Computers and Electrical Engineering 000 (2017) 1–10

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

Classification of collaborative behavior from free text interactions☆

Franco D. Berdun, Marcelo G. Armentano*, Luis Berdun, Martín Mineo

ISISTAN Research Institute (CONICET/UNICEN), Campus Universitario - Paraje Arroyo Seco - Tandil, Argentina

A R T I C L E I N F O

Article history:
Received 14 November 2016
Revised 17 July 2017
Accepted 20 July 2017
Available online xxx

Keywords:
Collaborative behavior
Interaction process analysis
Computer-supported cooperative work
Classification

A B S T R A C T

In a computer-supported collaborative work (CSCW) environment, a group of people can work together to fulfill a given assignment. In this context, collaborative problems might naturally arise. The detection of these problems is extremely important for teachers or team leaders to help the team members to improve their collaborative skills and the resolution of the collaboration problems. The observation and analysis of the interaction process of several groups is a time-consuming and difficult task for any teacher or team leader. In this article, we propose a multi-phase classification approach to automatically classify free text observed form the chat of a CSCW environment into different collaborative categories of behavior. We obtained promising results that can be used to help teachers or team leaders with the interaction process analysis in order to focus only in the resolution's actions of the problems that might arise.

1. Introduction

Currently, the large number and variety of platforms and resources focused on collaboration, along with the simplicity that exists in the context of Web 2.0 (blogs, wikis, collaborative documents, video and audio sharing platforms, etc.) offer new possibilities to enhance learning, and generate new challenges in the teaching and learning processes. It has been found, for example, that the use of a wiki in a classroom fosters collaborative learning among students [1]. It has also been shown that many collaborative activities based on Internet facilitate teamwork [2], social skills and basic computer skills [3].

These tools generate a large amount of data from the interactions among participants. These data can be used, for example, to analyze if any participant demonstrate collaborative problems during the performance of a task, or to improve and shape the learning process itself. Many tools have been developed to improve the learning processes; for example, intelligent assistants that are able to detect collaborative problems from the interaction among students in the context of a collaborative work platform either based on a work plan [4], or based on group interactions [5]. These intelligent assistants can alert teachers about the detected problems, so that they can intervene or make recommendations to students in order to take corrective actions.

Collaborative skills are specially important in CSCW environments where the participants of a shared task do not meet face-to-face. Then, in order to achieve an effective collaboration between participants, it is needed to identify which skills they have and which not. By identifying the collaborative skills of the participants, it is possible to build an intelligent

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Dr. S. Sioutas.

* Corresponding author.
E-mail address: marcelo.armentano@isistan.unicen.edu.ar (M.G. Armentano).

http://dx.doi.org/10.1016/j.compeleceng.2017.07.015

system that assist them to achieve a better collaboration experience. For example, in the software market, the increasing complexity of applications requires the coordinated efforts of one or more teams to carry out the execution of any project. Teamwork is an issue frequently cited among managers and business owners: the team members can often feel that the participant members outshine their ideas, or that teamwork slows processes, rather than contribute positively to draft.

In order to be able to detect collaborative problems, we need some characterization of collaborative abilities. In this work, we used the Interaction Process Analysis method (IPA), which is one of the most elaborated, better validated and most used since its formulation.

In this article, we propose a multi-phase classification approach that analyze the interactions among participants of a collaborative work and matches the free text obtained from the chat of the collaborative platform to the interactions scheme proposed by the IPA method. We evaluated our approach in the context of a classroom capstone project, obtaining significant improvements respect to previous approaches.

The rest of the article is organized as follows. In Section 2 we present some related work. Then, Section 3 we present the materials and methods used in the proposed approach. We describe the CSCW environment selected (Section 3.1), the method for classifying the interactions (Section 3.2) and our three-phases classification approach (Section 3.3). In Sections 4 and 5 we present the experiments performed, describing the dataset constructed, the experiment process and the results obtained. Finally, in Section 6 we present our conclusions.

2. Related work

The Interaction Process Analysis (IPA) for computer-mediated communication has been used in some research in the area of Computer Sciences. Birnholtz et al. [6], for example, presented an experimental study of two-people groups (dyads) editing documents together. The focus of the study, however, was not to automatically classify the interactions into IPA categories but to analyze group maintenance, impression management and relationship-focused behavior. Savolainen [7] studied the extent to which blogs are used as interactive forums in which people can ask and share information. Differently to our approach, the author only considered the four reactions of the IPA categorization: showing positive and negative reactions, and asking and answering questions. Experiments on eight blogs revealed that the emotional reactions exhibited by blog contributors were predominantly positive, suggesting that blogs are particularly strong in building supportive communities of interest. Additionally, the study revealed that the interactions mainly occurred through giving information and opinion, while the role of asking questions remained secondary.

There are some related work that rely on IPA classifications but, differently to our approach, the classifications are made manually by human experts. Kim et al. [8], for example, manually classified video recordings segments into IPA categories. After that the group behavior has been manually classified, authors analyze the effect of teamwork practice activity promoting design team creativity and investigate design team interaction from the perspective of personal creativity modes in student design teams. Löfstrand and Zakrisson [9] followed a similar approach. Authors manually mapped video recordings into IPA categories with the aim of identifying competitive and non competitive behavior. More recently, Delic et al. [10] presented a method aimed at observing and measuring the evolution of user preferences and actions in a tourism decision making task. Authors used audio recordings of group discussions and manually identified and categorized each unit of interaction according to the twelve categories of behavior. This literature review allows us to claim that the automation of the classification procedure presented in this article will make it possible to enhance existing systems and to develop software assistants aiming at helping teachers and team leaders in the study of group dynamics.

Attempts to reduce the effort needed to encode group dynamics into IPA categories introduced the concept of *sentence openers*. Sentence openers predefine a limited set of sentences that are allowed to start a communication act. Each sentence opener is directly matched to one IPA category. Approaches such as AcademicTalk [11], Group Leader Tutor [12] and EPSILON [13] are examples of the use of this approach. Classifiers based on sentence openers can suffer from a high error rates. For example Israel and Aiken [12] report a classification error of 25%. They found that this high error rate can be attributed to the fact that the content of many interactions do not relate to the sentence opener selected by the users. For example, the sentence opener “Alternatively” followed by the text “ask the teacher because I am not sure about this solution” might not lead to a correct classification. To improve the classification precision, the most recent version of Group Leader Tutor [12] also considers a set of keywords that must be present in the free-text part of the interactions in order to confirm that the sentence opener was correctly selected. AcademicTalk uses opener sentences in combination with rules that specify the appropriate kind of answers. More recently, Mansilla et al. [14] also use sentence openers in a CSCL environment to detect collaborative problems. Students interact in a forum-like interface, limited by sentence openers to create new messages. Based on the analysis done by a multi-agent system, the teacher is alerted when problems are detected.

Among the positive aspects of sentence openers, we highlight the fact that they enable an easy identification of the communicative intentions, with no need to use any natural language processing technique. Additionally, users might be more critical in the discussions, since the sentence openers give “permission” to challenge others contributions, leaving aside courtesy aspects. On the other hand, the speech is restricted since users need to select from a predefined set of sentences to start each communication. Users might not find a sentence opener that suites the idea they want to express, giving place to the misuse of the sentences or the omission of the interaction.

Regarding classification of *raw* text interactions into IPA categories, the work by Zhang and Zhang [15] classified the content of news group messages (asynchronous media) into IPA categories with a one-phase classification schema using support vector machines (SVM). Text messages were preprocessed by removing stopwords and performing stemming. Special characters and emoticons were also considered as part of the messages. Beside the content of the message, authors considered the kind of message, the position in the news group and the replies received. Differently to our approach, the performance of the classifier was measured at the *reaction* level of IPA taxonomy, that is each sentence was classified into four categories: question, answer, positive and negative. Authors performed experiments on two datasets obtaining 85.1/86.6% precision for the category "question", 85.1/91.5% precision for the category "answer", 70.3/70.7% precision for the category "positive", and 95.9/100% precision for the category "negative". A similar approach was considered in [16] in a CSCW environment (synchronous media). Interactions were also preprocessed by removing stopwords and performing stemming. Authors trained and compared different classifiers (SVM, Decision Trees and Naive Bayes) to identify the *reaction* level of IPA taxonomy (four classes) and also the category level (twelve classes). Results of this research were improved later in [17] by improving the preprocessing of the text interactions. The main problem in classifying the interactions into only four categories corresponding to the reactions instead of the categories identified by IPA is that reactions are not enough to detect collaborative problems.

The major difference of the approaches described in this section with the approach presented in this article is that we propose a three-phases classification scheme to better identify the IPA category from raw text interactions. We show in Sections 4 and 5 that the proposed classification scheme outperform the one-phase scheme used in previous approaches to the problem.

3. Materials and methods

In this section we describe the materials and methods used in our study. We start by selecting a CSCW environment to perform our experiments (Section 3.1) and a well established method for interactions analysis (Section 3.2). Then in Section 3.3 we describe how these components are related in our approach.

3.1. CSCW environment

We chose Google Docs as the environment for developing the collaborative work. Google Docs is an online tool that can be freely used by anyone with internet access to write documents collaboratively [18]. This tool provides several advantages towards our objective: real-time editing, creation of comments and notes and the availability of a free text chat for the participants to discuss the task being performed.

Several studies have demonstrated that by using this collaborative tool can led to an enhancement in the performance of a collaborative work. Zhou et al. [19] compare the performance of different groups working on a collaborative task. One of the groups used Google Docs and the other did not. Authors concluded that the tool had a good reception from the students, who showed a general trend to adopt the tool once introduced. Another study [20] concluded that students wrote longer essays and were able to work more effectively in a collaborative writing assignment when they used Google docs in comparison with Microsoft Word™. Brodahl et al. [21] analyzed characteristics of students that used online writing tools, concluding that high-competitive students, with positive attitude towards digital tools got better results. A more recent study [22] provided empirical evidence of the actual use of Google Docs for collaborative writing. The authors of the study found large amounts of simultaneous work in the document, that the students freely edited each other' s entries, in some cases making major moves, deletions, and additions, and that more balanced participation was correlated with higher rated quality of the documents.

3.2. Interactions Process Analysis (IPA)

Bales' Interaction Process Analysis (IPA) is one of the most used methods in the study of small group interactions [23]. The purpose of the IPA method is to identify and record the nature (not the content) of each separate act in ongoing group interaction. IPA assigns a score to each unit of interaction or communication act. These units are typically made up from one simple sentence expressing one idea.

IPA assumes that all the interactions among members are observable and that they can be analyzed in relation to its impact or effect on the rest of the members or on the system. IPA enables to codify all the interactions, both those related to the task being performed (task dimension) and also the relational content of the verbal interactions (social-emotional dimension). Interactions must be observed by someone outside the group under study.

IPA codifies each interaction into two main categories (social-emotional and task-oriented) that are then sub-classified into twelve sub-categories, that represent the main behavior that can be observed in the task-solving process of a group: C1: shows solidarity, raises others status, gives help, reward; C2: shows tension release, jokes, laughs, shows satisfaction; C3: agrees, shows passive acceptance, understands, concurs, complies; C4: gives suggestion, direction, implying autonomy for other; C5: gives opinion, evaluation, analysis, expresses feeling, wish; C6: gives orientation, information, repeats, clarifies, confirms; C7: asks for orientation, information, repetition, confirmation; C8: asks for opinion, evaluation, analysis, expression of feeling; C9: asks for suggestion, direction, possible ways of action; C10: disagrees, shows passive rejection, formality, withholds help; C11: shows tension, asks for help, withdraws out of field; C12: shows antagonism, deflates others status,

Table 1
IPA categories and problem thresholds.

Area	Reaction	Category	Problem	Upper threshold	Lower threshold
Social-Emotional	Positive	C1	Reintegration	5%	0%
		C2	Tension reduction	14%	3%
		C3	Decision	20%	6%
Task	Answer	C4	Control	11%	4%
		C5	Evaluation	40%	21%
		C6	Communication	30%	14%
	Question	C7	Communication	11%	2%
		C8	Evaluation	9%	1%
		C9	Control	5%	0%
Social-Emotional	Negative	C10	Decision	13%	3%
		C11	Tension reduction	10%	1%
		C12	Reintegration	7%	0%

defends or asserts self. These categories can be also classified according to the type of reaction that is expressed: positive, answers, questions, negative.

When applying IPA, the interactions of each member of the team are tagged with one of the twelve categories to obtain information about the problems that emerged in the group, the members that mainly caused those problems, the members that tried to solve those problems, the main kind of solution adopted, how the approach to the solution evolved in the group, etc.

Bales also defined a set of successive phases that are common to many group dynamics during a collaborative work. He established that collaborative problems arise when there exist inappropriate number of the different types of interactions en each phase and defines different ranges between which a number of interactions of a given type can be considered appropriate (Table 1).

3.3. Automatic detection of IPA categories from free text interactions

In order to determine the presence of collaborative problems, each interaction must be associated with one of the possible IPA categories. Then, the total number of interactions for each category is computed, both for the whole group and for each participant. The percentage of each type of interaction manifested by each participant is contrasted with the ranges proposed by Bales to detect possible collaborative problems (see Table 1). In a typical scenario, the task of assigning a IPA category to each interaction is performed by a human expert. This task is extremely difficult and time consuming. For this reason, an automatic classification of the interactions is desirable.

Machine Learning is an interdisciplinary research area that takes concepts from several fields such as statistics, artificial intelligence, information theory, biology, cognitive sciences and control theory. Machine learning aims at building computational models automatically from the domain observations. Many algorithms have been developed that are effective for particular learning tasks, and that let us to solve such diverse and complex problems as speech recognition, face recognition, spam detection, credit fraud detection, medical diagnosis and products recommendation, among many others.

Machine learning can be classified into three categories, according to the information that we can give to the learning algorithm: supervised learning, unsupervised learning and semi-supervised learning. Supervised learning, also known as classification, assumes that data given to the learning algorithm have been manually labeled with the correct category. In other words, the class for each training example is known beforehand. Unsupervised learning, assumes that training data have not been labeled with any class and aims to group learning examples to discover such classes. Finally, semi-supervised learning works with few classified examples and many unclassified examples. Our work falls into supervised learning as a classification problem: given a free text sentence, we seek to assign the most probable IPA category it represents.

Classification is the problem of finding a model or function that describes and distinguish classes or concepts in order to be able to use this model to predict the class of new instances [24]. Basically, classification can be divided into two stages: training and prediction. During the first stage, the algorithm learns the model parameters using a training dataset, which consists of previously labeled instances. Then, in the second phase, the generated model is able to predict the class of new instances from the learned parameters. In our work, the classifier is built from the analysis of a set of free text sentences given as training data, which are manually labeled by an expert who is responsible of specifying the IPA category for each of the training instances. After the model is built, the classifier is expected to be able to predict the most probable IPA class of new free text interactions.

4. Experiments

In this Section, we first described the dataset collected to perform the experiments (Section 4.1). Then, in Section 4.2, we present our approach for the automatic classification of free text interactions into IPA categories. Finally in Section 4.3 we present the results obtained and compared them to a classical classification approach.

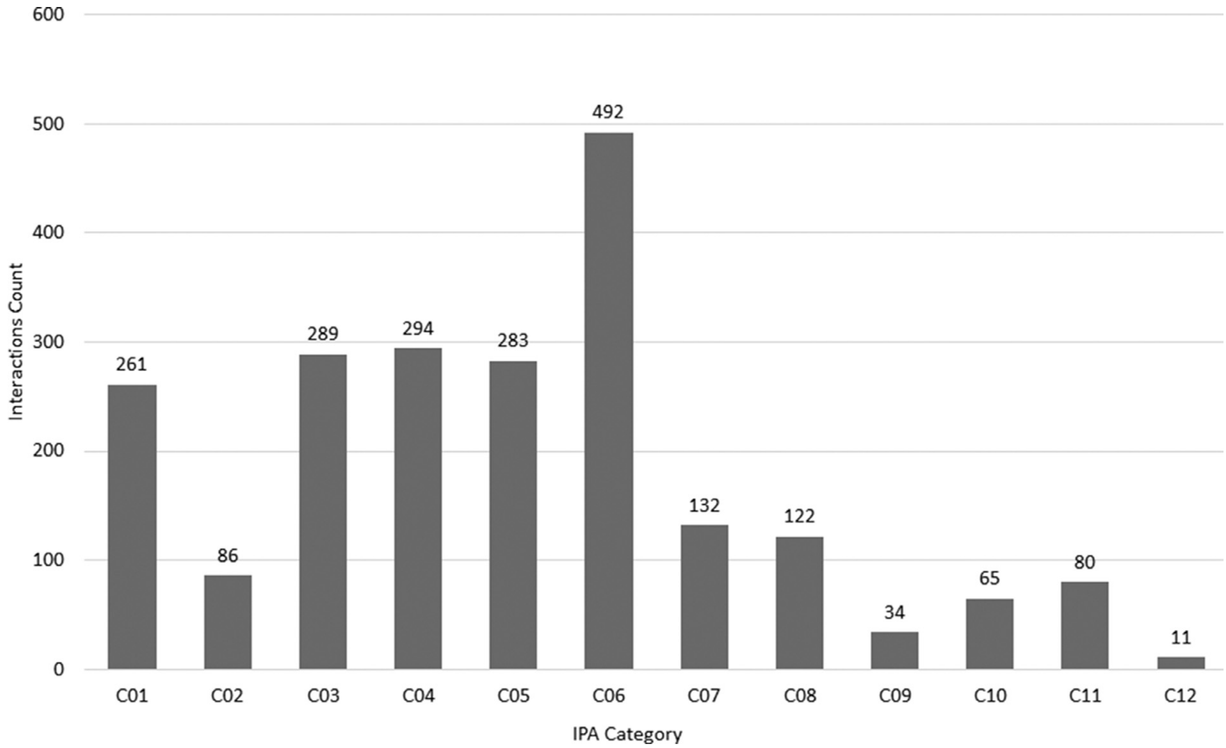


Fig. 1. Distribution of the interactions regarding the IPA categories.

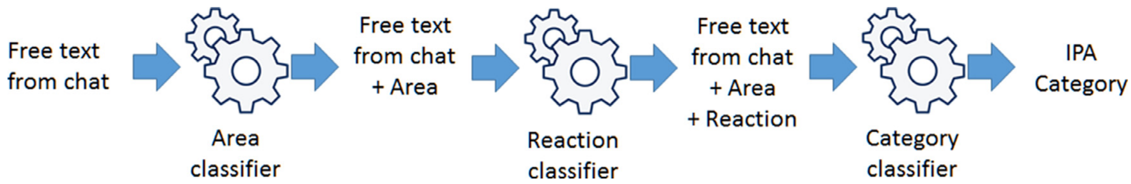


Fig. 2. Proposed three-phases classification approach.

4.1. Dataset

To perform the experiments we collected a set of interactions between students performing a collaborative assignment using Google Docs. Students belong to a system engineering course of the UNICEN University (Argentina). The 82 participants of the experiment were divided into 17 groups, with 5-6 members each. Each group were instructed to write, in a collaborative way, three documents with different assignments.

Data from the interaction between the students were collected and manually analyzed to assign the most related IPA category to each interaction. In this analysis step, invalid interactions and stop words were removed. We also applied stemming to reduce each word to its morphological root and performed a part-of-speech (POS) tagging to identify the syntactical function of each word in the interaction sentences. The total number of records in the dataset resulted in 2149 interactions. Fig. 1 shows the distribution of the interactions according to the 12 IPA categories.

4.2. Experiment process

Our aim is to build a classification model able to automatically classify the interactions between members of a group in the 12 IPA categories in order to reduce the high number of human-temporal resources needed to manually apply this method. We propose a classification approach divided in three phases. First, we train a classification model directly from the free text to identify the IPA area (social-emotional or task-oriented). In the second phase, we train a model with the text and the area detected in the previous phase to detect the type of reaction being expressed (positive, negative, question or answer). Finally, a third classifier is trained with the free text, the area, and type of reaction detected by the previous classifiers to assign one of the 12 IPA categories. Fig. 2 shows our classification approach.

Table 2
Accuracy (%) comparison between models.

Classifier	Our approach		Direct classification	
	Without PoST	With PoST	Without PoST	With PoST
NaiveBayes	57.88	54.07	33.47	33.89
Decision Tree	58.16	58.62	35.42	36.45
SVM	59.92	**60.67**	38.73	39.84

Table 3
Confusion matrix for our approach.

Classified as →	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
C1	**152**	11	98	0	0	0	0	0	0	0	0	0
C2	10	**68**	8	0	0	0	0	0	0	0	0	0
C3	44	3	**242**	0	0	0	0	0	0	0	0	0
C4	0	0	0	**144**	45	106	0	0	0	0	0	0
C5	0	0	0	62	**114**	108	0	0	0	0	0	0
C6	0	0	0	85	77	**330**	0	0	0	0	0	0
C7	0	0	0	0	0	0	**95**	31	6	0	0	0
C8	0	0	0	0	0	0	51	**59**	12	0	0	0
C9	0	0	0	0	0	0	15	13	**6**	0	0	0
C10	0	0	0	0	0	0	0	0	0	**47**	18	0
C11	0	0	0	0	0	0	0	0	0	32	**48**	0
C12	0	0	0	0	0	0	0	0	0	5	6	**0**
Precision (avg=0.60)	.74	.83	.69	.49	.48	.61	.59	.57	.25	.56	.67	0
Recall (avg=0.61)	.58	.79	.84	.49	.40	.67	.72	.48	.18	.72	.60	0
F-measure (avg=0.60)	.65	.81	.76	.49	.44	.64	.65	.52	.21	.63	.63	0
ROC area (avg=0.92)	.95	.97	.96	.87	.85	.88	.97	.96	.96	.98	.99	.97

Our hypothesis is that by performing the three-phases classification process we can improve the results obtained by a simple classifier that assigns one of the 12 IPA categories directly to the free text of each interaction.

We tested the performance of different classification algorithms using the classification approach proposed contrasting with a simple classification model: Decision Trees, Naive Bayes and Support Vector Machines. In this way, we seek to demonstrate that the proposed approach is able to obtain significant improvements compared to the direct application of the simple classification models.

4.3. Results

We compared the performance of the three classification algorithms both for the proposed three-phases approach and the state-of-the-art one-phase approach. We also compared the influence of the PoS tagging pre-processing method in the classification performance. Table 2 shows the results obtained.

We can observe that our three-phases approach obtained better results for all classification algorithms compared to the classical one-phase approach. The better classification accuracy was obtained by the three-phases SVM model with PoS tagging pre-processing. This model obtained an improvement of 20.83% with respect to the better classification results of the one-phase approach.

Tables 3 and 4 shows the confusion matrices for the best classification models of our approach and the classical approach. The main observation from these confusion matrices is that for our approach the miss-classified instances are always within the IPA reaction. For example, we can see that there are instances for categories C1, C2 and C3 that are not correctly classified, but the categories assigned by our approach are always C1, C2 and C3, which correspond to the "Positive" IPA reaction (R1). In other words, there is no cross-reaction miss-classifications. This fact does not occur for the single-phase classification approach, where we can observe that the instances that are not correctly classified by the model spread across all categories. We can also observe that the area under the ROC curve for our approach is 0.917, which indicates a really good classification model.

None of the approaches is able to identify instances with the class C12 (shows antagonism, deflates others status, defends or asserts self). This behavior can be attributed to the fact that there are only 0.5% (11 instances from 2149) in the dataset labeled with class C12.

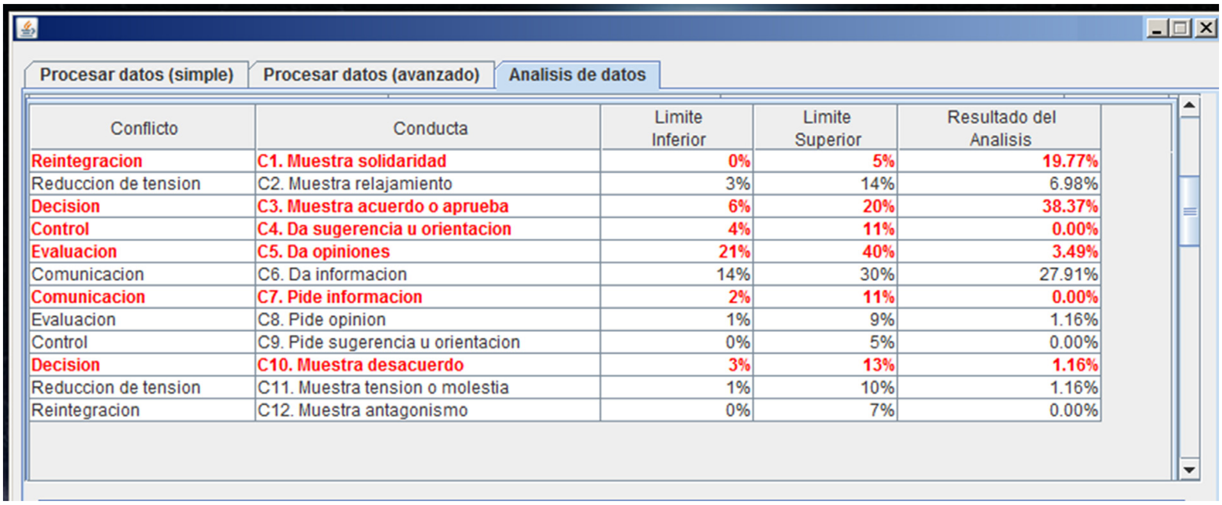
The experiments revealed that the proposed three-phases classification approach achieve very good levels of performance that significantly improve the classification methods used in previous approaches to the problem.

5. Application of the proposed approach to detect collaborative problems in a CSCW

The main goal of automating the process of classifying free text interactions into IPA categories of behavior is to assist the teacher or team leader in the detection of problems in the group dynamics. To this aim, we implemented a software

Table 4
Confusion matrix for the direct classification approach.

Classified as →	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
C1	**118**	8	72	15	11	26	5	3	0	1	2	0
C2	5	**67**	4	2	0	4	0	2	0	0	2	0
C3	26	3	**199**	16	19	16	1	8	1	0	0	0
C4	19	8	38	**106**	29	68	7	12	3	1	4	0
C5	14	3	44	40	**85**	64	13	15	0	3	3	0
C6	39	2	61	62	55	**228**	20	13	1	3	8	0
C7	16	2	26	15	11	37	**15**	5	2	1	1	1
C8	10	1	23	24	16	18	3	**26**	0	1	0	0
C9	6	0	3	6	3	4	3	5	**2**	2	0	0
C10	3	6	22	3	14	10	1	1	0	**3**	2	0
C11	5	14	16	9	5	16	2	3	0	2	**8**	0
C12	1	0	2	2	3	2	1	0	0	0	0	**0**
Precision (avg=0.38)	.45	.59	.39	.35	.34	.46	.21	.28	.22	.18	.27	0
Recall (avg=0.40)	.45	.78	.69	.36	.30	.46	.11	.21	.06	.05	.10	0
F-measure (avg=0.38)	.45	.67	.50	.36	.32	.46	.15	.24	.09	.07	.14	0
ROC area (avg=0.74)	.78	.90	.81	.75	.68	.71	.71	.68	.76	.74	.69	.62



Conflicto	Conducta	Limite Inferior	Limite Superior	Resultado del Analisis
Reintegracion	C1. Muestra solidaridad	0%	5%	19.77%
Reduccion de tension	C2. Muestra relajamiento	3%	14%	6.98%
Decision	C3. Muestra acuerdo o aprueba	6%	20%	38.37%
Control	C4. Da sugerencia u orientacion	4%	11%	0.00%
Evaluacion	C5. Da opiniones	21%	40%	3.49%
Comunicacion	C6. Da informacion	14%	30%	27.91%
Comunicacion	C7. Pide informacion	2%	11%	0.00%
Evaluacion	C8. Pide opinion	1%	9%	1.16%
Control	C9. Pide sugerencia u orientacion	0%	5%	0.00%
Decision	C10. Muestra desacuerdo	3%	13%	1.16%
Reduccion de tension	C11. Muestra tension o molestia	1%	10%	1.16%
Reintegracion	C12. Muestra antagonismo	0%	7%	0.00%

Fig. 3. GUI presented to the teacher highlighting the collaborative problems in the group.

assistant that use the classification model trained with the interactions captured in the first experiment to detect collaborative problems in the group dynamics. While the participants of a group work towards a common goal using a CSCW environment, the software assistant observes the textual interactions and classifies them into IPA categories by using the classification model trained in the first experiment. According to this classification, the assistants checks the upper and lower limits defined for each type of interaction described in Table 1. If the number of interactions is not within the thresholds recommended by Bales [23], a possible problem is highlighted to the teacher or team leader. Fig. 3 shows a screenshot of the assistant, as shown to the teacher/team leader. Problems detected by the assistant are highlighted in red to alert the teacher/team leader.

We performed a second experiment in a real context with the aim of comparing the number of problems detected by our approach with respect to the number of problems detected by a manual analysis of the interactions. In this second experiment, 82 students were divided into 12 groups of 6-7 members each to perform a first assignment. The same group of students were assigned to 12 different groups (with 6-7 members) to perform a second assignment. In total, we collected the interactions of 24 groups working in two assignments. Table 5 shows the number of problems of each type detected by our automatic classification approach and by a manual analysis of the interactions of the group. We can see that although the proposed approach achieves a classification accuracy of 60.67%, the number of problems detected by our approach is a good approximation to the manual analysis. Our approach was able to detect 80.49% of the existent problems (recall), with a false recovery rate of 16.1%. The false recovery rate implies that from the problems automatically detected by our approach, only 16.1% are false alarms, and 83.9% are real problems.

A further analysis of each type of problem detected showed that, in most groups:

- Reintegration problems were detected by high number of manifestations of C1 behavior (shows solidarity)
- Tension reduction problems arose due to low number of C11 behavior (shows tension)

Please cite this article as: F.D. Berdun et al., Classification of collaborative behavior from free text interactions, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.07.015

Table 5
Number of problems detected by the automatic approach with respect to the manual analysis.

Problem	Manual analysis	Problems detected	False positives
Reintegration	22	22	2
Tension reduction	20	13	3
Decision	22	22	2
Control	21	7	2
Evaluation	16	16	8
Communication	22	19	2
TOTAL	123	99	19

- Decision problems were consequence of a high number of C3 behavior (shows agreement) or low number of C10 behavior (express disagreement)
- Control problems were highlighted due to low number of observations of C4 behavior (asking suggestions)
- Evaluation problems were detected due to low number of observations of C5 behavior (giving opinions)
- Communication problems reflect the fact of high number of observations of C6 behavior (giving information)

From this second experiment, we can conclude that the students tended to show “good” behavior while they work together to complete the assignments. This is in concordance with the conclusions presented by Fahy [25], who also used IPA to compare groups working face to face with groups who interact by using and online platform. Results of Fahy’s experiments showed that online groups express less disagreement and in general less negative behavior face to face groups. The analysis of the problems detected in our experiment revealed a similar behavior in our groups of study.

6. Discussion and conclusions

In this article we presented a classification approach to identify automatically the Interaction Process Analysis (IPA) collaborative categories from free text interactions. This is an extremely difficult task for two main reasons: 1) the high number of classes into which an interaction can be classified, and 2) differently from other approaches that use predefined “opening sentences” that participants have to use to begin an interaction, we let participants to interact with no limitations using a chat window. Although a classification accuracy of 60.67% might seem low at first sight, it is 20.83% higher than the accuracy obtained by direct classification and 52.34% higher than a random classifier. Furthermore, since the incorrectly classified instances with our approach are always classified within the correct IPA reaction, a recommender system can be built to assist the teacher by limiting the set of possible categories he/she has to select to tag the interactions to speed up the time required to apply the IPA method. Finally, our approach is useful to detect collaborative problems arisen in a CSCW environment with a high ratio of accuracy.

As a future work, we will consider the collaborative problems detected with the approach presented in this article to investigate the constitution of better groups in order to improve the collaborative work experience.

References

[1] Lamb B. Wide open spaces: wikis, ready or not. EDUCAUSE Rev 2004;39(5):36–49.
[2] Goodhart GS, Blair R. Beyond crossroads: implementing mathematics standards in the first two years of college. College Math J 2008;39(4):324–6.
[3] Bottge BA, Rueda E, Kwon JM, Grant T, LaRoque P. Assessing and tracking students’ problem solving performances in anchored learning environments. Educ Technol Res Dev 2009;57(4):529–52.
[4] Casamayor A, Amandi A, Campo M. Intelligent assistance for teachers in collaborative e-learning environments. Comput Educ 2009;53(4):1147–54.
[5] Costaguta R, Garcia P, Amandi A. Using agents for training students collaborative skills. IEEE Lat Am Trans 2011;9(7):1118–24.
[6] Birnholtz J, Steinhardt S, Pavese A. Write here, write now!: an experimental study of group maintenance in collaborative writing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’13. New York, NY, USA: ACM; 2013. p. 961–70. ISBN 978-1-4503-1899-0
[7] Savolainen R. Asking and sharing information in the blogosphere: the case of slimming blogs. Libr Inf Sci Res 2011;33(1):73–9. ISBN 0740-8188
[8] Kim MS, Kim YS, Kim TH. Analysis of team interaction and team creativity of student design teams based on personal creativity modes. In: ASME 2007 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers; 2007. p. 55–67.
[9] Löfstrand P, Zakrisson I. Competitive versus non-competitive goals in group decision-making. Small Group Res 2014;45(4):451–64.
[10] Delic A, Neidhardt J, Nguyen TN, Ricci F. Research methods for group recommender systems. Proc RecTour 2016.
[11] Mcalister S, Ravenscroft A, Scanlon E. Combining interaction and context design to support collaborative argumentation using a tool for synchronous CMC. J Comput Assist Learn 2004;20(3):194–204. ISSN:1365-2729
[12] Israel J, Aiken R. Supporting collaborative learning with an intelligent web-based system. Int J Artif Intell Educ 2007;17(1):3–40.
[13] Goodman BA, Linton FN, Gaimari RD, Hitzeman JM, Ross HJ, Zarrella G. Using dialogue features to predict trouble during collaborative learning. User Model User-Adapt Interact 2005;15(1-2):85–134.
[14] Mansilla PS, Costaguta R, Schiaffino S. Multi agent model for skills training of CSCL e-tutors: modelo multi agente para el entrenamiento de habilidades de e-tutores de ACSC. In: Proceedings of the 7th Euro American Conference on Telematics and Information Systems. ACM; 2014. p. 30.
[15] Zhang C, Zhang C. Discovering users’ participant roles in virtual communities with the help of social interaction theories. PACIS 2005 Proc 2005:65.
[16] Cincunegui M, Berdun F, Armentano MG, Amandi A. Clasificación de conductas colaborativas a partir de interacciones textuales. Argentine Symposium on Artificial Intelligence (ASAI 2015)-JAIIO 44 (Rosario, 2015); 2015.
[17] Berdun F, Armentano MG, Amandi A. Inferencia de roles de equipo a partir de conductas colaborativas detectadas en interacciones textuales. Simposio Argentino de Inteligencia Artificial (ASAI 2016)-JAIIO 45 (Tres de Febrero, 2016); 2016.

Please cite this article as: F.D. Berdun et al., Classification of collaborative behavior from free text interactions, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.07.015

[18] Oishi L. Working together: google apps goes to school. Technol Learn 2007;27(9):46.
[19] Zhou W, Simpson E, Domizi DP. Google docs in an out-of-class collaborative writing activity. Int J Teaching Learn Higher Educ 2012;24(3):359–75.
[20] Apple K, Reis-Bergan M, Adams A, Saunders G. Online tools to promote student collaboration. In: Getting connected: Best practices for technology enhanced teaching and learning in high education; 2011. p. 239–52.
[21] Brodahl C, Hadjerrouit S, Hansen NK. Collaborative writing with web 2.0 technologies: education students' perceptions. J Inf Technol Educ 2011;10:73–103.
[22] Olson JS, Wang D, Olson GM, Zhang J. How people write together now: beginning the investigation with advanced undergraduates in a project course. ACM Trans Comput-Hum Interact 2017;24(1) 4:1–4:40. ISSN 1073-0516.
[23] Bales RF. Interaction process analysis; a method for the study of small groups. Addison-Wesley; 1950.
[24] Han J, Pei J, Kamber M. Data mining: concepts and techniques. Elsevier; 2011.
[25] Fahy PJ. Online and face-to-face group interaction processes compared using bales' interaction process analysis (IPA). Eur J Open, Distance, E-learn 2016;9(1).

Franco D. Berdun is a PhD student at UNICEN University, with a research grant of CONICET (National Council for Scientific and Technological Research). He received a System Engineer degree from UNICEN University, Argentina. His research interests include user modeling, collaborative learning and serious games. Contact him at franco.berdun@isistan.unicen.edu.ar.

Marcelo G. Armentano is a researcher in ISISTAN Research Institute (CONICET-UNICEN), Argentina and a professor at UNICEN University, Argentina. He received his PhD in Computer Sciences in 2008 and a Master in Systems Engineering in 2006 from the UNICEN. His research interests include user modeling, recommender systems and social network analysis. Contact him at marcelo.armentano@isistan.unicen.edu.ar.

Luis Berdun is a researcher in ISISTAN Research Institute (CONICET-UNICEN), Argentina and a professor at UNICEN University, Argentina. He received a PhD degree in Computer Science in 2009 and a Master in Systems Engineering in 2005 from the UNICEN. His research interests include intelligent aided software engineering, planning algorithms, Knowledge Management. Contact him at luis.berdun@isistan.unicen.edu.ar.

Martin Mineo is a System Engineering student at UNICEN University. He obtained a Analyst Programmer degree at UNICEN in 2009. His research interests include machine learning, user modeling and CSCW. Contact him at mmineo@alumnos.exa.unicen.edu.ar