

A GA-Based Role Assignment Approach for Web-based Cooperative Learning Environments

Yi-Chun Chang and Jian-Wei Li

Abstract—Web-based cooperative learning focuses on (1) the interaction and the collaboration of community members, and (2) the sharing and the distribution of knowledge and expertise by network technology to enhance learning performance. Numerous research literatures related to web-based cooperative learning have demonstrated that cooperative scripts have a positive impact to specify, sequence, and assign cooperative learning activities. Besides, literatures have indicated that role-play in web-based cooperative learning environments enhances two or more students to work together toward the completion of a common goal. Since students generally do not know each other and they lack the face-to-face contact that is necessary for the negotiation of assigning group roles in web-based cooperative learning environments, this paper intends to further extend the application of genetic algorithm (GA) and propose a GA-based algorithm to tackle the problem of role assignment in web-based cooperative learning environments, which not only saves communication costs but also reduces conflict between group members in negotiating role assignments.

Keywords—genetic algorithm (GA); role assignment; role-play; web-based cooperative learning.

I. INTRODUCTION

COLLABORATION learning has been defined as “small groups of learners working together as a team to solve a problem, to complete a task, to discover information, and to accomplish a common goal” [1]. The cooperative learning model requires dividing students into several groups so that the group members need to work together to accomplish a shared or a common goal rather than working alone [2], [3]. In such learning process the common goal involves some group tasks, and each group member is responsible for one or more group tasks [4]. Ellis and Whalen also indicated that the essential feature of cooperative learning is one student helping other students to be successful [5]. Furthermore, communication and information exchange among group members is the most important action among partners to complete group goals [6], [7].

With the popularity of the Internet, cooperative learning is not limited to traditional classroom teaching [8]. Web-based collaborative learning is an emerging research field that focuses

on 1. The interaction and the collaboration of community members, and 2. The sharing and the distribution of knowledge and expertise by network technology to enhance learning performance [9], [10]. Numerous research literatures related to web-based collaborative learning have demonstrated that collaboration scripts have a positive impact to specify, sequence, and assign collaborative learning activities [11].

Slavin have defined the roles as prescribed functions that guide collaboration between group members [12]. Morris *et al.* also indicate that the structure provided by roles facilitates members to concentrate on the task and the collaboration [13]. Besides, the mechanism of role-play in web-based cooperative learning environments enhances two or more students to work together toward the completion of a common goal [14]-[16].

As a result, group members can take more responsibilities to find more personal values and are more likely to become agents in their own learning [10], [17]. Hence, completing a common goal heavily depends on the individual contributions of group members [10]. In traditional classroom learning environments, group members can discuss with each other to negotiate the assignment of group roles. However, in web-based cooperative learning environments, students generally do not know each other [18] and they lack the face-to-face contact that is necessary for the negotiation of assigning group roles. Besides, in web-based cooperative learning environments, instructors are having difficulties to monitor the interaction and negotiation of whole groups. Therefore, web-based cooperative learning environments need to provide help for assigning group roles to individual members so that each member has individual responsibilities for completing a common goal.

The genetic algorithm (GA) approach was firstly proposed by Holland in 1975 [19], which has been shown to successfully optimize a wide range of objective functions [20]. The attractiveness of using GA is due to the following features: natural metaphor, simplicity, stochastic move, adaptively, positive feedback, and high quality solutions [20]. Thus, we intend to further extend the application of GA and propose a GA-based algorithm to tackle the problem of role assignment in web-based cooperative learning environments, which not only saves communication costs but also reduces conflict between group members in negotiating role assignments.

The rest of this paper is organized as follows: Section II briefly reviews related works. Section III presents the GA-based role assignment approach in detail. Finally, Section IV offers conclusions.

Yi-Chun Chang is with the Department of Computer Science and Information Engineering, Hungkuang University, No.34, Chung-Chie Rd., Sha Lu, Taichung, Taiwan 43302, R.O.C. (phone: 886-4-26318652 Ext. 4207; fax: 886-4-26324084; e-mail: changyc@sunrise.hk.edu.tw).

Jian-Wei Li is with the Department of Information and Communication Engineering, Chaoyang University of Technology, No. 168, Jifong East Road, Wufong Township, Taichung, Taiwan 41349, R.O.C. (phone: 886-4-23323000 Ext. 7833; fax: 886-4-23305539; e-mail: lijw@cyut.edu.tw).

II. RELATED WORKS

The GA approach was firstly proposed by Holland in 1975 [19], which has been widely applied to solve an optimization problem by a systematic way [20]. The standard GA consists of several executing steps, namely chromosome encoding and population initialization, fitness evaluation, selection, crossover, and mutation [19], [21]-[25], which are further reviewed within the following subsections.

A. Chromosome Encoding and Population Initialization

Chromosome encoding is for transferring a candidate solution into a chromosome, according to a given optimization problem. A chromosome consists of several genes. Depending on the specific optimization problem, a gene can be encoded as a binary bit, an integer or a real number. A population is composed of several chromosomes, while an initial one is usually randomized generated. Through a number of developed genetic evolutionary processes, the population evolves from one to next generation to improve the quality of chromosomes. The population size, i.e., the number of chromosomes in a population, is kept constant.

B. Fitness Evaluation

To evaluate the fitness of each chromosome, a fitness function needs to be defined according to the specific optimization problem. The fitness value derived from the fitness function can be used to determine which chromosomes are better solutions for the optimization problem.

C. Selection

According to the roulette selection [19], the chromosomes that have fitter fitness values possess higher probability to be selected to propagate offspring.

D. Crossover

It is a genetic evolutionary process, in which each individual has a chance to interchange gene information from two parent chromosomes. Crossover is performed with a chosen crossover probability. A random number can be generated between [0.0, 1.0] for each mating pair. If the random number is less than the crossover probability, the crossover is performed to propagate offspring. Otherwise, no crossover is performed. There are two broadly adopted crossover operators [26]. One is the single-point crossover, which yields offspring by interchanging all the genes after a random position from the parent chromosomes. The other operation is the two-point crossover,

which generates two random positions and interchanges the genes between the two positions from the parent chromosomes.

E. Mutation

Being a genetic evolutionary process, this operation creates a new gene that does not inherit from the parent chromosomes. Every bit in an offspring chromosome could mutate with a predefined mutation probability. In general, mutation is performed with a very low mutation probability. For performing mutation, two genes are randomly selected to exchange their positions within a certain chromosome.

F. Stop Criterion

In general, the stop criterion is set to 100 generations. That is, the steps B-E are iteratively performed 100 times. After 100 generations, a near optimal solution can be derived by the fitness function.

III. THE PROPOSED APPROACH

This section presents the GA-based role assignment approach, which includes A. The problem definition of the role assignment problem, and B. The execution steps of the proposed approach.

A. The Problem Definition

We consider the role assignment problem by exploring the following scenario: in network supported cooperative learning environments, a group is the combination of K members who possess M features to cooperatively complete a common goal. $F_{k,m}$, $1 \leq k \leq K$ and $1 \leq m \leq M$, denotes the m^{th} feature for the k^{th} member L_k . For the m^{th} feature of the k^{th} member L_k , $\{f_{k,m,1}, f_{k,m,2}, \dots, f_{k,m,N}\}$ denotes the set of N factors, where $f_{k,m,n}$, $1 \leq k \leq K$, $1 \leq m \leq M$ and $1 \leq n \leq N$, denotes the n^{th} factor, which is a binary value. If the k^{th} member L_k possesses the n^{th} factor for the m^{th} feature, $f_{k,m,n}=1$. Otherwise, $f_{k,m,n}=0$. Suppose that the cooperative script includes P roles. For the p^{th} role R_p , $1 \leq p \leq P$, $\{R_{p,1}, R_{p,2}, \dots, R_{p,M}\}$ denotes the set of M features, where $R_{p,m}$, $1 \leq p \leq P$ and $1 \leq m \leq M$, denotes the m^{th} corresponding feature for the p^{th} role R_p that needs to match with the m^{th} feature for a member. For the m^{th} corresponding feature of the role R_p , $\{r_{p,m,1}, r_{p,m,2}, \dots, r_{p,m,N}\}$ denotes the set of N factors, where $r_{p,m,n}$, $1 \leq p \leq P$, $1 \leq m \leq M$ and $1 \leq n \leq N$, denotes the corresponding n^{th} factor, which is a binary value. If the p^{th} role involves the n^{th} factor for the m^{th} feature, $r_{p,m,n}=1$. Otherwise, $r_{p,m,n}=0$. Table I defines the parameters that are instrumental to the problem description.

TABLE I
THE PARAMETERS USED BY THE PROPOSED APPROACH

Parameter	Range	Description
K	$2 \leq K \leq 6$	The number of members
L_k	$1 \leq k \leq K$	The k^{th} member
M	$M \geq 1$	The number of features
$F_{k,m}$	$1 \leq k \leq K$ and $1 \leq m \leq M$	The m^{th} feature for the member L_k
\vec{N}	$\{n_1, n_2, \dots, n_M\}$	The number of factors
$f_{k,m,n}$	$1 \leq k \leq K$, $1 \leq m \leq M$, and $1 \leq n \leq N$	The n^{th} factor of the m^{th} feature for the k^{th} member L_k : If the k^{th} member L_k possesses the n^{th} factor of the m^{th} feature, $f_{k,m,n}=1$. Otherwise, $f_{k,m,n}=0$.
P	$P \geq 2$	The number of roles
R_p	$1 \leq p \leq P$	The p^{th} role

$R_{p,m}$	$1 \leq p \leq P$ and $1 \leq m \leq M$	The m^{th} feature for the p^{th} role R_p
$r_{p,m,n}$	$1 \leq p \leq P$, $1 \leq m \leq M$, and $1 \leq n \leq N$	The n^{th} factor of the m^{th} feature for the p^{th} role R_p . If the p^{th} role R_p involves the n^{th} factor of the m^{th} feature, $f_{k,m,n}=1$. Otherwise, $f_{k,m,n}=0$.
$x_{k,p}$	$1 \leq k \leq K$ and $1 \leq p \leq P$	Decision variable: If the k^{th} member L_k is assigned the p^{th} role R_p , $x_{k,p}=1$. Otherwise, $x_{k,p}=0$.

Since $r_{p,m,n}$, $1 \leq p \leq P$, $1 \leq m \leq M$ and $1 \leq n \leq N$, is a binary value, $R_{p,m} = \{r_{p,m,1}, r_{p,m,2}, \dots, r_{p,m,N}\}$ is a binary string. The converted formula for converting a binary string into a decimal number is derived as

$$R_{p,m(\text{dec.})} = \sum_{n=1}^N r_{p,m,n} \times 2^{(N-n)}, \quad 1 \leq p \leq P \text{ and } 1 \leq m \leq M. \quad (1)$$

Similarly, for the k^{th} member L_k , $F_{k,m} = \{f_{k,m,1}, f_{k,m,2}, \dots, f_{k,m,N}\}$, $1 \leq p \leq P$ and $1 \leq m \leq M$, can be converted into a decimal number that is derived using

$$F_{k,m(\text{dec.})} = \sum_{n=1}^N f_{k,m,n} \times 2^{(N-n)}, \quad 1 \leq p \leq P \text{ and } 1 \leq m \leq M. \quad (2)$$

The objective function Z that defines the problem of role assignment can be derived as

$$\text{Min. } Z = \sum_{k=1}^K \frac{\sum_{p=1}^P \left(x_{k,p} \times \sum_{m=1}^M |F_{k,m(\text{dec.})} - R_{p,m(\text{dec.})}| \right)}{\sum_{p=1}^P x_{k,p}}, \quad (3)$$

subject to

$$\sum_{k=1}^K x_{k,p} \geq 1, \quad 1 \leq p \leq P. \quad (4)$$

$$\sum_{p=1}^P x_{k,p} \leq 2, \quad 1 \leq k \leq K. \quad (5)$$

Equation (3) indicates that the objective function, in which the smaller value is suitable for the problem solution in this paper. Constraint Equation (4) indicates that each role needs to be assigned to at least one member. Constraint Equation (5) indicates that the number of assigned role for each member needs to be less than two.

B. The Execution Steps of the Proposed Approach

The proposed approach contains several steps in finding the near optimal solution for the aforementioned definition problem. These steps are presented as follows:

Step 1. Normalization

The relevant parameters for deriving fitness function need to be normalized. The normalized formula for $R_{p,m}$, $1 \leq p \leq P$ and $1 \leq m \leq M$, denotes the m^{th} corresponding feature for the p^{th} role R_p , can be derived using

$$R_{p,m}' = \frac{R_{p,m(\text{dec.})} - \min_{k=1-K, p=1-P} (F_{k,m(\text{dec.})}, R_{p,m(\text{dec.})})}{\max_{k=1-K, p=1-P} (F_{k,m(\text{dec.})}, R_{p,m(\text{dec.})}) - \min_{k=1-K, p=1-P} (F_{k,m(\text{dec.})}, R_{p,m(\text{dec.})})}, \quad (6)$$

where $1 \leq p \leq P$ and $1 \leq m \leq M$.

For $F_{k,m}$, $1 \leq k \leq K$ and $1 \leq m \leq M$, denotes the m^{th} feature for the k^{th} member L_k , the normalized formula can be derived using

$$F_{k,m}' = \frac{F_{k,m(\text{dec.})} - \min_{k=1-K, p=1-P} (R_{p,m(\text{dec.})}, F_{k,m(\text{dec.})})}{\max_{k=1-K, p=1-P} (R_{p,m(\text{dec.})}, F_{k,m(\text{dec.})}) - \min_{k=1-K, p=1-P} (R_{p,m(\text{dec.})}, F_{k,m(\text{dec.})})}, \quad (7)$$

where $1 \leq k \leq K$ and $1 \leq m \leq M$.

Step 2. Chromosome Encoding and Population Initialization

The chromosome refers to a candidate solution, which is denoted as an KP -dimensional vector $[x_{1,1} x_{1,2} \dots x_{1,P} x_{2,1} x_{2,2} \dots x_{2,P} \dots x_{K,1} x_{K,2} \dots x_{K,P}]$. As Table I depicts, $x_{k,p}$, $1 \leq k \leq K$ and $1 \leq p \leq P$, is a binary value. If the k^{th} member L_k is assigned the p^{th} role R_p , $x_{k,p}=1$. Otherwise, $x_{k,p}=0$.

Step 3. Fitness Evaluation

The fitness function of the proposed measures the quality of chromosomes solutions which satisfy (3)-(5). However, the chromosomes generated by the proposed approach may violate one or more of Constraint Equations (3)-(5). To cope with this problem, we use the fitness function which incorporates penalty terms to evaluate a chromosome. This paper denotes the penalty terms corresponding to Constraint Equations (3)-(5) as follows:

- α penalty: According to Constraint Equation (4), this term penalizes cases in which a role is not be taken charge of by any member, which can be derived as

$$\alpha = \sum_{p=1}^P \max \left(0, \left(1 - \sum_{k=1}^K x_{k,p} \right) \right). \quad (8)$$

- β penalty: According to Constraint Equation (5), this term penalizes cases in which the number of the role assigned to a member exceed two, which can be derived as

$$\beta = \sum_{k=1}^K \max \left(0, \left(\sum_{p=1}^P x_{k,p} - 2 \right) \right). \quad (9)$$

Since the relevant parameters are normalized in Step 1, the objective function Z' is updated as follows,

$$\text{Min. } Z' = \sum_{k=1}^K \frac{\sum_{p=1}^P \left(x_{k,p} \times \sum_{m=1}^M |F_{k,m}' - R_{p,m}'| \right)}{\sum_{p=1}^P x_{k,p}}, \quad (10)$$

Hence, in this paper, the fitness function $F(x)$ for evaluating the fitness of a chromosome x is derived as

$$F(x) = Z' + \alpha + \beta. \quad (11)$$

A smaller fitness value derived by (11) represents a better chromosome.

Step 4. Selection

This paper adopts the roulette selection that a fitter chromosome is selected to propagate offspring. Given the set of the fitness value for Q chromosomes in a population is $\{f_1(x), f_2(x), \dots, f_q(x)\}, 1 \leq q \leq Q$. The probability $prob(f_i(x)), 1 \leq i \leq Q$, that the i^{th} chromosome is selected to propagate child chromosomes is defined as

$$prob(f_i(x)) = \frac{\frac{1}{f_i(x)}}{\sum_{j=1}^Q \frac{1}{f_j(x)}}, 1 \leq i \leq Q. \quad (12)$$

Step 5. Crossover

This paper adopts two-point crossover, while the crossover probability is set to 0.7.

Step 6. Mutation

For performing the mutation operation two genes are randomly selected to exchange their corresponding positions within a chromosome, and the mutation probability is set to 0.1.

Step 7. Stop Criterion

The evolution population is 100 generations.

IV. CONCLUSION

To save communication costs and then reduce conflict between group members for negotiating in web-based cooperative learning environments, this paper has proposed the GA-based approach to support role assignment. In future, we consider applying the proposed approach for actual learning environment to investigate the learning achievements and the satisfaction of students.

ACKNOWLEDGMENT

The research is supported by the National Science Council of the Republic of China under the grant number NSC101-2221-E-241-018 and NSC101-2221-E-324-043.

REFERENCES

- [1] Artz, A. F. and Newman, C. M., Cooperative learning, Mathematics Teacher. 1990, **83**, pp. 448-449.
- [2] Johnson, D. W., Johnson, R. T. and Smith, K. A., Cooperative Learning: Increasing College Faculty Instructional Productivity. 1991: School of Education and Human Development, George Washington University.
- [3] Anaya, A. R. and Boticario, J. G., Application of machine learning techniques to analyse student interactions and improve the collaboration process, Expert Systems with Applications. 2011, **38**(2), pp. 1171-1181.
- [4] Labidi, S. and Ferreira, J. S. Technology-assisted instruction applied to cooperative learning: the SHIECC project, in proceedings of 28th Annual Frontiers in Education Conference (FIE '98) 1998, 1, Tempe, AZ, USA. pp. 286-291.
- [5] Ellis, S. S. and Whalen, S. F., Cooperative Learning: Getting Started. 1990: Scholastic New York.
- [6] Johnson, D. W. and Johnson, F. P., Joining together: group theory and group skills (7th ed.). 2000: Allyn & Bacon, Boston.
- [7] Wardell, C. S. and Paschetto, G. Small group instruction in real-time over the web, in proceedings of the inter-service industry training simulation and education conference. 2001, Orlando, Florida. pp.
- [8] Chan, C., Co-regulation of learning in computer-supported collaborative learning environments: a discussion, Metacognition and Learning. 2012, **7**(1), pp. 63-73.
- [9] Lippinen, L., Hakkarainen, K. and Paavola, S., Practices and orientations of CSCL. In: J.W. Strijbos, P.A. Kirschner and R.L. Martens (Eds.), What we know about CSCL. 2004, Norwell, MA: Kluwer Academic. 31-50.
- [10] Wang, Q., Design and evaluation of a collaborative learning environment, Computers & Education. 2009, **53**(4), pp. 1138-1146.
- [11] Schellens, T., Keer, H., Wever, B. and Valcke, M., Scripting by assigning roles: Does it improve knowledge construction in asynchronous discussion groups?, International Journal of Computer-Supported Collaborative Learning. 2007, **2**(2-3), pp. 225-246.
- [12] Slavin, R. E., A Model of Effective Instruction, The Educational Forum. 1995, **59**(2), pp. 166-176.
- [13] Morris, R., Hadwin, A. F., Gress, C. L. Z., Miller, M., Fior, M., Church, H. and Winne, P. H., Designing roles, scripts, and prompts to support CSCL in gStudy, Computers in Human Behavior. 2010, **26**(5), pp. 815-824.
- [14] Fu, F. L., Wu, Y. L. and Ho, H. C., An investigation of cooperative pedagogic design for knowledge creation in Web-based learning, Computers & Education. 2009, **53**(3), pp. 550-562.
- [15] Gillies, R. M., The effects of cooperative learning on junior high school students during small group learning, Learning and Instruction. 2004, **14**, pp. 197-213.
- [16] Law, Y. k., Effects of cooperative learning on second graders' learning from text, Educational Psychology. 2008, **28**(5), pp. 567-582.
- [17] Milner-Bolotin, M., The effects of topic choice in project-based instruction on undergraduate physical science students' interest, ownership, and motivation. 2001: Unpublished doctoral dissertation, The University of TX at Austin.
- [18] Sun, P. C., Cheng, H. K., Lin, T. C. and Wang, F. S., A design to promote group learning in e-learning: Experiences from the field, Computers & Education. 2008, **50**(3), pp. 661-677.
- [19] Holland, J. H., Adaptation in Natural and Artificial System. 1975, The University of Michigan Press: Ann Arbor.
- [20] Hwang, G. J., Yin, P. Y., Wang, T. T., Tseng, J. C. R. and Hwang, G. H., An enhanced genetic approach to optimizing auto-reply accuracy of an e-Learning system, Computers & Education. 2008, **51**(1), pp. 337-353.
- [21] Mitchell, T., Machine Learning. 1997: McGraw-Hill Companies.
- [22] Rothlauf, F., Representations for Genetic and Evolutionary Algorithms. 2006: Springer Verlag.
- [23] Yao, X., Evolving Artificial Neural Networks, Proceedings of the IEEE. 1999, **87**(9), pp. 1423-1447.
- [24] Chang, Y. C., Hsieh, S. M. and Li, J. W. Application of the Genetic Algorithm in Customization Personalized E-Course, in proceedings of 2010 International Conference on System Science and Engineering (ICSSE). 2010, Taipei, Taiwan. pp. 221-226.
- [25] Chang, Y. C., Li, J. W. and Ku, H. H., A Genetic Algorithm (GA)-based Personalized Learning Service in Cloud Learning Environments, IEEE Technology and Engineering Education. 2012, **7**(2), pp. 28-32.
- [26] Syswerda, G. Uniform crossover in genetic algorithms, in proceedings of the third international conference on genetic algorithms and their applications. 1989, pp. 2-9.