learning Category	**URL for topic X**
Learning Objective	T0XF00XO00X
Theory	T0XF00XN00X
Example	T0XF00XC00X
Exercise	T0XF00XX00X
Activity	T0XF00XA00X
Glossary	T0XF00XG00X
Assessment	T0XF00XS00X

Figure 2 shows the structure of the learning material. The student can choose any learning goal which has fulfill the pre-requisites. There are several learning concepts to be learned in order to achieve the learning goal. The learning concept is presented into four categories; example, theory, activity and exercise. While using the system, if the student is not familiar with certain terminology or concept, the student can refer to the glossary provided as a help feature in the system. The URL for each category has been set as Table 2. The coding system is used in order to identify the navigation path that has been accessed by the student. Further, based on the number of URL on certain code, it will be easier to determine the coverage of certain category of the learning material.

For each type of student, different category of the learning material will be presented adaptively in different sequence by the system. To identify the student learning style and also the sequence of the learning structure, fuzzy logic technique is used to generate suitable rule of adaptation.

System Flowchart

There are two different flows designed for first time user and second time user as shown in Figure 3. The first time user has to fill in a registration form and a questionnaire. The information collected in the registration form is used to classify the knowledge level of the student explicitly. The result from the questionnaire is used to determine the student's personality factor value either as introvert, extrovert, intuitive and sensor. In order to determine the prominent learning style for a student, fuzzy logic technique is used to identify the student learning style either fall into one of these four types Introvert-Sensor, Extrovert-Sensor, Introvert-Intuition or Extrovert-Intuition.

Both data from the preliminary classification of student's knowledge level and the category of learner are used to determine the adaptive presentation for the learning material structures sequence. For example, if the student is fall into beginner student and is identified as an Introvert-Sensor category of learner, the presentation of the learning material structures is in the sequence of Example, Theory, Activity and Exercise as presented by the tree structure depicted in Figure 4. Figure 4 shows the tree structure for learning Sorting concept. Different category of learner will get different tree structure that display different sequence of learning structure.

The knowledge level of the student will be updated once the student logs off the system. In this situation Kohonen program will be executed and given the latest information on the student activities while learning on-line, the score for doing assessment and the interaction history as input vector. Meanwhile, second time user must login before using the system. The current knowledge level and the diagnosed learning style will be extracted from the database in order to display the learning material adapted to the status level and learning style.

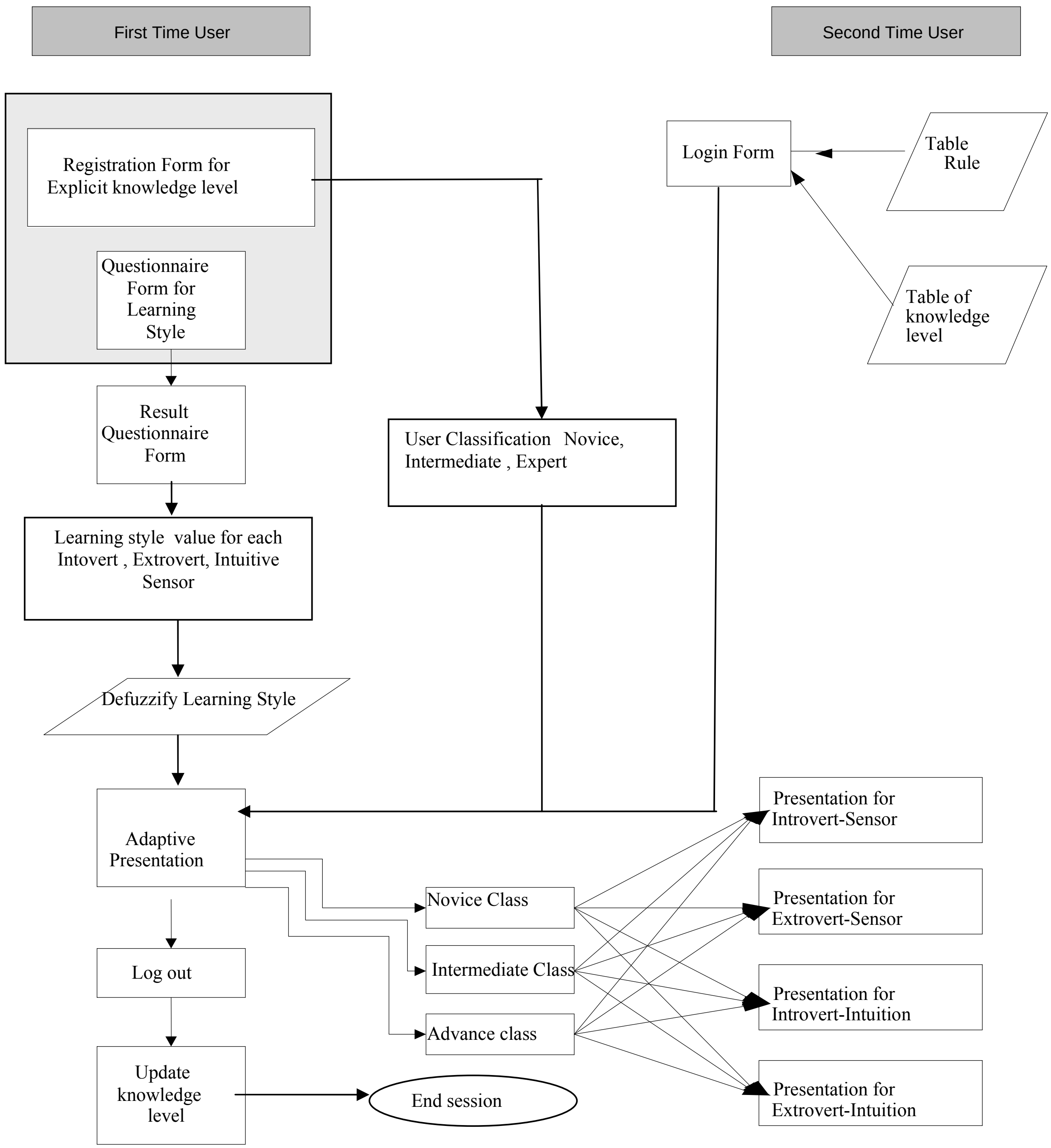


Figure 3: The system flow chart

Isihan

⇒ Objektif Pembelajaran

⇒ Contoh
  - Isihan Buih
  - Isihan Pilihan
  - Isihan Selitan
  - Isihan Cantum
  - Isihan Cepat

⇒ Teori
  - Pengenalan
  - Isihan Buih
  - Isihan Pilihan
  - Isihan Selitan
  - Isihan Cantum
  - Isihan Cepat

⇒ Aktiviti
  - Isihan Buih
  - Isihan Pilihan
  - Isihan Selitan
  - Isihan Cantum
  - Isihan Cepat
  - Kesimpulan

⇒ Latihan
  - Isihan Buih
  - Isihan Pilihan
  - Isihan Selitan
  - Isihan Cantum
  - Isihan Cepat

⇒ Penilaian pelajar

Figure 4: The tree structure for Introvert-Sensor category

Supervised Kohonen For Classifying Students

Supervised Kohonen is used to classify students based on their knowledge acquisition. It employs a teacher signal when processing nodes for which a target label exists. The idea is to assign weight vectors to the same class as the node that was mapped at this location. Training proceeds in a similar manner as the unsupervised case with the difference that the weight entries are rejected if they belong to a different class. Table 3 shows the criteria to classify learners into beginner, intermediate and advances. When the learner log into the learning system, the system will count the time he/she spent on learning a concept. The system suggests the time to be spent on each category of the learning material. The learning time is calculated based on the percentage actual learning time taken to finish learning from the estimated time as shown in Table 4.

Table 3: Criteria for learner's classification

Attribute	Beginner	Intermediate	Advanced
Learning Time, *t*	$t > 80\%$	$30\% \le t \le 80\%$	$t < 30\%$
Numb. Of Backtracking, *b*	$b > 4$	$2 \le b \le 4$	$b < 2$
Numb. Of Using Help, *h*	$h > 4$	$2 \le h \le 4$	$h < 2$
Score, *s*	$s < 30\%$	$30\% \le s \le 80\%$	$s > 80\%$

Table 4 : Calculation scheme for learning time

Part	Student's actual leaning time	Estimated learning time by SPATH
Theory	acTime1	esTime1
Exercise	acTime2	esTime2
Example	acTime3	esTime3
Activity	acTime4	esTime4
SUM for student's actual learning time (SUM_acTime)		acTime1+acTime2+acTime3+acTime4
SUM for estimated learning time (SUM_esTime)		esTime1+esTime2+esTime3+esTime4
Average for student A's learning time (Avg_acTime)		SUM_acTime/4
Average for estimated learning time (Avg_esTime)		SUM_esTime/4
Learning time, t		Avg_acTime/Avg_esTime X 100 %

The number of backtracking shows that the learner is not fully master the concept, lose direction or change his/her learning goal. The number of backtracking is defined by counting how many times the learner reopen any pages in particular concept. In this research, help function is a list of definition and explanation on terms used in the notes given. This attribute shows that the more help the learner gets, the more he/she is having a difficulty in understanding a concept. The number of getting help is defined by counting how many time the learner click on the help button in particular concept. To test the learner's level of mastering, the system provides an assessment at the end of the learning session. The score is calculated by the percentage of correct answers given.

All the data must be transformed into a standard format to get a valid and accurate classification. The transformation of the data is included in the pre-processing phase using a normalization method. In the training phase, input data is given to the Kohonen network. The weights are captured after completing the training phase. In this experiment the size of the training sample is 1050. In the testing phase, there is no target data is provided. We used 450 dataset to the network. The network classifies the data based on the weights and outputs are obtained. When the testing results were obtained, the percentage of the classification accuracy was calculated.

During training, the network learns the data and generates the weights by calculating the nearest distance to the real data presented. From 450 data presented, the network is able to classify 445 data correctly. From the result achieved, we concluded that Kohonen network is capable of classifying the learners' data into the right class. It gives more than 90% accuracy in both training and testing phase. The Kohonen's SOM is definitely a good tools to classify data into a number of groups without supervision. It will be very useful in this study because it can deal with more complex and bigger sample of data when it is applied to the real learners' data in the learning system's database.

Conclusion and Further Work

This paper has proposed a way to personalize the course content for adaptive hypermedia learning system, which aims to provide learners with a customized learning environment. It emphasizes the combination of pedagogical theories and artificial intelligent techniques. Supervised Kohonen has proved to be the suitable technique to classify the student accurately. Fuzzy logic is able to identify the personality factor of a student and the material presented will be based on the personality factor. The design of the learning material is also important in supplying the right material with the right learning style and knowledge level. For future work, we intend to implement web usage mining in order to measure the efficiency of the learning approach given to the student based on their learning style.

Acknowledgement

This research is conducted under IRPA grant, project vote 74129: The Development of Knowledge Management for Adaptive Hypermedia Learning System.

References

Bariah Yusob, Norreen Haron, Nor Bahiah Hj Ahmad, Shahliza Abd Halim. (2004). Individualizing the learning material and navigation path in an adaptive hypermedia learning system. 2nd National Conference On Computer Graphics & Multimedia, Kajang, Selangor , 8-10 December.

Bariah Yusob, Siti Mariyam Hj Shamsuddin, (2005). Supervised Kohonen Network for classification problems. Simposium Kebangsaan Sains Matematik ke-XIII UUM/PERSAMA/Kerajaan Negeri Kedah, 31 Mei - 2 Jun

Brusilovsky, P.(2001). Adaptive Hypermedia. *User Modeling and User-Adapted Instruction*, 11(1 – 2),
87-110.

Brusilovsky P, Christoph P. (2003). Adaptive and Intelligent Web-based Educational Systems, *International Journal of Artificial Intelligence in Education* ,13, 156–169.

Ford, N, Chen S.Y. (2000). Individual Differences, Hypermedia Navigation, and Learning: An Empirical Study. *Journal of Educational Multimedia and Hypermedia*, 9(4), 281-312.

Norreen Haron, Naomie Salim.(2005). Integrating personality factor (MBTI) and Fuzzy logic techniques in individualizing learning material of adaptive hypermedia learning system. To be presented at Konvensyen Tenologi Pendidikan Ke 18. 16-19 September, Marang, Terengganu.

Siti Zaiton, M.H., Norazah, Y., Paridah, S., & NorBahiah, H.A. (2001). Design and implementation of an adaptive web-based hypermedia learning system: a prototype. In: Proc. Of International Conference on Organizational Development and Leadership in Education, Kuala Lumpur.

Paridah, S., Nor Bahiah, H.A., & Norazah, Y. (2001). Neural Network Application on Knowledge Acquisition For Adaptive Hypermedia Learning. In: *Chiang Mai Journal Science*, Jan. 2001.

Papanikolaou, K.A., Grigoriadou, M., Kornilakis, H., Magoulas, G. D.: Personalizing the inter-action in a Web-based educational hypermedia system: the case of INSPIRE. *User- Modeling and User-Adapted Interaction* 13(3), 213-267

Sadler-Smith E., (1997) Learning Style: frameworks and instruments. *Educational Psychology,* 17, 51-63.

Zywno, M.S. and Waalen, J.K. (2002), The Effect of Individual Learning Styles on Student Outcomes in Technology-enabled Education, *Global Journal of Engenering. Education*, 6(1)

PAPER VII:

SUMMATIVE EVALUATION FOR THE USABILITY OF INTELLIGENT TUTORING SYSTEM (SPATH)DEVELOPMENT OF AN ADAPTIVE HYPERMEDIA LEARNING SYSTEM

Shahliza Abd Halim[1], Nor Bahiah Hj Ahmad[1], Norreen Haron[2] ,Bariah Yusob[2]

Department of Software Engineering
Faculty of Science Computer and Information Systems

University of Technology Malaysia, 81310 Skudai

[1]{shahliza, bahiah}@fsksm.utm.my

[2]{byvail, norreen.haron} @gmail.com

Abstract

The various combination of multidisciplinary areas in Intelligent Tutoring System (ITS) contributes to various methods of evaluation for ITS and therefore the suitable method of evaluation must be chosen wisely. This paper discusses the evaluation of an adaptive hypermedia learning system called SPATH. SPATH provides adaptation of learning content by personalizing the learning material structure based on student learning style adapted from student's personality factor (Myers-Briggs Type Indicator (MBTI) and by the students knowledge acquisition level. In this research, we use summative evaluation where we address the educational impact of this system on students and its practical acceptability in terms of usability. In the usability study we adopt the questionnaire technique where three key factors, learnability, efficiency and satisfaction in usability were measured. Preliminary results of the usability study revealed that this system has a high percentage in learnability and satisfaction factor.

1. Introduction

Most of the early research in adaptive educational hypermedia was inspired by the area of Intelligent Tutoring Systems (ITS). A combination of ITS and learning materials organized as hypermedia was the beginning for the research on adaptive hypermedia system (AHS)[1]. Due to the close relationships between ITS and AHS, the method of evaluation for ITS is considered as suitable evaluation for AHS. Among the advantage of evaluation is that it provides an opportunity to learn from mistakes and is capable of improving the ITSs life-span as well as their usability [2].

Adaptive Hypermedia Learning System (AHLS) has been identified as an effective strategy for solving many learning problems involved in large hypermedia, such as cognitive overload and user disorientation [3]. The idea of adaptive hypermedia is to adapt the course content accessed by a particular user to his/her characteristics. Most adaptive educational hypermedia system researches focuses on adapting to user features like user's goals/tasks, knowledge, background, hyperspace experience, preference and interests [1]. However, a web-based educational system must also include information about student learning styles to adapt optimally instructional materials to the student [3],[4],[5],[6]. Identification of the learner's learning style is the key to the development of a hypermedia course that addresses different learning styles [6].

This research has developed an adaptive hypermedia learning system called SPATH for learning Data Structures at Faculty of Computer Science and Information System in Universiti Teknologi Malaysia. We have explored the computational intelligence technique that effectively able to adapt the learning material based on the student learning style. In this case, we combined the students personality factor (Myers-Briggs Type Indicator (MBTI)) and fuzzy logic techniques to produce a dynamic course adaptation which will present the appropriate structure of the learning material to the student [7]. As for the classification of students based on their level of knowledge acquisition we used supervised Kohonen network for classification purpose[8].

There are three main components in SPATH, adaptive engine, domain model and user profile model. Adaptive engine interface enable the user to interact with the system. User will be presented with the learning material and navigation path adapted based on the learning style and information stored in the user profile model. Adaptive engine generate features of adaptation and a tree structure to present the learning material structure.

User profile model stores the information about learners. The information is extracted from both explicit and implicit user profile. The explicit information is the information that the learner gave while filling in the registration form and the questionnaire when they used the system for the first time. The implicit information is the information that the system collects while the user interact with the system. Domain model comprise two subcomponents, dynamic curriculum sequencing and adaptive navigation path. Using Fuzzy logic, the student learning style and the sequence of learning material structure based on their learning style has been determined. In this research, we have identified four types of student learning style, Introvert-Sensor, Extrovert-Sensor, Introvert-Intuition and Extrovert-Intuition based on personality factor MBTI. Table 1 shows the sequence of the learning structure based on the learning style identified.

Table 1. Example of 4 different situations of the learning material structure based on the learning style

Learning Style	Structure of learning material
Introvert-Sensor	Example Theory Activities Exercise
Extrovert-Sensor	Activities Exercise Example Theory
Introvert-Intuitive	Theory Example Exercise Activities
Extrovert-Intuitive	Exercise Activities Theory Example

The knowledge level of the student will be updated once the student logs off the system. In this situation Kohonen program will be executed and be given the latest information on the student activities while learning on-line, the score for doing assessment and the interaction history as input vector. Adaptive navigation path component further will assist the student while navigating through the system.

This paper discusses the evaluation of SPATH in order to measure the efficiency, learnability and the satisfaction of the system usability. As researchers have pointed out in [2,10] the need for the evaluation of Intelligent Tutoring Systems (ITS). Among the advantage of evaluation is that it provides an opportunity to learn from mistakes and is capable of improving the life-span ITSs as well as their usability [10]. The organization of this paper is as follows: section 2 and 3 discusses the evaluation for ITS in general and proposes a suitable method for the usability study of this system. Section 4 and 5 discusses the method and process done while performing the study and also the illustrations of the preliminary results. Lastly, in section 6 and 7 is the discussion and follows by conclusion and future work of this research.

2. Evaluation Methods for ITS

Evaluation for ITS is not an easy tasks as ITS is the combination of multidisciplinary areas including expert system, Computer Based Instruction (CBI), education, psychology and also Human Computer Interaction (HCI) [2,9,10]. [2] and [9] highlighted that there are few agreed upon standards within the ITS community to guide investigators who wish to evaluate ITS given the diversity of the evaluation methods.

There are two types of evaluation for ITS, formative and summative evaluation [2,9]. Formative evaluation mainly occurs during design and early development of a project. It frequently addresses the question of relationship between the architecture of ITS and its behavior. On the other hand, summative evaluation is concerned with the evaluation of a completed system and making of formal claims about the system. It answers the question regarding the educational impact of an ITS on students. However, these types of classification are still too broad where a lot of methods can fall in either one of these classes.

In [2], the various methods have been further classified, so that the method could be differentiated from a number of others on a scale between external evaluation (considering the whole system) and internal evaluation (testing a component of the system). In addition a method could be classified along a dimension consisting of exploratory research versus experimental research. Though the proposed classification provide a simple yet robust way to select evaluation methods, the classification needs future work to add other dimension of formative and summative evaluation to the classification chart.

Another research has been done on evaluation solely on the usefulness of web based learning environment [10]. The research takes into consideration the multidisciplinary evaluation framework for evaluating a web based learning system. The framework combines two main issues, usability and utility issues where utility is broken down into two parts, pedagogical usability and added value. In [10], it shows the importance of usability remains the same regardless of how much web based learning is used in the teaching as a whole. As for pedagogical usability, its importance gradually increases as the focus of teaching shifted from traditional teaching into more on web based teaching.

In this research, we use summative evaluation where we address the educational impact of this system on students. We also use the multidisciplinary evaluation framework where we focus on the usability study of this system as [10] clearly shows that that the importance of usability is consistent regardless of the focus of teaching either more on traditional or more towards web based teaching.

3. Usability Study for ITS

Usability is a quantitative and qualitative measurement of the design of user interface grouped into five key factors: learnability, efficiency, memorability, errors and satisfaction [11]. The setting of the usability study can vary. A usability laboratory can be used for a controlled experiment. A workplace test can be used to test the user in their normal work environment such as at their desk during a routine work day. There is also web-based usability testing referred to as remote usability testing where the user and experimenter are not physically located in close proximity of each other. The different categories of usability tests consist of performance measurement, thinking aloud protocol, coaching method, retrospective testing, constructive interaction, and questionnaires [11, 12]

Performance measurement takes place when quantitative measures are taken during the testing such as the number of tasks completed successfully by the user, length of time to complete the test tasks, number of errors, and time spent recovering from errors. Thinking aloud protocol exists when users vocalize their thoughts and therefore share their positive and negative interpretations of different website features. The coaching method enables the users to ask questions and receive answers which give researchers insight into the type of help documentation or better technology design needed. Questionnaires are also a form of testing as it provides an opportunity to gather more usability feedback from a user after a testing session [13].

In this research, questionnaires were used as a tool to gather feedback from the participants and this research is conducted in a controlled environment setting. Three out of the five key factors, learnability, efficiency and satisfaction of the students were the focus of the study. Learnability of the students is based on the ease of use of the students when working towards completing the task specified for them. Efficiency looks at how productive the students once having learned the software and the last attribute satisfaction is to study the students level of pleasure using the system.

4. Methods and Processes

For this research, we have performed summative evaluation whereby a total of 44 computer science students in Faculty of Computer Science and Information System, Universiti Teknologi Malaysia have participated in this study. The learning material for the participants to study, task and processes of the usability study is discussed in the following subsections.

4.1 Learning Material

The learning material used in this study is based on sorting techniques which is explained fully in Malay language. The topics involved are bubble sort, selection sort, insertion sort, merge sort and quick sort. In order to handle different types of learning styles, the learning material for this topic is structured into theory, exercise, example and activity [14]. The theory part consist of explanation and pictures describing the sorting technique, exercise drill the students on the sorting techniques and gives their score and hints for the answer of each questions. Example shows the simulation of the sorting technique. Students can view the simulated sorting activity based on the execution of the algorithm code line by line. In the activity session, students can generate data to be sorted randomly and view the simulation of the sorting technique.

The learning material structure will be adapted based on student learning style. Different category of learner will get different tree structure. Figure 1 shows the sequence of the learning material based on the Extrovert-Intuitive students. For each learning material structure, links are provided for each topic in sorting.

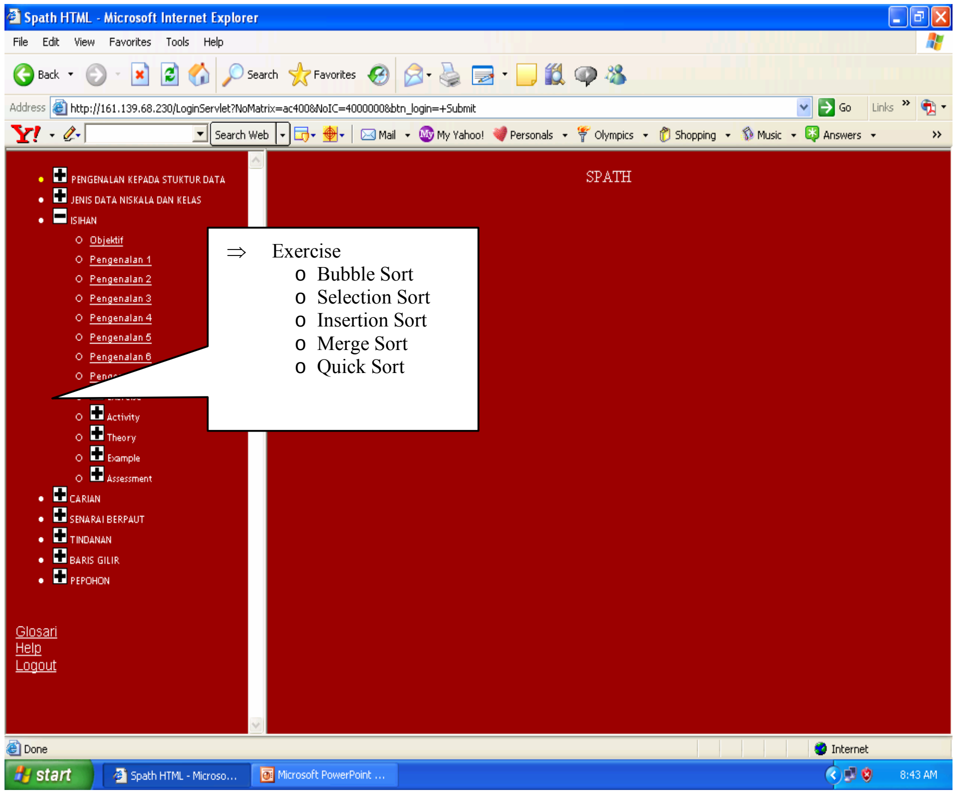


Fig. 1. The tree structure for learning sorting concept in support of Extrovert-Intuitive students

4.2 Task

The participants were given a task to study Sorting Techniques based on the learning materials discussed in section 4.1. As part of the study task, it was determined that the following steps would be the task process each participant would go through in the usability study:

- Hearing brief explanations about the system.
- Reading the task instructions.
- Reading short and simple user manual if they need further guidance about the system.

- Filling in the on-line questionnaire based on MBTI in order to determine the student's personality factor.
- Performing the task to study the learning materials.
- Completing the assessment on-line after they have completed the study session.
- Filling in questionnaires on learning satisfaction.
- Filling in questionnaire on usability of the system.

4.3 Processes

The entire task specified will be further refined and elaborated. The usability study is carried out in laboratory in a controlled environment where the students were required to complete their learning in the lab for two hours or until they finish their learning. Two experimenters were around to give explanations or help if needed. A desktop with Internet connection was provided to each participant to complete the study. The process of the study is as follows:

1. For this study, students were divided into two groups, the first group is required to learn based on the topic sequence provided and the second group can learn freely without following the sequence.
2. The students were given a short and simple user manual to guide them while using the system
3. Before the student can use the system, he/she has to register and fill in an online questionnaire. The information collected during registration were used to gain some personal information such as the user's name, id, cpa and programming knowledge status in order to initialize the student model.
4. On-line questionnaire based on MBTI is used to determine the student's personality factor value either as Introvert-Sensor, Extrovert-Sensor, Introvert-Intuition or Extrovert-Intuition. Figure 2 shows the sample of the on-line questionnaire. Data collected on the questionnaire were used to determine the learning style of the students using Fuzzy Logic.
5. The students started learning and they chose the learning material structure given to them either sequentially or freely.

6. After the students finished learning, they have to do the assessment on-line. The assessment is provided in order to measure the knowledge level of the students on the topic.
7. Upon the completion of learning and assessment, participants were required to fill up two questionnaires on learning satisfaction and usability of the system. The second questionnaire is adapted with slight changes from [15] with close ended and open ended questions concerning on user subjective satisfaction.
8. Analysis of the data collected during the study consists of data from on-line learning style, students learning satisfaction and usability of the system questionnaire. The results are discussed in the following sections.

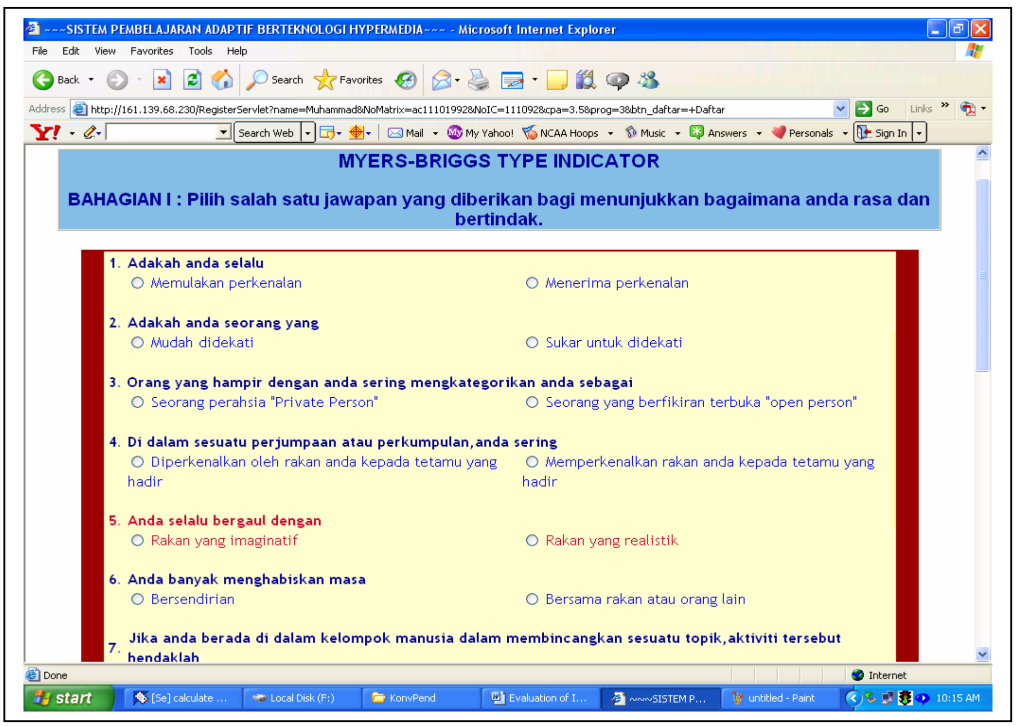


Fig. 2. Interface of the on-line questionnaire for personality factor based on Myers-Briggs Type Indicator (MBTI)

5. Results

The results elaborated in this section are mainly from the analysis of data collected from the two questionnaires based on the three key factors in usability study the learnability factors, efficiency factors and satisfaction factors. Apart from the three key factors, the students learning styles are determined based on the on-line questionnaire given in the system.

5.1 Learning Style Categorization

Based on the Myers Briggs Type Indicator (MBTI) personality factor, the students are categorized into four groups of learning styles. Figure 3 shows the number of students for each learning styles where 21 students are in Extrovert-Sensor group, 8 students are grouped in Introvert-Sensor, 4 students are in the Extrovert-Intuitive group and another 11 students 4 in the Introvert-Intuitive group.

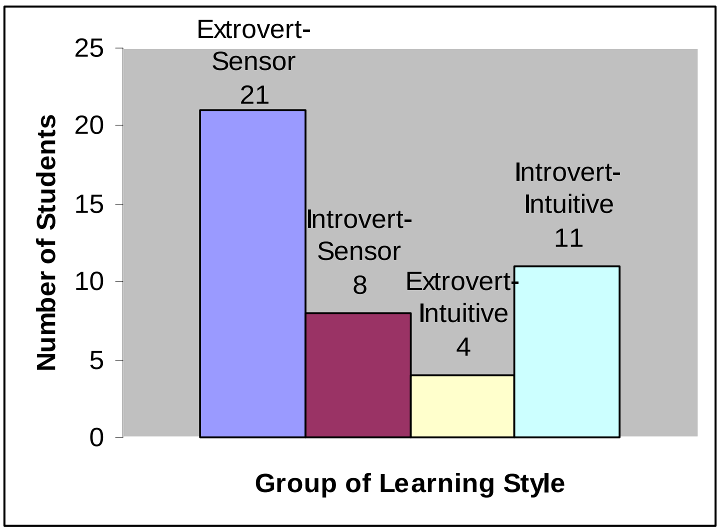


Fig. 3. Learning style of the participants

5.2 Learnability Factor

The learnability factor is measured quantitatively based on students' ease in completing the task of learning using the system. The learnability factor is measured based on these three factors:

1. The ease in learning based on personalization of the students learning style using the MBTI personality factor.
2. The ease in learning based on the learning structure given
3. The ease in learning based on the navigational path

Based on the questionnaire, 42 students were either strongly agreed or agreed on the effectiveness in learning based on MBTI learning style while only 2 students disagree in the learning style approach. The high learnability factor of this system is further proved by the number of students prefer the to learn based on the learning structure sequence given. The last element which contributes to the ease of using the system is the adaptive navigational paths provided by the system. 6 of the respondents strongly agree and another 30 respondents agree that the navigational paths provide ease of using the system. Only 8 of the respondents disagreed that annotated navigational paths provide ease of using this system. Table 2 shows the three features used to measure learnability factors of the system.

Table 2. Three features to measure learnability factor in using the system

	Strongly Agree	**Agree**	**Disagree**
Effectiveness of learning	4	38	2
Preferable learning structure sequence	7	27	10
Ease of using adaptive navigational path	6	30	8

5.3 Efficiency Factor

Efficiency factor of this system is measured by the productiveness of the students in terms of the score from the assessment they have completed in the system. The assessment has been done on-line by the students after they have completed their study using the system. Figure 8 shows that the number of he students with score between 80 to 100is quite high, 28 students while 11 and 2 students has either scored between 40 and 79 and between 0 and 39 respectively. The number of students with their respective assessment score is as shown in Table 3.

Table 3. Feature to measure efficiency factor in using the system

	Assessment Score		
	0-39 marks	40-79 marks	80-100 marks
No. of students	11	5	28

5.4 Satisfaction Factor

The final key attributes measured is the satisfaction factor of the user while using the system. The students satisfaction were determined based on attributes such as the familiarity and understandability of the terms used in the system, the easiness to understand the learning contents, the suitability of colors used in the system, the usefulness of the interactivity in the system and the overall satisfaction the user experienced when they use the system. Table 4 shows the number of students and their learning satisfaction in using the system.

The usability questionnaire have two open ended questions for the participants to make suggestions to enhance the application functionality and to express participants feeling while using the system

Table 4. Feature to measure satisfaction factor in using the system

	Strongly agree	**Agree**	**Disagree**
Participants learning satisfaction	4	38	2

Some excerpts of the feedback from the participants are:

1. The system is interesting because I can clearly see how the sort function working.
2. Interesting especially using the flash application but the yellow color for the background is giving me pressure.
3. Good, this is a new technique to improve the student in learning Data Structure.
4. Interesting because it is easy to understand the subject by using this system.
5. A little confusing at first but at the end I am amazed. It is easier to learn according to the learning style.

6 Discussion

From the results, system has achieved high percentage of learnability factor by the ease in using this system with the provided adaptive learning material sequence based on students learning style and also the ease in navigating of the system with the help of annotated links. Although half of the students have been given the task of learning freely regardless of the sequence of learning material, but more than half of the students (61.4%) have followed the sequence given by the system. This result showed that sequence of learning material structure suits their learning style thus helps them in their learning.

The second key attributes in this usability study is the efficiency factor. Efficiency is measured from the productiveness of the participants or in this study the assessment score of the students after learning using this system. Although only 11.4% of the students scored between 80 and 100, it is an acceptable figure considering students were expected to understand both the theory and algorithms of each sorting technique within two hours and also due to the detailed coverage of the assessment questions which really test the students in their understanding. The performance of the student can be improved if the web based usability study is adopted instead of controlled experiment. With web based usability study the participants can learn freely without time constraint [13].

The last key attributes, the satisfaction factor is measured based on the level of pleasure the participants experience while using this system. In this study, the results has showed high satisfaction factor whereby majority of the participants agree on the pleasurable attributes the system offer such as the understandable terms, the learning material structure and the interactivity in the system. Furthermore, the subjective feedback given by the participant demonstrate that learning material structure as in example and activity which shows the simulation of the algorithm and the programming really give them more understanding of the subject thus giving them satisfaction in using the system.

From this research point of view, the summative evaluation which addresses the educational impact of this system to student can be answered by the learnability factor which shows the suitability of the learning style in personalizing the learning material for the students. Another factor contributing to the educational impact of the system to the user is the learning material structure itself such as example and activity structures really provide the students satisfaction in learning using this system.

Even though the results of this study shows positive outcome in the usage of this system but there are also several weaknesses we encounter in conducting the usability study. The drawbacks are:

1. There are no comparison between students who have already learned data structure and students who are still learning this subject. If this type of comparison can be done, more interesting result can be analyzed.
2. In this research there is also no comparison done between the usages of this system with other similar system. In future, we decide to do a comparison study on the usage of this system with INSPIRE system [16].
3. There are comments from students particularly on the system choice of color which some of them thought as glaring and also mismatched. Other students were annoyed with the continuous system's automatic refresh where they have to reselect again the tree structure in order for all the learning material to be displayed again.

7 Conclusion and Future Work

This paper has discussed the evaluation of an adaptive hypermedia system by implementing the usability study. Three out of five key factors in the usability study has been measured. The method and processes of the usability study has also been elaborated. From the preliminary results and discussion it shows that this system has a high percentage for learnability and satisfaction factor thus these factors contributes to the answering of the question asked in summative evaluation: "What is the educational impact of this system to user?". For future work of this system, we intend to implement web usage mining in order to measure the efficiency of the learning approach given to the student based on their learning style.

Acknowledgement

This research is conducted under IRPA grant, project vote 74129: The Development of Knowledge Management for Adaptive Hypermedia Learning System.

References

1. P. Brusilovsky.. "Adaptive Hypermedia: From Intelligent Tutoring Systems to Web-Based Education (Invited Talk)". In G. Gauthier. C. Frasson and K. VanLehn(eds). Intelligent Tutoring Systems. Lecture Notes in Computer Science, Vol 1839: Springer Verlag, (2000) pp. 1-7.

2. Oppermann, Patel A. and Kinshuk.. "A Classification of Evaluation Methods for Intelligent
3. Tutoring Systems". (1999). Software Ergonomie '99.
4. P. Brusilovsky, C. Peylo. "Adaptive and intelligent web-based educational systems", International Journal of Artificial Intelligence in Education. (2003). (13) 156–169.
5. H. Alomyan. "Individual differences: implications for web-based learning design". International Education Journal.. 4(4). (2004) 188-196.
6. N. Ford, S.Y. Chen. "Individual differences, hypermedia navigation, and learning: An empirical study". Journal of Educational Multimedia and Hypermedia.. 9(4). (2000). 281-312.
7. G. Magoulas, K. Papanikolaou and M. Grigoriadou. "Adaptive web-Based learning: accommodating individual differences through system's adaptation". British Journal of Educational Technology.. (2003). 34(4).
8. Norreen Haron, Naomie Salim. "Integrating personality factor (MBTI) and Fuzzy logic techniques in individualizing learning material of adaptive hypermedia learning system". .(2005)Konvensyen Tenologi Pendidikan Ke 18. 16-19 September, Marang, Terengganu.
9. Bariah Yusob, Norreen Haron, Nor Bahiah Hj Ahmad, Shahliza Abd Halim. "Individualizing the learning material and navigation path in an adaptive hypermedia learning system. 2nd National Conference On Computer Graphics & Multimedia, Kajang, Selangor".(2004). 8-10 December.
10. M. Mark, J. Greer. "Evaluation Methodologies for Intelligent Tutoring Systems". Journal of Artificial Intelligence and Education, (19934) (2/3), 129-153.
11. K.Silius, AM. Tervakari. "An Evaluation of Usefulness of Web-Based Learning Environments. The Evaluation Tool into the Portal of Finnish Virtual University". International Conference on Network Universities and E-Learning, 8-9 May 2003, Valencia, Spain.
12. Nielsen, J.. Usability Engineering. Academic Press. (1993)
13. Lazar. User-Centered Web Development. Boston: Jones and Bartlett. (2001)
14. 13. D. Cartens, P. Patterson.. "Usability Study of Travel Websites". Journal of Usability Studies. Issue 1, Vol. 1, November 2005, (2005) pp 47-61.

15. Nor Bahiah Hj Ahmad, Shahliza Abd Halim, Bariah Yusob, Norreen Haron. “Development of an Adaptive Hypermedia Learning System”. Konvensyen Tenologi Pendidikan Ke 18. (2005) 16-19 September, Marang, Terengganu.
16. Granic, V. Glavinic and S.Stankov.. “Usability Evaluation Methodology for Web-based Educational Systems”. 8th ERCIM Workshop User Interfaces For All. (2004) 28-29 June
17. M. Grigoriadou, K. Papanikolaou, H. Kornilakis and G.Magoulas. “INSPIRE: An INtelligent System for Personalized Instruction in a Remote Environment”. Twelfth ACM Conference on Hypertext and Hypermedia (Hypertext'01). (2001) 14-18 August, Århus, Denmark, ,

DISCUSSION AND CONCLUSION

The fuzziness in individualizing learning style has motivates us in using Fuzzy Logic for adapting suitable learning materials to better suits student individual learning style. Another element of adaptation is the user model where in this research we use neural network to classify the students based on their knowledge acquisition captured in the user model. Our contributions on this research are inclusive of:

1. **The identification of learning content** (structure of learning material) for a given learner in online system where pedagogical structure is based on Myers-Briggs Type Indicator (MBTI) personality factor. Fuzzy sets and fuzzy rules acquired from standard of mastery where the criterion is referred based on lecturers experience and knowledge in student's learning style.

2. **The design and algorithm of neural network classifier** and also the **generation of rough set rules** for user model classification. The design of the neural network classifier includes training data consisting of input data and desired output data that will be used in classification using Kohonen neural Network.

3. **The combination of two artificial intelligence technique**s for user model classification using neural network and learning style determination with fuzzy logic. Both artificial intelligence techniques have been separately tested and Kohonen neural network has been proven to give more than 90% accuracy in classifying the learners data into categories. Fuzzy logic has also been verified to be an efficient way to reason on the students learning style based on the students personality especially when the students have equal learning style preference.

4. **The framework of the proposed AHLS** comprises of the architecture, learning structure model together with the design flowchart and also the development and implementation of the system has been explicitly represented for future reference.

5. **Summative evaluation** to address the educational impact of this system on students and its practical acceptability in terms of usability where three key factors i.e. learnability, efficiency and satisfaction were measured. Preliminary results of the usability study revealed that this system has a high percentage in learnability and satisfaction factor.

With the commercialization and also the virtualization trend in educational market, adaptive hypermedia learning system (AHLS) has big potential to be used as a learning material in private and government sector. Individualizing on-line education via AHLS can help students who are unable to attend courses in universities and also institutions specializing in distance learning. This research can be very useful to the researchers in e-learning field, educationist, private and government college universities and universities especially open universities. From the results and contribution, it is hoped that this research can contribute to the advancement of knowledge in the field of AHLS.

FUTURE WORK

This research can be further enhanced with the incorporation of more techniques or methods from both knowledge management and knowledge discovery areas. Among the proposed work that can be extended for further research are:

1. Open Adaptive Hypermedia System

Our adaptive hypermedia system works on closed materials whose adaptation functionalities and also learning materials are specifically tailored to our current Data Structure learning system. In order to commercialize the research product, the use of learning object has to be considered. With the learning object, the system can be more usable where not only data structure subject can be used for learning but in fact other kinds of subjects can be incorporated in the system. This research can lead towards better knowledge management methods and models as well as strategies to capture, reuse and maintain knowledge in the form of learning materials or user models or any other added resources.

2. Web Usage Mining

Web usage mining process is targeted on the extraction of user navigation patterns in web based environments. In this technique in knowledge discovery, usage patterns of various users and trends of usage are tracked and predictions are made about what users want. For educational on-line systems, web usage mining can be used to extracts students behavior patterns and tendencies in the use of educational site. By extracting access patterns that reveal students learning behavior, various types of analysis can be developed over the learning process and/ or learning environment. With this kind of knowledge discovery research, it provides tool for teachers and learning environment for designers to evaluate learning processes in web based courses.

PAPER I:

FUZZY LOGIC APPROACH TO EVALUATE STUDENT'S PREFERABLE LEARNING MATERIAL BASED ON STUDENT PERSONALITY FACTOR

Norreen Binti Haron
Software Engineering Lab
Department of Software Engineering
Faculty of Computer Science and Information System
Universiti Teknologi Malaysia
norreen.haron@gmail.com

Naomie Binti Salim
Faculty of Computer Science and Information System
Universiti Teknologi Malaysia
naomie@fsksm.utm.my

Abstract

AHLSs provide adaptation to user's learning styles. However, most of the AHLSs incorporate learning style based on the notion that each student has only one learning style which is not necessary true in the real life. Incorporating several learning styles of students into the system to allow for better matching of material to students cannot be done using crisp algorithms due to the fuzzy nature of the learning styles in each individual. In this research, a system based on fuzzy logic has been developed in which a student can have mixed traits of different styles, each with a certain percentage of membership, rather solely having one particular learning style. We also address problems in personalizing instructional material for representing learning material that matches the differences of students according to student's fuzzy membership to certain learning styles. There are four input and four output linguistic variables considered in this paper, where the inference rule of fuzzy reasoning consists of four antecedents and four consequents. The antecedents are represent based on the student's personality factor (Myers-Briggs Type Indicator (MBTI)); extrovert score, introvert score, sensor score and intuition score, and the consequents represented the student’s preferable learning material; (theory, example, exercise and activities). Based on fuzzy set and fuzzy rule theory, the vagueness and uncertainty intrinsically existing in the knowledge possessed by expert is computed whilst providing qualitative description of the input-output relationship using fuzzy linguistic terms. Triangle fuzzy set, Mamdani inference and center of gravity (COG) defuzzification techniques are used in the system. Comparison of the defuzzified values of the fuzzy rule base system and the value from conventional system shows that the fuzzy rule base system has better performance in structuring the learning material.

Keywords

Fuzzy logic, Fuzzy sets, Fuzzy ruled base, Learning style, AHLS.

1. Introduction

Research on learning has shown that student learns differently, they process and represent knowledge in different ways and they learn more effectively when taught with preferred methods. Information about learning style can help system become more sensitive to the differences students using the system.

Although learning style theory is widely accepted amongst educational theorists in the context of e-learning environments [1,2,3,4,5], there is still no research on the adaptation to individualize student’s learning material based on fact that students have more than one learning style in a certain degree. In particular the possibility of fluctuations in a learning style with changing tasks or content has not yet been addressed [6].

In this paper, a system based on fuzzy logic has been developed in which a student can have mixed traits of different styles, each with a certain percentage of membership, rather solely having one particular learning style. It aims to utilize the learning characteristics and provide a personalized learning environment, that exploit learning style and fuzzy logic techniques. We focus on using the fuzzy rule-based system that involved fuzzy sets and fuzzy logic. The learning style in this paper refer to the student’s personality factor; Myers-Briggs Type Indicator (MBTI) [7, 8]. Based on the MBTI theories, the fuzzy logic techniques are then use to classify the student’s preferable learning material.

In designing the fuzzy logic system, it is important to identify the main control variables and determine the term set which is at the right level of granularity for describing the values of each linguistic variable [9]. In this problem, the fuzzy system is represented by four input linguistic variables (or the antecedents), and four output linguistic variables (or the consequents). Also, each input and output may be represented by either a three-term-set or a five-term-set of linguistic values. After defining the fuzzy variables and its term sets, fuzzy rule base is then being constructed. The number of fuzzy rules being formed is directly related to the number of fuzzy term sets defined at the antecedents. At the end of the process, the crisps output shows the structure of the learning material, which is learning material that the student most preferable and which learning material that the student choose less attention.

2. Approach and Method

Fuzzy logic method, proposed by Zadeh [10], has proved to be very effective in handling the vagueness and uncertainty intrinsically existing in the knowledge possessed by people. Fuzzy rules based in fuzzy logic provide a qualitative description of the input-output relationship of a system using fuzzy linguistic term. Moreover, fuzzy linguistic rule appear close to human reasoning and in many real-world applications, and it is more adequate and flexible for knowledge representation than conventional crisp IF THEN rules. Thus, it is reasonable to use fuzzy logic system to classify students and determine the most suitable learning material for them.

Fuzzy logic techniques in this research are used to personalize the learning material where it denotes to structures of learning material. It calculates precisely the structures of learning material (theory, example, exercise and activities) that suit student's personality.

In this study, the fuzzy sets and fuzzy rules are constructed based on the standard of mastery or the criterion-referenced acquired from the human instructors' experience and knowledge about student's learning style.

The processes that implement in fuzzy logic system indicate 4 main stages; fuzzification, rule evaluation, aggregation and defuzzification. The fuzzification stage transforms crisp student personality, captured from pre-course questionnaires, into suitable linguistic values. The rule evaluation stage takes the fuzzified inputs, and applies them to the antecedents of the fuzzy rules. The aggregation stages use to combine all the fuzzy output derived from fuzzy rules, and last step, the defuzzification stage produces the output in crisp value. In this problem, the defuzzified value is derived based on the centre of gravity method. Figure 1 shows the flow of the fuzzy logic system.

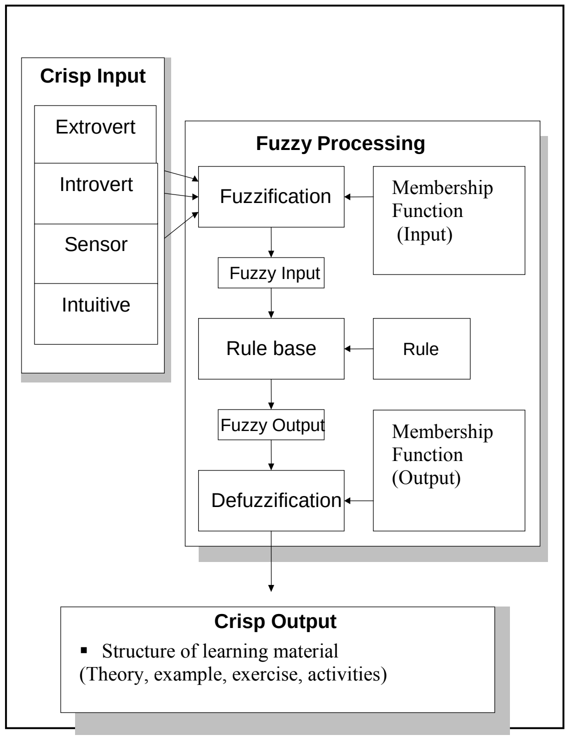


Figure 1: Flow of Fuzzy Logic System

In this paper, there are four inputs and four outputs of linguistic variables are being considered. The input linguistic variables are representing the student personality's scores; extrovert scores, introvert scores, sensor scores and intuitive scores.

The input expressed by:

$(x_a, x_b, x_c \text{ and } x_d)$. (1)

Whilst, the outputs are expressed by:

$(y_a, y_b, y_c \text{ and } y_d)$ (2)

that represents the student's acceptance level of learning material; theory, example, exercise and activities.

To identify student personality's scores as fuzzy numerical data in ranges (0.0, 1.0), all the score retrieve from pre-course questionnaire are gathered according to each personality and are calculated by dividing the total score of particular personality, s_i, with the total number of pre-course questionnaire answered for that particular personality, P, as follows:

$$X_a = \frac{\sum_{i=1}^{P} s_i}{P} \quad (3)$$

We also determine the membership function of fuzzy set in fuzzification stages. The membership functions of fuzzy set in this paper are based on Zimmerman [11] which is allows the element in set to have multiple degree of membership. For example:-

$$\tilde{A} = \{(x, \tilde{A}(x)) \mid x \text{ e } R\}) \quad (4)$$

where x={100, 90, 80, 70, 60, …, n}

then $\tilde{A}$ ={(100, 1), (90, 0.9), (80, 0.8),
(70,0.7), (60, 0.6),…, (n, $\tilde{A}(x)$)}

-R is student,

-x is student personality's value,

-$\tilde{A}$ (x) is membership function x in $\tilde{A}$,

-($\tilde{A}$) is a fuzzy set for personality's value x with student's acceptance level of theory learning material.

The membership value 1, appoint to the student personalities x=100, where the student's acceptance level of theory learning material at the highest level. The student's acceptance level will decrease when number in personality x become lower.

In the second stages, rule evaluation, the inference rule of fuzzy reasoning consists of multiple antecedents and multiple consequents are expressed as below:

Ri :

IF X_a is N_1 and X_b is N_2
and X_b is N_3 and X_d is N_3
THEN
Y_1 is M_1 and Y_2 is M_2
and Y_3 is M_3 and Y_4 is M_4

where Ri, (i = 1, 2, … n) is the rule number, Ni, (i =1, 2 and 3) are the membership functions of the antecedent part, Mi, (i =1, 2, 3, 4 and 5) are the membership functions of the consequent part. The example of fuzzy rules applied in this problem as show in Table 1.

Input:-
{E-extrovert, I-introvert, S-sensor, N-intuitive}
Output:-
{1-theory, 2-example, 3-excersice, 4-activity}.

Table 1: Example of Fuzzy Rules

Input- (student personality)				**Output - learning Material**			
E	**I**	**S**	**N**	**1**	**2**	**3**	**4**
Low	High	Low	High	VHigh	High	Low	VLow
Low	High	Med	Med	High	Med	Low	Vlow
Low	High	High	Low	Med	VHigh	VLow	Low
:	:	:	:	:	:	:	:
High	Low	High	Low	VLow	Low	High	VHigh

In this paper, the fuzzy set is expressed by a triangular function as triangular function can provide an adequate representation of the expert knowledge [12] and at the same time significantly simplifies the process of computation. The fuzzy set is expressed by three parameters {a, b, c} as shown in figure 2.

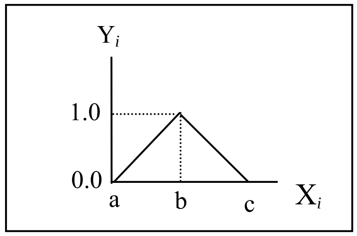


Figure 2- Triangular fuzzy set

The membership value is derived by the following formula as shown in Figure 3.

$$
Y_i = \begin{cases} \dfrac{X_i - a}{b-a} & \text{if } a <= X_i < b \\ 1 & \text{if } X_i = b \\ 1 + \dfrac{b - X_i}{c-b} & \text{if } b < X_i <= c \\ 0, & \textit{otherwise} \end{cases}
$$

Figure 3 - Triangular membership function

where a, b, and c are parameters that control the intervals of the triangle (as shown in figure 2), X_i is the input (i =1, 2, 3, 4), and Y_i is the output of fuzzification input i.

Overall process in second stages is expressed as following definition:-

$$\acute{L}i(z):=((R\alpha i = [Ai(x_{a1}) \square Bi(x_{b1})]) \square Li(z)) \quad \text{for } i= 1,2,3\ldots n \qquad (5)$$

Ĺi(z) is fuzzy output retrieved from selected rule (Rα) based on the rule evaluation , input data, antecedent (xa,xb,xc and xd) and output data, consequent, Li(z).

The maxima method in aggregation stage is based on the definition shown in equation (6):-

Ļ (z) = Ĺ1 (z) □ Ĺ2 (z)

= (Rα1 □ L1 (z)) □ (Rα2 □ L2 (z)) (6)

Ļ(z) are fuzzy set expressed from combination of all consequent values retrieved from selected rule, Ĺi(z), (i= 1,2,3…n), where Rαi is selected rule and Li(z), (i= 1,2,3…n), is fuzzy consequent value.

The defuzzified value of the fuzzy reasoning, are derived based on the Mamdani-style inference (centre of gravity-COG) as shown in equation (7) below:

$$COG = \frac{\sum_{x=a}^{c} Zb(x)\, x}{\sum_{x=a}^{c} Zb(x)} \qquad (7)$$

3. Experiment and Result

Table 2: The Comparison between Fuzzy rule base system and Conventional System

Data Input (student personality) %				**Fuzzy rule base system**				**Conventional System**			
E	I	S	N	Theory	Example	Exercise	Activity	Theory	Example	Exercise	Activity
30	70	60	40	0.554	0.683	0.317	0.446	0.000	1.000	0.000	0.000
46	54	40	50	0.526	0.470	0.500	0.428	1.000	0.000	0.000	0.000
50	30	80	20	0.187	0.569	0.25	0.750	0.000	0.000	0.000	1.000

Comparison of the defuzzified values of the fuzzy rule base system and the value from conventional system shows that the fuzzy rule base system present better performance in structuring the learning material. As shows in table 2, the conventional system can't structure the learning material and shows only one learning material to the user even though the user have mixed traits of different styles. Differ from fuzzy rule base system; the students can have different learning material according to their mixed traits of different styles, with a certain percentage of membership rather solely having one particular learning material to learn. Moreover, the result also shows that the learning material could be structure according to the defuzzified result.

4. Conclusion and Further Work

This paper has proposed a way to personalize the learning material for AHLS, which aims to provide learners with a customized learning environment. It emphasizes the combination of learning style theories and artificial intelligent techniques. Fuzzy logic techniques used to impart the learning content based on student's fuzzy personality data and instructional rules in order to support customisation that will allow learners to learn faster and understand the learning material much easier. Conversely, the challenge is to identify what those learning content (structure of learning material) for a given learner in online system based on student personality and which result is more precise to the learner's personality whether the traditional approach or using the fuzzy logic techniques.

For further work, the author suggested to test the evaluation and performance of this theory, the author will conduct two assessments; testing the performance of effectiveness and testing the accuracy of the system. In the first evaluation, the author divide the student into two groups where as both group have same knowledge level. The first group will use the system that proposes in this theory while other group will use a system without including this theory. Both groups will take a test before and after using both systems. The result from the test would show which of both group perform good result. This is base on the result from student's test after using the systems.

The second evaluation is to test the accuracy of the system. This test is use to test the precision of the system with the student choice of learning material. A prototype system will be build. In this prototype system, a sub topic of data structure subject will be show in different method whereas the method is base on the teaching style. Student must first answer a questionnaire before using this prototype system.

This questionnaire is use to identify the learning style of the student. In using the prototype system, student will be show several button that link to each of the teaching method. Based on the student questionnaire and student preferable of teaching method, the system will recognize the relationship between student learning style and the teaching style. The result will be use to compare with the result in fuzzy logic techniques. This comparison is use to find the precision of fuzzy logic techniques with student preferences.

5. References

[1] N. Stach, A. Cristea and P. De Bra. "Authoring of Learning Styles in Adaptive Hypermedia." in Procceding. of WWW Conference. 2004.

[2] J.E. Gilbert and C.Y. Han. "Adapting Instruction In Search Of a Significant Difference". Journal of Network and Computer applications. 1999. 22. 3.

[3] M. Kwok and C. Jones. "Catering For Different Learning Styles, Association for Learning Technology". -Association for Learning Technology Journal-(ALT-J). 1985. 3(1). 5-11.

[4] M. Grigoriadou, K. Papanikolaou, H. Kornilakis and G. Magoulas. " INSPIRE: An Intelligent System For Personalized Instruction In A Remote Environment". In Proceedings of 3rd Workshop on Adaptive Hypertext and Hypermedia . 2001. pp. 13-24.

[5] P. Paredes and P. Rodriguez. "Considering Learning Styles in Adaptive Web-based Education". The 6thWorld Multi Conference on Systemics, Cybernetics and Informatics. 2002. vol. II. pp. 481-485.

[6] C. Wolf. "iWeaver: Towards 'Learning Style'-based e-Learning in Computer Science Education. ACE2003". Conferences in Research and Practice in Information Technology. 2003.

[7] S. Carolyn. et. al. "Myers Briggs Type Preferences in Distance Learning Education". In International Journal of Educational Technology. 2001. 2(2).

[8] C. C. Bishop. et. al. "The Myers-Briggs Personality Type and Its Relationship to Computer Programming." Journal of Research on Computing in Education. 1994. 26(3). 358 -371.

[9] H. R. Berenji. “Fuzzy Logic Controllers In Yager R.R. & Zadeh L. A., An Introduction to Fuzzy Logic Applications in Intelligent System”. Kluwer Academic Publisher, 1992.

[10] L. A Zadeh. “Fuzzy Logic for Management of Uncertainty”. John Wiley & Sons, New York.1992.

[11] H.J. Zimmerman, “Fuzzy Sets Theory - and Its Applications”. Kluwer Publishing Company. 1985.

[12] M. Negnevitsky. “Artificial Intelligence. A Guide to Intelligent Systems”. First Edition, Pearson Education. 2002.

PAPER II:

INDIVIDUALIZING THE LEARNING MATERIAL AND NAVIGATION PATH IN AN ADAPTIVE HYPERMEDIA LEARNING SYSTEM

Bariah Yusob[1], Norreen Haron[2], Nor Bahiah Hj Ahmad[3], Shahliza Abd Halim[4]

[1] byvail@catcha.com , [2] n_reen@hotmail.com , [3] bahiah@fsksm.utm.my & [4] shahliza@fsksm.utm.my

Department of Software Engineering

Faculty of Computer Science and Information System

Abstract

This research aims to develop a prototype of an AHLS for teaching and learning Data Structure for university students. Our system integrates pedagogy of education and intelligence technique in the presentation of the learning material and the suggested navigation path. The system comprises of three main components, user profile model, domain model and adaptive engine. User profile model stores the learning activities, learning performances and interaction history of each student in the database. Adaptive navigation path will provide the annotated link based on the performance and the interaction history of each student. To reduce disorientation, each student will get different paths based on their level of knowledge acquisitions and learning style. Adaptive engine will determine the appropriate learning material and the navigation path based on the student's status that was retrieved from the user profile model.

The focus of this paper is on the use of computational intelligence technique in the classification of student models and in the adaptation of learning material and navigation path. Kohonen self-organizing maps is used to classify the student's status using simulated data. We found out that Kohonen was able to cluster students accurately based on the maps assigned. Meanwhile, the domain model focuses on the uses of fuzzy logic to dynamically adapt the choice of possible paths through the learning material based on the attributes captured in the student model. The material presented to the student is adapted based on the students learning style and

performance. By adapting the user both at presentation and navigation level, we hope that this study can solve disorientation and lost in hyperspace problem that usually occur in conventional hypermedia learning system.

Keywords: Adaptive hypermedia learning system, personalization, user profile, self-organizing map, learning style.

1.0 Introduction

The Adaptive Hypermedia Learning Systems HLS (AHLS) is a computer based learning system in which interactive and dynamic learning module is customized to each student. Research on learning has shown that each individual student learns differently and processes and represents knowledge in different ways. Therefore, it is important to diagnose the learning style because some students learn more effectively when taught with preferred methods. Information about the learning style can help system become more sensitive to the differences students using the system. Several systems adapting the learning style have been developed to date; however, it is not clear which aspects of learning characteristics are worth modelling, how the modelling can take place and what can be done differently for users with different learning style [Brusilovsky 2001]. There are serious consequences when student learning styles and teaching styles do not match. One of it, the students face on difficulties to understand what is being taught, this lead to decrease of student interest to continue study in the subject and the student need a long term to finish one lesson session [Hashim and Yaakub 2003].

This research aims to develop a prototype of an AHLS for teaching and learning Data Structure for university students. Our system integrates pedagogy of education and intelligence technique in the presentation of the learning material and the suggested navigation path. The system comprises of three main components, user profile model, domain model and adaptive engine. User profile model stores the learning activities, learning performances and interaction history of each student in the database. Adaptive navigation path will provide the annotated link based on the performance and the interaction history of each student. To reduce disorientation, each student will get different paths based on their level of knowledge acquisitions.

Adaptive engine will determine the appropriate learning material and the navigation path based on the student's status that was retrieved from the user profile model.

Domain model stores all the teaching materials including the learning objectives, lecture notes, examples, exercises and the answer for each question. To adapt to the user category, the flow of the learning material for each category of the student will be different. To assist the user in terms of navigation, an individualized navigation path is constructed for each student, suggesting the path or link that the student has already learned, forbidden (the prerequisite is not fulfill), ready to be learned and need revision. Each link for each node will be annotated in different colors suggesting the depth of knowledge that the user already acquired. This way, the user can choose which node has greater priority to be learned and thus enabling him to optimize the path he plan to explore while studying. By adapting the user both at presentation and navigation level, we hope that this study can solve disorientation and lost in hyperspace problem that usually occur in hypermedia learning system.

In this paper, a framework for learning path personalization in adaptive learning system is introduced. It aims to utilize the learning characteristics and to provide a personalized learning environment, that exploit pedagogical model and fuzzy logic techniques. The pedagogical model and learning style are referring to the student's personality factor; Myers-Briggs Type Indicator (MBTI) [Carolyn et al. 2001; Bishop et al. 1994]. Based on the MBTI theory, the fuzzy logic techniques are then use to classify learning material (structure of learning material, type of learning material and additional link).

Fuzzy set theory, proposed by Zadeh [Zadeh 1992], has proved to be very effective in handling the vagueness and uncertainty intrinsically existing in the knowledge possessed by people or implied in numerical data. Rules based on fuzzy logic provide a qualitative description of the input-output relationship of a system using fuzzy linguistic terms. Fuzzy linguistic rule appear close to human reasoning and in many real-world applications, and thus it is more adequate and flexible for knowledge representation than conventional crisp IF THEN rules. This is the major reason why fuzzy classification rules are adopted in this paper.

2.0 Approaches And Methods To Implement Adaptivity

In the construction of an AHLS, the first issue that needs to be considered is how to identify the user features and to develop the content that reflects the personality's principles. The proposed architecture in this paper is based on this question. The architecture involved three main phases, as can be seen in Figure 1.

Based on Figure 1, the user profile model stores the information about learners in the learning system. The profiles were extracted from both explicit and implicit user profile. The explicit information is the information that the learner gave willingly or directly and he/she is aware that the information is kept in the database. The implicit information is the information the system collects without the learner acknowledgement. It records the learner's activity and behavior as he/she navigates through the system.

In this work, we test the learner's knowledge by giving them some exercises to be completed after finishing a concept. We keep the score that represent the explicit data because learner has to finish the exercises to gain scores. The implicit data used are the learning time, number of backtracking and number of getting help. From these data, we use Kohonen network to classify the learners' into three categories as shown in Figure 1.The process of identifying the learner's learning features is difficult [Brusilovsky 2001]. Moreover, it is not clearly defined which aspect in learning features that really useful for learner's modelling. Besides, the process of developing and identifying the learner's attributes in the learner's model will take a very long time. Therefore, we use a simulated data that represent the actual learner's data.

In the second phase, the pedagogical framework is used as a guide in presenting a good learning strategy. A good learning strategy is influences by the combination of learning approach, method and techniques. This pedagogical framework is derived from pedagogical expert model. In this paper, the learning strategy is base on the MBTI personality factor, whilst the method and techniques are base on Howard Gardner theory [Dara-Abrams 2002; Gardner and Korth 2001] and Honey & Mumford theory [Schroeder 1993]. The MBTI personality factors that use in this paper indicate of four types; Extrovert (E); Introvert (I); Sensing (S) and Intuition (N). Extrovert and introvert are illustrating the student preferable condition in focusing attention while the sensing and intuition type, illustrating the student preferable way in taking information.

The extrovert students prefer and focus on the outer activity and are energize by interaction with others. They prefer to talk, participate and interaction with people. While the introvert students prefer and focus on the inner activity. They prefer reading, listening to others and writing. Sensing student prefer concrete information, facts and procedures. They are good in memorization and like to go systematically, they also learn best with instruction that allow them to use their sense. Intuition student prefer discovering possibilities and relationship. They also like courses that involve a lot of experimentation and experiences.

The pedagogical framework comprise the steps on how content is develop to reflect those personality principles. Table 1 shows the relationship between learning strategy and learning method.

Table 1: The Relationship between Learning Style and Method.

Learning Style (MBTI)	**Method (Howard Gardner and Honey & Mumford)**
Extrovert	1. Visual 2. Kinaesthetic 3. Interpersonal
Introvert	1. Verbal/Linguistic 2. Intrapersonal
Sensing	1. Verbal/Linguistic 2. Intrapersonal
Intuitive	1. Logical-mathematic 2. Kinaesthetic
Extrovert-Sensor	- Experiment
Extrovert-Intuitive	- Exercise
Introvert-Sensor	- Example
Introvert-Intuitive	- Theory

3.0 User Classification Based On Kohonen Network

Kohonen network has been widely used for the classification purposes and it produced an excellent results [Cho 1997; Vendlinski and Stevens 2000]. Basic Kohonen algorithm such as Vector Quantization or k-means clustering can be used as a simple classifier [Sarle 1994].

The Kohonen's self-organizing map (SOM) was introduced by Professor Teuvo Kohonen at the University of Helsinki in 1982. The idea was to create a neural network to represent the input space using the topological structure of a grid to store neighborhood relations. In contrast to most neural network methods that use the desired result to compute the weights of the network, SOMs need no reference output themselves (unsupervised learning).

A SOM defines a mapping of an n-dimensional input space *R* to an m-dimensional output space *C* (we use m=2). The output space consists of N_c neurons. They are embedded in the topological structure of *C*, which may be an m-dimensional grid or any other graph structure. To each neuron of the output space, a parametric weight vector in the input space $w_s = [\mu_{s1}, \mu_{s2}, \ldots, \mu_{sn}]^T \in R$ is associated.

Eq. 6 define the mapping $\varnothing$ from the input space *R* to the topological structure *C*:

$$\varnothing_w : R \rightarrow C \,,$$

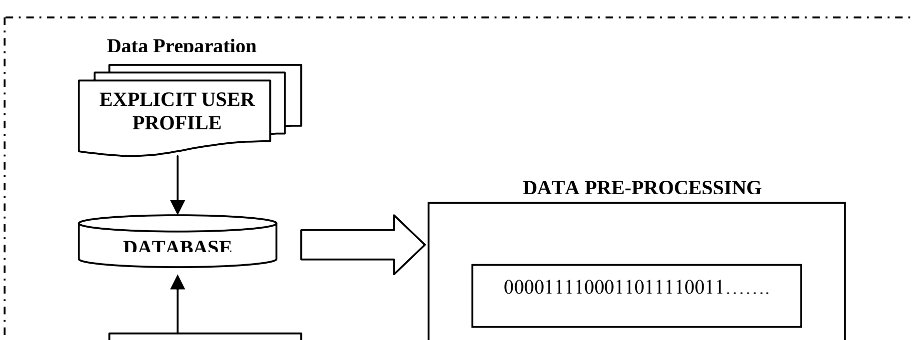
$$x \rightarrow \varnothing_w(x) \,, \qquad (1)$$

where,

$\varnothing_w(x)$ is defined as,

$$\varnothing_w(x) = \arg\min_i \{\| x - w_i \|\} \,. \qquad (2)$$

Every input sample is mapped to the neuron of the output layer whose weight vector is closest to the input.



$$d(x) = \sum_{x \in I} \sum_{i \in C} h_{ci} \| x - w_i \| \frac{1}{|I|}, \qquad (3)$$

Figure 1: The Architectures of the AHLS.

The goal of Kohonen learning process is to reduce the *distortion error:*

$$d(x) = \sum_{x \in I} \sum_{i \in C} h_{ci} \| x - w_i \| \frac{1}{|I|}, \qquad (3)$$

where,

I denotes the set of input samples,

h_{ci} denotes the neighbourhood relation in C between neuron i and the best matching one c.

3.1 Experiment with Self-Organising Maps (Kohonen)

Table 2 shows the attributes and the values defined as simulated data. To develop a simulated data, we use the criteria shown in Table 3. When the learner log into the learning system, the system will count the time he/she spent on learning a concept. The system suggests the time to be spent on each concept. Learning time is calculated based on the percentage the learner takes to finish learning from the suggested time. The learner's learning time is calculated as follows:

Suggested time for n concept = 1200 seconds

Total time spent by learner a = 900 seconds

$$\text{Percentage} = \frac{900}{1200}\times 100$$

$$t = 75.00.$$

Table 2: Simulated data

Attribute	Value
Learning time, t	0.00 – 100.00 %
Number of backtracking, b	0 – 5
Number of getting help function, h	0 – 5
Score, s	0.00 – 100.00 %

Table 3: Criteria for learner's classification

Attribute	Beginne	Intermediate	Advance

	r		d
Learning Time, t	$t > 80\%$	$30\% \le t \le 80\%$	$t < 30\%$
Numb. Of Backtracking, b	$b > 4$	$2 \le b \le 4$	$b < 2$
Numb. Of Using Help, h	$h > 4$	$2 \le h \le 4$	$h < 2$
Score, s	$s < 30\%$	$30\% \le s \le 80\%$	$s > 80\%$

The number of backtracking shows that the learner is not fully mastering the concept, lose direction or change his/her learning goal. The number of backtracking is defined by counting how many times the learner reopen any pages in particular concept. In this research, help function is a list of definition and explanation on terms used in the notes given. This attribute shows that the more help the learner gets, the more he/she is having a difficulty in understanding a concept. The number of getting help is defined by counting how many time the learner click on the help button in particular concept. To test the learner's level of mastering, the system provides an exercise at the end of the learning period. The score is calculated by the percentage of correct answers given. The SOM structure is defined as shown in Table 4.

All the data must be transformed into a standard format to get a valid and accurate classification. The transformation of the data is included in the pre-processing phase using a normalization method. We used a normalization method that was defined by [Rao 1995] as follows:

$$X_n = \frac{1}{\sqrt{\sum (x_n)^2}} \times x_n \qquad (4)$$

where, X_n is the input data for n.

Table 4: Parameter settings for the SOM training

Parameter	**Value**
Size	10x10
Dimensionality	2
Shape	Sheet
Map lattice	Rectangular
Neighbourhood	Gaussian
Learning rate	0.5
Iteration	5000
Size of training sample	1050
Size of testing sample	450

In the training phase, input data is given to the Kohonen network. The weights are captured after completing the training phase. The size of training sample is 1050. In the testing phase, there is no target data is provided. We used 450 dataset to the network. The network classifies the data based on the weights and outputs are obtained. When the testing results were obtained, the percentage of the classification accuracy was calculated.

3.2 Result of the Experiment

The training process does not consist of the class of learner. The map shown in Figure 2 is the mapping of the weights produced from the network learning through the data sample given. During the training, the network learns the data and generates the weights by calculating the nearest distance to the real data presented. The number 0, 1 and 2 are the representation of the classes defined during data simulation whereby 0 represents class for beginner, 1 for intermediate and 2 for advance.

```
2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 1
2 2 2 2 2 2 2 2 1 1
0 0 2 1 2 2 2 1 1 1
0 0 0 0 2 0 1 1 1 1
0 0 0 0 0 1 1 1 1 1
0 0 0 0 0 1 1 1 1 1
0 0 0 0 1 1 1 1 1 1
0 0 0 0 1 1 1 1 1 1
0 0 0 0 1 1 1 1 1 1
```

Figure 2: Result Map From Network Training

Table 5 shows the accuracy percentage of network classification in testing process. From 450 data presented, the networks are able to classify 445 data correctly. From the result shown, we conclude that Kohonen network is capable of classifying the learners' data into categories. It gives more than 90% accuracy in both training and testing phase. The Kohonen's SOM is definitely a good tool to classify data into a number of groups without supervision. It will be very useful in this study because it can deal with more complex and bigger sample of data when it is applied to the real learners' data in the learning system's database.

Table 5: Result from network testing

Numb. Of Correct Classification	Numb. Of testing data	Accuracy
445	450	98.89 %

4.0 Fuzzy Approach

In the second phase, fuzzy techniques are used to personalize the learning path where it denote to 3 output, structures of learning material, type of learning material and additional link. The first output is to calculate precisely the structures of learning material. The learning materials are structured in the form of theory, example, exercise and experiment that suit student's personality. Next stage is the calculation of the type of learning material that suit to the students whether the student prefer more visual or linguistic learning material. The last output is the additional link that link to e-mail and forum link. Student need to answer questionnaires before start learning process. This process is the same as the traditional process. The differences are in terms of the techniques in personalizing the student learning material. The traditional techniques totally use the result from the questionnaires to personalize the learning material. The highest result from the student's questionnaires will be considered as the student's personality and the learning method is referring to the personality. Meanwhile, the fuzzy system will calculate the student's personality and propose the suitable learning method to the student. The problem that occurs in traditional techniques is producing the suitable learning method once the students have equal result in their personality.

Fuzzy logic is computationally undemanding and is most suitable for processing imprecise input data, as it supports natural description of knowledge and reasoning in the form of imprecise concepts, operators and rules [Negnevitsky 2002]. In ITS, fuzzy logic techniques have been used due to their ability to handle imprecise information, such as student's actions, and to provide human descriptions of knowledge and of student's cognitive abilities [Stathacopoulou et al. 1999]. The fuzzy logic system consists of three main stages: Fuzzification, Rule Evaluation and Defuzzification. The fuzzification stage transforms crisp student's personality data, captured in the student database, into suitable linguistic values.

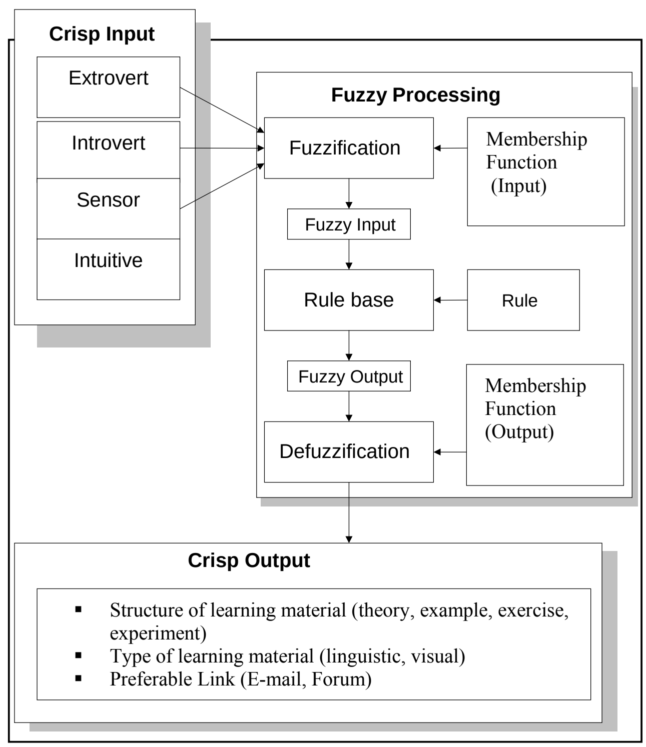


Figure 3. Architecture of fuzzy system

In this paper, the linguistic variables are based on the MBTI where it provides 4 linguistic variables; (Extrovert, introvert, sensor and linguistic). Whilst the fuzzy set that use for this paper is triangular fuzzy set where it hold 3 parameter {a, b, c} and can be seen in table 3. The rule evaluation stage takes the fuzzified inputs, and applies them to the antecedents of the fuzzy rules. The defuzzification stage produces the output in crisp value. In this problem, the defuzzified value is derived based on the maxima and sum method in Mamdani-style inference. The suitable defuzzified will be choose appropriately based on the result.

The triangular membership function can be specified by three parameters {*a, b, c*} as shown in figure 4: In figure 5, *a*, *b*, and *c* are parameters that control the intervals of the triangle, x_i is the input, and y_i is the output of fuzzification input *i*.

$$y_i = \begin{cases} \dfrac{x_i - a}{b-a} & \text{if } a <= x_i <= b \\ 1, & \text{if } x_i = b \\ 1 + \dfrac{b - x_i}{c-b} & \text{if } b < x_i <= c \\ 0, & \textit{otherwise} \end{cases}$$

Figure 4. Triangular membership function

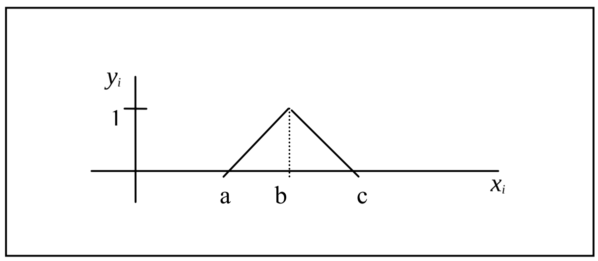


Figure 5. Triangular fuzzy set

The next process is the process to generate the fuzzy rule-base. This process classified the student learning material according to the student personality. The fuzzy rule based system is consisting of IF/THEN statement and combine with AND/OR operation as the example in figure 6.

First Rules:	IF (a is A1 AND (b is B3) AND (c is C1) AND (d is D3) THEN (e is E2), (f is F5), (g is G4), (h is H2), (i is I1), (j is J1) and (k is K3)

Figure 6: Fuzzy Rules Evaluation

In this figure a, b, c and d each of it represent the Student's Personality (extrovert, introvert, sensor and intuitive. Whereas (A1,A2,A3), (B1,B2,B3), (C1,C2,C3), (E1,E2,E3), (J1,J2,J3) and (K1,K2,K3) represent the input /output {High, Medium, Low}, and (F1,F2,F3,F4,F5), (G1,G2,G3,G4,G5), (H1,H2, H3, H4,H5) and (I1,I2,I3,I4,I5) represent the output. Table 6 illustrates more examples of the rules.

Table 6: Example of fuzzy rules

Input				Output						
Extrovert	Introvert	Sensor	Intuitive	Inter Personal	Theory	Example	Exercise	Experiment	Visual	Linguistic
Low	High	Low	High	Low	VHigh	High	Low	VLow	Low	High
Low	High	Med	Med	Low	High	Med	Low	Vlow	Med	High
Low	High	High	Low	Low	Med	VHigh	VLow	Low	Med	High
:	:	:	:	:	:	:	:	:	:	;
High	Low	High	Low	High	VLow	Low	High	VHigh	High	Low

The defuzzification process used in this paper is Mamdani inference-style where it involves difference operation of defuzzification. The best defuzzification operation will be selected. The criterion of selecting the best defuzzification is base to the most similar result that fulfils the expert expectation. Based on the result, the defuzzifications Centre of Area/Gravity (COG) have the most similarity. For example, see table 7 below.

5.0 Conclusion And Further Work

This paper has proposed a way to personalize the course content for AHLS, Which aims to provide learners with a customized learning environment. It emphasizes the combination of pedagogical theories and artificial intelligent techniques. It is important to note that for a given dataset and defined SOM properties, the SOM training process is dependent on the learning parameter settings.

Table 7: Comparison of Mamdani defuzzification operation with aggregation – max (input introvert =0.7, extrovert =0.3, sensor = 0.6, intuitive = 0.4)

	Exercise	Example	Theory	Experiment	Inter Personal
Mamdani Defuzzification (aggregation max)					
COA/G	0.317	0.683	0.446	0.554	0.354
Bisector	0.29	0.71	0.39	0.6	0.29
MOM	0.245	0.75	0.245	0.75	0.13
LOM	0.33	0.84	0.33	0.84	0.26

SOM	0.16	0.66	0.16	0.66	0

Further research is required to identify the most suitable parameter setting for real learners' data. Our future work will seek to apply different types of network structure such as mapping topology and lattice, and restructure the neighbourhood radius formulation to improve the Kohonen network.

In particular, for adapting the MBTI theories in AHLS, a specific pedagogical model must be prescribe. In this paper, we outline the pedagogical framework containing the approach, method and techniques that suit for AHLS. This first stage are use to describe how content is reflect to the MBTI personality in online system. Fuzzy logic techniques are then used to impart the learning content based on student's fuzzy personality data and instructional rules in order to support customisation that will allow learners to learn faster and understand the learning material much easier. Conversely, the challenge is to identify what those learning content (structure and type of learning material) for a given learner in online system based on student personality and which result is more precise to the learner's personality whether the traditional approach or using the fuzzy logic techniques. Fuzzy logic model provides an efficient way to reason the student's learning method based on the student's personality. For further research, the fuzzy logic model may need to hybrid with genetic algorithm for tuning the membership function and the scaling function for fuzzy input and output that result to better fuzzy logic techniques.

Acknowledgement

This research is conducted under IRPA grant, project vote 74129: The Development of Knowledge Management for AHLS.

6.0 References

BISHOP-CLARK, C. ET AL, 1994. The Myers-Briggs Personality Type and Its Relationship to Computer Programming. J*ournal of Research on Computing in Education* 26, 3, 358-371.

BRUSILOVSKY, P. 2001. Adaptive Hypermedia. User Modeling and User-Adapted Instruction, *In 2001 Kluwer academic Publishers,* 11, 1 – 2, 87-110.

CAROLYN, S. ET AL, 2001. Myers Briggs Type Preferences in Distance Learning Education. *In International Journal of Educational Technology* 2, 2.

CHO, S.B. 1997. Self-organizing Map with Dynamical Node-splitting: Application to Handwitten Digit Recognition. In *Neural Computation, 9*, 1345-1355.

DARA-ABRAMS, B. P. 2002. Applying Multi-Intelligent Adaptive Hypermedia to Online Learning. *In Proceedings of E-Learn2002, Montreal, Canada*.

GARDNER, B.S. AND KORTH, S. J., 2001. Using Psychologist Type Theory to Maximize Student Learning. *26th International Conference Learner-Centered Universities for the New Millennium.* Rand Afrikaas University, Johannesburg, South Africa.

HASHIM. S AND YAAKUB. R. 2003. *Terbitan Pertama Psikologi Pembelajaran dan Personaliti*. PTS Publications & Distributors Sdn. Bhd (Pelita).

NEGNEVITSKY, M. 2002. *Artificial Intelligence. A Guide to Intelligent Systems*, First Edition, Pearson Education.

RAO, V.B. 1995. *C++ Neural Networks and Fuzzy Logic*. MTBooks, IDG Books Worldwide.

SARLE, W.S. 1994. Neural Networks and Statistical Models. In *Proceedings of the Nineteenth Annual SAS Users Group Internationanl Conference*. Cary, NC: SAS Institute, 1538-1550.

SCHROEDER, C.C. 1993. *New-Students-New Learning Style.* 25(5): 21-26.

STATHACOPOULOU R., MOGOULAS G. D, GRIGORIADOU M. 1999, Neural Network-based fuzzy Modeling for the student in Intelligent Tutoring Systems, *Proceedings of the International Joint Conference on Neural Network, IJCNN'99*, 3517-3521.

VENDLINSKI, T. AND STEVENS, R. 2000. The Use of Artificial Neural Nets (ANN) to Help Evaluate Student Problem Solving Strategies In: *B. Fishman & S. O'Connor-Divelbiss (Eds.): Fourth International Conference of the Learning Science*. 108-114.

ZADEH, L. A., 1992. *Fuzzy Logic for Management of Uncertainty*. John Wiley & Sons, New York.

PAPER III:

INDIVIDUALIZING LEARNING MATERIAL OF ADAPTIVE HYPERMEDIA LEARNING SYSTEM BASED ON PERSONALITY FACTOR (MBTI) USING FUZZY LOGIC TECHNIQUES

Norreen Binti Haron
Faculty of Computer Science and Information System
Universiti Teknologi Malaysia
norreen.haron@gmail.com

Naomie Binti Salim
Faculty of Computer Science and Information System
Universiti Teknologi Malaysia
naomie@fsksm.utm.my

Abstract

The inflexible linking provided in conventional hypermedia learning system has some drawbacks that can cause teaching and learning to be less effective. Research on learning has shown that student learn differently since student process knowledge in different ways, with some students learning more effectively when taught with methods that suits their learning style. One solution to this problem is to develop an adaptive hypermedia learning system which basically incorporates intelligence and knowledge about the individual user learning style to assist learner to achieve learning objectives. Information about learning style can help system become more sensitive to the differences of students using the system. Domain modeling is also an important task in the development of an adaptive hypermedia learning system, as much semantic of the domain and support for the adaptive navigation have to be catered for and incorporated in the model. In this paper, a framework for individualizing the learning material structure in adaptive learning system is introduced. It aims to utilize the learning characteristics and provide a personalized learning environment that exploit pedagogical model and fuzzy logic techniques. The learning material consists of 4 structures; 1) theory, 2) example, 3) exercise and 4) activities. The pedagogical model and learning characteristics are based on the student's personality factor (Myers-Briggs Type Indicator (MBTI)), whilst the fuzzy logic techniques are used to classify the structure of learning material which is based on student's personality factors. This paper focuses on the use of fuzzy logic techniques for adaptation of the content to the user, allowing a learning system to dynamically adapt the choice of possible learning structure through the learning material based on the user's personality factor, with the hope to provide an adaptive hypermedia learning system that is user-customized to support faster and more effective learning.

Keywords

Adaptive Hypermedia system, Pedagogical Framework, Personality Factors (MBTI), Learning Styles, Fuzzy Logic.

1.0 Introduction

The adaptive hypermedia learning system (AHLS) is a computer based learning system in which interactive and dynamic learning module is customized to each student. Research on learning has shown that student learn differently and process knowledge in different ways. Information about learning style can help system become more sensitive to the differences of students that use the system. Several systems adapting different learning styles have been developed to date. However, it is not clear which aspects of learning characteristics are worth modeling, how the modeling can take place and what can be done differently for users with different learning style [1]. These problems may lead to students facing difficulties to understand what is being taught, decrease of students' interest to continue their study in the subject time taken to finish a particular lesson session [2].

Currently, the adaptation of student's learning style to learning is totally based on the dominant student learning style, where the dominant result is mainly stated as one particular student's preferable learning material, ignoring other learning styles that a student may also posses. In reality, a student's learning style can be of mixed traits, each with a certain percentage of membership to the student's overall style.

This paper tends to model the fuzziness in student's learning style and the appropriate learning material method suitable for student's fuzzy learning styles membership.

A framework for learning path personalization in adaptive learning system is introduced. It aims to utilize the learning characteristics and provide a personalized learning environment, that exploit pedagogical model and fuzzy logic techniques as shown in Figure 1 below. The pedagogical model and learning style refer to student's personality factor based on the Myers-Briggs Type Indicator (MBTI) [3, 4]. Based on the MBTI theory, fuzzy logic techniques are then used to classify learning material (structure of learning material).

Fuzzy set theory, proposed by Zadeh [5], has proved to be very effective in handling the vagueness and uncertainty intrinsically existing in the knowledge possessed by people or implied in numerical data. Rules based on fuzzy logic provide a qualitative description of the input-output relationship of a system using fuzzy linguistic terms. Fuzzy linguistic rule is closer to human reasoning and in many real-world applications, and thus it is more adequate and flexible for knowledge representation than the conventional crisp IF THEN rules which is the major reason for its adaptation in this research.

As shown in Figure 1, the architecture for the learning strategy is based on the MBTI personality factor whilst the learning method and techniques is based on the Honey & Mumford theory. The four MBTI personality types used in this research are; Extrovert (E); Introvert (I); Sensing (S) and Intuition (N). Extrovert and introvert represent the student's preferable condition in focusing attention. Sensing and intuition, illustrate the student's preferable way in taking information.

Extrovert students prefer and focus on the outer activities and are energized by interaction with others. They prefer to talk, participate and interact with people. The introvert students prefer and focus on the inner activities such as reading, listening to others and writing. Sensing type of students prefer concrete information, facts and procedures. They are good in memorization and like to go systematically, they also learn best with instruction that allow them to use their senses. Intuition student prefer discovering possibilities and relationship. They also like courses that involve a lot of experimentation and experiences [6, 7].

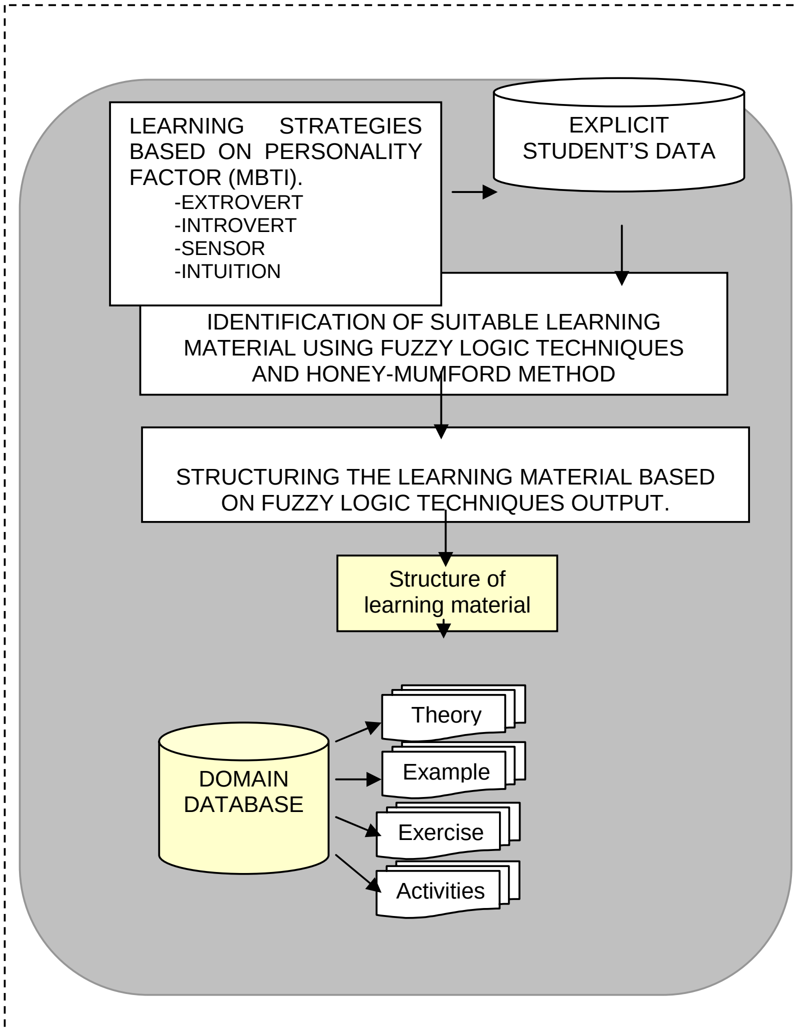


Figure 1: The Framework Of Fuzzy Logic Approach

2.0 Approach And Methods

The pedagogical framework consists of steps on how content is develop to reflect those personality principles. Table 1 below, shows the relationship between learning style and learning method.

Table 1: The Relationship Between Learning Style And Method.

Learning Style (MBTI)	Method (Honey & Mumford)
Extrovert-Sensor	Activity
Extrovert-Intuitive	Exercise
Introvert-Sensor	Example
Introvert-Intuitive	Theory

Fuzzy techniques are used to personalize the learning path where it has 4 outputs. The outputs indicate the structures of learning material (theory, example, exercise and activity) that suits student's personality, taking into account the most preferred learning material and the least preferred learning material.

Fuzzy logic is computationally undemanding and is most suitable for processing imprecise input data, as it supports natural description of knowledge and reasoning in the form of imprecise concepts, operators and rules [8]. In AHLS, fuzzy logic techniques have been used due to their ability to handle imprecise information, such as student's knowledge and their cognitive abilities [9]. Table 2 shows several examples of AHLS adapting intelligent techniques.

Table 2: Examples Of Previous Ahls Adapting Intelligent Techniques In User Modelling.

System	**Intelligent Techniques**	**Predict User Modeling**
KBS-Hyperbook [10]	Bayesian Network	User Knowledge Level
ALICE [11]	Fuzzy Logic Techniques	User Knowledge Level
iWeaver [12]	Bayesian Network	User Media Presentation
INSPIRE [13]	Neuro Fuzzy Techniques	User Knowledge Level

The processes implemented in fuzzy logic systems indicate 4 main stages; fuzzification, rule evaluation, aggregation and defuzzification. In this research, the fuzzification stage transforms crisp student personality, captured from pre-course questionnaires, into suitable linguistic values. The rule evaluation stage takes the fuzzified inputs, and applies them to the antecedents of the fuzzy rules. The aggregation stage combine all the fuzzy output derived from fuzzy rules, and the final stage, the defuzzification stage, produces the output in crisp value. In this problem, the defuzzified value is derived based on the maxima aggregation method in Mamdani-style inference. The suitable defuzzified value will be choosing appropriately based on the result. Figure 2 below shows the flow of the fuzzy logic system.

3.0 The Fuzzy Model

In this study, the fuzzy sets and fuzzy rules are constructed based on the standard of mastery or the criterion-referenced acquired from the human instructors' experience and knowledge about their students. In this problem, there are four input and four output linguistic variables being considered. The input linguistic variables represent the student's personality's scores: extrovert scores, introvert scores, sensor scores and intuitive scores. The fuzzy inputs are expressed by:

x_1, x_2, x_3 and x_4

Whilst, the outputs represents the student's acceptance level of learning material; theory, example, exercise and activities are expressed by:

y_a, y_b, y_c and y_d

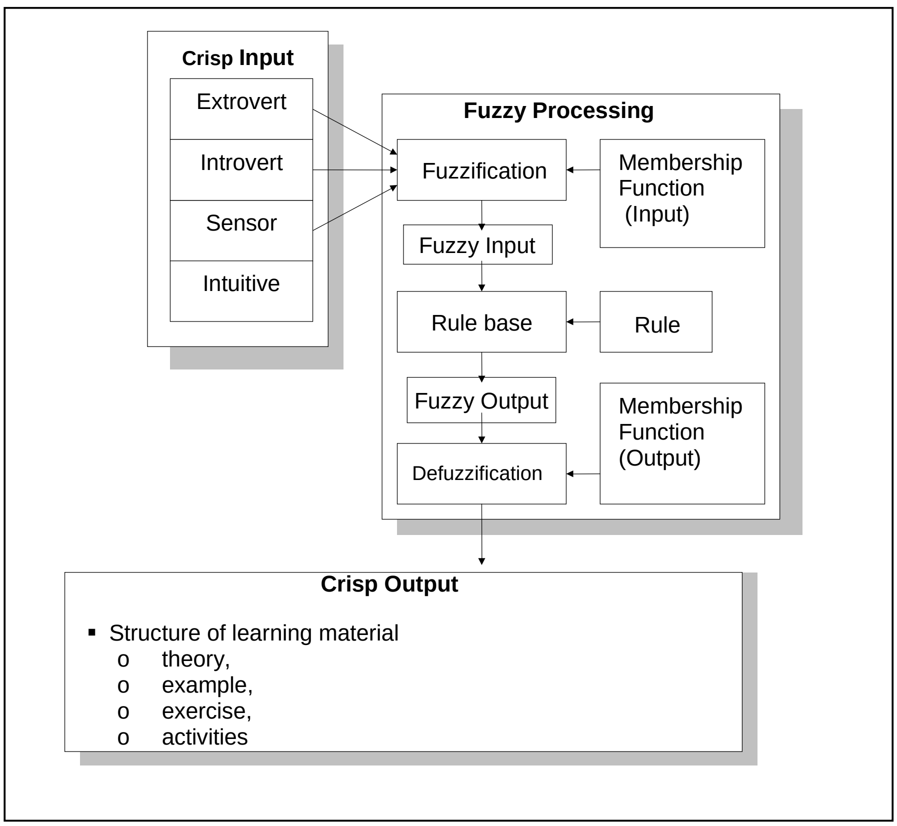


Figure 2: Flow Of Fuzzy System

In this paper, the fuzzy set is expressed by a triangular function as triangular function. This is chosen based on expert advises, interview session and survey session for the case at hand and at the same time significantly simplifies the process of computation. The fuzzy set is expressed by three parameters {a, b, c} as shown in figure 3 and figure 4 below:

$$y_i = \begin{cases} \dfrac{x_i - a}{b - a} & \text{if } a <= x_i <= b \\ 1, & \text{if } x_i = b \\ 1 + \dfrac{b - x_i}{c - b} & \text{if } b < x_i <= c \\ 0, & \textit{otherwise} \end{cases}$$

Figure 3: Triangular Membership Function

where a, b, and c are parameters that control the intervals of the triangle (as shown in figure 3 and figure 4), x_i is the input, and y_i is the output of fuzzification input i.

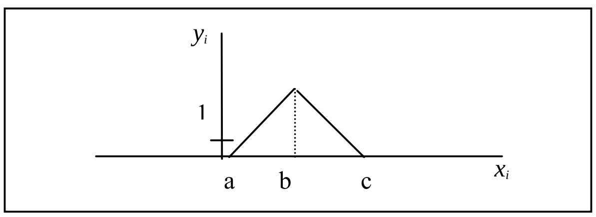
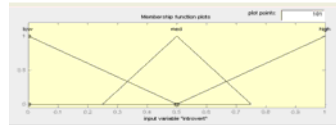
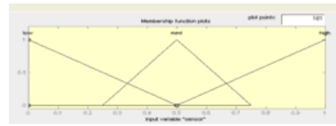
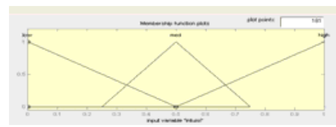
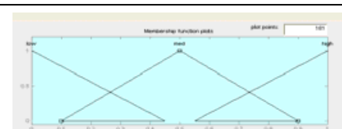
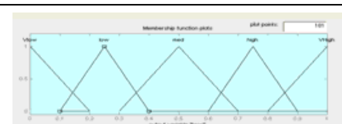
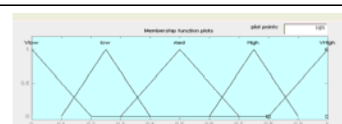
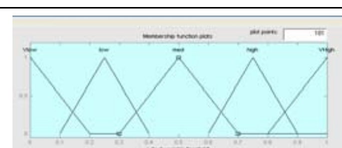


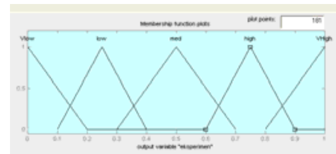
Figure 4: Triangular Fuzzy Set

Table 3 below show the fuzzification process that include the linguistic value, notation, numerical ranges and fuzzy set for the input and output of this study.

Table 3: Fuzzification Process

Linguistic Value	**Input / Output**	**Notation**	**Numerical Ranges (normalized) *Based on expert advise**	**Membership Function *Based on expert advise**
Extrovert	Input	Low Med High	[0, 0, 0.5] [0.25 0.5 0.75] [0.5 1 1]	

Introvert	Input	Low Med High	[0, 0, 0.5] [0.25 0.5 0.75] [0.5 1 1]	
Sensor	Input	Low Med High	[[0, 0, 0.5] [0.25 0.5 0.75] [0.5 1 1]	
Intuitive	Input	Low Med High	[0, 0, 0.5] [0.25 0.5 0.75] [0.5 1 1]	
Visual	Output	Low Med High	[0 0 0.45] [0.1 0.5 0.9] [0.55 1 1]	
Linguistic	Output	Low Med High	[0 0 0.45] [0.1 0.5 0.9] [0.55 1 1]	
Theory	Output	VLow Low Med High VHigh	[0.0, 0.0, 0.25] [0.2, 0.32, 0.45] [0.4, 0.5, 0.6] [0.55, 0.7, 0.85] [0.8, 1.0, 1.0]	
Excercise	Output	VLow Low Med High VHigh	[0.0, 0.0, 0.25] [0.2, 0.32, 0.45] [0.4, 0.5, 0.6] [0.55, 0.7, 0.85] [0.8, 1.0, 1.0]	
Example	Output	VLow Low Med High VHigh	[0.0, 0.0, 0.25] [0.2, 0.32, 0.45] [0.4, 0.5, 0.6] [0.55, 0.7, 0.85] [0.8, 1.0, 1.0]	

Activities	Output	VLow Low Med High VHigh	[0.0, 0.0, 0.25] [0.2, 0.32, 0.45] [0.4, 0.5, 0.6] [0.55, 0.7, 0.85] [0.8, 1.0, 1.0]	

The next stage is the process to generate the fuzzy rule-base. This process classifies the student's learning material according to the student's personality. The fuzzy rule based system consists of IF/THEN statement, multiple antecedents, multiple consequents and combined with AND/OR operation as expressed below:

R_i :

IF X_a is $N1$ AND X_b is N_2 AND X_c is N_3 AND X_d is N_3

THEN

Y_a is M_1 AND Y_b is M_2 AND Y_c is M_3 AND Y_d is M_4

where R_i, (i = 1, 2, … n) is the rule number, N_i, (i =1, 2 and 3) are the membership functions of the antecedent part, M_i, (i =1, 2, 3, 4 and 5) are the membership functions of the consequent part. The example of fuzzy rules applied in this problem as show in table 4.

Table 4: Example Of Fuzzy Rules

Input				**Output**					
Extrovert	**Introvert**	**Sensor**	**Intuitive**	**Theory**	**Example**	**Exercise**	**Activities**	**Visual**	**Linguistic**
Low	High	Low	High	VHigh	High	Low	VLow	Low	High
Low	High	Med	Med	High	Med	Low	Vlow	Med	High
Low	High	High	Low	Med	VHigh	VLow	Low	Med	High
:	:	:	:	:	:	:	:	:	;
High	Low	High	Low	VLow	Low	High	VHigh	High	Low

The defuzzification process used in this paper is Mamdani inference-style where it involves different operations of defuzzification. The best defuzzification operation will be selected. The criterion of selecting the best defuzzification is based on the most similar result that fulfils the expert expectation. Based on the result, the defuzzifications Centre of Area/Gravity (COG) have the most similarity.

4.0 Conclusion And Further Work

This paper has proposed a way to personalize the course content for adaptive hypermedia learning system, which aims to provide learners with a customized learning environment. It emphasizes on the combination of pedagogical theories and artificial intelligent techniques. In particular, for adapting the MBTI theories in adaptive hypermedia learning system, a specific pedagogical model must be prescribed. In this paper, we outlined the pedagogical framework containing the method and techniques suitable for adaptive hypermedia learning system. Fuzzy logic techniques are used to impart the learning content based on student's fuzzy personality data and instructional rules in order to support customization that will allow learners to learn faster and understand the learning material easier. The learning content (structure and type of learning material) for a given learner in online system based on student personality is identified. Fuzzy logic model provides an efficient way to reason the student's learning method based on the student's personality. We are in process of testing the fuzzy logic model by the students compared to crisp conventional method, based on two assessments; performance effectiveness and the accuracy of the system.

Acknowledgement

This research is conducted under IRPA grant, project vote 74129: The Development of Knowledge Management for Adaptive Hypermedia Learning System.

References

1. Brusilovsky, P. (2001). *Adaptive Hypermedia. User Modeling and User-Adapted Instruction.* Kluwer academic Publishers, Vol 11, No. 1 – 2, pp 87- 110.

2. Hashim, S, and Yaakub R. (2003). *Terbitan Pertama Psikologi Pembelajaran dan Personaliti.* PTS Publications & Distributors Sdn. Bhd (Pelita).
3. Carolyn, S. et. al. (2001). Myers Briggs Type Preferences in Distance Learning Education. *In International Journal of Educational Technology,* 2(2).
4. Bishop, C.C. et. al. (1994). The Myers-Briggs Personality Type and Its Relationship to Computer Programming, *Journal of Research on Computing in Education,* 26(3), pp. 358-371.
5. Zadeh, L. A. (1992). *Fuzzy Logic for Management of Uncertainty*. John Wiley & Sons, New York.
6. Gardner, B.S. and Korth, S. J. (2001). Using Psychologist Type Theory to Maximize Student Learning. *26th International Conference Learner-CenteredUniversities For the New Millennium.* Rand Afrikaas University, Johannesburg, South Africa.
7. Schroeder, C.C. (1993). New-Students-New Learning Style, 25(5), pp. 21-26.
8. Negnevitsky, M. (2002). *Artificial Intelligence. A Guide to Intelligent Systems*, First Edition, Pearson Education.
9. Stathacopoulou R., Mogoulas G. D, Grigoriadou M. (1999). Neural Network-based fuzzy modeling of the student in Intelligent Tutoring Systems, *In Proceedings of the International Joint Conference on Neural Network, IJCNN'99*, July 10-16, Washington DC, USA.
10. Henze. N., and Nejdl, W., (1999). Adaptivity in the KBS Hyperbook System, *In Proceedings of the 2nd Workshop on Adaptive Systems and User Modeling on the WWW*, May 11, Toronto.
11. Kavcic, A. (2001). Dealing with Uncertainty of User Knowledge: Fuzzy User Modelling in Educational Hypermedia, *Advances in Multimedia Communications, Information Processing and Education*, Invited Lecture, October, Leganés, Spain.
12. Wolf, C. (2002). iWeaver: towards an interactive web-based adaptive learning environment to address individual learning styles, online European Journal of Open and Distance Learning, (see www.eurodl.org). Retrieved December 17, 2003,

    web site: http://www.eurodl.org/materials/contrib/2002/2HTML/iWeaver

13. Kyparisia, A., et al. (2003). Personalizing the Interaction in a Web-based Educational Hypermedia System: the case of INSPIRE, *Journal ACM, User Modeling and User-Adapted Interaction.* 13(3), pp. 213-267.

PAPER IV:

ROUGH SET GENERATION FOR IDENTIFYING STATUS OF STUDENT'S KNOWLEDGE ACQUISITION

Nor Bahiah Hj Ahmad, Siti Mariyam Hj. Shamsuddin & Siti Zaiton Mohd Hashim
Soft Computing Research Group,
Faculty of Computer Science and Information System,
Universiti Teknologi Malaysia, Malaysia.
bahiah@fsksm.utm.my, mariyam@fsksm.utm.my, szaiton@fsksm.utm.my

Abstract

In this study, a generation of rough set rules are implemented in identifying the status of students knowledge acquisition in hypermedia learning. Along with this concept is the process of subdivision of the universal set of all possible categories of identifying the status into a number of distinguishable categories called elementary sets. Four attributes, cpa, learning time, number of backtracking, and scores obtained by the student have been collected in the student model and the decision of the students knowledge acquisition whether as good, average or poor are transformed into the information table. In this study, each object represents students's performance. The decision is the students's knowledge acquisition status, identified by adopting expert advice as poor, average and good. They are represented as Class 1, Class 2 and Class 3 respectively. We apply the rules to 76 unseen objects and the result shows that the total percentage of correct classification is 93.4211%.

Keywords: Rough set, rule generations, classification, hypermedia learning, knowledge acquisition.

1. Introduction

More and more hypermedia learning system have been developed to aid teachers and students in teaching and learning. However hypermedia system has some drawbacks that can cause teaching and learning to be less effective. First, the link structures between the nodes are static and cannot adapt to the student's background, understanding and desires. Secondly, when the system has thousands of non-linear links, the students surfing the information can get lost and confused of where they are, what actually they are searching for and where to go next. To address this problem, we are developing an adaptive learning system with hypermedia technology. Adaptivity is incorporated by presenting suitable learning materials, guidance and aid during learning session according to different categories of learners based on their majoring background and knowledge acquisition status [1].

Rough set theory, introduced by Zdzislaw Pawlak in the early 1980's is a new mathematical tool to deal with vagueness and uncertainty. The methodology is concerned with the classificatory analysis of imprecise, uncertain or incomplete information or knowledge or knowledge expressed in terms of data acquired from experience. Here, objects are perceived through the information that is available about them through the values for a predetermined set of attribute. The main advantage of rough set is that it requires no additional information to the data represented in table. It does not need any preliminary or additional information about data, such as probability distribution in statistic, grade of membership or the value of possibility in fuzzy set theory [2].

Objects characterized by the same information are indiscernible in view of the available information about them. The similarity/indiscernible relation generated in this way is the mathematical basis of the rough set theory. Any set of all indiscernible objects is called elementary set and form the basic granule of knowledge about the universe. Any union of some elementary sets is referred to as crisp set, otherwise a set is rough (imprecise, vague). Vague concept is characterized by the lower and upper approximation. The lower approximation consists of all objects which definitely belong to the set (members of the set), and upper approximation contain all objects which possibly belong to the set. The difference between the upper and lower approximation constitute the boundary region of the set.

In rough sets the issue of knowledge can be viewed as partition or classification. The knowledge representation system can be perceived as a data table, columns which are labeled as attributes, and rows are labeled as objects. Each row represents a piece of information about corresponding objects. The data table can be obtained as a results of measurements, observations or represents knowledge of an agent or a group of agents. Thus in the study, we present the classification of student's knowledge acquisition status using rough set.

2. Basic ideas of Rough Set and implementation

Rough Set theory offers some important techniques in managing an information system (IS), and consists of several steps leading towards the final goal of generating rules from information/decision systems. The main steps of the rough set approach described by [3] are:

- ❖ The mapping of information from the original database into the decision system format (Information System Table)
- ❖ Completion of data
- ❖ Discretisation of data
- ❖ Computation of reducts from data
- ❖ Derivation of rules from reducts
- ❖ Filtering of rules

2.1 Mapping of information and knowledge representation

The first step, mapping of information into a decision system, depends on the format of the original formation. In some cases, the data set is already on a format ready to be imported into a decision table. In other cases, attributes (including the decision attribute) and object must be identified before data can be placed in the decision table. For some types of data, such as time series, additional mapping of the data is necessary. In rough sets the issue of knowledge representation can be viewed as partition (classification). The knowledge representation system can be perceived as a data table, columns which are labeled as attributes, and rows are labeled as objects. Each row represents a piece of information about corresponding objects. The data table can be obtained as a results of measurements, observations or represents knowledge of an agent or a group of agents. Representation of knowledge in tabular form has great advantage especially for its clarity. The data table may be perceived as a set of propositions about reality, and can be viewed as a model for special logic, called decision logic. Rough sets offers its normal form representation of formulas and the second employing the concept of *indiscernibility* to investigate whether some formulas are true or not. Indiscernibility leads to simple algorithms for data reduction and analysis.

In this study, the information table represents input data gathered from any domain in students's model. This information table describes students's performance, and the decisions are represented as status of knowledge acquisition. Each student is characterized by the CPA earned, duration of learning period, number of times doing back tracking, and finally scores obtained from exercises and tests [4]. The decision for this information table is classified by the experts as good, average or poor. Table 1 shows some of the informations with related decisions for the domain described above.

Table 1. Information table with decision

No	CPA	Learnin g Time	# Back tracking	Score	D
1	Good	Good	Good	Poor	Good
2	Avg	Good	Avg	Good	Good
3	Good	Good	Good	Poor	Good
4	Poor	Good	Poor	Good	Poor
5	Good	Good	Good	Poor	Good
6	Poor	Good	Poor	Good	Poor
7	Poor	Good	Poor	Good	Poor
8	Good	Poor	Poor	Good	Avg
9	Good	Poor	Poor	Good	Avg

2.2 Indiscernibility relation

The definition of discernibility is given to explain the indiscernibility relation. The basic definition of discernibility is given as input an attribute and two attribute values, and returns true if it is possible for the two values to be different. The standard rough set theory uses ordinary definition of inequality:

$$discerns(a, a(x_i), a(x_j)) = (a(x_i) \neq a(x_j)) \qquad (1)$$

From Table 1, we consider attributes {CPA, Learning Time, #Backtracking, Scores } and generate classes of indiscernibility objects for the selected attributes.

E1 = {1,3,5}

E2 = {2}

E3 = {4,6,7}

E4 ={8,9}.

E1 to E4 classes contain objects of the same condition or attribute values i.e., set of objects that are indiscernible by attributes described above. Table 2 shows the indiscernibility relations obtained from 100 cases.

Table 2. Indiscernibility relation of students knowledge of acquisition

Class	CPA	Learning Time	# Back tracking	Score	D
E1 (50x)	Goo d	Good	Good	Poor	Good
E2 (5x)	Avg	Good	Avg	Good	Good
E3 (30x)	Poor	Good	Poor	Good	Poor
E4 (15x)	Goo d	Poor	Poor	Good	Avg

For futher explanation, let us transform the data in Table 2 in simpler numerical representation as shown in Table 3.

a – CPA (1:poor, 2:average, 3:good)

b – Learning Time (1:poor, 2:average, 3:good)

c - #Backtracking (1:poor, 2:average, 3:good)

d - Score (1:poor, 2:average, 3:good)

Table 3. Simplified representation of students knowledge of acquisition

Class	a	b	c	d	# Object
E1	3	3	3	1	50x
E2	2	3	2	3	5x
E3	1	3	1	3	30x
E4	3	1	1	3	15x

2.3 Discernibility matrix

A *discernibility matrix* can be created by using the discerns predicate in Equation (1) Given an IS $\mathrm{A} = (U, A)$ and $B \subseteq A$, the *discernibility matrix* of A is M_B, where each entry $m_B(i, j)$ consists of the attribute set that discerns between objects x_i and x_j where $1 < i, j < n = |U / IND(B)|$. Table 4 shows the discernibility matrix of Table 3.

$m_B(i, j) = \{a \in B: discerns(a, a(x_i), a(x_j))\}$ (2)

Table 4. The discernibility matrix

	E1	E2	E3	E4	*f*
E1	x	acd	acd	bcd	c v d
E2	acd	x	ac	abc	a v c
E3	acd	ac	x	ab	a
E4	bcd	abc	ab	x	b

2.4 Completion

Completion of data is a preprocessing step that is used for decision tables with missing values. One way of making a data set complete is to simply remove objects which have missing values. Another way of dealing with the problem is to fill out missing values using a mean value computed from other objects in the decision table.

2.5 Discretisation

Discretization refers to the process of arranging the attribute values into groups of similar values. Discretization of real value attributes is an important task in data mining, particularly the classification problem. Empirical results are showing that the quality of classification methods depends on the discrimination algorithm used in preprocessing step. In general, discrimination is a process of searching for partition of attribute domains into intervals and unifying the values over each interval. Discretization involves searching for "cuts' that determine intervals. All values that lie within each interval are mapped to the same value, in effect converting numerical attributes that can be treated as being symbolic. The search for cuts is performed on the internal integer representation of the input Decision System. In this study, we discretise the data using Boolean reasoning approach. Table 5 is some of the data of students level of knowledge acquisition and, Table 6 shows the result of this data after discretisation using Boolean reasoning technique accordingly.

Table 5. Data of students knowledge acquisition

CPA	TIME	SCORE	BACK TRACK	STATUS
3.5	82	70	1	3

2.8	56	88	2	3
3.7	84	50	1	3
2.4	61	89	5	1
3.7	75	55	1	2
2.6	55	90	4	1
1.8	61	92	5	1
3.7	120	91	7	2
3.8	102	88	6	2
3.8	66	87	0.3	3
2.8	86	73	2.2	2
2.5	102	59	3	1
2	112	54	3.6	1
2	114	53	3.8	1

Table 6. Data of students knowledge acquisition after discretization

1	0	0	0	1
2	0	0	0	1
2	0	0	0	1
3	0	1	0	2
2	0	1	0	3
3	0	1	0	2
3	0	1	0	2
2	0	1	0	3
3	0	0	0	3
3	1	1	0	1
2	1	1	0	3
3	0	1	0	2
3	1	1	1	2
3	1	1	1	2

2.6 Computation of reducts

A fundamental problem in information system/decision systems is whether the whole knowledge is always necessary to define some categories available in the knowledge considered. This problem arises in many practical applications and be referred as knowledge reduction. In reduction of knowledge the fundamental concept offered by RS is *reduct* and *core*. A reduct of knowledge is the essential part in which suffices to define all basic concepts occur in the considered knowledge, whereas the core is in a certain sense its most important part. Reducts are subsets of the original attributes with the original dependencies preserved. Rules can be generated using a minimal subset of attributes. The computation of dynamic reducts is similar to the computation of reducts. The difference is that dynamic reducts only preserve the original dependencies approximately, not accurately. The result is that attributes that have a minimal impact on the decision attributes are not used to generate rules. Thus, reducts with their cardinalities for students knowledge acquisition in this study are listed accordingly :

{backtrack}	1
{cpa, scored}	2
{cpa}	1
{time, backtrack}	2
{scored, backtrack}	2
{time, scored}	2
{cpa, backtrack}	2
{cpa, time}	2

2.7 Rule generation

Rules are generated from reducts. The rules may be of different types and on different formats, depending on the algorithms used. Measures of confidence and frequency usually accompany the rules. We listed some of the rules generated from our reducts on knowledge acquisition data below:

Rule 1:

backtrack({0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, 1}) => **status(3)** – this mean that the values given in the parentheses are the generated rules and imply that the status of students knowledge acquisition are good represented by numerical value 3, and the same for the others accordingly.

Rule 2:

cpa({2.8, 2.9, 3, 3.2, 3.4, 3.5, 3.8, 3.9}) AND **scored**({27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52}) => **status(3)**

Rule 3:

cpa({2.8, 2.9, 3, 3.2, 3.4, 3.5, 3.8, 3.9}) AND **scored**({100, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99}) => **status(3)**

Rule 4:

cpa({0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, 1, 1.1, 1.2, 1.4, 1.5, 1.6, 1.7, 1.9, 2, 2.4, 2.5}) => **status(1)**

Rule 5:

time({{12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 55, 56, 58, 60, 61, 66, 68, 70, 72, 74}, {75}}) AND **backtrack**({2.1, 2.2, 2.4, 2.5, 2.7, 2.9, 3, 3.6, 3.8, 3.9, 4.1, 4.2, 4.3, 4.4, 4.6, 4.7, 4.8, 4.9, 5.1, 5.2, 5.3, 5.4, 5.6, 5.7, 5.8, 5.9, 6.1, 6.2, 6.3, 7}) => **status(1)**

Rule 6:

scored({100, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99}) AND **backtrack**({2.1, 2.2, 2.4, 2.5, 2.7, 2.9, 3, 3.6, 3.8, 3.9, 4.1, 4.2, 4.3, 4.4, 4.6, 4.7, 4.8, 4.9, 5.1, 5.2, 5.3, 5.4, 5.6, 5.7, 5.8, 5.9, 6.1, 6.2, 6.3, 7}) => **status(1)**

Rule 7:

scored({53, 54, 55, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81}) AND **backtrack**({2.1, 2.2, 2.4, 2.5, 2.7, 2.9, 3, 3.6, 3.8, 3.9, 4.1, 4.2, 4.3, 4.4, 4.6, 4.7, 4.8, 4.9, 5.1, 5.2, 5.3, 5.4, 5.6, 5.7, 5.8, 5.9, 6.1, 6.2, 6.3, 7}) =>**status(2)**

Rule 8:

scored({27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52}) AND **backtrack**({2.1, 2.2, 2.4, 2.5, 2.7, 2.9, 3, 3.6, 3.8, 3.9, 4.1, 4.2, 4.3, 4.4, 4.6, 4.7, 4.8, 4.9, 5.1, 5.2, 5.3, 5.4, 5.6, 5.7, 5.8, 5.9, 6.1, 6.2, 6.3, 7}) =>**status(1)**

Rule 9:

time({{100, 102, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98}}) AND **scored**({53, 54, 55, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81}) => **status(2)**

Rule 10:

time({{12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 55, 56, 58, 60, 61, 66, 68, 70, 72, 74}, {75}}) AND **scored**({53, 54, 55, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81}) => **status(3**)

Rule 11:

cpa({2.8, 2.9, 3, 3.2, 3.4, 3.5, 3.8, 3.9}) AND **backtrack**({2.1, 2.2, 2.4, 2.5, 2.7, 2.9, 3, 3.6, 3.8, 3.9, 4.1, 4.2, 4.3, 4.4, 4.6, 4.7, 4.8, 4.9, 5.1, 5.2, 5.3, 5.4, 5.6, 5.7, 5.8, 5.9, 6.1, 6.2, 6.3, 7}) => **status(2)**

Rule 12:

cpa({2.8, 2.9, 3, 3.2, 3.4, 3.5, 3.8, 3.9}) AND **time**({{12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 55, 56, 58, 60, 61, 66, 68, 70, 72, 74}, {75}}) => **status(3)**

3. Classification

After a rule set has been generated, these rules should be tested on a data table. Preferrably, the rules should be tested on unseen objects, that is, objects that were not used in the rule generation process. After applying the rules on unseen objects, the predicted outcomes can be compared to the correct classification of the object.

In this study, each object represents students's performance as shown in Table 1. The decision is the students's knowledge acquisition status, identified by adopting expert advice as poor, average and good. They are represented as Class 1, Class 2 and Class 3 respectively. We apply the rules to 76 unseen objects. The result shows that 71 objects are correctly classified. The total percentage of correct classification is 93.4211%. The details are as shown in the Table 7.

Table 7. Result of students knowledge acquisition classification

		Predicted	Class	
Actual Class	1	2	3	
1	30	1	0	0.967742
2	0	20	1	0.952381
3	0		21	0.875
	1.0	0.833333	0.954545	0.934211

To see more closely on how the classifications are made, we present a few snapshots of the objects as shown in Table 8. For object 4, its value for attributes cpa, time, score and backtrack are 2.4, 61, 89 and 5 respectively (Refer Table 5 for original value of the object's attributes). When the rules are applied, it satisfies 3 of them, and this includes rule 4, rule 5 and rule 6 with classification identification as Class 1. Similarly for object 10 in which the attributes are represented as 3.8, 66, 87 and 0.3. It satisfies rule 1, rule 3 and rule 12 with 3 rules classify it as Class 3. Different rules may generate different classification results. When this is the case, the class predicted by most rules is chosen as illustrated by object 12. Its attribute values satisfy rule 1, rule 7 and rule 9 whereby rule 1 classifies it as Class 1, rule 7 and 9 classify it as Class 2. Hence the error exists. Overall results in this study shows that the error rate is less than 7%.

Table 8. Result of students knowledge acquisition classification

```
Object 4:   ok
        Actual=1  (1)
        Predicted=1  (1)
        Ranking=(1.0)  1  (1)
        3  rule(s)

Object 10: ok
        Actual=3  (3)
        Predicted=3  (3)
        Ranking=(1.0)  3  (3)
        3  rule(s)

Object 12: ERROR
        Actual=1  (1)
        Predicted=2  (2)
        Ranking=(0.727273)  2 (2)
```

2 rule(s)

Ranking=(0.272727) 1 (1)

1 rule(s)

4. Conclusion

Four attributes, (cpa, learning time, number of backtracking, and scores obtained by the student), that have been collected in the student model and the decision of the students knowledge acquisition whether as good, average or poor are transformed into the information table. Based on the chosen attributes, classes of indiscernibility objects and the discernibility matrix are generated followed by discretisation using boolean reasoning approach. Rules generation is done using a minimal subset of attributes that are collected after the reducts computation. The rules are tested on a data table of unseen objects. The criterion of choosing the best classification model is based on the highest percentage of classification toward the new unseen data. The number of the classification percentage varies based on the number of reduct obtained in the process.

5. References

1. P. Brusilovsky, "Adaptive Hypermedia : An Attempt to Analyse and Generalize". *4th Intern. Conf. On User Modeling*, Massachusettes, USA 1994.

2. Z. Pawlak, "Rough Set : Theoretical Aspect of Reasoning About Data. " *Kluwer Publications*, 1991.

3. R. Greathe. "Knowledge Extraction From Process Data : A Rough Set Approach to Data Mining on Time Series", *MSc. Thesis*, 1999.

4. Siti Zaiton Mohd Hashim, Norazah Yusof, Paridah Samsuri and Nor Bahiah Ahmad, "Design and implementationof an adaptive web-based hypermedia learning system : a prototype", *International Conference on Organizational Development and Leadership in Education*, Kuala Lumpur, 2001.

PAPER V:

STUDENT CLASSIFICATION USING NEURAL NETWORK IN ADAPTIVE HYPERMEDIA LEARNING SYSTEM: PRELIMINARY WORK

Bariah Yusob, Nor Bahiah Hj Ahmad, Siti Mariyam Hj Shamsuddin, Shahliza Abd Halim, Norazah Yusof

Faculty of Computer Science and Information System

University of Technology Malaysia, 81310 Skudai

byvail@catcha.com

Abstract

Traditional Adaptive Hypermedia Learning System will pose disorientation and lost in hyperspace problem. Adaptive Hypermedia Learning System is the solution whereby the system will personalize the learning module presentation based on student classification: advanced, intermediate and beginner. Classification of student is needed for the system to provide suitable learning module to each individual student by taking consideration of students' knowledge level and their performances as they go through the system. Three classifiers determine student knowledge level. The first classifier determines the student initial status from data collected from explicit data extraction technique. Second classifier identifies student's status from implicit data extraction technique, and the third classifier will be executed if the student has finished doing exercises. Implicit extraction technique includes process of gathering and analysis of students' behavior provided by web log data while they navigate through the system. Explicit extraction technique on the other hand is a process of collecting students' basic information from user registration data. Finally, based on the AHLS architecture, this information will be integrated into user profile to perform classification using simple backpropagation neural network.

Keywords: student classification, web log analysis, neural network

1.0 Introduction

Adaptive Hypermedia Learning System (AHLS) is an approach to overcome the problems with traditional static hypermedia learning system (Brusilovsky, 1994). AHLS manage to present interactive and dynamic interface to provide suitable learning module to each student. Learning module presented is based on students' knowledge and skill level together with their preferences that can be seen through analysis of their behavior as they navigate along the system.

Learning module presentation that meets those features must be created through one very popular concept: personalization (Wang, 2000). This concept has been used as a trend in electronic application in World Wide Web including e-commerce, e-medicine and others. Through personalization, users are no longer treated equally. The AHLS will recognize users and provide services according to their personal needs. For example, a beginner student will not be presented with learning module that is too complex and have difficult paths. If the student receive unsuitable module that does not match his/her level, he/she will not be able to follow the learning process and finally the objective of the module is not achieved.

To apply personalization concept, student's information must be analyzed and extracted to get useful information about the student and to identify type of the student, whether he/she is advanced, intermediate or beginner. The aim of this project is to integrate two techniques of user data extraction in AHLS to create user profile. These techniques are explicit and implicit data extraction. Explicit technique includes data given by students when they register into the system, and score they gain as they perform exercises after finish one topic. Implicit user profile record students' behavior through their interaction with the system. At this stage, the students are not aware that their movements are recorded in web log file. Data collected are navigation path, clickstreams, together with date and time the node is requested.

This paper presents the preliminary work in order to classify student by manipulating data extracted explicitly and implicitly. We will use data stored in the system for the neural network training process to perform the classification task.

Specifically, the paper will first present a model of student classification using neural network. Second section contains the process of analyzing students' activity in web log data, including navigation, number of using help, number of backtracking and others. There are three types of classifier for determining the student's status, which will be explained in detail in the next section.

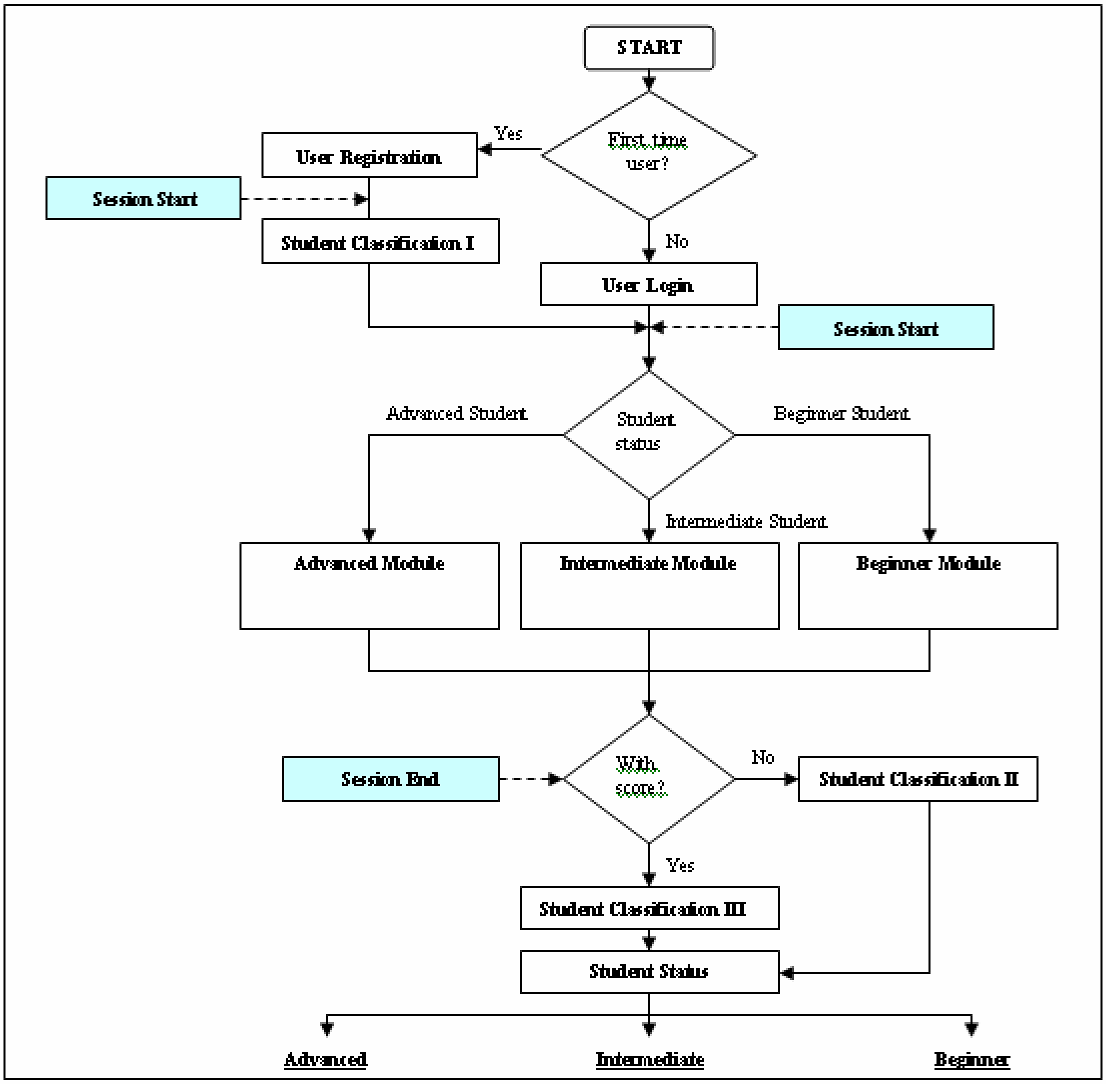


Figure 1: Workflow of Student Classification

2.0 Workflow Of Student Classification Using Neural Network

The system must first identify whether the student has register into the system before allowing them to start learning the module. For the classification purpose, there are situation need to be considered when student accessing the system. The situation includes

first time user, student without score value and student with score value. In the following, we will give an overview of each situation:

i) **Student who has not register into the system. It means that the student is a new user. The student must enter his/her information including name, metric number, ic number, cpa and programming knowledge. Once the student submit his/her registration data, the system will record the data into database and remember his/her metric number as a new user. Then, the system will execute first classifier to identify his/her initial status.**

ii) **The student has registered and will be asked to log into the system by entering his/her metric number so that the system could recognize him/her and will retrieve his/her previous status. Based on his/her last status, the suitable learning module will be presented. In this situation, there is possibility that the student simply finish learning the module without performing any exercises. Once he/she log out of the system, second classifier will identify his/her current status.**

iii) **When the user finished learning the module and has completed doing exercises, the score will be stored in system database. Third classifier will identify his/her current status using explicit and implicit data extraction.**

Based on Figure 1, the system must first identify students' status from the previous learning session. Once their status is identified, the learning module will be presented according to their status. After they finish their session, then the last classifier will be applied to categorize them into their current status to be used in the next session when they come back for learning new module.

3.0 Explicit Data Extraction: Classifier I

For the first time user, she/he has to register into the system by filling his/her information into the student registration form. Explicit data is a data provided interactively and willingly by the student.

3.1 Student Registration

Data collected from student registration process are student's name, metric card number, identity card (ic) number, cpa and programming knowledge (prog). Figure 2 shows the student registration form. Registration data will be stored in system's database. Once student submit his registration, system will assign a new session that recognize the student by his metric number, noMatrik. Data collected from student registration will be used in student classification phase to identify the status.

Figure 2: Student Registration Form

3.2 Training Data Sample

Before data can be fed into neural network for student classification, one set of training data is needed. Training data is used to train the network to perform classification into desired groups. To obtain the training sample, one program is developed to calculate the desired output values. The program will be using input data from cpa and prog.

Table 1: Input Data

Input, x	Value	Weight, w
cpa, x_1	0.00 – 4.00	$w_1 = 0.7$
prog, x_2	1,2,3	$w_2 = 0.3$

Where,

Output, $y = x_1*w_1 + x_2*w_2$

Student's *cpa* will hold the value between 0.00 – 4.00.
prog will hold value whether 1 for beginner, 2 for intermediate and 3 for advanced. Weight is given based on the priority between these two data. cpa carries more priority than prog value.

Classification I:

If y > 2.00 then
    Status = beginner
Else if 2.00 ≤ y ≤ 3.00 then
    Status = Intermediate
Else if y > 3.00 then
    Status = advanced

Data Representation:

Beginner : 00
Intermediate : 01
Advanced : 11

4.0 Implicit Data Extraction: Classifier Ii

Implicit data extraction is a process of analyzing web log data, which contain students' activity through their interaction with the system.

4.1 Web Log Analysis

Web log analysis is a process of extracting useful information about user's behavior recorded in web log server file. In this project, we perform implicit technique by collecting web log data and analyze it to get students' path navigation through the system.

Figure 3 shows lines from a file using the following fields: date, time, client IP address and URI stem (requested node). The original data show sequences of requests by IP address and request time. In order to analyze the log data, we first need to define an individual session.

A session is defined by a unique IP address and a unique request time. It begins when user login or when system come across an IP address. The request time is defined as the beginning of a session. System then keep tracking that individual's requests continuously, and define the end of that particular session for that individual to be when a subsequent request does not appear within an hour. Figure 4 shows the log data extracted from the original web log file into one session.

```
#Software: Microsoft Internet Information Services 5.0
#Version: 1.0
#Date: 2003-09-19 16:03:54
#Fields: date time c-ip cs-uri-stem cs(Referer)
2003-09-19 16:03:54 127.0.0.1 /iishelp/iis/htm/core/iiauths.htm
http://127.0.0.1/iishelp/iis/htm/core/iiabasc.htm
#Software: Microsoft Internet Information Services 5.0
#Version: 1.0
#Date: 2003-09-19 16:21:21
#Fields: date time c-ip cs-uri-stem
2003-09-19 16:21:21 127.0.0.1 /spath/
2003-09-19 16:21:48 127.0.0.1 /spath/fT02.htm
2003-09-19 16:21:48 127.0.0.1 /spath/banner.htm
2003-09-19 16:21:48 127.0.0.1 /spath/sisiMD.htm
2003-09-19 16:21:48 127.0.0.1 /spath/EFRONT1.gif
2003-09-19 16:21:48 127.0.0.1 /spath/T02.htm
2003-09-19 16:21:51 127.0.0.1 /spath/T02F001.htm
2003-09-19 16:21:54 127.0.0.1 /spath/T02F002.htm
2003-09-19 16:21:54 127.0.0.1 /spath/T02F001R001.jpg
2003-09-19 16:22:02 127.0.0.1 /spath/T02F003.htm
2003-09-19 16:22:03 127.0.0.1 /spath/T02F003N001.htm
2003-09-19 16:22:15 127.0.0.1 /spath/T02F003N003.htm
2003-09-19 16:22:15 127.0.0.1 /spath/T02F002N003R001.gif
2003-09-19 16:23:30 127.0.0.1 /spath/
2003-09-19 16:23:41 127.0.0.1 /spath/T02.htm
2003-09-19 16:23:43 127.0.0.1 /spath/T02F001.htm
2003-09-19 16:23:48 127.0.0.1 /spath/T02F002.htm
2003-09-19 16:23:48 127.0.0.1 /spath/T02F001R001.jpg
2003-09-19 16:23:53 127.0.0.1 /spath/T02F003.htm
2003-09-19 16:24:09 127.0.0.1 /spath/T02F003N001.htm
2003-09-19 16:24:16 127.0.0.1 /spath/T02F003N003.htm
2003-09-19 16:24:16 127.0.0.1 /spath/T02F002N003R001.gif
```

Figure 3: Example of Original Web Log Data

Data collected from web log data include date of request, time of request, IP address/username and page request. In this phase, we also need a training data before using it with neural network to perform student classification task. One more program is developed to calculate the desired output. Input for this program will be time and request. Figure 5 is the example of output of this program. Recommended learning time is 60 minutes.

```
2003-09-19 16:21:21 127.0.0.1 /spath/
2003-09-19 16:21:48 127.0.0.1 /spath/fT02.htm
2003-09-19 16:21:48 127.0.0.1 /spath/banner.htm
2003-09-19 16:21:48 127.0.0.1 /spath/sisiMD.htm
2003-09-19 16:21:48 127.0.0.1 /spath/EFRONT1.gif
2003-09-19 16:21:48 127.0.0.1 /spath/T02.htm
2003-09-19 16:21:51 127.0.0.1 /spath/T02F001.htm
2003-09-19 16:21:54 127.0.0.1 /spath/T02F002.htm
2003-09-19 16:21:54 127.0.0.1 /spath/T02F001R001.jpg
2003-09-19 16:22:02 127.0.0.1 /spath/T02F003.htm
2003-09-19 16:22:03 127.0.0.1 /spath/T02F003N001.htm
2003-09-19 16:22:15 127.0.0.1 /spath/T02F003N003.htm
2003-09-19 16:22:15 127.0.0.1 /spath/T02F002N003R001.gif
2003-09-19 16:23:30 127.0.0.1 /spath/
2003-09-19 16:23:41 127.0.0.1 /spath/T02.htm
2003-09-19 16:23:43 127.0.0.1 /spath/T02F001.htm
2003-09-19 16:23:48 127.0.0.1 /spath/T02F002.htm
2003-09-19 16:23:48 127.0.0.1 /spath/T02F001R001.jpg
2003-09-19 16:23:53 127.0.0.1 /spath/T02F003.htm
2003-09-19 16:24:09 127.0.0.1 /spath/T02F003N001.htm
2003-09-19 16:24:16 127.0.0.1 /spath/T02F003N003.htm
2003-09-19 16:24:16 127.0.0.1 /spath/T02F002N003R001.gif
```

Figure 4: User Session 1

```
Date       Time     ipaddr    request                              Visited  TimeTaken
2003-09-19 16:21:21 127.0.0.1 /spath/                              0        00:27
2003-09-19 16:21:48 127.0.0.1 /spath/fT02.htm                      0        00:00
2003-09-19 16:21:48 127.0.0.1 /spath/banner.htm                    0        00:00
2003-09-19 16:21:48 127.0.0.1 /spath/sisiMD.htm                    0        00:00
2003-09-19 16:21:48 127.0.0.1 /spath/EFRONT1.gif                   0        00:00
2003-09-19 16:21:48 127.0.0.1 /spath/T02.htm                       0        00:03
2003-09-19 16:21:51 127.0.0.1 /spath/T02F001.htm                   0        00:03
2003-09-19 16:21:54 127.0.0.1 /spath/T02F002.htm                   0        00:00
2003-09-19 16:21:54 127.0.0.1 /spath/T02F001R001.jpg               0        00:08
2003-09-19 16:22:02 127.0.0.1 /spath/T02F003.htm                   0        00:01
2003-09-19 16:22:03 127.0.0.1 /spath/T02F003N001.htm               0        00:02
2003-09-19 16:22:15 127.0.0.1 /spath/T02F003N003.htm               0        00:00
2003-09-19 16:22:15 127.0.0.1 /spath/T02F002N003R001.gif           0        00:01
2003-09-19 16:23:16 127.0.0.1 /spath/T02F003N003help.htm           0        00:14
2003-09-19 16:23:30 127.0.0.1 /spath/                              1        00:03
2003-09-19 16:23:33 127.0.0.1 /spath/T02F003help.htm               0        00:08
2003-09-19 16:23:41 127.0.0.1 /spath/T02.htm                       1        00:02
2003-09-19 16:23:43 127.0.0.1 /spath/T02F001.htm                   1        00:05
2003-09-19 16:23:48 127.0.0.1 /spath/T02F002.htm                   1        00:00
2003-09-19 16:23:48 127.0.0.1 /spath/T02F001R001.jpg               1        00:05
2003-09-19 16:23:53 127.0.0.1 /spath/T02F003.htm                   1        00:16
2003-09-19 16:24:09 127.0.0.1 /spath/T02F003N001.htm               1        00:07
2003-09-19 16:24:16 127.0.0.1 /spath/T02.htm                       1        00:00
2003-09-19 16:24:16 127.0.0.1 /spath/T02F003N003.htm               1        00:

Count of backtracking = 9
Count of help       = 2
Learning time               = 105s --> 1.75 min

Recommended learning time = 60 min

Percentage        = 2.9167%
```

Figure 5: The Example of Output of Program To Calculate Training Data For Classifier II

Based on Table 2 below, there are a few criteria identified to be an input data for student classification II. These include learning time, number of backtracking, and number of using help (Hashim et al., 2001).

We identify students' status by comparing the time taken by student with the time recommended by system. If the student took more than the recommended time, this indicates that he/she is a slow learner. On the other hand, if he/she took less time, he/she is a fast learner.

In this research, we assume that a beginner student will backtrack the relevant materials he/she has gone through earlier as shown in Table 2. The more the student use help, shows that the student is having problem understanding the module. So, we consider number of using help as one criteria to classify student's status.

Table 2: Criteria to Classification II

Criteria	Beginner, α	Intermediate, β	Advanced, γ
Learning Time, x_1	$x_1 > 100\%$	$80\% \leq x_1 \leq 100\%$	$x_1 < 80\%$
Freq. Of Backtracking, x_2	$x_2 > 5$	$3 \leq x_2 \leq 5$	$x_2 \leq 1$
Freq. Of Using Help, x_3	$x_3 > 5$	$3 \leq x_3 \leq 5$	$x_3 \leq 1$

Mathematically,

$$\alpha(x_i) \in \begin{Bmatrix} x_1 > 100\% \\ x_2 > 5 \\ x_3 > 5 \end{Bmatrix} \quad \beta(x_i) \in \begin{Bmatrix} 80\% \leq x_1 \leq 100\% \\ 3 \leq x_2 \leq 5 \\ 3 \leq x_2 \leq 5 \end{Bmatrix} \quad \gamma(x_i) \in \begin{Bmatrix} x_1 < 80\% \\ x_2 \leq 1 \\ x_3 \leq 1 \end{Bmatrix}$$

Data Representation:

Beginner : 00

Intermediate : 01

Advanced : 11

5.0 Integration Of Explicit And Implicit Techniques: Classifier III

For the third classifier, we integrate the explicit and implicit data extraction to obtain input data for the student classification.

5.1 Training Data Sample

Students' behavior data are collected from analysis of web log data similar to the second classification process. The data considered are the requested pages and time of request. In the explicit user data extraction process, we collect students' score value from system's database as an additional input data into our program to calculate the desired output data for the training sample. The question and answer session is used to test the student's understanding of the material being learned. Table 3 shows the possible criteria for input data to calculate the desired output.

Table 3: Criteria To Classification III

Criteria	Beginner, α	Intermediate, β	Advanced, γ
Learning Time, x_1	$x_1 > 100\%$	$80\% \leq x_1 \leq 100\%$	$x_1 < 80\%$
Freq. Of Backtracking, x_2	$x_2 > 5$	$3 \leq x_2 \leq 5$	$x_2 \leq 1$
Freq. Of Using Help, x_3	$x_3 > 5$	$3 \leq x_3 \leq 5$	$x_3 \leq 1$
Score, x_4	$x_4 < 60\%$	$60\% \leq x_4 \leq 800\%$	$x_4 > 80\%$

Mathematically,

$$\alpha(x_i) \in \begin{Bmatrix} x_1 > 100\% \\ x_2 > 5 \\ x_3 > 5 \\ x_4 < 60\% \end{Bmatrix} \quad \beta(x_i) \in \begin{Bmatrix} 80\% \leq x_1 \leq 100\% \\ 3 \leq x_2 \leq 5 \\ 3 \leq x_2 \leq 5 \\ 60\% \leq x_4 \leq 800\% \end{Bmatrix} \quad \gamma(x_i) \in \begin{Bmatrix} x_1 < 80\% \\ x_2 \leq 1 \\ x_3 \leq 1 \\ x_4 > 80\% \end{Bmatrix}$$

Data Representation:

Beginner	: 00
Intermediate	: 01
Advanced	: 11

6.0 Conclusion And Future Work

Data preprocessing is implemented through the process of extracting information from registration data (explicit technique) and web log analysis (implicit technique) to develop a complete the student profile that use the system. We defined three types of classifiers to be used in classifying the students' status based on input data collected from registration form, web log data or combination of the web log data and the students' score while doing the exercises.

The results from this phase is a set of training data including input data and desired output data that will be used in classification with neural network for future work. Interesting issue here concern the session identification using student's metric number. In this paper we consider IP address as a unique session identifier. For the future research, we will look at integration of metric number variable and IP address and request time to identify unique session.

Acknowledgements

This research is conducted under IRPA grant, vote project 74129: The Development of Knowledge Management For Adaptive Hypermedia Learning System, Faculty of Computer Science, UTM Skudai.

References

1. **Bailey, D. And Thompson, D. (1990).** ***How To Develop Neural-Network Applications*****. In AI Expert. June 1990. 38-45.**

2. **Brusilovsky, P. (1999).** ***Adaptive and Intelligent Technologies For Web-Based Education*****. In C. Rollinger and C. Peylo (eds) Kunstliche Intelligenz, Special Issue on Intelligent Systems and Teleteaching. 4. 19-25.**

3. **Buser, D., Kauffman, J., Llibre, J.T., Francis, B., Sussman, D., Ullman C. And Duckett J. (2000).** ***Application, Sessions And Cookies*****. Wrox Press Ltd. (2000). Beginning Active Server Pages 3.0. 289-330.**

4. **Catledge, L.D. and Pitkow, J.E. (1995).** ***Characterizing Browsing Strategies in The World Wide Web*****. Georgia Institute of Technology, Atlanta, USA.**

5. **Cooley, R., Mobasher, B. And Srivastava, J. (1997). Web Mining: Information and Pattern Discovery on The World Wide Web. Departmanet of Computer Science, University of Minnesota, USA.**

6. **Cripps, A. (1996).** ***Using Artificial Neural Nets to Predict Academic Performances*****. In Association of Computing Machinery, Inc.(ACM). 33-37.**

7. Hashim, S.Z.M., Yusof, N., Samsuri, P. and Ahmad, N.B. (2001). *Design and Implementation of An Adaptive Web-Based Hypermedia Learning System: A Prototype*. In Proceedings of The International Conference on Organizational Development and Leadership in Education. Kuala Lumpur, Malaysia.

8. **Jeon-Slaughter, H. (2002).** ***Report of Web Log Data Mining*****. In E-Journal User Study: Recent Findings. Stanford University Libraries.**
   URL: http://ejust.stanford.edu/logdata.pdf

9. Moore, D.J., Hobbs, D.J., Mullier, D. And Bell, C. (1997). *Interaction Paradigms With Educational Hypermedia*. In In Proceedings of 23rd EUROMICRO Conference'97 New Frontiers of Information Technology. Budapest, Hungary.

10. **Mullier, D.J. (2000). *The Application of Neural Network And Fuzzy Logic Techniques To Educational Hypermedia*. Ph.D Thesis. Leeds Metropolitan University, UK.**

11. Mullier, D.J. (1995). *Using A Neural Network To Model Hypermedia Browsing – An Alternative To Traditional Intelligent Tutoring Methods*. Faculty of Information and Engineering Systems, Leeds Metropolitan University, UK.

12. **Quentin-Baxter, M. (1993). *Hypermedia Learning Environments Limit Access To Information*. Faculty of Medicine Computing Centre, University of Newcastle Upon Tyne, UK.**
URL: http://decweb.ethz.ch/WWW7/1941/com1941.htm

13. **Wang, Y. (2000). *Web Mining and Knowledge Discovery of Usage Patterns*. In CS 748T Project Report (Part I).**
URL: http://db.uwaterloo.ca/~tozsu/courses/cs748t/surveys/wang.pdf

14. Yusof, N., Salim, A., Samsuri, P., Hashim, S.Z.M. and Ahmad, N.B. (2000). *SPAtH: An Adaptive Web-Based Learning System*. In Proceedings of IASTED International Conference. Banff, Canada.

15. **Zaiane, O.R. and Luo, J. (2001). *Towards Learners' Behaviour in Web-Based Distance Learning Environment*. Department of Computing Science, University of Alberta, Canada.**

PAPER VI:

DEVELOPMENT OF AN ADAPTIVE HYPERMEDIA LEARNING SYSTEM

Nor Bahiah Hj Ahmad, [1]Shahliza Abd Halim, [2]Bariah Yusob, [3]Norreen Haron

Department of Software Engineering
Faculty of Science Computer and Information Systems

University of Technology Malaysia, 81310 Skudai

bahiah@fsksm.utm.my
[1]shahliza@fsksm.utm.my
[2]byvail@gmail.com
[3] norreen.haron@gmail.com

Abstract

This paper presents the architecture and design for the development of an adaptive hypermedia learning system for teaching and learning Data Structure. The learning system developed integrates pedagogy of education and intelligence technique in the presentation of the learning material and the suggested navigation path. The system comprises of three main components, user profile model, domain model and adaptive engine. The focus of this research is on the use of computational intelligence technique in the classification of student models and in the adaptation of learning material and navigation path. Both learning style and knowledge acquisition have been considered as feature of adaptation in the learning system. Supervised Kohonen network is used to classify the students into advance, intermediate and beginner based on their knowledge acquisition and performance. Meanwhile, fuzzy logic is used to dynamically adapt the choice of possible paths through the learning material based on the students learning style. The integration of the two intelligence technique is able to personalize the user both at the presentation and navigation level. This study is also expected to solve disorientation and lost in hyperspace problem that usually occur in conventional hypermedia learning system.

Introduction

The advancement in educational technology has transformed the processes of teaching and learning in higher education from traditional method into computer based approach. Web-based learning environment are now used extensively as components of course delivery in higher learning institutions. Academics are placing more course material on-line to supplement their lecture based delivery method. However, research showed that traditional web-based learning system is not always suitable for all learners. Not all learners is able to work independently in web-based learning systems. Student's individual differences such as student's background, learning styles, cognitive styles and prior knowledge possessed may need more attention and support from instructional designers. The flexibility of linking in hypermedia learning environment also impose problems related to cognitive overload and lost in hyperspace which caused confusions and frustrations among learners. Therefore, adaptive learning system is important in providing students the capability to tailor the learning material presented based on the information stored in the student profile.

The purpose of this research is to develop an adaptive hypermedia learning system for learning Data Structures at Faculty of Computer Science and Information System in Universiti Teknologi Malaysia. We have explored the computational intelligence technique that effectively able to adapt the learning material based on the student learning style. In this case, we combined the students personality factor (Myers-Briggs Type Indicator (MBTI)) and fuzzy logic techniques to produce a dynamic course adaptation which will present the appropriate structure of the learning material to the student [(Norreen & Naomie (2005)]. We also have experiment supervised Kohonen network in order to classify the student based on their level of knowledge acquisition [Bariah et al. (2004)]. Besides that, we follow the software engineering principle in developing the architecture design and the development of the system. Special attention also have been made in the design of the structure of the learning material in order to present the material that can imitate the way human teacher teach in the classroom.

This paper will describe how the adaptive hypermedia learning system has been developed. First, we discusses the adaptive hypermedia learning system concepts, followed by features of adaptation and system development. The architecture of the system and the design of the course material are explained in the system development phase. Supervised Kohonen that has been used for classifying students based on their knowledge level is discussed in the subsequent section.

Adaptive Hypermedia Learning System

Adaptive hypermedia learning system (AHLS) can be defined as the technology that allows personalization for each individual user in hypermedia learning environment. In this case, adaptive refers to the ability of the website to change its behaviour or responses in reaction to the way it is used. AHLS is the solution to the conventional hypermedia learning system which allows freedom to the user to navigate through the system according to their preferences and paces. AHLS on the other hand will present the learning material based on user preferences, goal, and knowledge acquisition and user characteristics. The adaptation of AHLS is implemented through a decision making and personalization engine which adapts the content according to a user model.

AHLS system mostly comprises of three main components, user profile model, domain model and adaptive engine [Brusilovsky (2001)]. User profile model stores the learning activities, learning performance and interaction history of each student in the database. The profiles were extracted from both explicit and implicit user profile. The explicit information is the information that the learner gave willingly or directly and he/she is aware that the information is kept in the database. The implicit information is the information the system collects without the learner acknowledgement. It records the learner's activity and behavior as he/she navigates through the system.

Adaptive navigation path provide the annotated link based on the interaction history of each student. To reduce disorientation, each student will get different paths based on their level of knowledge acquisitions. Adaptive engine will determine the appropriate learning material and the navigation path based on the student's status that was retrieved from the user profile model. Domain model stores all the teaching materials including the learning objectives, lecture notes, examples, exercises and the answer for each question.

Features Of Adaptation

Research activity in the e-learning domain that apply adaptive features in web-based learning has been very intense. Methods and techniques of adaptive hypermedia had been introduced by Brusilovsky. Information that usually used are user's goals/tasks, knowledge, background, hyperspace experience, and preferences [Brusilovsky (2001)]. At least two more items can be added to this list; the user's interests and individual traits [Brusilovsky,2003)]. User interests are not a new feature of the user to be modeled. User's individual traits is a group name for user features that together define a user as an individual. Examples are personality factors such as introvert or extravert, cognitive factors, and learning styles. Like user background, individual traits are stable features of a user that either cannot be changed at all, or can be changed only over a long period of time. [Kobsa et al. (1999)] suggested distinguishing adaptation to user data, usage data, and environment data. User data comprise various characteristics of the users, usage data comprise data about user interaction with the systems and environment data comprise all aspects of the user environment that are not related to the users themselves.

Among other variables that influence the success of learning is learning styles [Ford & Chen (2000)]. It is important to diagnose the students learning style because some students learn more effectively when taught with preferred methods. Information about the learning style can help system become more sensitive to the differences of students using the system. Understanding learning styles can improve the planning, producing, and implementing of educational experiences, so they are more appropriately tailored to students' expectations, in order to enhance their learning, retention and retrieval [Zywno & Waalen (2002)]. Sadler-Smith (1997) identified four broad categories of learning style in an attempt to acknowledge and accommodate the range of aspects of individual differences referred in the educational psychology literature in an holistic way. Table 1 listed all the categories.

[Papanikolaou et al.(2003)] analyzed learners' studying behavior such as time spent and hits on resources and navigation traces by the different learning style categories to provide evidence about the way learners that belong to different learning style categories select and use educational.

Table 1 : Learning style categories [Sadler-Smith (1997)]

Learning Style Categories	Method of learning
cognitive personality	field dependence and field independence
Information-processing style	(converger, diverger, accommodator, assimilator) or activist, reflector, theorist, pragmatist
Instructional preferences	individual's propensity to choose or express a liking for a particular instructional technique or combination of techniques
Approaches to Studying	Deep, surface, strategic approach, lack of direction

Based on the literature done, our research has focused on the design of adaptation based on the learning style information and level of knowledge acquisition. To determine the level of knowledge acquisition, we have selected the attributes of adaptation based on the work done by [Papanikolaou et al.(2003)] and [Paridah et. al. (2001)]. The attributes selected are the learning time, number of backtracking, number of getting help function and the score earned while doing the exercise. Meanwhile, the pedagogical and learning style refer to student's personality factor based on (MBTI) as explained in [Norreen & Naomie (2005)]. We also have identified the structure of the learning material that the system should offer to learners with different styles and characteristics.

System Development

The adaptive hypermedia learning system developed in this research will be used by students taking Data Structure subject. This course is offered to second year computer science students in Faculty of Computer Science and Information System. The main objective of this course is to introduce the data structure concepts and to enable the students to apply the concepts in programming. Students are expected to master both the theory of the data structures and also the programming part. From observations, we found out that weak students usually have problems in understanding the theory and programming part. They need more explanation, more exercise and more practical in the computer laboratory compared to other students. Meanwhile, moderate students can understand the theory part easily but have difficulty in the programming part. Advance students don't have much problem in understanding the theory part and able to implement the program without many difficulties. In average, we conclude that some students need more explanation in the theory part, need to do a lot of exercises and need more practices in the programming part compared to others.

The overall architecture for the Adaptive Hypermedia System is shown in Figure 1. There are three main components in this system, adaptive engine, domain model and user profile model. Adaptive engine interface enable the user to interact with the system. User will be presented with the learning material and navigation path adapted based on the learning style and information stored in the user profile model. Adaptive engine generate features of adaptation, a tree structure to present the learning material structure.

User profile model stores the information about learners. The information are extracted from both explicit and implicit user profile. The explicit information is the information that the learner gave while fill in the registration form and the questionnaire while using the system for the first time. The implicit information is the information that the system collects while the user interact with the system. The user activity such as the URLs accessed and the behavior while he navigates through the system is recorded. The user behaviour being considered is the usage of help properties, such as referring to glossary, using back and previous button and the learning time taken while learning a concept. The information captured in the user profile will be normalized and given as input to supervised Kohonen self-organizing maps to identify the class of student, whether as beginner, intermediate and advanced.

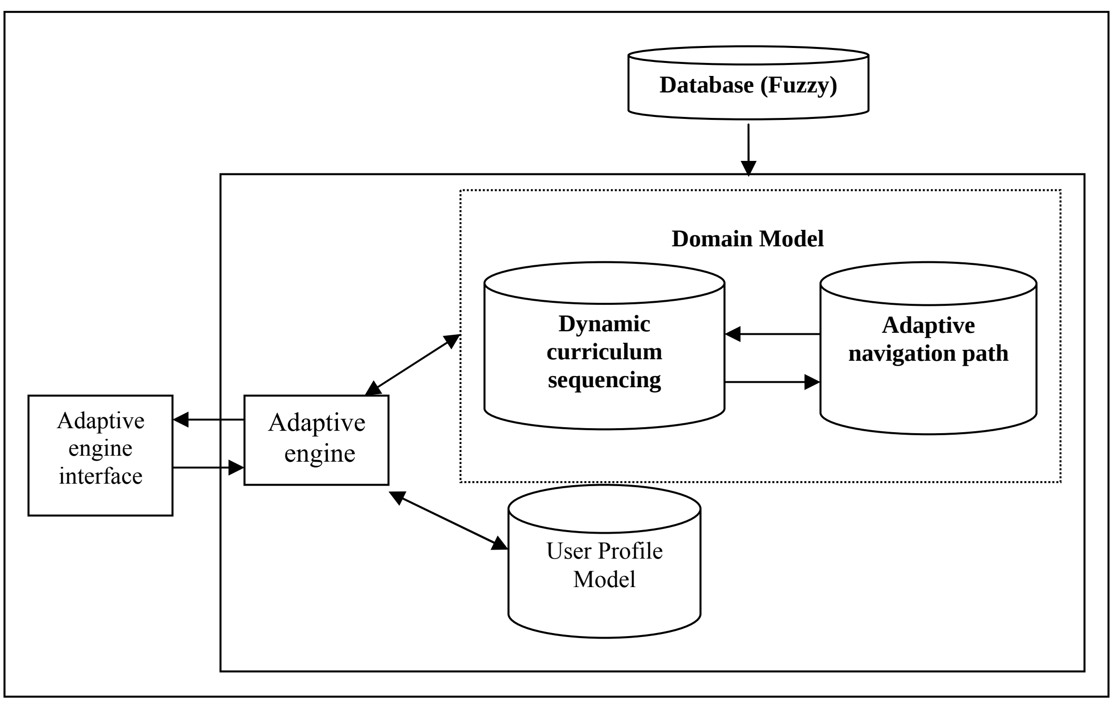


Figure1: Architecture of an adaptive hypermedia system

Domain model comprise two subcomponents, dynamic curriculum sequencing and adaptive navigation path. Dynamic curriculum sequencing will determine the student learning style and the sequence of learning material structure based on their learning style. In this research, we have identified four types of student learning style, Introvert-Sensor, Extrovert-Sensor, Introvert-Intuition and Extrovert-Intuition based on personality factor MBTI. Adaptive navigation path component further will assist the student while navigating through the system. The link showed to each student will be annotated with different colors such as yellow, red and green to signify the material that the student have already learned, forbidden because of prerequisite is not fulfilled and ready to be learned. The link annotation can help prevent the student from disorientation while navigates.

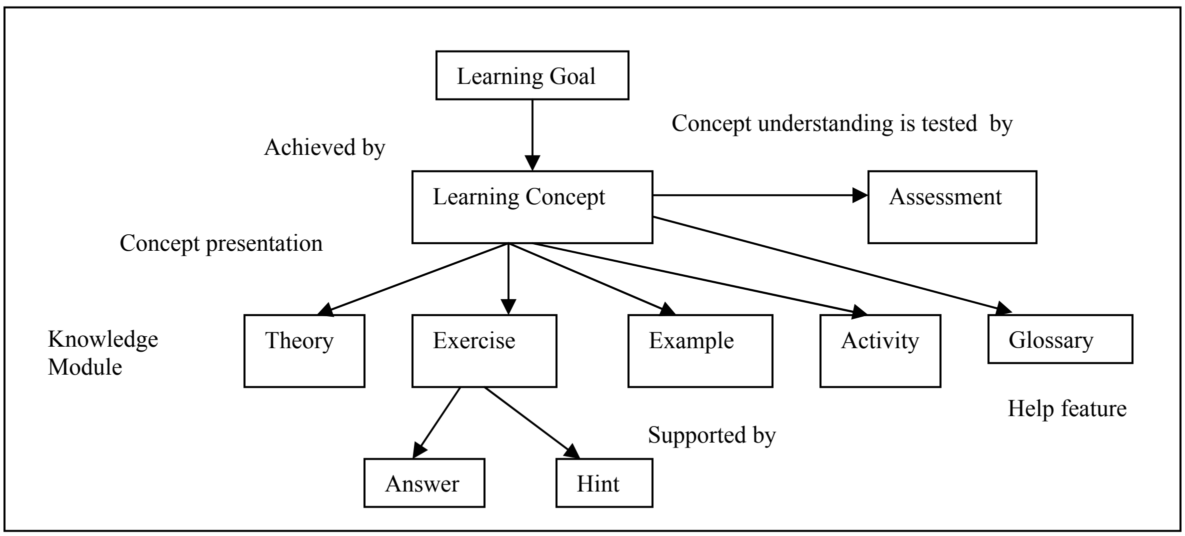


Figure 2: The structure of the learning material

Table 2 : The coding scheme for the URL for each category

Learning Category	**URL for topic X**
Learning Objective	T0XF00XO00X
Theory	T0XF00XN00X
Example	T0XF00XC00X
Exercise	T0XF00XX00X
Activity	T0XF00XA00X
Glossary	T0XF00XG00X
Assessment	T0XF00XS00X

Figure 2 shows the structure of the learning material. The student can choose any learning goal which has fulfill the pre-requisites. There are several learning concepts to be learned in order to achieve the learning goal. The learning concept is presented into four categories; example, theory, activity and exercise. While using the system, if the student is not familiar with certain terminology or concept, the student can refer to the glossary provided as a help feature in the system. The URL for each category has been set as Table 2. The coding system is used in order to identify the navigation path that has been accessed by the student. Further, based on the number of URL on certain code, it will be easier to determine the coverage of certain category of the learning material.

For each type of student, different category of the learning material will be presented adaptively in different sequence by the system. To identify the student learning style and also the sequence of the learning structure, fuzzy logic technique is used to generate suitable rule of adaptation.

System Flowchart

There are two different flows designed for first time user and second time user as shown in Figure 3. The first time user has to fill in a registration form and a questionnaire. The information collected in the registration form is used to classify the knowledge level of the student explicitly. The result from the questionnaire is used to determine the student's personality factor value either as introvert, extrovert, intuitive and sensor. In order to determine the prominent learning style for a student, fuzzy logic technique is used to identify the student learning style either fall into one of these four types Introvert-Sensor, Extrovert-Sensor, Introvert-Intuition or Extrovert-Intuition.

Both data from the preliminary classification of student's knowledge level and the category of learner are used to determine the adaptive presentation for the learning material structures sequence. For example, if the student is fall into beginner student and is identified as an Introvert-Sensor category of learner, the presentation of the learning material structures is in the sequence of Example, Theory, Activity and Exercise as presented by the tree structure depicted in Figure 4. Figure 4 shows the tree structure for learning Sorting concept. Different category of learner will get different tree structure that display different sequence of learning structure.

The knowledge level of the student will be updated once the student logs off the system. In this situation Kohonen program will be executed and given the latest information on the student activities while learning on-line, the score for doing assessment and the interaction history as input vector. Meanwhile, second time user must login before using the system. The current knowledge level and the diagnosed learning style will be extracted from the database in order to display the learning material adapted to the status level and learning style.

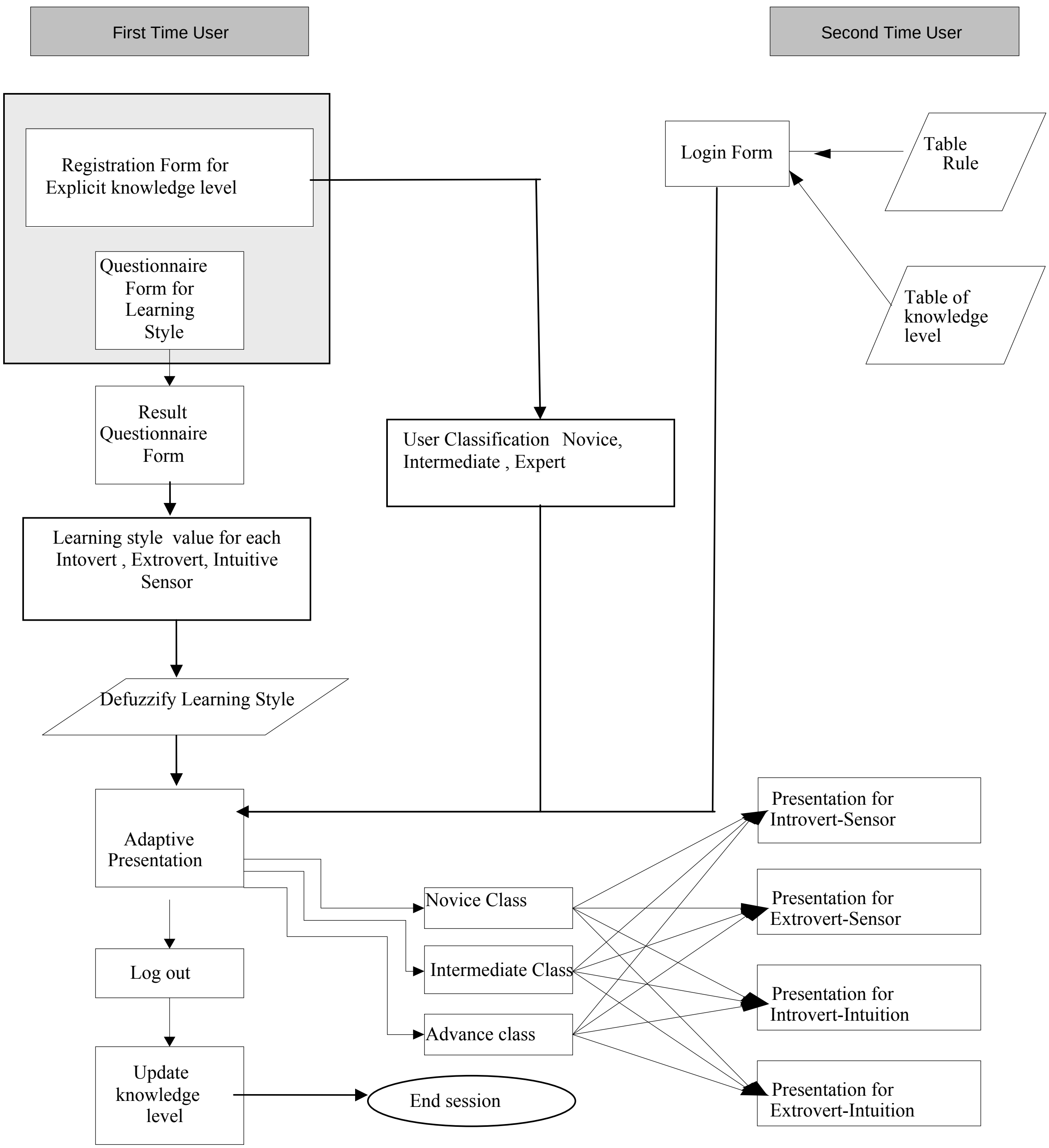


Figure 3: The system flow chart

Isihan

- ⇒ Objektif Pembelajaran
- ⇒ Contoh
  - o Isihan Buih
  - o Isihan Pilihan
  - o Isihan Selitan
  - o Isihan Cantum
  - o Isihan Cepat
- ⇒ Teori
  - o Pengenalan
  - o Isihan Buih
  - o Isihan Pilihan
  - o Isihan Selitan
  - o Isihan Cantum
  - o Isihan Cepat
- ⇒ Aktiviti
  - o Isihan Buih
  - o Isihan Pilihan
  - o Isihan Selitan
  - o Isihan Cantum
  - o Isihan Cepat
  - o Kesimpulan
- ⇒ Latihan
  - o Isihan Buih
  - o Isihan Pilihan
  - o Isihan Selitan
  - o Isihan Cantum
  - o Isihan Cepat
- ⇒ Penilaian pelajar

Figure 4: The tree structure for Introvert-Sensor category

Supervised Kohonen For Classifying Students

Supervised Kohonen is used to classify students based on their knowledge acquisition. It employs a teacher signal when processing nodes for which a target label exists. The idea is to assign weight vectors to the same class as the node that was mapped at this location. Training proceeds in a similar manner as the unsupervised case with the difference that the weight entries are rejected if they belong to a different class. Table 3 shows the criteria to classify learners into beginner, intermediate and advances. When the learner log into the learning system, the system will count the time he/she spent on learning a concept. The system suggests the time to be spent on each category of the learning material. The learning time is calculated based on the percentage actual learning time taken to finish learning from the estimated time as shown in Table 4.

Table 3: Criteria for learner's classification

Attribute	Beginner	Intermediate	Advanced
Learning Time, *t*	$t > 80\%$	$30\% \leq t \leq 80\%$	$t < 30\%$
Numb. Of Backtracking, *b*	$b > 4$	$2 \leq b \leq 4$	$b < 2$
Numb. Of Using Help, *h*	$h > 4$	$2 \leq h \leq 4$	$h < 2$
Score, *s*	$s < 30\%$	$30\% \leq s \leq 80\%$	$s > 80\%$

Table 4 : Calculation scheme for learning time

Part	Student's actual leaning time	Estimated learning time by SPATH
Theory	acTime1	esTime1
Exercise	acTime2	esTime2
Example	acTime3	esTime3
Activity	acTime4	esTime4
SUM for student's actual learning time (SUM_acTime)		acTime1+acTime2+acTime3+acTime4
SUM for estimated learning time (SUM_esTime)		esTime1+esTime2+esTime3+esTime4
Average for student A's learning time (Avg_acTime)		SUM_acTime/4
Average for estimated learning time (Avg_esTime)		SUM_esTime/4
Learning time, t		Avg_acTime/Avg_esTime X 100 %

The number of backtracking shows that the learner is not fully master the concept, lose direction or change his/her learning goal. The number of backtracking is defined by counting how many times the learner reopen any pages in particular concept. In this research, help function is a list of definition and explanation on terms used in the notes given. This attribute shows that the more help the learner gets, the more he/she is having a difficulty in understanding a concept. The number of getting help is defined by counting how many time the learner click on the help button in particular concept. To test the learner's level of mastering, the system provides an assessment at the end of the learning session. The score is calculated by the percentage of correct answers given.

All the data must be transformed into a standard format to get a valid and accurate classification. The transformation of the data is included in the pre-processing phase using a normalization method. In the training phase, input data is given to the Kohonen network. The weights are captured after completing the training phase. In this experiment the size of the training sample is 1050. In the testing phase, there is no target data is provided. We used 450 dataset to the network. The network classifies the data based on the weights and outputs are obtained. When the testing results were obtained, the percentage of the classification accuracy was calculated.

During training, the network learns the data and generates the weights by calculating the nearest distance to the real data presented. From 450 data presented, the network is able to classify 445 data correctly. From the result achieved, we concluded that Kohonen network is capable of classifying the learners' data into the right class. It gives more than 90% accuracy in both training and testing phase. The Kohonen's SOM is definitely a good tools to classify data into a number of groups without supervision. It will be very useful in this study because it can deal with more complex and bigger sample of data when it is applied to the real learners' data in the learning system's database.

Conclusion and Further Work

This paper has proposed a way to personalize the course content for adaptive hypermedia learning system, which aims to provide learners with a customized learning environment. It emphasizes the combination of pedagogical theories and artificial intelligent techniques. Supervised Kohonen has proved to be the suitable technique to classify the student accurately. Fuzzy logic is able to identify the personality factor of a student and the material presented will be based on the personality factor. The design of the learning material is also important in supplying the right material with the right learning style and knowledge level. For future work, we intend to implement web usage mining in order to measure the efficiency of the learning approach given to the student based on their learning style.

Acknowledgement

This research is conducted under IRPA grant, project vote 74129: The Development of Knowledge Management for Adaptive Hypermedia Learning System.

References

Bariah Yusob, Norreen Haron, Nor Bahiah Hj Ahmad, Shahliza Abd Halim. (2004). Individualizing the learning material and navigation path in an adaptive hypermedia learning system. 2nd National Conference On Computer Graphics & Multimedia, Kajang, Selangor , 8-10 December.

Bariah Yusob, Siti Mariyam Hj Shamsuddin, (2005). Supervised Kohonen Network for classification problems. Simposium Kebangsaan Sains Matematik ke-XIII UUM/PERSAMA/Kerajaan Negeri Kedah, 31 Mei - 2 Jun

Brusilovsky, P.(2001). Adaptive Hypermedia. *User Modeling and User-Adapted Instruction*, 11(1 – 2),
87-110.

Brusilovsky P, Christoph P. (2003). Adaptive and Intelligent Web-based Educational Systems, *International Journal of Artificial Intelligence in Education* ,13, 156–169.

Ford, N, Chen S.Y. (2000). Individual Differences, Hypermedia Navigation, and Learning: An Empirical Study. *Journal of Educational Multimedia and Hypermedia*, 9(4), 281-312.

Norreen Haron, Naomie Salim.(2005). Integrating personality factor (MBTI) and Fuzzy logic techniques in individualizing learning material of adaptive hypermedia learning system. To be presented at Konvensyen Tenologi Pendidikan Ke 18. 16-19 September, Marang, Terengganu.

Siti Zaiton, M.H., Norazah, Y., Paridah, S., & NorBahiah, H.A. (2001). Design and implementation of an adaptive web-based hypermedia learning system: a prototype. In: Proc. Of International Conference on Organizational Development and Leadership in Education, Kuala Lumpur.

Paridah, S., Nor Bahiah, H.A., & Norazah, Y. (2001). Neural Network Application on Knowledge Acquisition For Adaptive Hypermedia Learning. In: *Chiang Mai Journal Science*, Jan. 2001.

Papanikolaou, K.A., Grigoriadou, M., Kornilakis, H., Magoulas, G. D.: Personalizing the interaction in a Web-based educational hypermedia system: the case of INSPIRE. *User- Modeling and User-Adapted Interaction* 13(3), 213-267

Sadler-Smith E., (1997) Learning Style: frameworks and instruments. *Educational Psychology*, 17, 51-63.

Zywno, M.S. and Waalen, J.K. (2002), The Effect of Individual Learning Styles on Student Outcomes in Technology-enabled Education, *Global Journal of Engenering. Education*, 6(1)

PAPER VII:

SUMMATIVE EVALUATION FOR THE USABILITY OF INTELLIGENT TUTORING SYSTEM (SPATH)DEVELOPMENT OF AN ADAPTIVE HYPERMEDIA LEARNING SYSTEM

Shahliza Abd Halim[1], Nor Bahiah Hj Ahmad[1], Norreen Haron[2] ,Bariah Yusob[2]

Department of Software Engineering

Faculty of Science Computer and Information Systems

University of Technology Malaysia, 81310 Skudai

[1]{shahliza, bahiah}@fsksm.utm.my

[2]{byvail, norreen.haron} @gmail.com

Abstract

The various combination of multidisciplinary areas in Intelligent Tutoring System (ITS) contributes to various methods of evaluation for ITS and therefore the suitable method of evaluation must be chosen wisely. This paper discusses the evaluation of an adaptive hypermedia learning system called SPATH. SPATH provides adaptation of learning content by personalizing the learning material structure based on student learning style adapted from student's personality factor (Myers-Briggs Type Indicator (MBTI) and by the students knowledge acquisition level. In this research, we use summative evaluation where we address the educational impact of this system on students and its practical acceptability in terms of usability. In the usability study we adopt the questionnaire technique where three key factors, learnability, efficiency and satisfaction in usability were measured. Preliminary results of the usability study revealed that this system has a high percentage in learnability and satisfaction factor.

1. Introduction

Most of the early research in adaptive educational hypermedia was inspired by the area of Intelligent Tutoring Systems (ITS). A combination of ITS and learning materials organized as hypermedia was the beginning for the research on adaptive hypermedia system (AHS)[1]. Due to the close relationships between ITS and AHS, the method of evaluation for ITS is considered as suitable evaluation for AHS. Among the advantage of evaluation is that it provides an opportunity to learn from mistakes and is capable of improving the ITSs life-span as well as their usability [2].

Adaptive Hypermedia Learning System (AHLS) has been identified as an effective strategy for solving many learning problems involved in large hypermedia, such as cognitive overload and user disorientation [3]. The idea of adaptive hypermedia is to adapt the course content accessed by a particular user to his/her characteristics. Most adaptive educational hypermedia system researches focuses on adapting to user features like user's goals/tasks, knowledge, background, hyperspace experience, preference and interests [1]. However, a web-based educational system must also include information about student learning styles to adapt optimally instructional materials to the student [3],[4],[5],[6]. Identification of the learner's learning style is the key to the development of a hypermedia course that addresses different learning styles [6].

This research has developed an adaptive hypermedia learning system called SPATH for learning Data Structures at Faculty of Computer Science and Information System in Universiti Teknologi Malaysia. We have explored the computational intelligence technique that effectively able to adapt the learning material based on the student learning style. In this case, we combined the students personality factor (Myers-Briggs Type Indicator (MBTI)) and fuzzy logic techniques to produce a dynamic course adaptation which will present the appropriate structure of the learning material to the student [7]. As for the classification of students based on their level of knowledge acquisition we used supervised Kohonen network for classification purpose[8].

There are three main components in SPATH, adaptive engine, domain model and user profile model. Adaptive engine interface enable the user to interact with the system. User will be presented with the learning material and navigation path adapted based on the learning style and information stored in the user profile model. Adaptive engine generate features of adaptation and a tree structure to present the learning material structure.

User profile model stores the information about learners. The information is extracted from both explicit and implicit user profile. The explicit information is the information that the learner gave while filling in the registration form and the questionnaire when they used the system for the first time. The implicit information is the information that the system collects while the user interact with the system. Domain model comprise two subcomponents, dynamic curriculum sequencing and adaptive navigation path. Using Fuzzy logic, the student learning style and the sequence of learning material structure based on their learning style has been determined. In this research, we have identified four types of student learning style, Introvert-Sensor, Extrovert-Sensor, Introvert-Intuition and Extrovert-Intuition based on personality factor MBTI. Table 1 shows the sequence of the learning structure based on the learning style identified.

Table 1. Example of 4 different situations of the learning material structure based on the learning style

Learning Style	Structure of learning material
Introvert-Sensor	Example Theory Activities Exercise
Extrovert-Sensor	Activities Exercise Example Theory
Introvert-Intuitive	Theory Example Exercise Activities
Extrovert-Intuitive	Exercise Activities Theory Example

The knowledge level of the student will be updated once the student logs off the system. In this situation Kohonen program will be executed and be given the latest information on the student activities while learning on-line, the score for doing assessment and the interaction history as input vector. Adaptive navigation path component further will assist the student while navigating through the system.

This paper discusses the evaluation of SPATH in order to measure the efficiency, learnability and the satisfaction of the system usability. As researchers have pointed out in [2,10] the need for the evaluation of Intelligent Tutoring Systems (ITS). Among the advantage of evaluation is that it provides an opportunity to learn from mistakes and is capable of improving the life-span ITSs as well as their usability [10]. The organization of this paper is as follows: section 2 and 3 discusses the evaluation for ITS in general and proposes a suitable method for the usability study of this system. Section 4 and 5 discusses the method and process done while performing the study and also the illustrations of the preliminary results. Lastly, in section 6 and 7 is the discussion and follows by conclusion and future work of this research.

2. Evaluation Methods for ITS

Evaluation for ITS is not an easy tasks as ITS is the combination of multidisciplinary areas including expert system, Computer Based Instruction (CBI), education, psychology and also Human Computer Interaction (HCI) [2,9,10]. [2] and [9] highlighted that there are few agreed upon standards within the ITS community to guide investigators who wish to evaluate ITS given the diversity of the evaluation methods.

There are two types of evaluation for ITS, formative and summative evaluation [2,9]. Formative evaluation mainly occurs during design and early development of a project. It frequently addresses the question of relationship between the architecture of ITS and its behavior. On the other hand, summative evaluation is concerned with the evaluation of a completed system and making of formal claims about the system. It answers the question regarding the educational impact of an ITS on students. However, these types of classification are still too broad where a lot of methods can fall in either one of these classes.

In [2], the various methods have been further classified, so that the method could be differentiated from a number of others on a scale between external evaluation (considering the whole system) and internal evaluation (testing a component of the system). In addition a method could be classified along a dimension consisting of exploratory research versus experimental research. Though the proposed classification provide a simple yet robust way to select evaluation methods, the classification needs future work to add other dimension of formative and summative evaluation to the classification chart.

Another research has been done on evaluation solely on the usefulness of web based learning environment [10]. The research takes into consideration the multidisciplinary evaluation framework for evaluating a web based learning system. The framework combines two main issues, usability and utility issues where utility is broken down into two parts, pedagogical usability and added value. In [10], it shows the importance of usability remains the same regardless of how much web based learning is used in the teaching as a whole. As for pedagogical usability, its importance gradually increases as the focus of teaching shifted from traditional teaching into more on web based teaching.

In this research, we use summative evaluation where we address the educational impact of this system on students. We also use the multidisciplinary evaluation framework where we focus on the usability study of this system as [10] clearly shows that that the importance of usability is consistent regardless of the focus of teaching either more on traditional or more towards web based teaching.

3. Usability Study for ITS

Usability is a quantitative and qualitative measurement of the design of user interface grouped into five key factors: learnability, efficiency, memorability, errors and satisfaction [11]. The setting of the usability study can vary. A usability laboratory can be used for a controlled experiment. A workplace test can be used to test the user in their normal work environment such as at their desk during a routine work day. There is also web-based usability testing referred to as remote usability testing where the user and experimenter are not physically located in close proximity of each other. The different categories of usability tests consist of performance measurement, thinking aloud protocol, coaching method, retrospective testing, constructive interaction, and questionnaires [11, 12]

Performance measurement takes place when quantitative measures are taken during the testing such as the number of tasks completed successfully by the user, length of time to complete the test tasks, number of errors, and time spent recovering from errors. Thinking aloud protocol exists when users vocalize their thoughts and therefore share their positive and negative interpretations of different website features. The coaching method enables the users to ask questions and receive answers which give researchers insight into the type of help documentation or better technology design needed. Questionnaires are also a form of testing as it provides an opportunity to gather more usability feedback from a user after a testing session [13].

In this research, questionnaires were used as a tool to gather feedback from the participants and this research is conducted in a controlled environment setting. Three out of the five key factors, learnability, efficiency and satisfaction of the students were the focus of the study. Learnability of the students is based on the ease of use of the students when working towards completing the task specified for them. Efficiency looks at how productive the students once having learned the software and the last attribute satisfaction is to study the students level of pleasure using the system.

4. Methods and Processes

For this research, we have performed summative evaluation whereby a total of 44 computer science students in Faculty of Computer Science and Information System, Universiti Teknologi Malaysia have participated in this study. The learning material for the participants to study, task and processes of the usability study is discussed in the following subsections.

4.1 Learning Material

The learning material used in this study is based on sorting techniques which is explained fully in Malay language. The topics involved are bubble sort, selection sort, insertion sort, merge sort and quick sort. In order to handle different types of learning styles, the learning material for this topic is structured into theory, exercise, example and activity [14]. The theory part consist of explanation and pictures describing the sorting technique, exercise drill the students on the sorting techniques and gives their score and hints for the answer of each questions. Example shows the simulation of the sorting technique. Students can view the simulated sorting activity based on the execution of the algorithm code line by line. In the activity session, students can generate data to be sorted randomly and view the simulation of the sorting technique.

The learning material structure will be adapted based on student learning style. Different category of learner will get different tree structure. Figure 1 shows the sequence of the learning material based on the Extrovert-Intuitive students. For each learning material structure, links are provided for each topic in sorting.

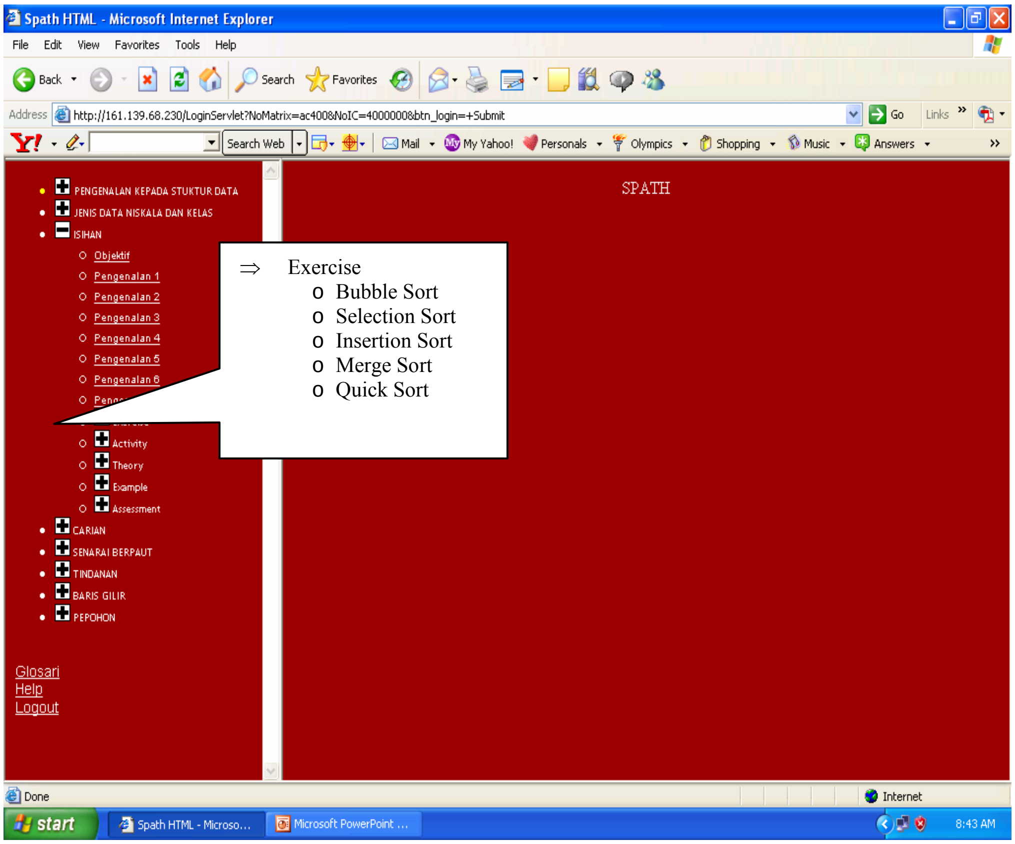


Fig. 1. The tree structure for learning sorting concept in support of Extrovert-Intuitive students

4.2 Task

The participants were given a task to study Sorting Techniques based on the learning materials discussed in section 4.1. As part of the study task, it was determined that the following steps would be the task process each participant would go through in the usability study:

- Hearing brief explanations about the system.
- Reading the task instructions.
- Reading short and simple user manual if they need further guidance about the system.

– Filling in the on-line questionnaire based on MBTI in order to determine the student's personality factor.

– Performing the task to study the learning materials.

– Completing the assessment on-line after they have completed the study session.

– Filling in questionnaires on learning satisfaction.

– Filling in questionnaire on usability of the system.

4.3 Processes

The entire task specified will be further refined and elaborated. The usability study is carried out in laboratory in a controlled environment where the students were required to complete their learning in the lab for two hours or until they finish their learning. Two experimenters were around to give explanations or help if needed. A desktop with Internet connection was provided to each participant to complete the study. The process of the study is as follows:

1. For this study, students were divided into two groups, the first group is required to learn based on the topic sequence provided and the second group can learn freely without following the sequence.
2. The students were given a short and simple user manual to guide them while using the system
3. Before the student can use the system, he/she has to register and fill in an online questionnaire. The information collected during registration were used to gain some personal information such as the user's name, id, cpa and programming knowledge status in order to initialize the student model.
4. On-line questionnaire based on MBTI is used to determine the student's personality factor value either as Introvert-Sensor, Extrovert-Sensor, Introvert-Intuition or Extrovert-Intuition. Figure 2 shows the sample of the on-line questionnaire. Data collected on the questionnaire were used to determine the learning style of the students using Fuzzy Logic.
5. The students started learning and they chose the learning material structure given to them either sequentially or freely.
6. After the students finished learning, they have to do the assessment on-line. The assessment is provided in order to measure the knowledge level of the students on the topic.

7. Upon the completion of learning and assessment, participants were required to fill up two questionnaires on learning satisfaction and usability of the system. The second questionnaire is adapted with slight changes from [15] with close ended and open ended questions concerning on user subjective satisfaction.
8. Analysis of the data collected during the study consists of data from on-line learning style, students learning satisfaction and usability of the system questionnaire. The results are discussed in the following sections.

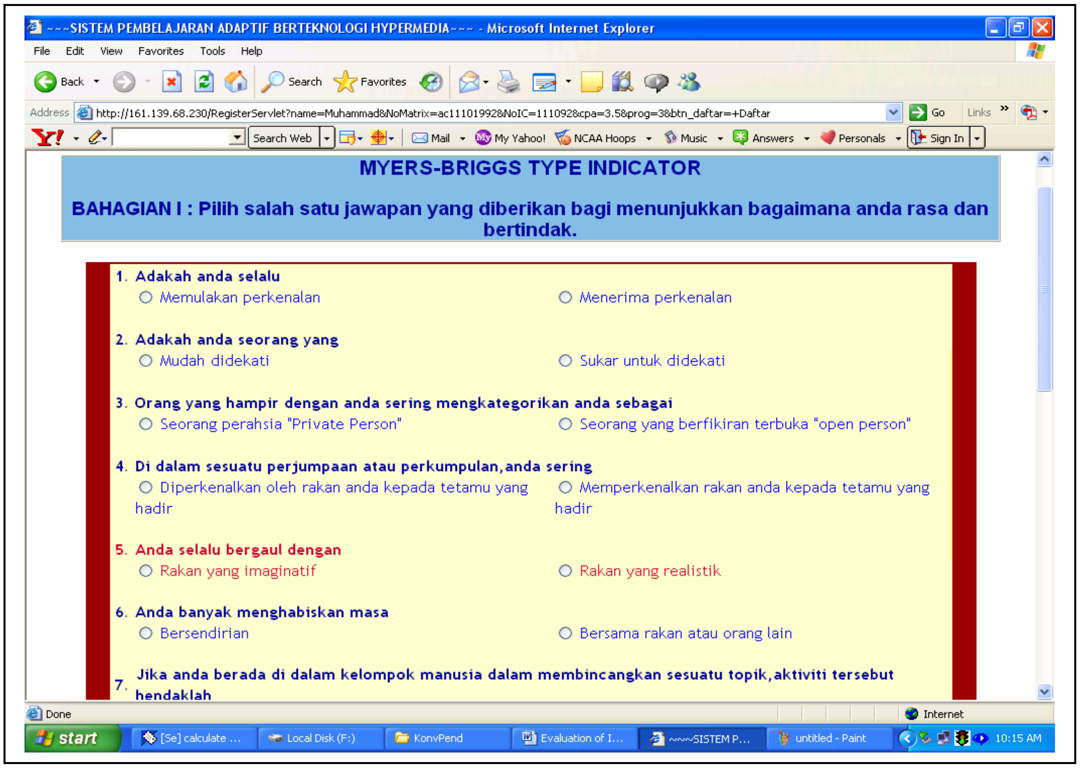


Fig. 2. Interface of the on-line questionnaire for personality factor based on Myers-Briggs Type Indicator (MBTI)

5. Results

The results elaborated in this section are mainly from the analysis of data collected from the two questionnaires based on the three key factors in usability study the learnability factors, efficiency factors and satisfaction factors. Apart from the three key factors, the students learning styles are determined based on the on-line questionnaire given in the system.

5.1 Learning Style Categorization

Based on the Myers Briggs Type Indicator (MBTI) personality factor, the students are categorized into four groups of learning styles. Figure 3 shows the number of students for each learning styles where 21 students are in Extrovert-Sensor group, 8 students are grouped in Introvert-Sensor, 4 students are in the Extrovert-Intuitive group and another 11 students 4 in the Introvert-Intuitive group.

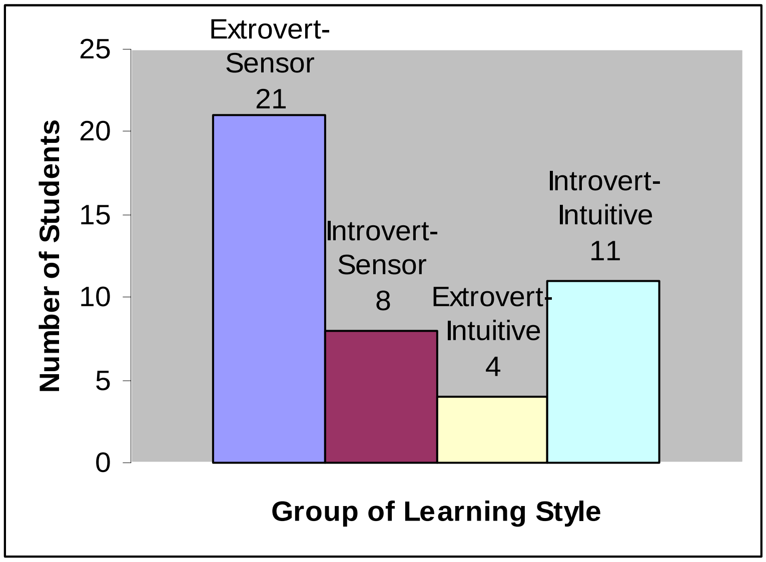


Fig. 3. Learning style of the participants

5.2 Learnability Factor

The learnability factor is measured quantitatively based on students' ease in completing the task of learning using the system. The learnability factor is measured based on these three factors:

1. The ease in learning based on personalization of the students learning style using the MBTI personality factor.
2. The ease in learning based on the learning structure given
3. The ease in learning based on the navigational path

Based on the questionnaire, 42 students were either strongly agreed or agreed on the effectiveness in learning based on MBTI learning style while only 2 students disagree in the learning style approach. The high learnability factor of this system is further proved by the number of students prefer the to learn based on the learning structure sequence given. The last element which contributes to the ease of using the system is the adaptive navigational paths provided by the system. 6 of the respondents strongly agree and another 30 respondents agree that the navigational paths provide ease of using the system. Only 8 of the respondents disagreed that annotated navigational paths provide ease of using this system. Table 2 shows the three features used to measure learnability factors of the system.

Table 2. Three features to measure learnability factor in using the system

	Strongly Agree	**Agree**	**Disagree**
Effectiveness of learning	4	38	2
Preferable learning structure sequence	7	27	10
Ease of using adaptive navigational path	6	30	8

5.3 Efficiency Factor

Efficiency factor of this system is measured by the productiveness of the students in terms of the score from the assessment they have completed in the system. The assessment has been done on-line by the students after they have completed their study using the system. Figure 8 shows that the number of he students with score between 80 to 100is quite high, 28 students while 11 and 2 students has either scored between 40 and 79 and between 0 and 39 respectively. The number of students with their respective assessment score is as shown in Table 3.

Table 3. Feature to measure efficiency factor in using the system

	Assessment Score		
	0-39 marks	40-79 marks	80-100 marks
No. of students	11	5	28

5.4 Satisfaction Factor

The final key attributes measured is the satisfaction factor of the user while using the system. The students satisfaction were determined based on attributes such as the familiarity and understandability of the terms used in the system, the easiness to understand the learning contents, the suitability of colors used in the system, the usefulness of the interactivity in the system and the overall satisfaction the user experienced when they use the system. Table 4 shows the number of students and their learning satisfaction in using the system.

The usability questionnaire have two open ended questions for the participants to make suggestions to enhance the application functionality and to express participants feeling while using the system

Table 4. Feature to measure satisfaction factor in using the system

	Strongly agree	Agree	Disagree
Participants learning satisfaction	4	38	2

Some excerpts of the feedback from the participants are:

1. The system is interesting because I can clearly see how the sort function working.
2. Interesting especially using the flash application but the yellow color for the background is giving me pressure.
3. Good, this is a new technique to improve the student in learning Data Structure.
4. Interesting because it is easy to understand the subject by using this system.
5. A little confusing at first but at the end I am amazed. It is easier to learn according to the learning style.

6 Discussion

From the results, system has achieved high percentage of learnability factor by the ease in using this system with the provided adaptive learning material sequence based on students learning style and also the ease in navigating of the system with the help of annotated links. Although half of the students have been given the task of learning freely regardless of the sequence of learning material, but more than half of the students (61.4%) have followed the sequence given by the system. This result showed that sequence of learning material structure suits their learning style thus helps them in their learning.

The second key attributes in this usability study is the efficiency factor. Efficiency is measured from the productiveness of the participants or in this study the assessment score of the students after learning using this system. Although only 11.4% of the students scored between 80 and 100, it is an acceptable figure considering students were expected to understand both the theory and algorithms of each sorting technique within two hours and also due to the detailed coverage of the assessment questions which really test the students in their understanding. The performance of the student can be improved if the web based usability study is adopted instead of controlled experiment. With web based usability study the participants can learn freely without time constraint [13].

The last key attributes, the satisfaction factor is measured based on the level of pleasure the participants experience while using this system. In this study, the results has showed high satisfaction factor whereby majority of the participants agree on the pleasurable attributes the system offer such as the understandable terms, the learning material structure and the interactivity in the system. Furthermore, the subjective feedback given by the participant demonstrate that learning material structure as in example and activity which shows the simulation of the algorithm and the programming really give them more understanding of the subject thus giving them satisfaction in using the system.

From this research point of view, the summative evaluation which addresses the educational impact of this system to student can be answered by the learnability factor which shows the suitability of the learning style in personalizing the learning material for the students. Another factor contributing to the educational impact of the system to the user is the learning material structure itself such as example and activity structures really provide the students satisfaction in learning using this system.

Even though the results of this study shows positive outcome in the usage of this system but there are also several weaknesses we encounter in conducting the usability study. The drawbacks are:

1. There are no comparison between students who have already learned data structure and students who are still learning this subject. If this type of comparison can be done, more interesting result can be analyzed.
2. In this research there is also no comparison done between the usages of this system with other similar system. In future, we decide to do a comparison study on the usage of this system with INSPIRE system [16].
3. There are comments from students particularly on the system choice of color which some of them thought as glaring and also mismatched. Other students were annoyed with the continuous system's automatic refresh where they have to reselect again the tree structure in order for all the learning material to be displayed again.

7 Conclusion and Future Work

This paper has discussed the evaluation of an adaptive hypermedia system by implementing the usability study. Three out of five key factors in the usability study has been measured. The method and processes of the usability study has also been elaborated. From the preliminary results and discussion it shows that this system has a high percentage for learnability and satisfaction factor thus these factors contributes to the answering of the question asked in summative evaluation: “What is the educational impact of this system to user?”. For future work of this system, we intend to implement web usage mining in order to measure the efficiency of the learning approach given to the student based on their learning style.

Acknowledgement

This research is conducted under IRPA grant, project vote 74129: The Development of Knowledge Management for Adaptive Hypermedia Learning System.

References

1. P. Brusilovsky.. “Adaptive Hypermedia: From Intelligent Tutoring Systems to Web-Based Education (Invited Talk)”. In G. Gauthier. C. Frasson and K. VanLehn(eds). Intelligent Tutoring Systems. Lecture Notes in Computer Science, Vol 1839: Springer Verlag, (2000) pp. 1-7.
2. Oppermann, Patel A. and Kinshuk.. “A Classification of Evaluation Methods for Intelligent
3. Tutoring Systems”. (1999). Software Ergonomie ’99.
4. P. Brusilovsky, C. Peylo. “Adaptive and intelligent web-based educational systems”, International Journal of Artificial Intelligence in Education. (2003). (13) 156–169.
5. H. Alomyan. “Individual differences: implications for web-based learning design”. International Education Journal.. 4(4). (2004) 188-196.
6. N. Ford, S.Y. Chen. “Individual differences, hypermedia navigation, and learning: An empirical study”. Journal of Educational Multimedia and Hypermedia.. 9(4). (2000). 281-312.

7. G. Magoulas, K. Papanikolaou and M. Grigoriadou. “Adaptive web-Based learning: accommodating individual differences through system’s adaptation”. British Journal of Educational Technology.. (2003). 34(4).
8. Norreen Haron, Naomie Salim. “Integrating personality factor (MBTI) and Fuzzy logic techniques in individualizing learning material of adaptive hypermedia learning system”. .(2005)Konvensyen Tenologi Pendidikan Ke 18. 16-19 September, Marang, Terengganu.
9. Bariah Yusob, Norreen Haron, Nor Bahiah Hj Ahmad, Shahliza Abd Halim. “Individualizing the learning material and navigation path in an adaptive hypermedia learning system. 2nd National Conference On Computer Graphics & Multimedia, Kajang, Selangor”.(2004). 8-10 December.
10. M. Mark, J. Greer. “Evaluation Methodologies for Intelligent Tutoring Systems”. Journal of Artificial Intelligence and Education, (19934) (2/3), 129-153.
11. K.Silius, AM. Tervakari. “An Evaluation of Usefulness of Web-Based Learning Environments. The Evaluation Tool into the Portal of Finnish Virtual University”. International Conference on Network Universities and E-Learning, 8-9 May 2003, Valencia, Spain.
12. Nielsen, J.. Usability Engineering. Academic Press. (1993)
13. Lazar. User-Centered Web Development. Boston: Jones and Bartlett. (2001)
14. 13. D. Cartens, P. Patterson.. “Usability Study of Travel Websites”. Journal of Usability Studies. Issue 1, Vol. 1, November 2005, (2005) pp 47-61.
15. Nor Bahiah Hj Ahmad, Shahliza Abd Halim, Bariah Yusob, Norreen Haron. “Development of an Adaptive Hypermedia Learning System”. Konvensyen Tenologi Pendidikan Ke 18. (2005) 16-19 September, Marang, Terengganu.
16. Granic, V. Glavinic and S.Stankov.. “Usability Evaluation Methodology for Web-based Educational Systems”. 8th ERCIM Workshop User Interfaces For All. (2004) 28-29 June
17. M. Grigoriadou, K. Papanikolaou, H. Kornilakis and G.Magoulas. “INSPIRE: An Intelligent System for Personalized Instruction in a Remote Environment”. Twelfth ACM Conference on Hypertext and Hypermedia (Hypertext'01). (2001) 14-18 August, Århus, Denmark.