Analysing student interaction processes in order to improve collaboration. The DEGREE approach

Email {bbarros,felisa}@ieec.uned.es

Abstract: Computer mediated collaborative learning allows the recording of a large amount of data about the interaction processes and the task performance of a group of students. This empirical data is a very rich source to mine for a variety of purposes. Some purposes are of practical nature like, for instance, the improvement of peer awareness on the on-going work. Other purposes are of a more long-term and fundamental scope such as to understand socio-cognitive correlations between collaboration and learning. Manual approaches to fully monitor and exploit these data are out of the question. A mixture of computational methods to organise and extract information from all this rough material together with partial and focused in-depth manual analysis seems a more feasible and scalable framework. In this paper we present an approach to characterise group and individual behaviour in computer-supported collaborative work in terms of a set of attributes. In this way a process-oriented qualitative description of a mediated group activity is given from three perspectives: (i) a group performance in reference to other groups, (ii) each member in reference to other members of the group, and (iii) the group by itself. In our approach collaboration is conversation-based. Then we propose a method to automatically compute these attributes for processes where joint activity and interactions are carried out by means of semi-structured messages. The final set of attributes has been fixed through an extensive period of iterative design and experimentation. Our design approach allows us to extract relevant information at different levels of abstraction. Visualization and global behavior analysis tools are discussed. Shallow analyses as presented in this paper are needed and useful to tackle with a large amount of information, in order to enhance computer-mediated support.

INTRODUCTION

Collaborative learning research has paid close attention to the study of pupils interactions during peer-based work in order to analyse and identify the cognitive advantages of joint activity (Dillenbourg, Baker, Blaye, & O’Malley, 1996). As Crook (1996) points out, the benefit of the collaborative approach for learning lies in the processes of articulation, conflict and co-construction of ideas occurring when working closely with a peer. Participants in a problem-solving situation have to make their ideas explicit (assertions, hypothesis, denials..) to other collaborators, disagreements prompt justifications and negotiations, helping students to converge to a common object of shared understanding.

The computer provides opportunities to support and enhance this approach in a number of ways, for instance offering computer-based problem spaces for jointly creating and exploiting structures of common knowledge and shared reference. Moreover, networks make possible opening the collaborative framework to distributed communities providing remote access to these spaces as well as computer-mediated communication to support interpersonal exchange and debate. An increasing number of collaborative learning environments for open and closed virtual groups have been built for a range of learning tasks (Scardamalia & Bereiter, 1994) (Edelson, Pea & Gomez, 1996 ) (Wan & Johnson, 1994) (Suthers & Jones, 1997), and
experiences of use are reported from school to university level (Bell & Davies, 1996), (Collis, 1998).

**Figure 1.** Representation of an activity as the Activity Theory

The socio-cultural theory proposes the Activity Theory (Nardi, 1996) for representing the activities of groups where Technology plays a role as mediator. Within this theory, an analysis model was developed for identifying and representing the human and artificial elements involved in joint tasks (Engeström, 1987). This socio-cultural framework provides the concept of *activity* as a unit of analysis, with a rich internal structure to make the context of a situation explicit, specially the interlinks between the individual and social levels stressing the role of the tools as mediating artifacts. Within an activity different related elements are represented (figure 1): the community involved and the social norms that govern it, the division of labour to be followed, the tools to be used for working, the subject of the activity, the object of the activity and, finally, the outcome produced by the group. An important mediational tool is language, and in particular for learning tasks, tutor student dialogue as well as peer dialogue.

Studies on educational dialogue have been mainly concerned with tutorial dialogue either for descriptive or prescriptive purposes. A number of proposals have paid attention to analyse natural language dialogues in collaborative learning situations, applying or extending NL discourse models (Katz et al. 1999, Ravenscroft et al. 1999, Pilkington et al. 1999). Usually the framework includes the computer playing the role of a facilitating tutor, with students performing a learning task. Students can express themselves either by means of (restricted) natural language utterances, using a semi-formal interface based on a rich predefined interaction language or a representation specially designed for reifying a particular type of learning situation such as knowledge negotiation, argumentation or explanation.

In this paper, we focus on the analysis of peer computer mediated interaction in collaborative learning processes. Although there is an increasing amount of work on this subject, analysis studies have been performed mainly manually (see for instance (Dillenbourg & Baker, 1996)). Most of the automatic analysis of computer mediated peer interactions have been up to now based on quantitative terms, with a few attempts to carry out automatic natural language processing (Henri, 1992).

In our approach, collaboration is conversation-based. Conversation consists of turn taking, where each contribution both specifies and grounds some content (Bobrow, 1991). The type of contributions and their constraints can be defined to establish an explicit structure for a conversation. For learners the benefit is twofold (1) they receive some support for the process of categorising and organising their ideas when contributing to the debate and (2) further inspection of the collaborative process is facilitated because the system can take into account the type of units to automatically structure students’ discussions. Thus, we propose to use semi-structured interfaces for our collaborative learning tasks, so that students have to select a communication mode when producing a contribution. The system is domain independent and the conversational structure for each task is defined in a declarative way.
Analysing student interaction processes in order to improve collaboration

Based on a generic conversational representation we have defined a model to characterise in qualitative terms a group interaction process. Collaboration is described in terms of a set of attribute-value pairs. These values are dynamically computed so that an analysis of the student interactions can be automatically performed. Furthermore, our system includes an advisor to generate feedback alerting students and tutors of potential problems and/or suggesting further actions.

The following section gives an overview of our system architecture. The configuration level is further detailed in section 3 to provide the background for the analysis level presented in section 4. The advisor module is described in section 5. We finish with a summary and some conclusions.

THE COLLABORATIVE ENVIRONMENT

Our system is based on some basic principles related to the CSCL (Computer Supported Collaborative Learning) paradigm (Koschmann, 1996). These principles are (1) joint construction of a problem solution, (2) coordination of group members for planning the tasks, (3) semi-structuration of the interaction mechanisms (4) and focus on both the learning process and the learning result, and therefore, explicit representation of the production and interaction processes.

![Figure 2. The four levels of the architecture of DEGREE](image)

The architecture of the system, named DEGREE acronym of Distance Environment for GRowup ExperiencEs, is organised into four levels: configuration, performance, analysis and organisation level (Barros, 1999) as shown in figure 2:

- **Configuration level.** Once the teacher(s) have planned an experience of collaborative learning, on this level they can configure and install automatically the environment needed to support the activities of groups of students working together. The environment will provide the resources needed for carrying out joint tasks. In the configuration level teachers specify tasks, resources and groups, either by starting from the scratch or reusing generic components.

- **Performance level.** This is the level where a group of students can carry out collaborative activities with the support of the system. Activities may involve a variety of tasks with associated shared workspaces. Collaboration is conversation-based. The system manages the users interventions, named *contributions*, supporting the co-construction of a solution in a collaborative argumentative discussion process. All the events related to each group and experience are recorded. They can be analysed and reused for different purposes in the analysis and organisation levels.
Barros and Verdejo

- **Analysis Level.** In this level we analyse the user's interaction and make interventions in order to improve them. We offer tools for quantitative and qualitative analysis for observing and analysing the process of solving a task in the performance level. In the analysis level we propose a way of observing and value the users attitudes when they are working together. We offer the possibility of intervention by sending messages to the group or to individuals explaining how to improve different points of their work. Finally, we register the messages and the moment when we make this intervention and analyse the improvements.

- **Organisation level.** Here we gather, select and store the results of collaborative learning experiences and the processes. The information is structured and valued for searching and reusing purposes. This information is stored as cases forming an Organisational Learning Memory. We offer functionalities for defining, searching, collecting different cases, and for defining links to work material in the configuration level for related tasks. For more information about this level, consult (Verdejo & Barros, 1998).

By means of the Activity Theory we globally describe the activities supported by each level in figure 3. A brief explanation is given below:

- **CONFIGURING AN EXPERIENCE:** One or more teachers (the community) starting from some learning goals (the object) will proceed to configure a collaborative learning experience (the outcome). For doing so, they will use the configuration level subsystem, which supports the definition of the experience's components. Cases from prior experiences could be seen also as mediational tools. An experience is defined through a four-step process, fully supported by the configuration system: group definition (roles are specified according to the collaboration model and the learning tasks to perform), conversational structure definition, the shared work spaces definition and, finally, the activities integrating the experience. Then the defined experience and the supporting environment will be mediational tools for other system activities. More specifically, the latter for the "Carrying out an experience" (labelled as G128 in figure 3) and the former for the "Evaluating an experience" (labelled as G243 in figure 3) activities.

- **CARRYING OUT AN EXPERIENCE:** A students' group possibly involving a teachers' group (the community) carries out an experience using a tailored computer framework for distance collaborative work (a mediational tool) to reach a set of common learning goals. An experience is organised into activities which, in turn, can be divided into tasks and subtasks (work division). Students agree to participate and share the responsibility of the process and its results assuming a particular role and operating with the conversational rules stated for argumentative discussion. A task is considered as finished when there is a jointly elaborated and agreed document. The experience's outcome includes a final document for each task as well as a representation of the elaboration process. This outcome will be the object of other activities: "Managing cases" (labelled as G236 in figure 3) and "Analysing an experience" (labelled as G246 in figure 3).

- **ANALYSING AN EXPERIENCE.** This activity usually involves teachers looking at student actions and interactions while performing a learning experience (community). Tools used to carry out this activity are, on the one hand the DEGREE analysis environment, providing graphic tools and inference methods, and on the other, all the data about the experience definition included in the configuration level. This activity comprises analysing, advising and feedback monitoring. The analysis task is made automatically by applying qualitative evaluation rules, while advise and monitoring are supported by the system but performed under teacher control. The object of the analysis is the recorded data of an experience including all the information about actions and interactions performed by participants. The outcome of this activity will be used for defining cases for the Organisational Memory (shown as G250 in figure 3). It is also used as feedback for the on-going experiences in order to improve students' attitudes (labelled as G247 in figure 3).
Analysing student interaction processes in order to improve collaboration

- **MANAGING CASES.** Teachers (the community), create and add cases from the experiences carried out with the system. For doing so, they use a DEGREE tool allowing them to visualise, select, annotate and assess the results and processes to create new cases. The outcome is a library of cases, which could be used as a tool for defining new experiences (labelled as ② in figure 3).

![Diagram of activities supported by DEGREE](image-url)

**Figure 3.** Activities supported by DEGREE
MODELLING THE EXPERIENCE COMPONENTS IN THE CONFIGURATION LEVEL

The DEGREE configuration level offers tools for defining the components to support a collaborative experience and installing a working environment for one or more groups. The outcome of the configuration activity is an application allowing users to carry out the defined learning activities through asynchronous communication via Internet. The configuration level is also distributed and lets several users (usually teachers) to define an experience.

As we said before each experience can be organised as a set of activities. An activity is designed to reach some pedagogical objectives and in turn, can be structured in a set of tasks.

Each collaborative learning task is supported by a workspace, defined by the following elements: (1) roles as profiles for group of users (without identifying them explicitly); this is a way to specify the division of labour for a particular task (2) the outcome including the type of document to create, structured in topics; (3) the mediational tools for instance the available multimedia material (papers and other sources to consult) and the interaction mode (4) the rules for carrying out the task: type of dialogue processes to achieve the subtasks, for example argumentative coedition finished with explicit agreement, (5) and the relation to the other tasks, if any, specified in terms of temporal restrictions.

Our underlying model for a conversational structure is a labelled oriented graph. This mechanism allows structuring a group conversation in a generic way. The nodes represent contribution types and the edges represent the allowed links between the different types of contributions while a conversation is going on. More formal models for communication are proposed in the Distributed Agents stream (Cohen and Levesque 1995, Finin et al. 1994). However, for our framework, a categorisation of types and conversational moves are enough to support the communication requirements. The content is to be interpreted by the participants. A semi-formal approach allows us to achieve a generic intermediate level, both useful and flexible to accommodate as much or as little structure as needed for a particular kind of application. Some works have been recently reported on semi-structured communication for student interaction in collaborative learning. Task-related categories and their effect on directing student focus has been discussed by Baker & Lund (1996). More generic conversational categories for collaborative learning are proposed by Soller et al. (1999). They present a sentence opener-based interface and some conclusions derived from a preliminary evaluation study. We very much agree with the idea of exploiting this knowledge for enabling an ICLS system to dynamically analyse group activity and support group interaction. In fact we have presented in (Barros & Verdejo 1999) a qualitative automatic analysis of learners’ interactions, in this paper we describe how our system has been extended to provide feedback to learners, either automatically or under teachers’ control.

Let us look at an example. Suppose we want to define a conversation structure for argumentative coediting as interplay of proposals/contraproposals between peers, refining ideas until agreement is reached. The conversational structure shown in figure 4 captures our conversation expectations for such a kind of collaborative task, where the node labels stand for the categorisation that is shown in table 1.

![Figure 4](image_url). An example of conversational structure.
Analysing student interaction processes in order to improve collaboration

**Table 1.** Interpretation of each contribution type of the conversational structure in figure 4.

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>P: PROPOSAL &lt;text&gt;</td>
<td>initial conversational-unit, a text is proposed opening the co-authoring debate related to a task/subtask</td>
</tr>
<tr>
<td>CN: CONTRAPROPOSAL &lt;ref&gt; &lt;text&gt;</td>
<td>intermediate conversational-unit for arguing to a proposal by means of an alternative text. The result is another proposal that can be totally new or an extension to the proposal P (&lt;ref&gt; is P)</td>
</tr>
<tr>
<td>CO: COMMENT &lt;ref&gt; &lt;text&gt;</td>
<td>intermediate conversational-unit for commenting another conversational-unit (&lt;ref&gt; can be P, CN, Q, CO, CL)</td>
</tr>
<tr>
<td>CL: CLARIFICATION &lt;ref&gt; &lt;text&gt;</td>
<td>intermediate conversational-unit for answering a question or for explaining something (&lt;ref&gt; can be P, CN or Q)</td>
</tr>
<tr>
<td>Q: QUESTION &lt;ref&gt; &lt;text&gt;</td>
<td>intermediate conversational-unit for asking for more information about another conversational-unit (&lt;ref&gt; can be P, CN, CO or CL)</td>
</tr>
<tr>
<td>A: AGREE &lt;ref&gt;</td>
<td>final conversational-unit to agree upon (&lt;ref&gt; can be P or CN)</td>
</tr>
</tbody>
</table>

Contribution types and conversational structure are defined in the configuration level, and therefore can be different for the various tasks and activities. Figure 5 shows another example for a different collaborative task, where agreement was not required.

![Diagram](image)

**Figure 5.** Second example of conversational structure.

The learner interface provided by the system for the elaboration space (corresponding to argumentative coedition as defined in figure 4) is shown in figure 6. This interface is organised in three areas. The upper area is a menu for accessing activity workspaces. In the left area, we can see the task schema, which has been automatically updated by the system. On the right side, after clicking on one of the contribution names, appears the whole content of the contribution and below the list of contributions related with this one according to the conversational structure of the task. Underneath there is a form for replying to the contribution. Labels on the window explain each element of the interface.
Figure 6. Interfaces of the elaboration space of the collaborative environment of the Performance level.

In our approach, each contribution type indicates a position in the conversation. So, in order to analyse the group interaction, the conversational graph specification includes the definition of values for a set of attributes for each type of contribution. We have proposed four attributes taking values in a range of (-10,10). This numerical range is used for internal computation. As we will see later (figures 11 and 12), all these values are mapped to a reduced qualitative scale for information purposes.

- **Initiative**, indicates the degree of involvement and responsibility required to produce a contribution.
- **Creativity**, relates to the degree of originality required to produce this type of contribution
- **Elaboration**, qualifies the workload needed for making a contribution.
- **Conformity**, establishes the degree of agreement of a contribution in relation to another selected and linked contribution. For instance making a contraproposal indicates a low degree of conformity to the proposal.

Table 2 shows the feature structure associated to the conversational graph presented in figure 4.
Analysing student interaction processes in order to improve collaboration

Table 2. Conversational graph and a table with the attributes and values for each contribution type

<table>
<thead>
<tr>
<th>CONTRIBUTION</th>
<th>initiative</th>
<th>creativity</th>
<th>elaboration</th>
<th>conformity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(J=(1,k))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proposal (P)</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>-10</td>
</tr>
<tr>
<td>Contraproposal (CN)</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>-10</td>
</tr>
<tr>
<td>Question (Q)</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Comment (CO)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Clarification (CL)</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Agreement (A)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>

ANALYSING AND SUPERVISING THE GROUP INTERACTION IN THE ANALYSIS LEVEL

The system records all the accesses and the actions performed by users when they are solving the tasks. The information automatically registered includes: the user identification, the time and date, the host computer, the learning experience, the group, the activity, the task, the workspace, and the type of action. A relational database stores this information so a variety of queries combining a selection of criteria can be easily formulated. With these facilities for accessing and manipulating the information, the system carries out an analysis in three modes. Figure 7 presents the analysis modules and their relation with the configuration and performance levels.

![Figure 7](image-url)

The analysis modes are the following:
- **Quantitative analysis.** Queries to this module cover a variety of parameters, the output can be represented in textual or graphical mode.
- **Qualitative analysis** considers the experience configuration, the process and the result, plus subjective knowledge from the observer about the interpretation of the observed values. The analysis is performed in three dimensions: globally comparing a group with other groups, an individual comparison of each student with his colleagues, and a group task analysis. We also consider the possibility of providing feedback by sending messages about the analysis results, highlighting what can be improved in the group activity.
- **Feedback analysis** uses the output of the advisory system and graphically represents the evolution of the group for a selected subset of the observed variables.
Quantitative analysis

A web-based interface is provided for selecting different and combined options for a query, which can be shown in textual or graphical mode. It is possible to visualise user interaction in different ways, for example:

- Evolution of the number of user contributions in an experience during a period, using a graphical display.
- Plot the number of hourly accesses for a group in an activity.
- Number of contributions by user, for all the group members in an activity.
- Number of contributions by user, by type, for each substask of a task.
- Contributions of a group, related to a workspace task, by type of contribution, displayed as a chart bar. Figure 8 shows an output example. The discussion process is summarised in term of the type units, in this case proposal, contraproposal and agreement have been the backbone of the process. Only a few comments, questions and clarifications have been performed. This graphic together with the previous question provides a good overview of the dynamics of a groupwork.
- Number of contributions, by member, for an activity (all the workspaces supporting the activity).
- Evolution of a discussion for a substask, graphically displayed (figure 9). We can observe here a first proposal followed by a contraproposal and a comment from the same author. A turn taking happens with a peer contraproposal and then two contraproposals plus a comment from the first author, a question from his peer follows, and after some time without receiving a response this student takes the initiative, making a new contraproposal. In this way they finally reach an agreement.

Figure 8. Number of contributions added by a group during the development of a learning task

Qualitative Analysis

Criteria and methods to evaluate whether a collaborative learning process has been carried out is a controversial and open question in the field. But at least, from a practical point of view, we need to identify, even roughly, if and when students have been addressing each other and working together. For this purpose we can exploit some of the evidence provided by the student contributions while performing a task.

We propose three kinds of analysis to characterise:

- Group behaviour compared to other groups performing the same task.
Analysing student interaction processes in order to improve collaboration

- Group behaviour in itself
- Individual student behaviour compared to the rest of the group members

**Figure 9.** Evolution of discussion for a subtask beginning at the top left (X is time evolution and Y is turn-taking)

Data for the process analysis of a learning experience comes from two sources (see figure 7):
- The experience configuration: task, conversation structure, roles and groups
- The learners’ contributions organised by the system as tree-like structures for each task workspace, and the set of messages interchanged through the coordination space associated to each activity.

We will express the result of the analysis in terms of a set of features (attribute-value pairs), subjectively established but tested and refined through an extensive experimental period of system use. Details about the number of students involved and the formative evaluation carried out are given in (Verdejo & Barros, 1999).

**Analysing group behaviour**

Figure 10 shows the attributes considered for the comparison of group behaviour. Values for these attributes are either (i) calculated from data about the task definition and the process performed or (ii) concluded from fuzzy inference using a set of evaluation rules, for those attributes appearing with a dark background in figure 10. Next we will describe these features in detail.

From this data we consider and compute the following attributes:
- For each elaboration workspace
  - *Mean number of contributions*: number of contributions from the group/number of group members
  - *Mean size of contributions*: mean size of contribution contents (in characters)
  - *Depth of the Contribution’s tree*: Maximum depth of the trees related to the workspace
  - *Interactivity*: Percentage of contributions responded or linked to other contributions made by a student, other than the contributor.

The values $V_{ai}$ for each $i$ attribute: Initiative, Creativity, Elaboration and Conformity are computed by the formula $V_{ai} = \sum_{j=1}^{k} (N_j \times V_{ij})$ where $N_j$ is the number of contributions for each type $j$, and $V_{ij}$ the value of the attribute for that type of contribution as is defined in table 2.

- For the coordination space, where off-task comments are produced, we consider at the moment the total number of messages captured by the variable “coordination messages”.

---

**Diagram:**

```
```

```
```

```
```

```
```

```
```

```
```
Barros and Verdejo

In the qualitative analysis the results are inferred by fuzzy inference methods. The knowledge base is composed of the definition of the attributes of the sets and the rules relating them.

Each attribute has been modelled as a set of linguistic labels. Each of these linguistic labels corresponds to a possible value for this attribute in a range of possible values. For example, for the variable $M_{Contributions\,Number}$ we have considered three linguistic labels little, appropriate and much, defining for each one the appropriate range values.

As the scales have to be comparable and adjusted to a task, the range defining each attribute is not the same for all the tasks, but calculated for each task related to a mean reference point (MRP). For instance, for a normal task 50 contributions is a typical number of contributions, so 55 could be in the "adequate" range linguistic label value for a normal task, although for a long task this could be in the "low" range. Mean reference points can be dynamically calculated taking into account all the similar tasks performed in all the learning experiences carried out with the system.

The rules are generated by combining the attributes. They consider all the possible combinations of attributes, although it is not necessary to generate all of them. It is feasible to generate rules that include others, considering fewer attributes in the antecedent. Figure 10 shows a way of combining the attributes for generating new ones (with a dark background).

![Figure 10. Relationship among attributes used for making the group behaviour analysis](image)

The logical product of each rule is inferred to arrive at a combined magnitude for each output membership function, by the MAX-MIN method. Then a defuzzification process is carried out for the output generation. The defuzzification function is performed by mapping the magnitudes to their respective output trapezoidal membership functions and computing the fuzzy centroid of the composite area of the member functions. In order to perform fuzzy inference, numerical values of the computed attributes have to be first of all mapped from their numerical scale to a linguistic label expressing their degree of membership. This is performed by a fuzzification function.

Figure 11 shows the results of the group behaviour analysis for a particular group learning experience. Here appears the name of each attribute, then, its linguistic labels, and below them (in white) the value of the attribute. The next column, on the right, indicates the list of attributes used in case of inference, and none, if the attribute value is calculated directly from the experience interaction data.
Analyzing student interaction processes in order to improve collaboration

**Figure 11.** Results of group behaviour for an experience, comparing each task with other groups performing the same task

**Individual analysis**

This analysis is done for each student in a group for a learning task, the method is the same as for the group behaviour, but in this case we consider a different set of attributes and rules. Some attributes are computed from the experience data for each student in the group taking into account the mean value for the group. Related to each student, the following attributes are considered: the number of contributions that he added, the mean size of his contributions, and the mean of Initiative, the Creativity, Elaboration and Conformity of his contributions. The new attributes for the individual analysis are:

- Number of contributions from this student that were answered by others
- Number of contributions authored by others that have been answered by this student
- Number of contributions authored by the student that were continued or answered by himself
- Number of contributions in the stage of “proposal” that were elaborated by the student

---

### Global Analysis for experience ENTORN02

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Values</th>
<th>Inferred From</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argumentation</td>
<td>low, mod,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>DepthTree, Interactivity, Initiative, Work</td>
</tr>
<tr>
<td>Collaboration</td>
<td>weak, norm, poor, weak, poor</td>
<td>Argumentation, Coordination, Work, Collaboration</td>
</tr>
<tr>
<td>Conformity</td>
<td>low, intermediate, high</td>
<td></td>
</tr>
<tr>
<td>Cooperation</td>
<td>low, low, moderate, high</td>
<td>Argumentation, Conformity, Creativity</td>
</tr>
<tr>
<td>Coordination</td>
<td>little, enough, much</td>
<td>Argumentation, Coordination, Messages, Initiative</td>
</tr>
<tr>
<td>Creativity</td>
<td>low, moderate, high</td>
<td></td>
</tr>
<tr>
<td>DepthTree</td>
<td>low, moderate, high</td>
<td></td>
</tr>
<tr>
<td>Elaboration</td>
<td>low, moderate, high</td>
<td></td>
</tr>
<tr>
<td>Initiative</td>
<td>low, intermediate, high</td>
<td></td>
</tr>
<tr>
<td>Interactivity</td>
<td>low, mod, high</td>
<td></td>
</tr>
<tr>
<td>MContributionNumber</td>
<td>little, appropriate, much</td>
<td></td>
</tr>
<tr>
<td>MContributionSize</td>
<td>short, average, long, VeryLong</td>
<td></td>
</tr>
<tr>
<td>Work</td>
<td>little, moderate, high</td>
<td>MContributionNumber, MContributionSize, Elaboration</td>
</tr>
</tbody>
</table>
Barros and Verdejo

All these attributes are combined to infer new ones in order to obtain an individual evaluation, as shown in figure 12. Results are displayed in a similar way as in the global analysis case, an example is shown in figure 13.

**Figure 12.** Relationship among the attributes in individual analysis

*Group task behaviour summary*

This analysis focuses on two aspects: distribution of work between group members and evolution of group activity in a period of time. We use the same data as before, (i.e. definition of the task and process contribution trees), but we consider that during an argumentative discussion for solving a group task there are different stages. We have identified three main stages:

- **Propose**, contributions that involve providing relevant information with ideas that can be in the final result of the task.
- **Argue**, contributions that involve the discussion of ideas, their refinement or the clarification of things that can help to advance the group discussion.
- **Agree**, contributions that manifest implicit agreement with the ideas of colleagues. These contributions are also used for finishing the discussion.

These stages can be repeated and interleaved in a collaborative process. A contribution type can only belong to one stage or category. The categories for the example in figure 4 are shown in Table 3.

**Table 3.** Mapping between categories of the conversational graph in figure 4 and stages in the collaborative process

<table>
<thead>
<tr>
<th>CONTRIBUTION</th>
<th>STAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposal</td>
<td>propose</td>
</tr>
<tr>
<td>Contraproposal</td>
<td>propose</td>
</tr>
<tr>
<td>Question</td>
<td>argue</td>
</tr>
<tr>
<td>Comment</td>
<td>argue</td>
</tr>
<tr>
<td>Clarification</td>
<td>argue</td>
</tr>
<tr>
<td>Agreement</td>
<td>agree</td>
</tr>
</tbody>
</table>
Analysing student interaction processes in order to improve collaboration

Figure 13. Displaying of the results of individual analysis.

The group task analysis is made by matching, classifying each contribution written by the users in each stage. Conclusions are reached in this case by comparing computed data from each user with the rest of the group. Considering how much data was written by each user in each stage, we can know if there was a distribution of work in the tasks, or if someone has worked more that the others or, otherwise, they have worked equally proposing and interchanging ideas, as we wished.

Figure 14 shows an example of results obtained from a particular task and a group of students. On the left we can observe the evolution of the task performed in three periods of activity, the first being where more contributions occurred. On the right, for each subtask, a summary of the student’s participation is given. The result in this case may suggest that collaboration, but also some division of labour, has really happened.
INFLUENCING THE COLLABORATIVE ACTIVITY

The results of analysing group and individual processes are a data source to support pedagogical decisions. At the moment we experiment the effect of sending messages about the collaborative aspects. The purpose of this feedback is to help students reflect on the process and to give them the opportunity of improving their way of working together.

The advisor module

Messages are generated by an advisor module, a mechanism based on the attributes described in the section above. The advisor uses the results of the qualitative analysis as input and generates advisory-messages for the attributes with poor values and reward-messages for those with good results. Messages are generated by a straightforward inference mechanism with rules following the pattern "if attribute-value then message". The text associated with a kind of message can be customised for each experience. The system differentiates between messages for individuals and messages for a group. Messages can be sent automatically related to time intervals for an active task or on demand, either by email or published in a private space. The advisor also offers the possibility of selecting only some messages for sending them to a specific group.

Figure 14. Intensity and evolution of work on the left, and participation by subtask on the right

Figure 15. Snapshot of the advisor’s results including some messages for a student.
Analysing student interaction processes in order to improve collaboration

Figure 15 shows a snapshot of the process of generating messages for a student in an intermediate state of her work. Messages cover every relevant attribute-value pair representing either a feature of her work that could be improved or an opportunity to stimulate the student. Positive reinforcement is known to have greater effect, therefore messages follow this style. As pointed out before, the system allows both sending the messages automatically or under the teacher's supervision. In the first case, predefined messages are as shown in figure 15. A predefined message starts giving an evaluation and then proposing ways to improve that attribute. In the latter, the teacher can select and/or modify some of the messages to suit his or her view and feelings of the student's progress, and how to suggest further improvement.

When the advisor generates a list of messages for one or more users, it saves the data, the addressee and the attributes analysed as well as their values in a database. In the cases where the content of the messages was modified for a particular user or group, the text of those messages would be relevant as well and hence is also stored. This data can be further exploited to carry out a study of the evolution of the whole process, as we will describe in the next section.

<table>
<thead>
<tr>
<th>Date</th>
<th>Attribute</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>990502</td>
<td>answeredOthers</td>
<td>Try proposing contributions of your own, proposing ideas about how to solve the task.</td>
</tr>
<tr>
<td>990502</td>
<td>answeredOthers</td>
<td>Keep the good work! You are causing the other participants to answer and thus making the discussion richer</td>
</tr>
<tr>
<td>990502</td>
<td>propose=low</td>
<td>It is important for the task to propose new ideas. They would make the group's work improve and the discussion would be richer</td>
</tr>
<tr>
<td>990516</td>
<td>NContributions/Number = appropriate</td>
<td>Keep contributing that way and check out the elaboration of your messages and how these are moving to further discussion.</td>
</tr>
<tr>
<td>990516</td>
<td>answeredOthers</td>
<td>You have not improved this aspect since my last message. Remember: Keep up the good work! You are causing the other participants to answer and thus making the discussion richer</td>
</tr>
</tbody>
</table>

**Figure 16.** Representation of the evaluation for one task in an experience, for four chosen variables: answered contributions, proposed contributions, initiative and number of contributions.
Feedback analysis

We are looking for ways to improve peer interaction. To give feedback directly to students is one aspect but this should be complemented with the opportunity to evaluate in which way the feedback affects positively their task performance. This involves (1) monitoring student’s behaviour before and after feedback is provided, (2) be able to suggest possible correlation where changes are observed, and finally (3) judging whether the feedback has been appropriate or not. This latter task is a complex one, likely to be done by the tutor. We will then consider how to support monitoring and detection of changes in student behaviour.

A relatively simple approach is to consider the mapping of attributes' values at the periods when messages were sent. The required computation and the visual representation is done by the feedback analysis module. In this module, the numeric values of the observed attributes (without defuzzification) are normalised (in the range from 1 to 10) and then represented graphically. Figure 16 plots the evolution of four attributes related to the evaluation of an individual student: answered contributions, proposed contributions, initiative and number of contributions. Data in this example comes from the experiences mentioned before. During this period feedback was generated on five occasions. We can observe increasing and decreasing intervals for some attributes, whilst others remain invariant. A change in an attribute can be considered as a potential feedback effect. Linked to dates on the Y axis are the messages sent. The next step is to consider whether messages related to the attribute were sent. This test is done automatically, and potential candidates are then filtered. For the remaining ones, the tutor, looking at the message contents has to confirm whether or not this change could reflect a feedback. For invariant attributes a similar process is carried out. First a test is run to discard cases where no messages related to the attribute were sent. Then, when messages were sent but no effect was observed, the tutor should look at the content and decide to which extent feedback has been ineffective. In this example, we can see that after the first feedback, with messages that suggest to making proposals, an increase of this activity is visible, while the feedback suggesting comment on peer proposals also seems to have no effect.

The advisor component can be improved in a number of ways, for instance, the knowledge in the rules for sending messages could be increased by taking into account the history of previous messages and their effect. The status of the task, whether it is starting, in progress, starving or to complete could also be a further source of information.

Nevertheless, before going further in this direction we plan to carry out a new set of formative evaluation experiences to study whether the current amount of automatic feedback is enough to stimulate participation. At this level the main purpose is to provide individual, positive incentive to participate rather than generate a deep diagnosis of the group performance.

CONCLUSIONS

Computer mediated collaborative learning allows the recording of a large amount of data about the interaction processes and the task performance of a group of students. This empirical data is a very rich source to mine for a variety of purposes. Some of practical nature like, for instance, the improvement of peer awareness on the on-going work. Others of a more long-term and fundamental scope such as to understand socio-cognitive correlations between collaboration and learning. Manual approaches to fully monitor and exploit these data are out of the question. A mixture of computable methods to organise and extract information from all this rough material together with partial and focused in-depth manual analysis seems a more feasible and scalable framework.

In this paper, we have presented first an approach to characterise group and individual behaviour in computer-supported collaborative work in terms of a set of attributes. In this way a process-oriented qualitative description of a mediated group activity is given from three perspectives: (i) a group performance in reference to other groups, (ii) each member in reference to other members of the group, and (iii) the group by itself. Then we have proposed a method to automatically compute these attributes for processes where joint activity and interactions are carried out by means of semi-structured messages. The final set of attributes has been fixed
Analysing student interaction processes in order to improve collaboration through an extensive period of iterative design and experimentation. We do not make theoretical claims about this particular set, the only worth of this proposal is practical value and therefore it is open to further refinement. The method uses the feature structures associated to the conversational structure of shared workspaces as data, therefore the attributes considered can be easily redefined. Moreover, in our system these are specified in a declarative way when configuring the computer environment for a learning experience.

The results of the analysis processes can be used in many ways for pedagogical purposes. We have explored the utility of an advisor, a mechanism to be used automatically or on demand. Taking into account the contributions of the individual learner, recommendations are sent to them in order to improve their participation in the joint task. Then a further analysis module allows the teacher to monitor the effectiveness of the feedback generated by the system, so that information is provided to sustain a human intervention.

Empirical testing has been a driving force in designing the DEGREE system. Each idea has been worked out from the learners, teachers and educational designers’ perspective. For instance in the case of the analysis level some experiences explicitly included a peer reflection activity, making the system analysis outcome accessible either within or among the groups of students performing similar tasks. Furthermore, we have found that analysis results are a quite useful input for educational designers to evaluate whether the definition of a collaborative task has been adequate. To combine process-oriented automatic evaluation together with manual evaluation of the final product has proved to be, for human tutors, a reasonable approach to deal with assessment matters in our distance learning courses.

A comprehensive theoretical perspective for analysing collaborative learning has been and would be a trend for the research community, partly depending on the development of computable methods for analysing process-oriented data.

Natural language processing would allow a categorisation of contributions from a content analysis, but current NLP techniques require expensive resources and processing. Considering semi-structured interactions is a non-intrusive way to break this complexity, at least for a broad range of tasks. Our design approach allows extracting relevant information at different levels of abstraction. Visualisation and global behaviour analysis tools are just two of them, supporting a variety of automatic and manual feedback. A fine-grained study of collaborative interactions would require more extensive modelling not only dealing with task and communication aspects but also with the learner’s beliefs. Nevertheless shallow analyses as presented in this paper are needed and useful to tackle a large amount of information, in order to enhance computer mediated support.

ACKNOWLEDGEMENTS

The work presented here has been partially funded by CICYT, The Spanish Research Agency, project TEL97-0328-C02-01. We would like to thank our UNED, UCM and UPM students, for their participation in the experiences carried out while designing the system prototypes.

REFERENCES

Barros and Verdejo

New York.
(http://www.kie.berkeley.edu/KIE/info/publications/AERA96/KIE_Instruction.html)


Analyzing student interaction processes in order to improve collaboration
