

# Students Group Formation Based on Case-based Reasoning to Support Collaborative Learning

Taís Borges Ferreira, Márcia Aparecida Fernandes  
Faculty of Computer Science  
Federal University of Uberlândia  
João Naves de Ávila 2121, Uberlândia, Brazil  
taisbferreira@ufu.br, marcia@ufu.br

**Abstract**—The group formation has been widely investigated since it is a crucial aspect to perform collaborative work. However, there is no consensus about the best set of metrics or how to combine student’s characteristics to improve group interactions, so it has been considered a challenge. Aiming to cope with that, this work proposes the use of case-based reasoning to suggest groups for collaboration based on the metrics and previous groups’ performances stored in a case base. We gathered data from students working on collaborative tasks to build a case base and ran a grouping experiment in a class of undergraduates to verify the effectiveness of the proposal. The results evidenced that grouping based on the Big Five improved students’ interactions.

**Index Terms**—Group formation, Big-Five Personality Traits, Case-Based Reasoning, Collaborative Learning

## I. INTRODUCTION

THREE attributes are linked to effective learning according to cognitive theory: active learning and knowledge building, cooperation and teamwork, and the use of learning based on problem-solving [1]. Collaborative learning meets these three attributes since it involves social processes through which a small group of students interact and work together to reach a shared goal [1] [2].

Some theories also emphasize learning as a social process that occurs more effectively through interpersonal interactions in a cooperative context [3]. That happens because cooperating with other people allows people to question their initial understanding of a topic and observe different points of view, which motivates them to learn. On the other hand, if a group is not able to interact and work together, learning via collaboration is not possible.

As a group composition influences how people work together to achieve a goal, this is one of the most important aspects to be considered before starting a collaborative activity [4] [5]. Silva et al. [6], for example, also point out that automatically formed groups achieve better performance than those randomly generated. Therefore, it is necessary to employ a strategy of group formation that can support collaboration, as grouping students careless can trigger undesirable situations, such as, social isolation.

In computer-supported collaborative learning, the educational benefits are strongly related to strategies that motivate students to interact with the group [7]. Moreover, the group

composition is crucial to trigger productive interactions between the peers [4] and eliminate conflicts that hinder the collaboration.

Grouping students to work collaboratively is not a task to be addressed by just employing a clustering technique to form homogeneous or heterogeneous groups according to given criteria [8]. Grouping strategy must also combine students’ characteristics with the other grouping criteria and, also, allow forming good groups, whether the groups are homogeneous, heterogeneous, or hybrid.

The Big Five personality traits, one of the most used ways to obtain students’ characteristics, are correlated with learning gains [9]. Groups systematically formed based on students’ Big Five personality traits can improve significantly group outcomes [10]. Thus, they can support group formation, the understanding of how each trait affects the learning process and can provide holder to students [9], [11].

Each Big Five personality trait represents a set of behavioral tendencies that can influence group outcomes [12], [13]. Aiming to take advantage of Big Five personality traits to form groups to support collaborative learning, this work presents the development of a Case-Based Reasoning (CBR) system to cope with the processes of forming groups in which students will be able to interact and work together.

By querying the case base (CB) to form groups, one can expect to have new groups replicating the previous groups’ compositions perceived as good. Through the CBR cyclic operation, each new group is converted into a case that will update the CB with their performance in a collaborative task, so it can be used later to influence new groupings, making the whole grouping process more precise and efficient.

Although all the effects of personality traits on the group’s performance may not be known, by searching a new solution and trying to adapt previous solutions (CBR approach), one can expect the new groups to present good performance too. Intending to verify this hypothesis, we did a group formation experiment by querying the CB.

The paper is divided into 7 sections. The second section presents a review of related work. Section III explains how the proposed CBR works. Section IV explains how the CB was built. The results of changing the similarity metrics to query the CB are in Section V, and the results obtained by grouping students based on the CB are in Sec-

tion VI. Section VII is the conclusion and future perspectives.

## II. LITERATURE REVIEW

The personality trait inventory used in this work was the well known Big Five [14]. The Big Five is a hierarchical organization composed of five basic dimensions of personality that comprehend a large number of human behavior [15]. Although Big Five was usually employed to support psychology studies, it is also a tool used to detect the student's affective state in Computer-Supported Collaborative Learning (CSCL) [13].

The behavior tendency associated with each trait comes up according to the score obtained in that trait. Those people that score high in Neuroticism will tend to be anxious, wary, concerned about social rules. A high level of Extraversion characterizes people that are gregarious, talkative, and usually show positive emotions. High levels of Openness characterizes the tendency of being curious, inquisitive, interested in new ideas. High scores in Agreeableness are related to being cooperative, warm-hearted, and agreeable. Conscientiousness relates to people that are goal-oriented and well-organized. However, as the traits are bipolar, a person who scores low in a trait will tend to show opposite characteristics.

Spoelstra et al. [16] and Altapoulou et al. [9] suggested that conscientiousness influences how productive groups tend to be. Spoelstra et al. [16] proposal considers as good to form productive and learning groups only individuals high scored in conscientiousness. On the other hand, Altapoulou et al. [9] suggest that those with a high degree of conscientiousness should be distributed among the groups, so they could positively influence the group's ability to meet the deadline effectively and efficiently.

Regarding extraversion, Roberts et al. [17] pointed out extraversion as linked to the intensity of activity in an individual's social network and Neuroticism, associated with social isolation in new groups. Extroverted people tend to act as the link between other people, improving interaction. However, as Altapoulou et al. [9] suggested in their study, a group composed only by highly extroverted individuals may negatively influence student's learning gain because they tend to distract from social interaction.

According to Bozionelos [18], extremes of conscientiousness may be considered inadequate to form social ties. Low conscientiousness leads to irresponsibility and too much conscientiousness can induce excessive preoccupation with activities and neglect social relations. Despite the theoretically poor profile for forming social ties, those who scored high in neuroticism had the same success as those with low scores in the network resource sharing. Although high neuroticism relates to very anxious people that are likely to avoid approaching other people since they are concerned about what other think, they tend to act according to the established norms and make efforts to maintain the social ties that have already been established.

The results obtained in the cited studies suggest ways to combine individuals by considering their traits. For instance, since extroverted students are known to be more likely to

neglect the shared goal due to social interaction, groups with all extroverted individuals should be avoided. However, the combination of traits changes the degree of influence of a particular trait in the group [19]. So, forming groups using a base of good groups seem to better than selecting members by evaluating their scores in some specific trait [11].

As for the distribution of student characteristics in a group, Santos et al. [20] showed that heterogeneous groups work better than those where there are similar students. They realized that homogeneous groups take time to collaborate effectively. Ruterfoord's case study [21] also indicated heterogeneous groups in personality as better. In a homogeneous group in terms of personality traits, all members will have the same social skills and weaknesses, without any member to balance these characteristics. If the group is heterogeneous, there are distinct characteristics and greater variability of strengths and weaknesses, making the group able to manage possible issues.

In a recent review of group formation [22], they found evidence that spontaneous groups might take advantage of self-motivation. However, since they are made of participants that share similar interests and points of view, it can lead to unsatisfactory results. Although grouping for collaborations has been widely researched, there is not a consensus about what leads to better results: homogeneous or heterogeneous groups based on student's personality traits[7].

Although Mujkanovic and Bollin [10] concluded that systematically constructed groups show significant improvements of the group outcomes and the composition of the groups are directly related to those outcomes, it is still hard to determine precisely how much personality influences group results. All those results were motivation to build a case base (CB) and use a Case-Based Reasoning (CBR) algorithm to support the new group's suggestion based on the experience of previous groups. So we do not need to program certain types of group configuration. CBR will find them.

In summary, the CBR process can be seen as a cycle involving the tasks of retrieve, reuse, review, and retention of cases. Retrieve seeks to similar cases and in the reuse task, the retrieved cases are adapted to solve a new case. Review task evaluates new cases and, if the solution found is effective, in the retention, it is included as a new case or used to update an existing case [23]. This behavior will be used to update the case base and provide more accurate information to form new groups.

Regarding the similarity metrics used to retrieve similar cases to be used as solutions, Stahl [24] divides the calculation of similarity between case and query, in two steps: the local similarity and global similarity. The global similarity is the function that aggregates the calculated local similarities and can be, for example, a weighted sum of the results. The local similarity concerns the calculation of the distance between the pairs of elements that make up the case and the consultation, to define how close the pairs are.

Besides the functions, to be able to calculate the similarity between a case and a query, it needs to provide a strategy to deal with unknown values. There are two well-known

strategies in the literature. The first one considers the distance between an unknown value and any other value as always 0.5 [25], hence the similarity will also be 0.5. The second strategy [26] is to assign distance 0 when both query and case values are unknown, and distance 1, when one of them is known. Thus, if both are unknown, similarity will be 1. When one of them is known, the similarity will be 0.

Approaches related to similarity also include techniques to automatically learn weights of attributes and the distance function, for example [27]. However, even though we ran some experiments to select the similarity metrics to work with the multi-object case (discussed in the next section), learn the similarity automatically was not the focus of this study. Thus, related to similarity, we adopted a well-known strategy to deal with some peculiarities of the case, such as the unsorted list of students that represent the case.

While talking about the implication for Education, Kolodner [28] mentions CBR as a model that can provide, for instance, suggestions about how a student will be able to have richer learning experiences by giving them the chance of applying what they have already learned. This statement reinforces the idea of using CBR to explore data related to previous group cases to extract information that will help students to take advantage of collaboration.

CBR has been successfully applied in many other areas. For instance, recommendation systems to recommend new items or products to a client ([29], [30]), stress monitoring [31], systems to decide the best assembly sequence ([32], [33]), algorithms to make decisions and find solutions in an environmental emergency scenario ([34]), etc. Although that, there are just a few applications of CBR in educational contexts and for collaboration [35].

Aiming to explore this gap, Cocea and Magoulas [35] modeled the student behavior using a CBR that was also a source of information to feed a clustering approach to form groups. Similarly, in our proposal, CBR models and updates information of students and groups. In a literature review conducted by Costaguta [36], among all the approaches of group formation researched, they found and presented only one work that applies CBR to form groups: the Cocea and Magoulas proposal. That evidences the lack of approaches involving CBR to form groups.

### III. THE PROPOSED CASE-BASED REASONING

The use of Case-Based Reasoning (CBR) for supporting group formation improves the ability to build good groups. In addition, it allows identifying groups that need to be undone to avoid poor group performance. The cyclic operation of CBR contributes to the evolution of the case base (CB), due to adding and updating cases, therefore it is expected that the solution quality will be more accurate, as the CBR cycle runs. To meet this goal, a case structure and a CBR operation are specified in the following subsections.

#### A. Case Structure

A case is an object composed of two types of information: group member characteristics and group metrics. The case

structure in the CB can be seen in Figure 1, where the component Group corresponds to group member characteristics and contains the Big Five personality traits (Openness (O), Conscientiousness (C), Extraversion (E), Agreeableness (A) and Neuroticism (N)), for each group member. The group is an unordered list of students and each student an array of personality traits.

The component Group metrics contain the number of group members (size), deadline, and group metrics. The deadline corresponds to a string indicating how much time the group can spend to solve the task and submit a solution. Its values can be C (class time), which means the task should be done during the class, or W (week) when the task is to be completed in extra-class meetings. The group performance metrics are a list of attributes to classify the performance of the groups. Group performance metrics contains:

- Everybody Contributes (EC) measures group members' contributions to solving a task and if it was a significant contribution to the solution.
- Task Completed (TC) indicates whether the group was able to complete the task within the given deadline or not.
- Grade (G) assigned to the group by the teacher.
- Interactions (I) among group members, by using communication tools or face-to-face.

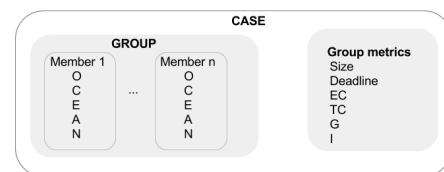


Fig. 1. The structure of a case.

The case structure was designed to represent three different problems related to grouping. The first one, type 1 problem, is to determine the most likely performance of a known group. Then, the case description contains student characteristics of the group and the solution will be the group performance metrics and the group quality. Therefore, the similarity will rely on the similarity between the case and the query groups. Case solution must fit the specified group size and, consequently, solutions containing fewer students are discarded. The deadline can be also used as a filter.

Type 2 problem is to discover the best set of students that are more likely to perform well, which means to find the best partners to form a group given the personality traits of the students. The case description of type 2 is composed of the personality traits of a student. The case solution will be the list of students that could be grouped with a specific student. Task deadline and group size are filters. By performing that kind of query, the group performance is not part of the case description or solution, but thresholds to guarantee the solution is based only on cases that represent good groups.

The type 3 problem is related to find the characteristics of a task that could help students to solve a collaborative task and perform well. The case description is given by the group performance metrics and group composition, and the solution is the deadline. So, the similarity for type 3 must consider both, performance metrics and group composition. The solution is intended to support collaborative tasks. However, these tasks involve other information than their deadline, which is not currently implemented.

### B. Similarity Metrics Between Query and Case

To help teachers group their students and take advantage of collaborative tasks by querying to case base, one needs to define suitable similarity metrics to compare the query with the description of the possible solutions in the case. Possible solutions are the most similar cases in the case base, according to a similarity metric. So, the similarity metrics must be defined based on the case structure.

As the case in our proposed CBR is an object that contains two components, the case-query similarity ( $S$ ) depends on group metrics similarity ( $GMS$ ) and group similarity ( $GS$ ) that corresponds to the average of the students' similarity ( $SS$ ). Similarity calculus also depends on those problem representation types defined in the previous section. Thus, the query type is a parameter considered for querying the CBR.

The final similarity value is  $S = GS$  for type 1 and 2, and  $S = (GS + GMS)/2$ , if the type is the third one. However, since a case (or query) is an object and the similarity between two objects is given by the aggregation of attributes' similarities in a single value, to calculate  $GMS$ ,  $SS$  and  $GS$ , it is necessary to calculate the attribute similarity.

Therefore, let  $a_c$  and  $a_q$  be a pair of corresponding attributes, the first one from the case and the second one from query, the attribute similarity, based on the distance value, can be calculated using one of following functions: Threshold (Eq. (1)), Linear ((2)), Exponential (Eq. (3)) and Sigmoid (Eq. (4)). Then, with the pair distance, given by the linear distance  $d_a = |a_c - a_q|$ , one of these functions is used to convert it into a similarity. The similarity value belongs to interval  $[0, 1]$ , where the value 0 means the attributes are completely different and 1 means they are 100% equal.

$$\text{Threshold: } \text{sim}(a_c, a_q) = \begin{cases} 1, & d_a \leq t \\ 0, & \text{otherwise} \end{cases}, \quad (1)$$

where  $t$  is the threshold of  $d_a$  below which similarity will be 1 (100%). The value of  $t$  was defined as 0. So, if  $d_a = 0$ , the similarity will be 100%, and if  $d_a >= 0$ , will be 0%.

$$\text{Linear: } \text{sim}(a_c, a_q) = \frac{\max - d_a}{\max - \min} \quad (2)$$

where  $\max$  is the maximum possible value of the attribute  $a$  and  $\min$ , is the minimum. If  $a$  is a student's personality trait,  $\max = 1$  and  $\min = 0$ .

$$\text{Exponential: } \text{sim}(a_c, a_q) = e^{d_a * \alpha} \quad (3)$$

where  $\alpha = -1$  and  $d_a \in [0, 1]$ . As the similarity function needs to return a value between 0 and 1, due to the characteristics of Exponential function, if  $\alpha = 1$ , for instance, the result will be something between 1 and 2.7. In case of  $d_a \in \{0, 0.5, 1\}$ , by using  $\alpha = -1$ , it will return similarity into the desired interval.

$$\text{Sigmoid: } \text{sim}(a_c, a_q) = \frac{1}{1 + e^{\frac{d_{if}(c_i, q_i) - \theta}{\alpha}}}, \quad (4)$$

where  $\alpha = 0.01$  and  $\theta = 0.5$ . The  $\theta$  is the value of the central point, making the curve turning point. As the difference is normalized between  $[0, 1]$ , the central value was defined as 0.5. If the curve amplitude were to big, the values of similarities will be also out of the interval. Using  $\alpha = 0.01$  and  $\theta = 0.5$ , it will return values into the interval of  $[0, 1]$ .

Finally, to perform the aggregation of attribute similarities and obtain the value of  $GSM$  or  $SS$ , two aggregation functions were implemented, Minkowski (Eq. (5)) and Simple Matching (Eq. (6)). The  $GMS$  and  $SS$  calculus are quite similar, since group metrics and students are represented by arrays that contain the respective attributes. According to [24],  $GMS$  and  $SS$  are calculated by using a local and global similarities, that correspond to our attribute and object similarities, respectively.

$$\text{Minkowski: } \left( \frac{\sum_{i=1}^n \text{sim}(a_c, a_q)^p}{n} \right)^{\frac{1}{p}}, \quad (5)$$

where  $n$  is the number of attributes. The numeric values of Eq. (5) are aggregated by using exponentiation and radication according to the value of  $p$ . When  $p = 1$ , the Minkowski equation corresponds to the formula of Manhattan distance and, for  $p = 2$ , Euclidean distance.

$$\text{Simple Matching: } \frac{\#equal}{n} \quad (6)$$

where  $n$  is the number of attributes and  $\#equal$  is the number of attributes having equal value. Thus, the similarity calculated with Simple Matching is the rate of equal attributes of the compared objects. This is a type of aggregation function that works better with binary or categorical values because values can be only equal or different. Working with numerical values, it would be better to consider the similarity as the rate that indicates how close their values are, even though they are not equal.

As previously,  $GS$  is the average of student's similarity ( $SS$ ). On the other hand, a group is an unsorted list of students, which means that the first student in the list of a query can be similar to the second (or any other) in the case. Moreover, if the case and the query are equal,  $GS$  must be 100% even though the students are presented in different orders. The strategy adopted to associate each student in the query with the most similar in the case, in order to obtain the correct pair of students to calculate the  $SS$ , is described by the following steps.

- 1) Calculate the similarity between all pairs of students and make a list.

- 2) Select the pair  $P(s_c, s_q)$  with the highest similarity in the list.
- 3) Discard every other value calculated with  $s_c$  or  $s_q$
- 4) If the list still has items, return to step 2.

Step 1 does not really require a significant computational cost. The groups are formed by up to 5 members and building the list takes constant time,  $\theta(1)$ , in the worst-case. To avoid the combinatorial search, a greedy strategy selects the pair with the highest similarity.

### C. CBR Operation

Having the case structure and similarity metric defined, which is given by the similarity  $S$ , defined in the previous section, the four tasks of the proposed CBR are summarized in Figure 2. When a new problem arrives, it is converted to case structure, becoming a new case to be used for querying the CB and retrieving the most similar cases. Once a retrieved case is taken as a solution to the new problem, this solution is revised, and, finally, the group feedback is evaluated to decide if this solution (case) should be added or not to CB. Each task in Figure 2 is described as follows.

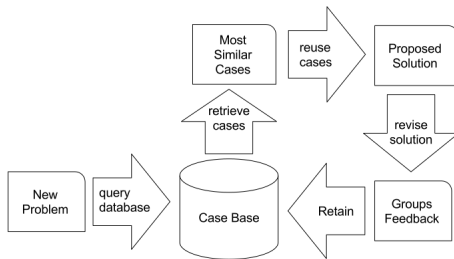


Fig. 2. Case-Based Reasoning operation.

**Retrieve Task:** To retrieve cases as solutions, one should provide data on the case structure elements and, optionally, minima thresholds for each group metric. These thresholds will work as filters for querying the CB. For example, by providing  $G = 0.6$  and  $I = 0.8$  as thresholds in the query, the query result will be only those cases where  $G \geq 0.6$  and  $I \geq 0.8$ . The retrieving process will then bring all the cases that fit the given thresholds and order them according to their similarity with query. Those showing high similarity with the query are the most suitable cases to solve it.

**Reuse Cases:** A query retrieves from CB all suitable cases, given thresholds. The similarity ( $S$ ) between each retrieved case and query is calculated. The most similar cases are selected and applied to solve the new problem. For example, if the query was built taking into account only students' characteristics, without any group performance information, similar cases will be those where the group formation approximates to the group in the query. Then, the most similar cases returned can be applied to predict group performance. This approach allows predicting if the group is more likely to succeed or fail before students are grouped, due to the considerations of their characteristics provided in the query.

**Revise Solution:** The revision of a solution occurs after the groups worked in collaborative tasks. Based on group's evaluation performing those tasks, we use their performance to evaluate the solution and to check how well it worked. We can calculate how many of the new groups worked well calculating each group performance, using the four group performance metrics: everybody contributes to the group solution ( $EC$ ), group's grade ( $G$ ), members' communication level using communication tools or face-to-face ( $I$ ) and task completed ( $TC$ ). The groups were classified as a good or poor, using the weighted average ( $WA$ ) =  $(0,4 * EC) + (0,1 * G) + (0,3 * I) + (0,2 * TC)$ .

Once the weighted average is calculated, the group performance ( $GP$ ) is classified as GOOD, if  $WA > 0.5$ , and POOR, otherwise. The higher weights were applied to  $EC$  and  $I$  because they are the metrics that indicate how many group members are interacting to solve the proposed task. They may also point, for instance, when there is some student isolated and not working with the group. On the other hand,  $TC$  and  $G$  are also metrics that measure group effectiveness. Thus, in addition to interaction metrics, a good group should be able to reach a satisfactory grade and finish the proposed task.

**Retain Task:** The common CBR operation usually cut off solutions that are not classified as good ones in the revision step. In our proposal, poor solutions are as useful as good solutions. The last ones help to form new good groups and the former ones help to identify group compositions that may not succeed. For this reason, both good and poor groups are inserted in the CB in Retain Task.

The way a revised case is inserted in the CB depends on the combination of students in the revised case. If it is a combination that is already represented in a case of the CB, it will be used to update an existing case. If there is no other case with such a combination of students' characteristics in CB, the case will be inserted as a new case.

## IV. BUILDING THE CASE BASE

The CB was formed from real cases. Data were collected in 4 classes composed of students enrolled in Computer Science, Information Management, and Business Management courses. The evaluation activity proposed by the teacher and data collection were different in each class, but all groups were evaluated according to the metrics ( $EC$ ,  $TC$ ,  $G$ , and  $I$ ) in the previous section. In two classes, the collaborative task involved a shared writing tool. The tool records all user activities in its logs. The logs, recorded during task solution, were applied to assess the metrics  $EC$ ,  $TC$ , and  $I$ . In the third class, the teacher used Moodle's chats and forums. Thus, group metrics were assessed by analyzing students' activity in Moodle. In the fourth class, no collaborative tool was used and the metrics  $EC$ ,  $TC$ , and  $I$  were evaluated based on the teacher's report.

Students' personality traits were calculated using the 44-item Big Five inventory [37], translated to Brazilian Portuguese by [38]. Although the collaborative activity was compulsory and part of their evaluative activities, answering the Big Five inventory was not. Thereat, in some groups, some

students have not answered the Big Five inventory and their personality traits were unknown. Students were allowed to choose their groups, limited to 5 members per group. As a result, 24 groups were formed: 2 groups of 2, 13 groups of 3, 2 groups of 4, and 7 groups of 5 students. One student decided to work alone. A total of 87 students was involved in group activities.

The formed CB was mainly composed of individuals that have received medium or high scores in openness, conscientiousness, and agreeableness. It has been at a certain level expected since the activities related to Management and Computing involves being organized, goal-oriented, and good group workers. Undergraduate students are also expected to be more open since it is less likely that a person not interested in new knowledge to join a graduation course. Extraversion, however, was mostly medium or high and just a small number of students low on extraversion, a feature that is likely to happen among students. So, the CB may not be as comprehensive as it must be to cover the possibilities of group combination.

Regarding the characteristics of the groups and the performance observed, the combination of individuals in groups leads to good results. On the other hand, the results regarding the collaboration to solve the task were not good when the time for task resolution was very long (W). Of all of which students had a one-week deadline to solve the work, only the 3-member group interacted and worked collaboratively. So, if the task deadline is long, the group must have 3 or 2 students.

Among the groups with a shorter deadline, some characteristics of the individuals seemed to intervene in group interactions. In the class where groups were mainly composed of students with high and medium scores in all traits except Neuroticism, the only poor group, regarding member interactions and students grades, was the one composed by students having a medium degree of Neuroticism. This suggests that having individuals with high and low Neuroticism is positive for collaboration.

In the class where students had a week to solve the task, and almost all groups failed, although the best group was the one composed by students medium scored in Neuroticism, all students were also high for Openness. The degree of Openness can increase the degree of engagement of individuals, especially if a task is perceived as interesting and lead to the gain of new knowledge. In the class where there was no middle score student in Neuroticism, the combination of a high, middle, and low level of Conscientiousness make them show a high level of interaction.

## V. EFFECT OF CHANGING THE SIMILARITY METRICS

The CBR proposed supports the change of the functions involved in the similarity metric calculation. To test the effect of each function we used the same query to retrieve cases from the CB, built as in Section IV, but changing the combination of the function involved in similarity calculus. A different function can increase (or decrease) the distance between opposites (H and L) and the similarity between close values,

such as M and L. So, the results of each function used to calculate the similarity and its effects according to the case structure defined could be evaluated.

Table I shows the effects on similarity value (column Similarity) obtained using one of the functions available to calculate attribute similarity (AS) together with one of the functions to calculate object similarity (OS). For example, selecting the Threshold function as attribute similarity and the 3-degree Minkowski function (Minkowski  $p=3$ ) as object similarity, the similarity obtained was 71.42%. The presented values (column Similarity) correspond to the similarity between a case, composed of a group of 3 known students, and a query composed of only 1 known student.

A known student is the one that we know the personality traits. The unknown are those that we do not know the personality traits. As a result, by querying the CB, the known student will be compared to a known student of the case, using one of the available AS functions, and the AS of the unknown students will be calculated using the defined strategy to deal with unknown values. As a result, the similarity will be the mean of the OS calculated for each student (known or unknown).

According to the results, the Simple Matching function did not affect the similarity value, even if different functions are used to calculate AS. One can notice by comparing the results in Table I. Even if the function to calculate AS was changed to Threshold, Linear, Exponential, or Sigmoid, the similarity obtained will be the same: 40%.

It is important to notice that the representation of the students' characteristics is numeric. The levels L, M, and H, used to represent their level in each personality trait, in the database corresponds to 0, 0.5, and 1, respectively. As a result, 0.5 is closer to 1 than 0. The Simple Matching only counts attributes with identical values and ignores the degree of proximity between numerical values, no matter how close the values are.

If the values being compared are slightly different the similarity will always be 0 and will not affect the similarity value calculated. So, Simple Matching is not a good choice to deal with numeric values like those stored in our CB. On the other hand, the Minkowski function works by changing the similarity smoothly as according to how close or distant are the values of each attribute in the case and the query. That implies that the Minkowski function will work better than the Simple Matching.

According to the results, an increase in the Minkowski degree ( $p$ ) also increases the similarity. However,  $p=3$  might not be a good choice, because it also increases the number of cases retrieved with higher similarity value, even though they are not too similar. The use of the function degree equals 2 or 3 to compare the members of a group caused similarity between medium score (M) and extreme scores (H or L) to increase. Although the medium scored individuals tend to be a bit similar to those that have a high or low score in personality traits, the proximity caused by  $p=2$  can be suitable to retrieve similar cases that will be used as alternative solutions, when

the solution is nonexistent in the CB.

The results in Table I corresponds to the similarity of the same pair of query and case retrieved according to the functions selected to calculate AS and OS. Despite that, we also performed other queries and compared them with each case on the CB. The effect of changing the functions to calculate similarity was the same observed and described in above.

TABLE I  
THE SIMILARITY BETWEEN A CASE AND A QUERY OBTAINED BY CHANGING THE ATTRIBUTE SIMILARITY (AS) AND OBJECT SIMILARITY (OS)

AS	OS	Similarity
Threshold	Minkowski p=1	65.00%
	Minkowski p=2	69.72%
	Minkowski p=3	71.42%
	Simple Matching	40.00%
Linear	Minkowski p=1	70.00%
	Minkowski p=2	71.10%
	Minkowski p=3	71.89%
	Simple Matching	40.00%
Exponencial	Minkowski p=1	71.07%
	Minkowski p=2	71.73%
	Minkowski p=3	72.26%
	Simple Matching	40.00%
Sigmoid	Minkowski p=1	70.00%
	Minkowski p=2	71.10%
	Minkowski p=3	71.89%
	Simple Matching	40.00%

## VI. GROUP SUGGESTION BASED ON CBR QUERIES

In the previous experiments, we did not set up query thresholds, since the goals were to observe the CBR behavior and the similarity changes when the functions involved were alternated. This time, the goal was to suggest good groups. By grouping students based on querying the built CB, it was expected to form new good groups reflecting the CB.

The grouping experiment was conducted in a class of Data and Business Information. To determine student's personality traits, we asked them to answer the 44-item Big Five Inventory, a questionnaire translated to Portuguese and validate in Brazil by Andrade [38]. Using the questionnaire answers, we calculate the value of each personality trait and fill in the array of characteristics that represent each student.

After a previous group activity, in which students could choose their partners, the teacher proposed 3 more activities, but now, using the suggestions based on CBR queries. As it was intended to form groups, the traits of each student in the class were used to build queries of type 2. The result of every type 2 query is a list of students that could work well with the student given in the query. So, type 2 was set up before performing the queries.

All the queries were also set up with the following parameters: grade higher than 60% of the total grade ( $G > 0.6$ ), the interaction between students greater than 50% ( $I > 0.5$ ), and everybody contributes to solving a task ( $EC = 1$ ). To calculate the similarity, the function Linear was chosen to calculate attribute similarity and Minkowski with  $p = 2$

to calculate object similarity. The strategy adopted to deal with unknown values in all queries performed was to consider unknown values as 50% similar.

The cases retrieved meeting all the restrictions were ordered according to the similarity between the cases and the query. Then, groups were formed by grouping together those students whose characteristics brought the same case as a solution. That is, together they will be similar to the case used to group them. There was no automation of the grouping based on the query solution when the experiments were conducted. The list of suggestions was manually done.

The eight resulting groups were suggested to the teacher. In Table II, the column "Group" identifies the suggested groups. For each student, their score in each Big Five personality trait is shown in columns O, C, E, A, and N. Students' characteristics in three of the suggested groups were quite similar to the cases used to group them.

The other groups were also similar but, for some traits, the differences were bigger. However, they were formed preserving the patterns observed as good while building the CB. For example, except for a group, there is no suggested group composed only by a middle-scored individual in Neuroticism. Almost every group has more one member with high in Openness, considered good for collaboration.

Groups with all members high scored in Consciousness and Extroversion were also avoided. The exception is the group H. Each of these traits isolated can increase the probability of certain undesirable situations. High Extroversion, for example, may lead to distraction with social interactions and high Conscientiousness to isolation due to excessive focus on goals. However, both simultaneously can reduce the possible negative effects.

The day the collaborative activities were applied, some students were absent or arrived late. Thus, the teacher changed the suggested groups a bit. For the first activity, the teacher decided to remove the absent students without making any other changes. Thus, students in groups D, G, and H worked with fewer members. The group adopted by the teacher in the first activity and the performance observed according to the level interaction ( $I \in [0, 5]$ ) are shown in Table III.

All groups in the first activity showed good interaction ( $I = 5$ ). The exception was the group G that lost the student 32, the only member with L score for Extroversion. In the following activities, the groups were also changed because of some students not present in the first activity, showed up for the second one, for example. The group that shown poor interaction in activity 1 was modified by removing the student 31. Most of the groups remained the same formation in the three activities. In the second and third activities, all groups had a good level of interaction ( $I=5$ ).

According to the report on collaborative activities sent by the teacher, all groups were able to complete the proposed activities. Moreover, the groups formed by means the CBR recommendation were more efficient and better regarding the iteration, when compared to the groups formed by students themselves. Despite the questioning about the group forma-

TABLE II  
GROUPS SUGGESTIONS BASED ON THE QUERIES.

Group	Student	O	C	E	A	N
A	1	H	H	L	M	M
	2	M	H	M	M	M
	3	H	H	M	M	H
B	4	H	H	M	H	L
	5	H	H	L	M	H
	6	M	M	L	M	L
	7	M	M	M	M	L
C	8	H	M	M	M	H
	9	H	M	M	H	L
	10	M	H	L	M	M
	11	M	M	M	M	L
D	12	M	M	M	H	L
	13	H	M	M	H	L
	14	H	M	M	H	L
	15	H	M	M	M	H
E	16	H	L	M	M	M
	17	H	M	M	H	L
	18	M	M	M	H	M
	19	H	M	M	M	M
F	20	H	M	M	M	M
	21	M	M	H	H	H
	22	H	H	M	H	H
	23	H	M	M	M	M
G	24	M	H	H	H	M
	25	M	M	L	H	M
	26	M	M	H	H	M
	27	M	M	H	H	M
H	28	M	H	M	H	M
	29	M	H	M	H	M
	30	M	M	M	M	L
	31	M	M	M	M	M
H	32	M	M	L	M	M
	33	L	H	H	H	M
	34	H	H	H	H	L
	35	M	H	H	H	L
	36	H	H	H	H	L
	37	H	H	H	H	L

tions, the students' perception of their performance was positive considering the new grouping. Also, groups that worked together on previous activity, in general, continued to show good interaction in the following activities.

Groups B, D, E, and H remained unchanged in the three activities and showed positive results, which points as positive such combinations of characteristics. They were groups of students with extreme scores (H or L) or groups of students with scores M, combined with H or L for the trait of Neuroticism. Group G, the group that had poor interaction in the first activity, was mainly formed by students with score M in most of the traits. By removing a member, it reduced the number of members having M score for Consciousness, Ability, and Neuroticism, which seems to improve the interaction.

Even after the change to deal with students' presence/absence on the day activities 1, 2, and 3 were applied, the groups also suffered alterations that, however, did not negatively influence the interaction and did not result in bad groups. The changes made by the teacher bear some resemblance to previous activity groups and base cases. However, they also showed differences in some traits and therefore could also be used to popular the case base as new examples of good groups.

A student is represented by an array of 5 personality traits

TABLE III  
FIRST ACTIVITY GROUPS AND INTERACTION.

Group	Student	O	C	E	A	N	I
A	1	H	H	L	M	M	5
	2	M	H	M	M	M	
	3	H	H	M	M	H	
B	4	H	H	M	H	L	5
	5	H	H	L	M	H	
	6	M	M	L	M	L	
	7	M	M	M	M	L	
C	8	H	M	M	M	H	
	9	H	M	M	H	L	5
	10	M	H	L	M	M	
	11	M	M	M	M	L	
D	12	M	M	M	H	L	
	13	H	M	M	H	L	
	14	H	M	M	H	L	5
	15	H	M	M	M	H	
E	16	H	L	M	M	M	
	19	H	M	M	M	M	5
	20	H	M	M	M	M	
	21	M	M	H	H	H	
F	22	H	H	M	H	H	
	23	H	M	M	M	M	
	24	M	H	H	H	M	5
	25	M	M	L	H	M	
G	26	M	M	H	H	M	
	27	M	M	H	H	M	
	28	M	H	M	H	M	1
	29	M	H	M	H	M	
H	30	M	M	M	M	L	
	31	M	M	M	M	M	
	33	L	H	H	H	M	5
	34	H	H	H	H	L	
H	35	M	H	H	H	L	
	36	H	H	H	H	L	

that can assume 3 different values (L, M, and H), so they can be represented in 243 different ways. That means we have about 59,000 possibilities for 2-member groups. As the inclusion policy adopted in this work is to include as a new case every new group that has a configuration not found in the case base (there is no case matching under 100% of similarity), which means that, in the worst case, the base will reach around 59,000 2-member cases.

It can get worst if we think about groups formed by 3, 4, and 5 students. Therefore, as the case base grows bigger the strategies deal with such a huge amount of registers are a necessity. The data we worked on until now is not too big so we can search for all the cases to find a suitable one. Despite that, it is an issue that should be addressed in a future version of the CBR presented in this paper.

## VII. CONCLUSION

In this work, we proposed the use of case-based reasoning to support the creation of groups to work in collaborative tasks. It is a crucial aspect when it comes to performing collaborative work. The way individuals are grouped to work together can influence their interactions or lead to an undesirable situation, such as isolation in a group. Due to that relevance, group formation has been widely studied in the context of Collaborative Learning. However, there is no consensus on what are the best set of metrics that improve group quality



and how to combine students' characteristics to improve group interactions. Therefore, this issue is still considered a challenge.

Many known works suggested clustering algorithms to form groups, but it is not just a problem of clustering students together according to their similarities or differences. It needs to consider the combination of the different attributes of the groups to reach good results. Aiming to cope with this issue, we proposed the case-based reasoning to recommend suitable groups for collaborative work. The basic four operations of case-based reasoning allow the teacher to use previous knowledge on group performance, according to the characteristics of their students, as well to form new groups.

In this work, we used students' characteristics (Big-Five personality traits) and group attributes to represent a case. To populate the case base with real cases, we collect data from undergraduate students' groups working on collaborative tasks. Next, with this case base, we ran experiments to form groups in a different class, aiming to verify the effectiveness of the proposal. Based on the group's performance and results reported by the class teacher, it evidenced that personality traits influence the interaction level in a group. Furthermore, the results demonstrated grouping based on the Big-Five personality traits improved students' interactions in that class.

We also tested the effect of changing similarity metrics employed to retrieve a solution from the case base. Considering the current configuration of our case base, with all attributes represented as numeric values, the Minkowski (with  $p=2$ ) fits well the role of object similarity metric. As well, the linear function as the similarity between attributes. Despite that, a broader set of functions to calculate similarity metrics supports the inclusion of new group attributes and new types of case representation.

As future work, we plan to include new metrics known as influencing groups' outcomes, for instance, student role and test its effectiveness in interactions, student's motivation, and academic performance of the students in a group. Furthermore, we plan to update and enlarge the case base with new data from groups working collaboratively. That might cause the base to grow big and make the search in all case base registers computationally expensive. Thus, to define an indexing technique or a strategy for selecting relevant cases and removing the unnecessary ones is another further work that must be done. Consequently, retrieving cases to suggest a group formation will keep reasonable computing time, even if the case base grows bigger.

#### REFERENCES

- [1] M. Alavi, "Computer-mediated collaborative learning: An empirical evaluation," *Journal MIS Quarterly*, vol. 18, pp. 159–174, 1994. doi: <https://doi.org/10.2307/249763>
- [2] A. Hron and H. F. Friedrich, "A review of web-based collaborative learning: factors beyond technology," *Journal of Computer Assisted Learning*, vol. 19, pp. 70–79, 2003. doi: <https://doi.org/10.1046/j.0266-4909.2002.00007.x>
- [3] L. Vygotsky, *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press, 1978.
- [4] I. Magnisalis, S. Demetriadis, and A. Karakostas, "Adaptive and intelligent systems for collaboration learning support: A review of the field," *IEEE Transactions on Learning Technologies*, vol. 4, pp. 5–20, 2011. doi: <https://doi.org/10.1109/TLT.2011.2>
- [5] S. Manske, T. Hecking, I. Chounta, and H. Hoppe, "Using differences to make a difference: a study on heterogeneity of learning groups," in *Proceedings of International Conference on Computer Supported Collaborative Learning*, 2015. doi: <https://doi.org/10.22318/csc12015.191>. ISSN 1573-4552 pp. 182–189.
- [6] F. E. O. Silva, C. L. R. Motta, and F. M. Santoro, "Team composer: Assembling groups through social matching," in *Proceedings of the International Conference on Computer Supported Cooperative Work in Design*, 2010. doi: <https://doi.org/10.1109/CSCWD.2010.5471990> pp. 128–133.
- [7] R. C. D. Reis, C. L. Rodriguez, G. C. Chalco, P. A. Jaques, I. I. Bittencourt, and S. Isotani, "Relação entre os estados afetivos e as teorias de aprendizagem na formação de grupos em ambientes cscl," in *Anais do XXVI Simpósio Brasileiro de Informática na Educação*, 2015. doi: <http://dx.doi.org/10.5753/cbie.sbie.2015.1012> pp. 1012–1021.
- [8] J. V. M. Chaves, "Automatic group formation," Ph.D. dissertation, Faculdade de Engenharia da Universidade do Porto, 7 2019.
- [9] P. Altanopoulou and N. Tselios, "How does personality affect wiki-mediated learning?" in *Proceedings of International Conference on Interactive Mobile and Communication Technologies and Learning*, 2015. doi: <https://doi.org/10.1109/IMCTL.2015.7359546> pp. 16–18.
- [10] A. Mujkanovic and A. Bollin, "Personality-based group formation: A large-scale study on the role of skills and personality in software engineering education," in *OCCE 2018: Empowering Learners for Life in the Digital Age*, 2019. doi: [https://doi.org/10.1007/978-3-030-23513-0\\_21](https://doi.org/10.1007/978-3-030-23513-0_21) pp. 207–217.
- [11] T. Ferreira, J. Buiar, M. Fernandes, A. Pimentel, and O. Luiz, "Detecção automática de traços de personalidade e recomendação de agrupamento com o modelo big five," in *XXIX Simpósio Brasileiro de Informática na Educação (Brazilian Symposium on Computers in Education)*, 2018. doi: <http://dx.doi.org/10.5753/cbie.sbie.2018.1643> pp. 1643–1652.
- [12] M. A. G. Peeters, C. G. Rutte, H. F. J. M. van Tuijl, and I. M. M. J. Reyman, "The big five personality traits and individual satisfaction with the team," *Small Group Research*, vol. 37(2), p. 187–211, 2006. doi: <https://doi.org/10.1177/1046496405285458>
- [13] R. C. D. Reis, S. Isotani, C. L. Rodriguezac, K. T. Lyraa, P. A. Jaques, and I. I. Bittencourt, "Affective states in computer-supported collaborative learning: Studying the past to drive the future," *Computers and Education*, vol. 120, pp. 29–50, 2018. doi: <https://doi.org/10.1016/j.compedu.2018.01.015>
- [14] L. R. Goldberg, "Language and individual differences: The search for universal in personality lexicons," *Review of personality and social psychology*, vol. 2, pp. 141–166, 1981.
- [15] R. R. McCrae and O. John, "An introduction to the five-factor model and its applications," *Journal of Personality*, vol. 60, pp. 175–215, 1992. doi: <https://doi.org/10.1111/j.1467-6494.1992.tb00970.x>
- [16] H. Spoelstra, P. Van Rosmalen, T. Houtmans, and P. Sloep, "Team formation instruments to enhance learner interactions in open learning environments," *Computers in Human Behavior*, vol. 45, pp. 11–20, 2015. doi: <https://doi.org/10.1016/j.chb.2014.11.038>
- [17] S. G. B. Roberts, R. Wilson, P. Fedurek, and R. I. M. Dunbar, "Individual differences and personal social network size and structure," *Personality and Individual Differences*, vol. 44, pp. 954–964, 2008. doi: <https://doi.org/10.1016/j.paid.2007.10.033>
- [18] G. Bozionelos, "The relationship of the big-five with workplace network resources: More quadratic than linear," *Personality and Individual Differences*, vol. 104, pp. 374–378, 2017. doi: <https://doi.org/10.1016/j.paid.2016.08.036>
- [19] T. Ferreira and M. Fernandes, "Detecção de traços de personalidade em textos para apoiar a formação de grupos para colaboração," in *Proceedings of Brazilian Symposium on Computers in Education*, 2017. doi: <http://dx.doi.org/10.5753/cbie.sbie.2017.1627> pp. 1627–1636.
- [20] O. C. Santos, A. Rodriguez, E. Gaudio, and J. G. Boticario, "Helping the tutor to manage a collaborative task in a web-based learning environment," in *Supplementary Proceedings of International Conference on Artificial Intelligence in Education*, vol. 4, Sidney, Austrália, 2003, pp. 153–162.
- [21] R. H. Rutherford, "Using personality inventories to form teams for class projects – a case study," in *Proceedings of SIG-ITE'06 Proceedings of the 7th conference on Information Technology Education*. Canterbury, United Kingdom: ACM, 2006. doi: <https://doi.org/10.1145/1168812.1168817> pp. 73–76.

- [22] S. Borges, R. Mizoguchi, I. I. Bittencourt, and S. Isotani, "Group formation in cscl: A review of the state of the art," *Higher Education for All. From Challenges to Novel Technology-Enhanced Solutions*, vol. 832, pp. 71–88, 2018. doi: [https://doi.org/10.1007/978-3-319-97934-2\\_5](https://doi.org/10.1007/978-3-319-97934-2_5)
- [23] J. L. Kolodner, "An introduction to case-based reasoning," *Artificial Intelligence Review*, vol. 6, pp. 3–34, 1992. doi: <https://doi.org/10.1007/BF00155578>
- [24] A. Stahl, "Learning of knowledge-intensive similarity measures in case-based reasoning," Ph.D. dissertation, Departamento de Ciência da Computação da Universidade de Kaiserslautern, 10 2003.
- [25] F. Ricci and P. Avesani, "Learning a local similarity metric for case-based reasoning," in *International Conference on Case-Based Reasoning (ICCBR): Case-Based Reasoning Research and Development*, vol. 1010. Sesimbra, Portugal: Springer, 1995. doi: [https://doi.org/10.1007/3-540-60598-3\\_27](https://doi.org/10.1007/3-540-60598-3_27) pp. 301–312.
- [26] J. Surma and K. Vanhoof, "Integrating rules and cases for the classification task," in *International Conference on Case-Based Reasoning (ICCBR): Case-Based Reasoning Research and Development*, vol. 1010. Berlin, Heidelberg: Springer, 1995. doi: [https://doi.org/10.1007/3-540-60598-3\\_29](https://doi.org/10.1007/3-540-60598-3_29) pp. 325–334.
- [27] P. Perner, "Case-based reasoning - methods, techniques, and applications," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications (CIARP 2019)*, vol. 11896, 2019. doi: [https://doi.org/10.1007/978-3-030-33904-3\\_2](https://doi.org/10.1007/978-3-030-33904-3_2) pp. 16–30.
- [28] J. L. Kolodner, "Educational implications of analogy a view from case-based reasoning," *American Psychologist*, vol. 52, pp. 57–66, 1997. doi: <https://doi.org/10.1037/0003-066X.52.1.57>
- [29] V. Gupta and S. K. Sahana, "Nudge-based hybrid intelligent system for influencing buying decision," *Advances in Computational Intelligence. Advances in Intelligent Systems and Computing*, vol. 988, no. 1, pp. 165–174, 2020. doi: [https://doi.org/10.1007/978-981-13-8222-2\\_14](https://doi.org/10.1007/978-981-13-8222-2_14)
- [30] J. W. Chang, M. C. Lee, and T. I. Wang, "Integrating a semantic-based retrieval agent into case-based reasoning systems: A case study of an online bookstore," *Computers in Industry*, vol. 78, pp. 29 – 42, 2016. doi: <https://doi.org/10.1016/j.compind.2015.10.007> Natural Language Processing and Text Analytics in Industry.
- [31] S. Begum, M. U. Ahmed, P. Funk, and R. Filla, "Mental state monitoring system for the professional drivers based on heart rate variability analysis and case-based reasoning," in *2012 Federated Conference on Computer Science and Information Systems (FedCSIS)*, 2012, pp. 35–42.
- [32] S. Chen, J. Yi, H. Jiang, and X. Zhu, "Ontology and cbr based automated decision-making method for the disassembly of mechanical products," *Advanced Engineering Informatics*, vol. 30, no. 3, pp. 564 – 584, 2016. doi: <https://doi.org/10.1016/j.aei.2016.06.005>
- [33] Y. Qin, W. Lu, Q. Qi, X. Liu, M. Huang, P. J. Scott, and X. Jiang, "Towards an ontology-supported case-based reasoning approach for computer-aided tolerance specification," *Knowledge-Based Systems*, vol. 141, pp. 129 – 147, 2018. doi: <https://doi.org/10.1016/j.knosys.2017.11.013>
- [34] D. Wang, K. Wan, and W. Ma, "Emergency decision-making model of environmental emergencies based on case-based reasoning method," *Journal of Environmental Management*, vol. 262, p. 110382, 2020. doi: <https://doi.org/10.1016/j.jenvman.2020.110382>
- [35] M. Cocea and G. D. Magoulas, "User behaviour-driven group formation through case-based reasoning and clustering," *Expert Systems with Applications*, vol. 39, p. 8756–8768, 2012. doi: <https://doi.org/10.1016/j.eswa.2012.01.205>
- [36] R. Costaguta, "Algorithms and machine learning techniques in collaborative group formation," in *MICAI 2015: Advances in Artificial Intelligence and Its Applications*, 2015. doi: [https://doi.org/10.1007/978-3-319-27101-9\\_18](https://doi.org/10.1007/978-3-319-27101-9_18) pp. 249–258.
- [37] O. John and S. Srivastava, "The Big Five trait taxonomy: History, measurement, and theoretical perspectives," *Handbook of personality: Theory and research*, vol. 2, pp. 102–138, 1999.
- [38] J. M. Andrade, "Evidências de validade do inventário dos cinco grandes fatores de personalidade para o brasil," Ph.D. dissertation, Instituto de Psicologia - Universidade de Brasília, 7 2008.