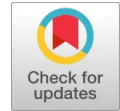# Developing an Intelligent Quiz System Based on Genetic Algorithm

Van Kiet Huynh, Ba Lam To

Abstract: In the normal quiz system, all questions are the same. However, the questions of an intelligent quiz system have different difficulty level, different response time, etc. This paper proposed a novel solution based on genetic algorithm (GA) to build an intelligent quiz system. This solution is experimented with website system via PHP and MySQL. The question bank for experiment is input into the system by lecturers. The experimental result obtains very good with the accuracy rate of 95%.

Index Terms: Evolutionary Algorithms, Genetic Algorithm (GA), Intelligent Quiz System, Learning Outcome (LO).

I. INTRODUCTION

The quiz system is increasingly popular in universities. However, the questions of the current quiz system are at the same difficulty level and response time. Current training programs are built based on learning outcomes (LOs). The program must be designed with the learning outcomes of the graduate. Each subject and lesson must include the learning outcomes of the program. Learning outcomes are often aims, goals or objectives of the program. Learning outcomes are the achievements of the learner which teacher need to help the learner to achieve [1, 2]. Each question of the quiz system should have different difficulty levels and different response time. Moreover, each question must cover some learning outcomes. Genetic algorithms are called evolutionary computation, optimization algorithm. They imitate the biological processes of reproduction and natural selection to find for the best optimal solutions [3, 4]. Genetic algorithm is built from Charles Darwin's theory of natural evolution. Genetic algorithm search the best individuals from a population by natural selection. The process of natural selection selects the best individuals from a population. These individuals crossover to produce offspring which inherit the characteristics of the parents and the new generation is formed from new offspring. When parents are strong, offspring have a better chance at surviving and better strong. The above process is repeated until the optimum condition is reached and a generation with the fittest individuals will be found [5, 6]. Genetic algorithm includes five phases such as initializing population, calculating fitness function, selecting the best individuals, crossover to reproduce next generation and mutating some individuals to make diversity within the population [7].

A population includes individuals. Each individual has a set of characteristics known as Genes. Each individual is a solution to the problem. Calculating fitness function for each individuals to determine how fit an individual. An individual with high fitness score will be selected for reproduction. Selecting the best parents based on their fitness scores, these parents help to reproduction better offspring.

The selected parents will be mated in pairs by choosing at random from within the genes. This process is called crossover phase. Mutating some individuals to make diversity within the population and prevent premature convergence. The algorithm terminates if the optimal condition has met. The result of the genetic algorithm is a set of solutions to problem. In an intelligent quiz system, each question has different difficulty level, different response time and some learning outcomes. In order to generate an exam which meet the requirements of lecturers, an intelligent quiz system has to use evolutionary algorithms to find an optimal solution. This paper proposed a new solution based on genetic algorithm to generate an exam in the intelligent quiz system.. The remainder of the paper is organized as follows. The system design is discussed in Section 2. In section 3, the numerical results of experiment are illustrated. Finally, Section 4 concludes this paper and figures out the future works.

II. SYSTEM DESIGN

A. Database Model

Database model using for the intelligent quiz system includes the following main tables:

1) Table "Subject"

This table contains information about subjects. Table structure is described in Figure 1.

Figure 1. Structure of Table "Subject"

2) Table "Learning Outcomes"

This table contains information about learning outcomes. Each learning outcome belongs to a subject. Table structure is described in Figure 2.

Figure 2. Structure of Table "Learning Outcomes"

3) Table "Question"

This table contains information about questions. Each question belongs to a subject. Table structure is described in Figure 3.

Figure 3. Structure of Table "Question"

4) Table "Answer"

This table contains information about answers. Each answer belongs to a question. Table structure is described in Figure 4.

Figure 4. Structure of Table "Answer"

5) Table "Question_LO"

This table describes the relationship between question and learning outcomes. Each question includes some learning outcomes. Table structure is described in Figure 5.

Figure 5. Structure of Table "Question_LO"

B. Proposed Algorithm

In the proposed system, there are three options to support lecturers choose to generate an exam. In the first option, lecturers choose exam time and difficulty level. All generated questions include whole learning outcomes of subject.

In the second option, lecturers choose exam time. All questions of generated exam have different difficulty levels. All generated questions include whole learning outcomes of subject.

In the third option, lecturers choose exam time and

percentage of each difficulty level. All generated questions include whole learning outcomes of subject.

Three options to support lecturers choosing conditions is described in Figure 6.

Figure 6. Three options

The proposed algorithm is done in 2 steps. The first step, querying database to extract questions required, the second step uses GA algorithm to generate the questions of exam aim to meet the optimal condition. First, this algorithm queries database to extract the questions meet the requirements. For example, the first option, extracting all questions of chosen subject with chosen difficulty level.

Second, the algorithm model is described in Figure 7. This algorithm is implemented with the following phase:

- Initial population phase: This phase creates a 2-dimensional array (MxN), M is the number of individuals, N is the number of extracted questions. M is the rows of array, N is the columns of array. Randomly generating values of 1 or 0 for elements in the array.

- Calculating fitness function phase: this phase calculates the fitness function of each individual. The fitness function of each individual is calculated by sum of response time of each question. When the value of the elements in each individual is 1, adding the response time of the corresponding question to the fitness function of that individual.

- Optimized phase: This phase determines the optimal condition for stopping the algorithm. The optimal condition is "the value of fitness function of certain individual equals to the exam time and whole learning outcomes of all questions are also whole learning outcomes of subject".

- Next generation selection phase: This phase retains half of the individual (M/2) with the most optimal fitness function (nearly equal to the exam time), eliminating half of the remaining individuals (M/2).

- Crossover phase: From half of the individual (M/2) with the most optimal fitness function, this phase create next population (M individuals) via crossover. Each pair of parents will be mated randomly to create new offspring. a new population generation is formed from these new offspring.

- Mutation phase: This phase performs mutations for some genes of certain new offspring with a low random probability. Mutating some individuals to make diversity within the population and prevent premature convergence.
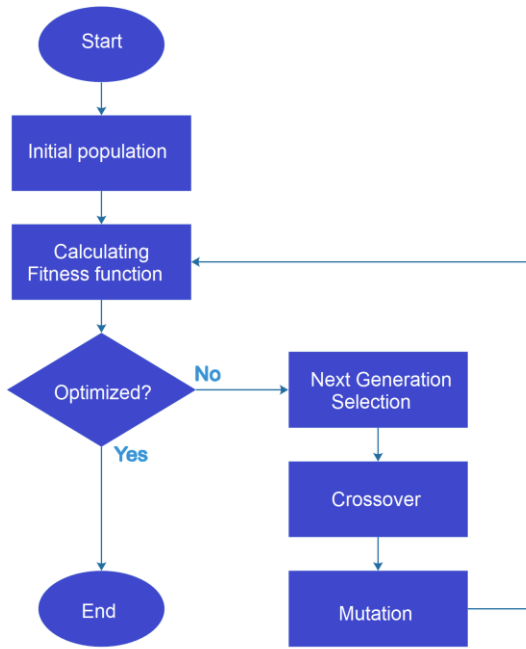
Figure 7. Proposed Algorithm

III. EXPERIMENTAL RESULTS

A Website system is developed to conduct some experiments. This Website is designed and programmed via PHP and MySQL [8]. The optimal condition of these experiments is "the value of fitness function of certain individual equals to the exam time and whole learning outcomes of all questions are also whole learning outcomes of subject". If there is no fitness function value of individuals equal to the exam time, error rate less than 5% between the value of fitness function of certain individual and the exam time is the optimal condition.

A. The first Experiment

In this experiment, lecturers choose the first option with the exam time of 30 minutes. The exam generated with the exam time of 30 minutes and "Easy" level is described in Figure 8 in Vietnamese. The exam generated with the exam time of 30 minutes and "Medium" level is described in Figure 9 in Vietnamese. The exam generated with the exam time of 30 minutes and "Hard" level is described in Figure 10 in Vietnamese.

Figure 8. Exam time of 30 minutes and "Easy" level

Figure 9. Exam time of 30 minutes and "Medium" level

Figure 10. Exam time of 30 minutes and "Hard" level

In this experiment, lecturers generated 50 exams with "easy" level and the exam time of 30 minutes for different subjects, 50 exams with "medium" level and the exam time of 30 minutes for different subjects, 50 exams with "hard" level and the exam time of 30 minutes for different subjects. The experimental results are shown in Table 1.

Table 1. The experimental result of the first experiment

Level	The number of exams with exact time	The number of exams with error 5% time	Accuracy Rate
Easy	48	2	96%
Medium	49	1	98%
Hard	49	1	98%

As a result of table 1, it found that the accuracy rate of generating exams achieves 97.3%.

B. The second Experiment

In this experiment, lecturers generated 50 exams with the exam time of 30 minutes for different subjects, 50 exams with the exam time of 40 minutes for different subjects, 50 exams with the exam time of 50 minutes for different subjects, 50 exams with the exam time of 60 minutes for different subjects. The experimental results are shown in Table 2.

Table 2. The experimental result of the second experiment

Exam time (minutes)	The number of exams with exact time	The number of exams with error 5% time	Accuracy Rate
30	50	0	100%
40	48	2	96%
50	48	2	96%
60	47	3	94%

As a result of table 2, it found that the accuracy rate of generating exams achieves 96.5%.

The exam generated with the exam time of 30 minutes is described in Figure 11 in Vietnamese. The exam generated with the exam time of 40 minutes is described in Figure 12 in Vietnamese.

Figure 11. Exam time of 30 minutes

Figure 12. Exam time of 40 minutes

C. The third Experiment

In this experiment, lecturers generated 50 exams with 50% "easy" level, 30% "medium" level, 20% "hard" level and the exam time of 30 minutes for different subjects; 50 exams with 20% "easy" level, 50% "medium" level, 30% "hard" level and the exam time of 30 minutes for different subjects; 50 exams with 30% "easy" level, 20% "medium" level, 50% "hard" level and the exam time of 30 minutes for different subjects. The experimental results are shown in Table 3.

Table 3. The experimental result of the third experiment

Level	The number of exams with exact time	The number of exams with error 5% time	Accuracy Rate
50% Easy, 30% Medium 20% Hard	48	2	96%
20% Easy, 50% Medium 30% Hard	48	2	96%
30% Easy, 20% Medium 50% Hard	47	3	94%

AS a result of table 3, it found that the accuracy rate of generating exams achieves 95.3%.

The exam generated with the exam time of 30 minutes and 50% Easy, 30% Medium, 20% Hard level is described in Figure 13 in Vietnamese. The exam generated with the exam time of 30 minutes and 20% Easy, 50% Medium, 30% Hard level is described in Figure 14 in Vietnamese. The exam generated with the exam time of 30 minutes and 30% Easy, 20% Medium, 50% Hard level is described in Figure 15 in Vietnamese.

D. Discussion

Through three above-mentioned experiments, the accuracy rate of generating exams are over 95%. the accuracy rate of experiments is described in Figure 16. It found that the proposed algorithm is very good for this solution.
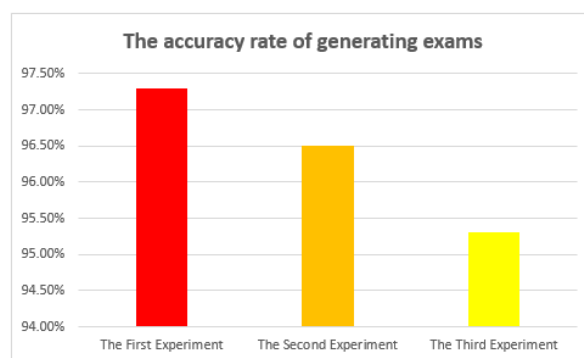
Figure 16. The accuracy rate of generating exams

IV. CONCLUSION

This paper proposed a novel solution based on genetic algorithm to develop an intelligent quiz system. This solution is experimented with website system via PHP and MySQL. The experimental result obtains very good with the accuracy rate of 95%. In the near future, this work will be improve the proposed algorithm at crossover and mutation phase to increase the accuracy rate. Moreover, we also experiment with bigger datasets to evaluate the proposed algorithm exactly and overcome its disadvantages.

ACKNOWLEDGMENT

Developing an Intelligent Quiz System Based on Genetic Algorithm

Thí sinh tham gia thi : Kiệt Lu	Kiểu thành viên : Admin
Tổng số câu hỏi: 16	Thời gian thi : 30 phút
Thời gian còn lại : 29 : 56	

Cú pháp để tạo bảng có cột định danh sau đây đúng hay sai: Tên cột [kiểu dữ liệu] IDENTITY(số_bắt_đầu,chỉ_số_tăng)
1) ○ Đúng
2) ○ Sai

Cú pháp để tạo bảng có giá trị mặc định sau đây đúng hay sai: Tên cột [kiểu dữ liệu] DEFAULT giá_trị
1) ○ Đúng
2) ○ Sai

Cho Q(ABCDEGH) với tập phụ thuộc hàm F={B-->A, DA-->CE, D-->H, GH-->C, AC-->D}.Tính bao đóng của X={BD} dựa trên F ?
1) ○ {BDAC}
2) ○ {BDACE}
3) ○ {BDCE}
4) ○ {BDACEH}
5) ○ {BDACEGH}

Nếu BN-->CI là 1 trong những phụ thuộc hàm của lược đồ quan hệ U(B,N,C,I). Vậy câu nào dưới đây luôn đúng?
1) ○ BN là khóa chính của U
2) ○ NB là khóa chỉ định của U
3) ○ Không có 2 bộ trên lược đồ quan hệ U có cùng giá trị C và I
4) ○ Không có 2 bộ trên lược đồ quan hệ U có cùng giá trị B và N
5) ○ Tất cả các câu đều đúng

Figure 13. Exam time of 30 minutes and 50% Easy, 30% Medium 20% Hard level

Thí sinh tham gia thi : Kiệt Lu	Kiểu thành viên : Admin
Tổng số câu hỏi: 16	Thời gian thi : 30 phút
Thời gian còn lại : 29 : 58	

Câu lệnh nào sau đây không phải để tạo cấu trúc bảng?
1) ☐ SET UP TABLE
2) ☐ MAKE TABLE
3) ☐ CREATE TABLE
4) ☐ NEW TABIE

1 lược đồ quan hệ là:
1) ☐ 1 tập giá trị dữ liệu
2) ☐ 1 sơ đồ quan hệ thực thể
3) ☐ Tên quan hệ
4) ☐ 1 tập các thuộc tính

Câu lệnh nào sau đây giới hạn cột trả về ?
1) ○ Select
2) ○ Order by
3) ○ From
4) ○ Where

Thuật ngữ "data" đề cập tới:
1) ☐ Hình ảnh và đối tượng
2) ☐ Ký tự
3) ☐ băng hình, giọng nói và ảnh động
4) ☐ Số

Cho phụ thuộc hàm (UO-->AXS và U-->S) và lược đồ quan hệ D=(U,O,A,X,S).Chúng ta có thể suy ra:
1) ○ UO là khóa của D

Figure 14. Exam time of 30 minutes and 20% Easy, 50% Medium 30% Hard level

Thí sinh tham gia thi : Kiệt Lu	Kiểu thành viên : Admin
Tổng số câu hỏi: 15	Thời gian thi : 30 phút
Thời gian còn lại : 29 : 57	

Câu lệnh nào sau đây không phải để tạo cấu trúc bảng?
1) ☐ SET UP TABLE
2) ☐ MAKE TABLE
3) ☐ CREATE TABLE
4) ☐ NEW TABIE

SQL là 1 ngôn ngữ thủ tục
1) ○ Đúng
2) ○ Sai

Khi ta ánh xạ 1 thuộc tính đa trị của thực thể E từ mô hình ER sang mô hình quan hệ.Chúng ta sẽ tạo:
1) ○ 1 cột trong quan hệ tương ứng với thực thể E
2) ○ Nhiều quan hệ, mỗi 1 quan hệ tương ứng với mỗi giá trị riêng biệt của thuộc tính đó.
3) ○ 1 quan hệ mà có chứa 1 khóa ngoại và 1 cột cho thộc tính đó
4) ○ Không có câu nào đúng

Cho I(D,V,C,Y,J) với 1 tập phụ thuộc hàm M{DV-->CJ, V-->Y, Y-->D}.Phụ thuộc hàm nào có thể suy ra từ O ?
1) ☐ J-->YCDJ
2) ☐ DJ-->Y
3) ☐ DY-->CJ
4) ☐ JY-->YCJ

Figure 15. Exam time of 30 minutes and 30% Easy, 20% Medium, 50% Hard level

REFERENCES

1. Guide to AUN-QA Assessment at Programme Level, ASEAN University Network (AUN), Version No. 3.0, October 2015.
2. Learning Outcomes Available:
https://eua.eu/component/tags/tag/77-learning-outcomes.html
3. Haupt, R. L., & Haupt, S. E. (2004).Practical Genetic Algorithms(2nd ed.). Hobo-ken: Wiley.
4. Simon, D. (2013).Evolutionary Optimization Algorithms: Biologically-Inspired andPopulation-Based Approaches to Computer Intelligence.Hoboken: Wiley.
5. R. Oftadeh et al. (2010), "A novel meta-heuristic optimization algorithm inspired by group hunting of animals: Hunting search", 60, 2087–2098.
6. Hasançebi, O., Kazemzadeh Azad, S. (2015), "Adaptive Dimensional Search: A New Metaheuristic Algorithm for Discrete Truss Sizing Optimization", *Computers and Structures*, 154, 1–16.
7. Genetic Algorithm. Available:
https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3

8. Laura Thompson and Luke Welling, PHP and MySQL Web Development, Addison-Wesley Professional; 4 edition (October 11, 2008).

AUTHORS PROFILE

Van Kiet Huynh received B.Sc. degree in Computer Science from Ho Chi Minh City University of Transport, Viet Nam. Currently, He is preparing to get a M.Sc degree in Computer Science. His current research interests include optimal algorithm, Soft Computing and Artificial Intelligence.

Ba Lam To is a lecturer in the Faculty of Information Technology, Ho Chi Minh City University of Transport, Vietnam. He received his B.Eng. degree in Information Technology from HCM City University of Technology in 2006 and received Master and Ph.D degree in 2008 and 2012 respectively from Université Pierre et Marie Curie (UPMC -Paris 6). His current research interests include Cloud Computing, Routing in Cognitive Radio Network, Network Security, Intelligent transportation system.