



# A genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics

Julián Moreno<sup>a,\*</sup>, Demetrio A. Ovalle<sup>a</sup>, Rosa M. Vicari<sup>b</sup>

<sup>a</sup> Universidad Nacional de Colombia, Medellín, Colombia

<sup>b</sup> Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

## ARTICLE INFO

### Article history:

Received 3 December 2010

Received in revised form

7 September 2011

Accepted 11 September 2011

### Keywords:

Collaborative learning

Group formation

Multi-objective optimization

Genetic algorithms

## ABSTRACT

Considering that group formation is one of the key processes in collaborative learning, the aim of this paper is to propose a method based on a genetic algorithm approach for achieving inter-homogeneous and intra-heterogeneous groups. The main feature of such a method is that it allows for the consideration of as many student characteristics as may be desired, translating the grouping problem into one of multi-objective optimization. In order to validate our approach, an experiment was designed with 135 college freshmen considering three characteristics: an estimate of student knowledge levels, an estimate of student communicative skills, and an estimate of student leadership skills. Results of such an experiment allowed for the validation, not only from the computational point of view by measuring the algorithmic performance, but also from the pedagogical point of view by measuring student outcomes, and comparing them with two traditional group formation strategies: random and self-organized.

© 2011 Elsevier Ltd. All rights reserved.

## 1. Introduction

Collaborative learning (CL) may be concisely defined as the synergy of individual works within a group by means of discussion and joint knowledge construction. Although there are many studies which confirm that CL is useful in the educational context (Dewiyanti, Brand-Gruwel, Jochems, & Broers, 2007; Francescato, Mebane, Porcelli, Attanasio, & Pulino, 2007; Francescato, Porcelli, Mebane, Cuddetta, & Renzi, 2006; Gokhale, 1995; Krejins, Kirschner, & Jochems, 2002; Nielsen, Hvas, & Kjaergaard, 2009; Prichard, Stratford, & Bizo, 2006; Solimeno, Mebane, Tomai, & Francescato, 2008; Tolmie et al., 2010; Veermans & Cesareni, 2005; Zurita & Nussbaum, 2004), there are few studies which focus on one key issue: the method to be used for group formation.

Having adequate groups allows for a good interaction among the members and is fundamental in order to obtain appropriate learning results. Even so, group formation is made in many cases without any criterion at all or by using simple random selection (Huxland & Land, 2000) which could lead to a well-known phenomenon: just a few groups are able to achieve high performance whereas the others are far from reaching the expected goals. This may occur when such a selection gathers students together, resulting in “segregated” groups where all members exhibit desirable or undesirable characteristics. To avoid such a problem, it is important to use group formation methods that not only seek the general performance of each group but also seek adequate results for individuals with different characteristics. In other words, the ideal situation should be having groups with members who are as similar among themselves as possible (inter-homogeneous), but also empowering the students’ individual differences inside such groups (intra-heterogeneous).

Even though at first sight this task may seem quite simple, it is actually very complex for several reasons. On one hand, it depends on how many student characteristics are considered. To illustrate such an issue with a simple example, one may imagine a classroom where several groups are required. In such a case, a simple method for achieving a certain homogeneity (academically speaking) will be to sort the students from high to low according to their grades and to assign one student to each group sequentially. This however is not a guarantee for developing a successful activity because, as mentioned in Liu and Tsai, 2008, having groups with high-achievement peer members is not the only requirement for adequate group work. Then, one consideration is: What may happen if for each student other characteristics are

\* Corresponding author. Tel.: +57 4 425 5221; fax: +57 4 425 5365.

E-mail address: [jmoreno1@unal.edu.co](mailto:jmoreno1@unal.edu.co) (J. Moreno).

considered besides grades? And furthermore, what might happen if such characteristics are not directly proportional among each other? In these cases, the distribution of students among the groups is not trivial, but requires some sort of intelligent search which allows for finding a solution that satisfies the required conditions. On the other hand, the combinatory explosion related to the number of students and required groups is a relevant issue. The relationship among these variables and possible grouping alternatives is factorial, making this an NP-hard problem. As mentioned in [Hwang, Yin, Hwang, and Tsai \(2008\)](#), while the number of students and characteristics under consideration become very large in a course, it is almost impossible for an educator to organize a set of learning groups that meets multiple criteria.

Several studies deal precisely with the group formation problem but considering a limited number of particular student characteristics. For example, the studies presented in ([Graf & Bekele, 2006](#); [Hwang et al., 2008](#); [Lin, Huang, & Cheng, 2010](#)) focus on two characteristics. The first one uses a particle swarm optimization approach and considers student understanding levels and interests. The second one uses an enhanced genetic algorithm approach and considers the number of already known concepts of a certain course domain and the scores of a pre-test (although authors mention that other characteristics could be considered making some modifications). The third one uses an ant colony optimization approach and considers general student performance and an estimation of their personality traits. [Wang, Lin, and Sun. \(2007\)](#) focus on three characteristics which refer to three particular thinking styles and also use a genetic algorithm approach. A study presented by [Bekele \(2005\)](#) considers a larger number of characteristics: gender, group work attitude, interest in mathematics, achievement motivation, self-confidence, shyness, English and mathematics performance; and uses a vector space model, experimenting with three different algorithms to compose groups.

In contrast to those studies, we propose a generic group formation method for solving the problem described, considering an arbitrary number of student characteristics. In order to do that, we translate it into a multi-objective optimization problem in a combinatorial scenario, where each objective consists of reaching the highest similarity level possible with regard to each student characteristic between the mean of each group and the mean of the total students' sample. One way to solve such an optimization problem is using an exhaustive search but, as mentioned before, this approach is not always feasible depending on the number of students and groups. In these cases, heuristic search methods may be a good alternative because even if they do not guarantee an optimum solution, they do try to find a satisfactory one, using a considerably lower computational effort. Some of these methods are: local search, simulated annealing, taboo search, and genetic algorithms ([Glover & Kochenberger, 2003](#); [Reeves, 1993](#); [Resende & Pinho De Sousa, 2004](#)). From these, the last one was the approach selected for this research due to the authors' previous experience in optimization problems.

The rest of the paper is organized as follows: Section 2 presents a brief introduction to genetic algorithms, describing their general procedure. Section 3 describes the proposed method focusing on its mathematical and algorithmic formulation, whereas Section 4 presents its application in a particular study case. Section 5 shows the results obtained from such a study; and finally, Section 6 presents the concluding remarks and future work.

## 2. Genetic algorithms

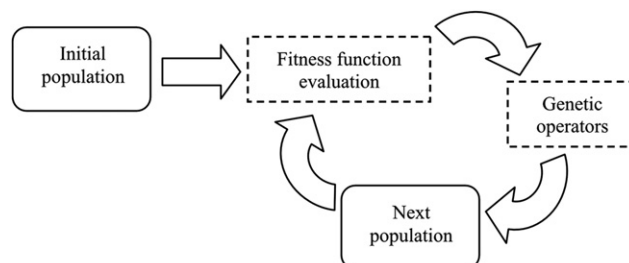
Genetic algorithms (GAs) are considered to be a computational family of models inspired by the evolution principles of Darwin and are often viewed as function optimizers, although the range of problems to which they have been applied is quite broad ([Whitley, 1994](#)). The common feature of these algorithms is that they encode the potential solutions of the problem they face through a chromosome-like data structure, generally an array, and apply recombination operators looking for the preservation of critical information which guides the process to a satisfactory solution ([Goldberg, 1989](#)).

A general schema of a GA is presented in [Fig. 1](#) in which it can be observed that the starting point is an initial population of individuals. An individual is understood to be a feasible solution (not necessarily a good one) and is represented as a chromosome composed of genes where each gene makes reference to a portion or sequence of such a solution. Later, those individuals are evaluated using a certain fitness function, and several genetic operators are applied in order to obtain a new population (the next generation). The aim of such operators is to preserve the chromosomes, or a portion of them, which represent better solutions following the natural selection principles: More suitable individuals have more chances to pass on their genes than those who are not. This process is repeated until a stopping criterion is met.

A more detailed explanation of these processes, as well as their application to the problem of interest, is presented in the next section.

## 3. Method formulation

Considering the GA principles as well as the nature of the problem of interest, the method proposed for composing groups is described in detail in this section. Such a method is based on a previous work presented in ([Moreno, Rivera, & Ceballos, 2010](#)) intended for grouping elements (not necessarily students) homogeneously.



**Fig. 1.** General schema for GAs.

### 3.1. Student representation

Since the idea is to consider not only one, but several student characteristics, each student  $n$  may be represented by using an array, being  $M$  the number of characteristics:

$$E_n = \{C_1, C_2, \dots, C_M\} \quad (1)$$

These characteristics could have different natures. They could be, for example, biographic (age, gender, etc.), academic (grades, pre-tests, self-evaluation, etc.), cognitive (learning styles, intelligence types, etc.), personality traits (leadership, shyness, level of motivation, etc.), and/or many others.

In the array representation, each characteristic  $m$  ( $1 \leq m \leq M$ ) must be a numerical value in a pre-defined range. This does not mean that categorical attributes cannot be considered, because in such a case, a previous numerical discretization process may be used. For instance, if an attribute takes the categorical values *high*, *medium*, and *low*, they may be changed for the numbers 1, 2, and 3, respectively. This way the total set of elements could be represented using an  $N \times M$  matrix,  $N$  being the number of students. Once the elements are organized onto this matrix, all data must be carried to the same scale in order to avoid perturbations in the fitness function calculation. A simple way to do this is using the formula:

$$C^* = \frac{C - C_{\min}}{C_{\max} - C_{\min}} \quad (2)$$

In this way, all data fit a 0–1 range,  $C_{\min}$  and  $C_{\max}$  being the minimum and maximum values, respectively, of the corresponding characteristic in the available dataset.

### 3.2. Individual representation

As was mentioned previously, an individual in the context of GA is understood to be a feasible solution for the problem analyzed, and is represented by a chromosome-like data structure. In the grouping problem, an individual corresponds to a defined setting of groups, each of which has a maximum number of students. In most works in GA, the data structure used is an array where each position corresponds to a part (a gene) of the solution. In the method presented in this paper, a matrix approach is proposed instead, where the amount of rows corresponds to the number of groups desired  $G$ , and the amount of columns corresponds to the maximum size of each group  $N/G$ . In this way, each gene that composes the encoding of the whole solution (which is the chromosome) contains the identifier of a student, and its position inside the matrix defines the group to which it belongs. This representation, besides the simplicity of its formulation and its clear interpretation, eases the use of the crossover genetic operator that is proposed later in this paper.

In the aforementioned grouping problem, as well as in other combinatorial problems, a chromosome cannot have repeated genes. This means that each group element must be placed in one and only one position of the chromosome. For instance, if there are 12 students, and 3 groups are needed, each one would contain exactly 4 different students. In such a case, a feasible individual—if the 12 students are numbered sequentially—may have students 1 to 4 in the first row (or group); 5 to 8 in the second row; and 9 to 12 in the third row.

### 3.3. Fitness value

Since the main goal of this work is to obtain homogeneous groups regarding to the total sample of students, it is necessary to define a measure of such homogeneity for each individual. For doing so, it is necessary to calculate, in first place, the total mean ( $TM$ ) of each characteristic for all students:

$$TM = \{\overline{C_1}, \overline{C_2}, \dots, \overline{C_M}\} \quad (3)$$

Then, for each group  $g$  ( $1 \leq g \leq G$ ) of each individual, the mean of each attribute must be calculated. As each individual  $i$  is represented by a matrix  $X^i$ , such an individual's means ( $IM$ ) may be obtained as:

$$IM_g^i = \{\overline{X_{g,1}^i}, \overline{X_{g,2}^i}, \dots, \overline{X_{g,M}^i}\} \quad (4)$$

Later, the sum of the squared differences ( $D$ ) with regard to the  $M$  characteristics between each group  $g$  of the individual  $i$  and the whole sample is calculated as:

$$D^i = \sum_{g=1}^G \left[ \left( \overline{C_1} - \overline{X_{g,1}^i} \right)^2 + \left( \overline{C_2} - \overline{X_{g,2}^i} \right)^2 + \dots + \left( \overline{C_M} - \overline{X_{g,M}^i} \right)^2 \right] \quad (5)$$

**Table 1**  
Example elements.

ID	$C_1$	$C_2$	$C_3$
1	10	0.8	120
2	20	0.6	200
3	10	0.9	0
4	50	0	560
5	30	0.3	800
6	30	1	600

**Table 2**  
Scaled values.

ID	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>
1	0	0.8	0.15
2	0.25	0.6	0.25
3	0	0.9	0
4	1	0	0.7
5	0.5	0.3	1
6	0.5	1	0.75

The lower the value of this measure—with a minimum of zero—the more similar each group of such an individual would be on average with regards to the whole sample of students. In order to clarify this measure, as well as all the concepts explained so far, consider the following example, where there are six students and three characteristics, as presented in Table 1.

After scaling these values according to the procedure described at the end of Section 3.1, Table 2 is obtained.

Now, suppose that two groups are needed to be formed, each one with two students. Two possible individuals for this case are presented in Table 3.

Applying (3), it is possible to obtain  $TM = \{0.375, 0.6, 0.475\}$ , whereas  $IM_g^i$  is calculated according to (4) using Tables 2 and 3 as follows:

$$IM_g^1 = \begin{Bmatrix} 0.083 & 0.767 & 0.133 \\ 0.667 & 0.433 & 0.817 \end{Bmatrix} \text{ and } IM_g^2 = \begin{Bmatrix} 0.167 & 0.667 & 0.383 \\ 0.583 & 0.533 & 0.567 \end{Bmatrix}$$

Finally, when using (5), the fitness values  $D^1 = 0.4592$  and  $D^2 = 0.1125$  are obtained. When this happens, one may conclude that the grouping represented with individual 2 is more inter-homogeneous than individual 1, i.e., with this distribution, all groups reflect the whole sample of students more accurately when looking all characteristics together.

### 3.4. Initial population and evolution

In the example presented on the left part of Table 3 (individual 1), a trivial group formation is shown, assigning each element sequentially to a group according to their identifiers. In other words, the first  $N/G$  elements (in this case three students) belong to group 1, the next  $N/G$  belong to group 2, and so on. Even though this grouping method is valid, the idea of the initial population is to generate  $k$  individuals randomly by using the matrix representation described in Section 3.2 and fulfilling the restriction of an unrepeated element per position. Once the initial population is obtained, and according to Fig. 1, the evolution process starts passing from one generation to the next using the genetic operators which will be described later, until a certain fitness value is fulfilled, or until  $h$  generations are reached.

### 3.5. Genetic operators

The three basic genetic operators are: selection, crossover, and mutation (Reza Hejazi & Saghafian, 2005). The first operator consists of selecting the most suitable individuals so their chromosomes or a portion of them are preserved for the next generation (according to the principle of survival of the fittest). The second operator consists of generating children from suitable parents, somehow merging their chromosomes (according to the principle of suitable genes preservation). The last operator consists of changing one or more genes of the individuals in the next generation, hoping that such changes in the chromosome level may work in favor of the population.

One way to perform the selection of individuals is by using the mechanism known as *roulette wheel* (Weise, 2008), in which the individuals from a generation may be imagined as portions of a roulette, as shown in Fig. 2, where the area of each portion is proportional to the fitness value  $f$  of the individual  $X$ . Once the portions, or selection probabilities, are defined for each individual, the roulette is spun and when it stops, the corresponding individual is selected. This procedure is repeated until a percentage  $\alpha$  of the generation is selected and their chromosomes are “cloned” onto the next one. Even though Fig. 2 represents the case in which fitness function is maximized, an analogous process may be used in the case of minimization, which is the case in the method proposed.

Now, in order to perform the genes crossover between the chromosomes of several individuals to produce one or several “children”, there are several operators. Some of the most known operators for problems with discrete and unrepeated genes are: two-point crossover (2X), linear order crossover (LOX), partially-mapped crossover (PMX), cycle crossover (CX), C1 operator, NABEL operator, edge recombination crossover operator, multi-parent crossover (MPX), and longest common subsequence crossover (LCSX) (Reza Hejazi & Saghafian, 2005).

The crossover operator selected is a modification of C1. The original version of the C1 operator chooses a crossover point (fixed or random) between the chromosomes of two parents, merges the first segment of the first parent with the second segment, but in the order in which the corresponding genes appear in the second parent, and vice versa. The proposed modification consists of taking advantage of the matrix representation of individuals and using multiple random crossover points (as many as the number of rows), as illustrated in Fig. 3.

**Table 3**  
Example individuals.

Individual 1			Individual 2		
1	2	3	1	3	5
4	5	6	2	4	6

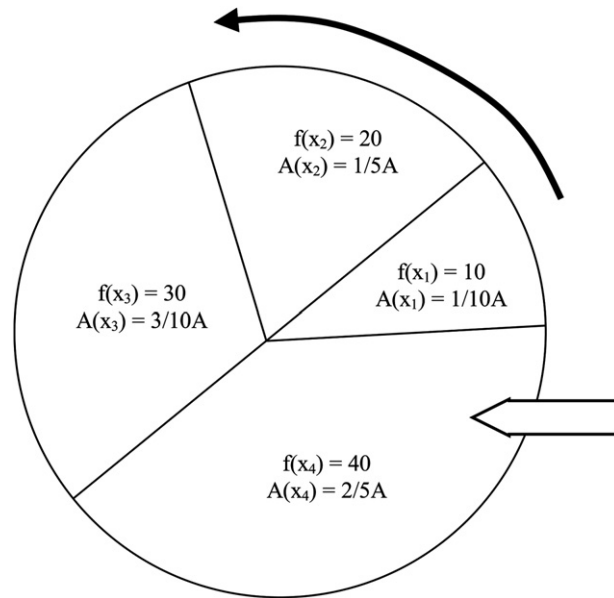


Fig. 2. Roulette wheel example. Adapted from Weise (2008).

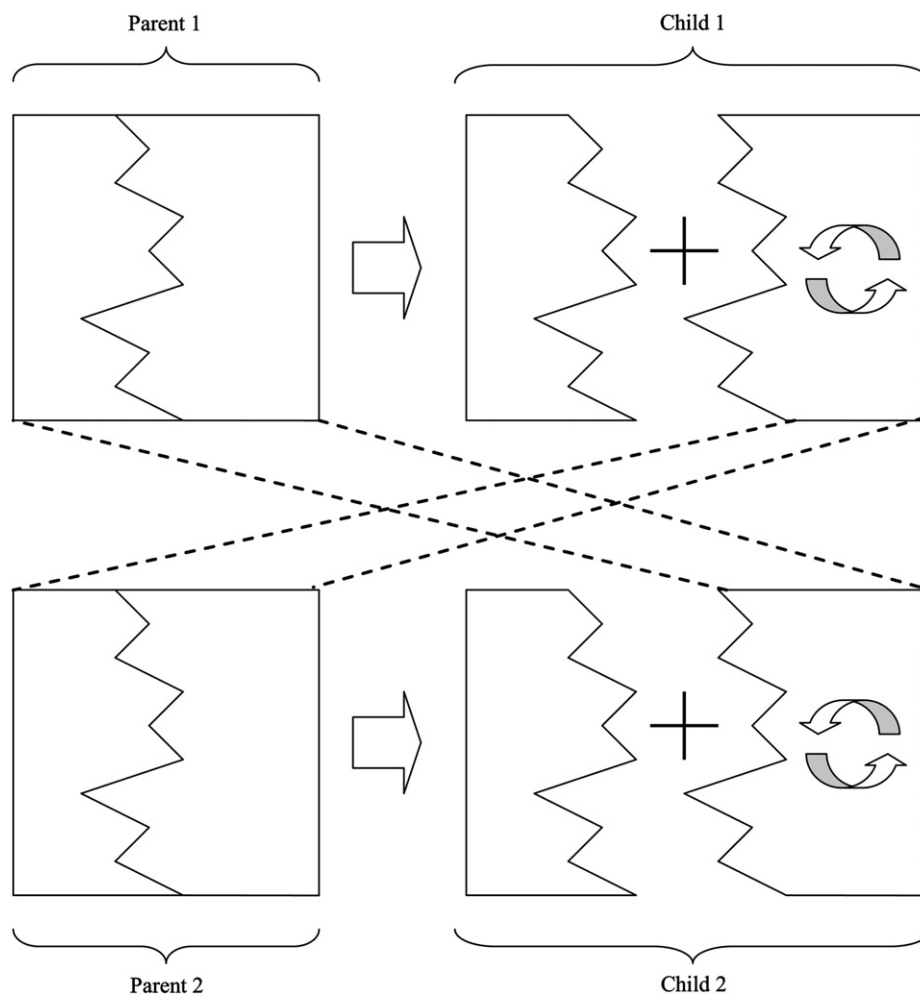


Fig. 3. Crossover operator.

**Table 4**  
Crossover example.

$X^1$				$X^2$				$X^{1*}$				$X^{2*}$			
9	4	12	15	15	7	2	12	9	4	15	12	15	7	12	1
1	3	11	6	4	13	8	14	1	13	3	16	4	11	14	8
14	8	2	13	3	10	16	1	14	8	2	6	3	10	16	2
7	10	16	5	9	6	5	11	7	10	5	11	9	6	13	5

For instance, if the crossover point between two individuals  $X^1$  and  $X^2$  is the array  $R_j$  ( $0 \leq R_j \leq N/G$ ), the students  $X_{j,b}^1$  ( $R_j + 1 \leq b \leq N/G$ ) would be re-sorted according to their positions in  $X^2$  producing the first child, and it would do similarly with  $X^2$  to produce the second one. Table 4 illustrates this procedure with an example using  $R_j = \{2,1,3,2\}$ , where  $X^1$  and  $X^2$  are the parents and  $X^{1*}$  and  $X^{2*}$ , the corresponding children.

In order to keep the number of individuals among generations constant, the remaining  $(1-\alpha)$  percentage of each new generation is obtained through the crossover operation previously described, selecting the parents with the roulette wheel mechanism. Once the individuals of the next generation are defined, the mutation process starts. The individuals which are going to mutate, as well as the genes that change, are selected randomly using two parameters  $\theta$  and  $\phi$ . For instance, if these parameters take the values 0.15 and 0.05, which means that each individual would have a 0.15 mutation probability and, for a selected individual, each of its genes would have a 0.05 probability of change. Considering the grouping problem nature and the representation of the individuals proposed in this paper, it is important to highlight that such a mutation consists of exchanging the gene's value with another -also randomly-, but located in a different row.

### 3.6. Algorithm performance

In order to evaluate the performance of the algorithm proposed in this section, a comparison was done using a dataset of simulated values (uniformly distributed) for only two characteristics. The other competing algorithms are an exhaustive method and a random method. The exhaustive method basically enumerates all possible groups guaranteeing that the optimal fitness value is found. For the GA, 200 generations and 20 individuals per generation were configured, meaning that up to 4000 different solutions were analyzed. To be fair with the random method, the same number of candidate solutions as those used for the GA was configured. In the case of the exhaustive method, the number of analyzed solutions is not fixed. In fact, as mentioned in Section 1, it increases in a factorial proportion with the number of students. Thus, with 6 students, the number of analyzed solutions is 720; with 9 students such a number rises to 362,880; with 12 students such a number rises to approximately  $4.79 \times 10^6$ , and so on.

The results for the different numbers of students are presented in Table 5, where the time is measured in seconds. For the sake of simplicity, the number of members per group was three in all cases. The three methods were implemented in the Java™ programming language using the Netbeans™ environment and all runs were done on a laptop computer with an Intel Core 2 Duo 2.4 Ghz processor and 2 Gb RAM.

There are no surprises when analyzing the execution time. With the exhaustive method, such time increases dramatically, whereas with AG and random methods, it increases almost linearly. Here it is important to clarify that, although with these two methods the number of analyzed solutions does not vary with regard to the number of students, the number of calculations needed to evaluate the fitness function for each one of them does.

As was also expected, the exhaustive method obtains better fitness values as the number of students increases because with more students, more groups can be formed, and variance in the measured variables decreases. This does not happen with the random method because, as the number of possible solutions increases, the chances of finding good ones by mere good luck decrease. What is important to say about the AG approach is that, even if it does not reach optimal values in the fitness values, it is able to obtain adequate solutions in reasonable execution times even when the number of students is large. Having said so, it is also important to mention that small differences in such fitness values are practically irrelevant for all practical purposes.

## 4. Research study

To validate the method proposed, a research study was conducted with freshman students in engineering programs at the National University of Colombia, Medellín Campus, considering certain characteristics. Such a study used as reference point two traditional group formation strategies: random and self-organized. As many educators may know, the two strategies have certain advantages and disadvantages. The first strategy allows for mixing all students together with the hope of reaching heterogeneity inside the groups. This hope however is not always fulfilled because such a strategy may obtain groups where all members exhibit desirable or

**Table 5**  
Performance comparative between exhaustive, random, and proposed method.

Students	Exhaustive		Random		Proposed	
	time	fitness	time	fitness	time	fitness
6	0.004	0.0426	0.038	0.0498	0.088	0.0445
9	0.558	0.0385	0.044	0.185	0.103	0.0411
12	956.021	0.0291	0.053	0.2172	0.134	0.038
24	N/A	N/A	0.098	0.4012	0.205	0.0327
60	N/A	N/A	0.202	0.6719	0.399	0.0316
180	N/A	N/A	0.553	1.2073	0.661	0.0307

undesirable characteristics. The second strategy allows—in theory—having groups with high empathy among the members, but may entail two undesirable consequences: it usually prevents students from working with classmates who do not belong to their regular study groups; and it may produce groups formed not by pedagogical criteria, but by other criteria of a different nature (friendship, for example).

#### 4.1. Participants and instruments

A sample of 135 students from a computer programming course was selected for the research. The CL activity consisted of generating groups of five students each, and giving them an assignment about a specific course subject. Such an assignment had a set of problem-based exercises that each member of the group must solve individually. Later they had to discuss with their teammates to define the group solutions they must deliver to the instructor. The solution for each problem could be that of a particular student (if he/she convinces the others members that it is the best) or any other which they define jointly after discussion. All groups were required to deliver their final solutions as well as the individual ones (in order to validate everyone's participation).

All the students were distributed at the same time and, except in the case of the self-organized groups, the students were only informed about the distribution, not about the distribution mechanisms. Such a distribution was originally done in the following way: 9 were formed using the proposed method, 9 were formed randomly, and the remaining 9 were formed allowing students to organize themselves. However, in two of the randomly formed groups there were personal problems among some students and they ended up rearranging those groups themselves. This occurred before starting the CL activity, so it affected the number of groups per alternative but not the activity development. In summary, there were 9 groups formed with the proposed method, 7 random groups, and 11 self-organized groups.

After the development of the learning activity, each of the 27 assignments was graded with a quantitative value varying from zero to five with one decimal digit. This was done without the course instructor knowing from which one of the three kinds of groups they belonged to, thus avoiding a biased assessment. Two weeks later, an individual test was performed on all students (the first exam of the course) in which the same concepts were evaluated, and it was graded in the same way.

#### 4.2. Considered characteristics

For the particular research study presented in this section, three student characteristics were used based on a literature review and the particular features of the population studied, that is, freshman students with limited experience in teamwork: 1) an estimate of the student related subject knowledge levels, 2) an estimate of the student communicative skills, and 3) an estimate of the student leadership skills. In other words, the main goal of the grouping method in this case is to obtain inter-homogeneous groups in terms of having members with the ability of explaining concepts to colleagues and solving misconceptions (the first characteristic); others with the ability of encouraging and facilitating discussions within the group (the second characteristic); and finally, others with the ability of solving conflicts and coordinating activities when needed (the third characteristic).

To obtain such estimates, three questionnaires were performed for all students, one questionnaire for each characteristic. The first questionnaire was graded from zero to 20 with no decimal points and consisted of a set of multiple-choice single-answer questions about logical reasoning and other related course topics. The second questionnaire was based on an on-line psycho-technical test and consisted of a set of statements about communication issues. An example of those instances is: "I use modulation and the volume of my voice to reinforce what I'm saying", where students should answer with integer values varying from 1 to 4 according to how strongly they felt related with relation to the statements. Finally, the third questionnaire was based on the situational leadership theory (Hersey & Blanchard, 1977) and consisted of a set of situations and alternatives in which the student must select the alternative which describes what he/she usually does in similar scenarios. In all three cases, the outcomes of the questionnaires were scaled in a zero to one range, as explained in Section 3.1, giving an approximate quantitative measure about how "good" each student is in relation to each characteristic.

#### 4.3. Algorithm settings

For the first genetic operator, the selection, it was defined that 40% of the individuals from one population would be cloned into the next, based on the roulette wheel mechanism. The crossover operator was used for the remaining 60% of the new population, maintaining the population number constant. The mutation operator was applied for the entire population in each generation with a mutation probability for each individual  $\theta = 0.2$ , and a mutation probability for each chromosome  $\phi = 0.15$ .

To determine the number of individuals in each population, as well as the maximum number of generations, several runs were made, varying the aforementioned parameters. We will consider the runs in the next section.

**Table 6**  
Summary of the algorithm execution results.

Generations	20 individuals		40 individuals		60 individuals		80 individuals		100 individuals	
	time	fitness	time	fitness	time	fitness	time	fitness	time	fitness
200	0.36	0.0962	0.54	0.0637	0.81	0.0387	1.05	0.0361	1.37	0.0362
400	0.51	0.075	1.08	0.0497	1.6	0.0309	2.12	0.0275	2.84	0.0283
600	0.76	0.0638	1.62	0.0452	2.44	0.0276	3.12	0.0242	4.35	0.0222
800	1.03	0.0583	2.16	0.0414	3.26	0.0267	4.08	0.0228	5.8	0.0209
1000	1.27	0.0561	2.76	0.0409	3.99	0.0257	5.2	0.0213	7.04	0.0203
2000	2.62	0.0554	5.46	0.0398	8.06	0.025	10.61	0.0207	13.66	0.0199
5000	6.54	0.054	13.8	0.0395	19.75	0.0247	26.78	0.0205	34.85	0.0197



**Table 7**  
Summary of the research study results.

	Amount of students	First questionnaire		Group grades		Individual grades	
		Mean	SD	Mean	SD	Mean	SD
Proposed method	45	14.65	3.67	4.53	0.53	4.29	0.71
Random	35	15.30	4.02	4.28	0.53	4.08	0.56
Self-organized	55	14.18	3.45	4.10	1.04	4.04	0.65

## 5. Results and discussion

Results presented in this section are analyzed from two different points of view. First, considering that the grouping problem was translated into a multi-objective optimization problem in a combinatorial scenario, the performance of the method proposed was evaluated for different parameter configurations. Second, considering the research study described in the previous section, the usefulness of the group formation process, taking the selected characteristics along with the groups' distribution itself into account, was validated.

### 5.1. Execution time and fitness value

Several runs were made using the same hardware and software configuration described in Section 3.6, varying the parameters mentioned in Section 4.3. The summary of these runs—20 for each scenario—are presented in Table 6: mean execution time (in seconds), and mean fitness value of the best individual. Variances of such runs are not presented because they were not significant compared to fitness values which means that the initial population in the GA procedure does not have a significant impact on the final results.

As expected, the fitness value improves when the number of individuals per population increases (there is more genetic variance, and therefore more probability of obtaining better individuals). This however, increases execution time as well, so a value of 80 was selected considering that from such a point, the fitness value seems to gain non-significant improvement. When increasing the number of generations, the fitness value improves as well, albeit it does so in a more marginal manner, whereas the execution time difference is practically constant. According to these results, a value of around 1000 for this parameter is recommended for groups with a number of students similar to the one of the study case.

### 5.2. Strategy outcomes

The results of the research study are summarized in Table 7. Here, the grades for the three grouping alternatives are presented, distinguishing the three moments mentioned in Section 4.1: the individual questionnaire performed prior the collaborative learning activity, the group grades obtained during the activity, and the individual grades of a later individual exam.

Before comparing the activity outcomes, and with the aim of verifying that there were no prior differences, all students from the groups were analyzed. When doing this, instead of counting with a pre-test of the corresponding course subjects, we used data from the first questionnaire described in Section 4.2, which may be a good descriptor of the required abilities. Using these data, a hypothesis test was performed using the *t*-statistic for the comparison of two independent samples. The null hypothesis establishes that means are equal and consequently, alternative hypothesis establishes that means are different. The corresponding *t*-statistics and *P*-values are presented in Table 8, as well as the ones used for analyzing group and individual grades. Such values were obtained using Statgraphics® Centurion software. When comparing the three kinds of groups, *P*-values for all pairs are higher than  $\alpha = .05$ . According to these values, alternative hypothesis were accepted with a 95% significance level, meaning that there was no statistical evidence for determining a significant difference among means for the first questionnaire. In other words, all students from three kinds of groups were on average in similar conditions.

Now, in order to determine whether there was a significant difference in the learning processes between groups during the experiment due to the collaborative activity, an analogous hypothesis test was performed, but this time using the group grades dataset. When comparing the groups formed with the proposed method against the ones formed randomly, the *P*-value is lower than 0.05, so it could be said, with a 95% significance level, that the null hypothesis is rejected in favor of the alternative one. In other words, the means are different to each other with a 0.25 difference in favor of the proposed method. When comparing these same groups with the self-organized ones, the *P*-value is also lower than 0.05, leading to a similar conclusion, but this time with a 0.43 difference in favor of the proposed method.

When comparing the random groups with the self-organized ones, the *P*-value is higher than 0.05. Then, it could be said, with a 95% significance level, that the null hypothesis is accepted, so there is not a significant statistically difference among means in this case. It could be noted, however, that there is a 0.18 difference in favor of random groups, which we did not initially expect. This phenomenon may have occurred because such a distribution was made by the instructor, so such a situation could put more pressure on the students of such groups with regard to the groups which were chosen freely.

**Table 8**  
*t*-statistics and *P*-values.

	First questionnaire		Group grades		Individual grades	
	<i>t</i>	<i>P</i> -value	<i>t</i>	<i>P</i> -value	<i>t</i>	<i>P</i> -value
Proposed method/Random	−0.7451	0.4587	2.0930	0.0398	1.4789	0.1432
Proposed method/Self-organized	0.6545	0.5144	2.6715	0.0091	1.8193	0.0722
Self-organized/Random	−1.3601	0.1786	−1.0817	0.2825	0.3101	0.7573



Finally, when doing the same test, but regarding to the grades of the later exam with the aim of evaluating individual learning, corresponding *P*-values in the three comparisons were higher than 0.05, so the alternative hypothesis were accepted. In other words, there was no statistical evidence to determine a significant difference among means between the three kinds of groups. One may note, however, that there is a 0.11 difference in favor of the proposed method compared with the random method, as well as a 0.15 difference compared with the self-organized method.

## 6. Concluding remarks

The work presented in this paper aims to be a relevant contribution in the CL context because the research focuses on a key element: learning group formation method. The main goal is to obtain inter-homogeneous groups, which are as similar as possible to the general characteristics of the total sample of students, but also considering the heterogeneity inside each one.

Although in this paper the experimentation was limited to a few characteristics, the proposed method allows for working with an unlimited number of characteristics from different natures, and therefore can be considered to be generic. In order to achieve this goal, the method uses an array representation for each student, and then translates the grouping problem into a multi-objective optimization problem under a combinatorial scenario, which is solved with a genetic algorithm-based approach. Such an algorithm defines the fitness value as the mean square difference between the whole sample and each group, and it uses an iterative process which is based on the evolution of the species principles, where different solutions are obtained looking for the optimum value.

Of course, before jumping to proclaim the benefits of the proposed method, it is important to highlight that, as in any other algorithmic problem, there are some approaches which are more adequate than others depending on the given conditions. The computational performance experiments presented in Section 3.6 using simulated data for two student characteristics demonstrated that when the number of students is very low (up to 12), an exhaustive method would be the best option. Exceptions for this rule are the on-line and on-demand collaborative learning activities. In these cases, the limit number should be 9 or 10 because the response time becomes highly relevant. This would depend on the hardware and software configuration in which the algorithm is run. All of the tests presented were performed using an “average” desktop computer; however, with a higher-performance computer, only the execution times would decrease (although not dramatically, we expect), neither the fitness values nor the algorithm complexity order.

Now, from a pedagogical point of view, the results obtained from a controlled experiment with freshman students using three characteristics show that groups formed with the proposed method produced better outcomes in terms of observable variables (in this case, grades) than those formed with traditional methods like random assignment and self organization. This was true when comparing the mean grades of the collaborative activity and a later individual exam. However, only in the first case, such a comparison was statistically significant according to a hypothesis test using the *t*-statistic for two independent samples.

Such an experiment not only demonstrated that proposed method achieves the goal of obtaining inter-homogeneous and intra-heterogeneous groups, but also that the characteristics considered positively affect the development of the activities within the collaborative learning context. On this point, it is important to highlight that the incorporation of particular student characteristics is a decision which must be taken carefully, preferably with some pedagogical guidance.

As future work, we expect to repeat the described experiment under different conditions in face-to-face courses as well as in virtual environments which promote computer-supported collaborative learning (CSSL). One of such changes would be using previous student grades to measure not only the results as a group, but also the result that collaborative learning has over low-performance students. Another work pending is to incorporate additional student characteristics, exploiting the multi-criteria nature of the method. Age and gender, for example, have been reported in several studies. Other characteristics, in which we are personally interested, are emotional factors like the empathy among students and their motivation regarding the learning activities.

## Acknowledgments

This research was partially supported by the *Departamento Administrativo de Ciencia, Tecnología e Innovación (Colciencias)*, Colombia, under the program *Doctorados Nacionales 2009*.

## References

- Bekele, R. (2005). *Computer-assisted learner group formation based on personality traits*. Ph.D. Dissertation, University of Hamburg. In: <http://www.sub.uni-hamburg.de/opus/volltexte/2006/2759/>, last access: 25.08.10.
- Dewiyanti, S., Brand-Gruwel, S., Jochems, W., & Broers, N. (2007). Students experiences with collaborative learning in asynchronous Computer-Supported Collaborative Learning environments. *Computers in Human Behavior*, 23(1), 496–514.
- Francescato, D., Mebane, M. E., Porcelli, R., Attanasio, C., & Pulino, M. (2007). Developing professional skills and social capital through computer supported collaborative learning in university contexts. *International Journal of Human-Computer Studies*, 65(2), 140–152.
- Francescato, D., Porcelli, R., Mebane, M., Cuddetta, M., & Renzi, P. (2006). Evaluation of the efficacy of collaborative learning in face-to-face and computer-supported university contexts. *Computers in Human Behavior*, 22(2), 163–176.
- Glover, F., & Kochenberger, G. (2003). *Handbook of Metaheuristics*. Kluwer Academic Publishers.
- Gokhale, A. (1995). Collaborative learning enhances critical thinking. *Journal of Technology Education*, 7, 1.
- Goldberg, D. (1989). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley Longman Publishing.
- Graf, S., & Bekele, R. (2006). Forming heterogeneous groups for intelligent collaborative learning systems with ant colony optimization.. In *Proceedings of intelligent tutoring systems 8th international conference* (pp. 217–226), Jhongli, Taiwan, June 26–30.
- Hersey, P., & Blanchard, K. (1977). *Management of organizational behavior: Utilizing human resources* (3rd ed.). Prentice Hall.
- Huxland, M., & Land, R. (2000). Assigning students in group work projects: can we do better than random? *Innovations in Education and Training International*, 37(1), 17–22.
- Hwang, G.-J., Yin, P.-Y., Hwang, C.-W., & Tsai, C.-C. (2008). An enhanced genetic approach to composing cooperative learning groups for multiple grouping criteria. *Educational Technology & Society*, 11(1), 148–167.
- Krejins, K., Kirschner, P., & Jochems, W. (2002). The sociability of computer-supported collaborative learning environments. *Educational Technology & Society*, 5(1), 8–22.
- Lin, Y.-T., Huang, Y.-H., & Cheng, S.-C. (2010). An automatic group composition system for composing collaborative learning groups using enhanced particle swarm optimization. *Computers & Education*, 55(4), 1483–1493.
- Liu, C.-C., & Tsai, C.-C. (2008). An analysis of peer interaction patterns as discoursed by on-line small group problem-solving activity. *Computers & Education*, 50(3), 627–639.
- Moreno, J., Rivera, J. C., & Ceballos, Y. F. (2010). Agrupamiento homogéneo de elementos con múltiples atributos mediante algoritmos genéticos. *Dyna*, 164, 246–254.

- Nielsen, T., Hvas, A. E., & Kjaergaard, A. (2009). Student team formation based on learning styles at university start: does it make a difference to the student? *Reflection Education*, 5(2), 85–103.
- Prichard, J., Stratford, R., & Bizo, L. (2006). Team-skills training enhances collaborative learning. *Learning and Instruction*, 16(3), 256–265.
- Reeves, C. (1993). *Modern heuristic techniques for combinatorial problems*. New York: John Wiley & Sons.
- Resende, M., & Pinho De Sousa, J. (2004). *Metaheuristics: Computer decision-making*. Kluwer Academic Publishers.
- Reza Hejazi, S., & Saghafian, S. (2005). Flowshop-scheduling problems with makespan criterion: a review. *International Journal of Production Research*, 43(14), 2895–2929.
- Solimeno, A., Mebane, M., Tomai, M., & Francescato, D. (2008). The influence of students and teachers characteristics on the efficacy of face-to-face and computer supported collaborative learning. *Computers & Education*, 51(1), 109–128.
- Tolmie, A., Topping, K., Christie, D., Donaldson, C., Howe, C., Jessiman, E., et al. (2010). Social effects of collaborative learning in primary schools. *Learning and Instruction*, 20(3), 177–191.
- Veermans, M., & Cesareni, D. (2005). The nature of the discourse in web-based Collaborative Learning Environments: case studies from four different countries. *Computers & Education*, 45(3), 316–336.
- Wang, D.-Y., Lin, S. S. J., & Sun, C.-T. (2007). DIANA: a computer-supported heterogeneous grouping system for teachers to conduct successful small learning groups. *Computers in Human Behavior*, 23(4), 1997–2010.
- Weise, T. (2008). *Global optimization algorithms - Theory and application*. University of Kassel. <http://www.it-weise.de/projects/book.pdf> last access: 25.08.10.
- Whitley, D. (1994). A genetic algorithm tutorial. *Statistics and Computing*, 4, 65–85.
- Zurita, G., & Nussbaum, M. (2004). MCSCL: Mobile Computer Supported Collaborative Learning. *Computers & Education*, 42(3), 289–314.