AIED 05 WORKSHOP 6

Amsterdam, 18-22 July, 2005

# Representing and Analyzing Collaborative Interactions: What works? When does it work? To what extent?

AIED 05 Amsterdam WORKSHOP

12th International Conference on Artificial Intelligence in Education, Amsterdam, the Netherlands

# AIED 2005 – Workshop 6

# Representing and Analyzing Collaborative Interactions: What works? When does it work? To what extent are methods re-usable?

Rosemary LUCKIN, Sadhana PUNTAMBEKER*
*IDEAS Lab,*
*Human Centred Technology Group*
*Department of Informatics, University of Sussex*
*Brighton, BN1 9QH, UK*
**The School of Education,*
*University of Wisconsin,*
*Madison. USA*

**AIED 2005 - Workshop - Representing and Analyzing Collaborative Interactions: What works? When does it work? To what extent are methods re-usable?**

Intrduction ........................................................................................................ 2

**Full Papers** ....................................................................................................... 5

Representing talk and action in collaborative activities for video analysis - Katerina AVRAMIDES, Hilary SMITH, Rosemary LUCKIN, Geraldine FITZPATRICK................ 6

Extending an Online Individual ScientificProblem-Solving Environment to Support and Mediate Collaborative Learning - Alessandra GIORDANI1, Luca GEROSA, Amy SOLLER, Ron STEVENS ............................................................................... 11

TagHelper: An application of text classification technology to automatic and semi-automatic modeling of collaborative learning interactions - Carolyn ROSÉ & Pinar DONMEZ .......................................................................................................... 22

Characteristics of Asynchronous Online Mathematics Help Environments: Do They Provide Conditions for Learning? - Dragana MARTINOVIC31 ......................................... 32

Analysing Learner Interaction in an Adaptive Communication Tool - Agoritsa GOGOULOU, Evangelia GOULI and Maria GRIGORIADOU ....................................... 39

Knowledge Development through Collaborative Work Supported by Interactive Technologies - Jorge Ferreira FRANCO, Roseli de Deus LOPES..................................... 49

**Short Papers** .................................................................................................... 59

An annotation scheme to analyse discussions in online classrooms - Aiman BADRI, Floriana GRASSO, and Paul LENG ..........59.................................. 59

Iterative Development of a Methodology for Analysing Collaborative Interactions.- Lucinda KERAWALLA, Nicola YUILL, Darren PEARCE, Rosemary LUCKIN and Amanda HARRIS.63.. ................................................................................... 63

Towards adaptable interaction analysis tools in CSCL - José Antonio Marcos[1] and Alejandra Martinez and Yannis Dimitriadis ..................................................... 67

Insightful Problem Solving Promoted by Collaboration: The Effect of Role Reversals between Trial and Observation - Taro IZAWA1, Sachiko KIYOKAWA ......................... 71

Adaptive Visualisation Support for Self-managed Learning Groups - Peter REIMANN, Judy KAY, Kalina YACEF, Peter GOODYEAR ............................................................ 75

# AIED 2005 - Workshop

# Representing and Analyzing Collaborative Interactions: What works? When does it work? To what extent are methods re-usable?

Rosemary LUCKIN, Sadhana PUNTAMBEKER
*IDEAS Lab,*
*Human Centred Technology Group*
*Department of Informatics, University of Sussex*
*Brighton, BN1 9QH, UK*
*The School of Education,*
*University of Wisconsin,*
*Madison. USA*

## 1 Description

The purpose of this workshop is to explore and increase our understanding about what methods can be used to investigate the collaborative learning interactions that technology can engender. This increased understanding will inform both the ways in which technology is developed, the manner in which it is integrated into learning contexts and the methods used to evaluate its impact. The workshop will bring together researchers who collect and analyze process data from contexts where learners are collaborating through, with or around technology. This will include methods for documenting and visualizing collaborative interactions, and for analyzing the relationship between successful learning and collaboration. As the range of technological artifacts that are available to support the process of collaborative learning increases we need to reflect upon the methods we have developed for evaluation and assess the extent to which we understand the nature of successful collaborative interactions and their relationship to individual learning? We need to assess the extent to which new technologies require new approaches and consider how we can expand and adapt the methods we have.

The increased emphasis upon pervasive technology for learning brings into sharp relief the importance of the context in which learners is interacting. The advances made in understating the ways in which communities of learners can be established and supported through networked technologies likewise requires us to increase the breadth of the data we encompass in our evaluations. But have our methods kept up with these developments? This workshop will help us to explore the extent to which our methods match our needs. It will foster the discussion and dissemination of successful methodologies and provide a forum for the development of innovative new tools. We invited submissions of research papers (maximum length 8 pages, font size no smaller than 11 pt) describing approaches to the representation and/or analysis of collaborative interactions.

## 2 Background

The workshop will encompass work that aims to evaluate the cognitive, meta cognitive, social and affective aspects of collaborative interactions. It will welcome both quantitative and qualitative approaches to the analysis of data collected through different media including video, audio, interaction logging, screen capture, written texts. Key questions that will be addressed in this workshop are:

- What are the characteristics that we should be looking for in our data? What are the features of successful collaborative interactions?
- What data should we capture and how should we capture it? What are the tools that are available to us and how should we decide which to select?
- How transferable are approaches that have been shown to be effective in one context or with one particular data set to another context or data set?
- What are the criteria that need to be used for the selection of a coding scheme or approach?
- How useful are visualizations to the process of data analysis
- How can we combine and/or integrate the analysis of data about what individuals contribute to an interaction with data about the group as a whole
- Can we extend techniques used for desktop computing to other forms of interface such as tangible or ambient technologies?
- How can we document that collaborative interactions have helped create communities of learners?
- What defines a successful community of learners?

This workshop will provide an opportunity for researchers with different backgrounds, using various approaches to documenting and assessing collaborative interactions to discuss answers to questions these questions and more. This list is not exhaustive and we welcome the identification of further issues that need to be addressed. The results of the workshop will help in a more rigorous understanding of the ways to document and analyze collaborative learning and inform the design of the next generation of tools. The workshop will build upon the work completed at the "Documenting Collaborative Interactions: Issues and Approaches" (Puntambekar & Luckin) workshop held at CSCL 2002

## 3 Workshop Organisers

Rosemary Luckin <rosel@sussex.ac.uk>
IDEAs Laboratory, Department of Informatics, University of Sussex,

Sadhana Puntambekar <puntambekar@education.wisc.edu>
The School of Education, University of Wisconsin, Madison. USA

## 4 Programme Committee

Soller, Amy <asoller@ida.org>
ITC-IRST, Italy

Julita Vassileva <jiv@cs.usask.ca>
University of Saskatchewan, Canada

Alejandra Martinez Mones <amartine@infor.uva.es>
Universidad de Valladolid, Spain

Patrick Jermann <patrick.jermann@epfl.ch>
CRAFT - Centre de Recherche et d'Appui pour la Formation et ses Technologies
Switzerland

Rachel Pilkington <r.m.pilkington@bham.ac.uk>
School of Education, University of Birmingham
Edgbaston, Birmingham, UK

Marta Rosatelli <martacr@inf.ufsc.br>
UniversidadeFederal de Santa Catarina, Brazil

Judith Good <judithg@sussex.ac.uk>
University of Sussex, UK

Nicola Yuill <nicolay@sussex.ac.uk>
University of Sussex, UK

# Full Papers

# Representing talk and action in collaborative activities for video analysis

Katerina AVRAMIDES, Hilary SMITH*,
Rose LUCKIN, Geraldine FITZPATRICK*
IDEAS Lab, Interact Lab*
Human Centred Technology Group
Department of Informatics, University of Sussex
Brighton, BN1 9QH, UK

**Abstract.** This paper introduces a collaboratively developed methodology for the qualitative analysis and visual representation of video data. The data to be analysed were video recordings made by secondary school students whilst engaged in a collaborative, in-the-field, data collection process using hand-held data logging equipment for monitoring Carbon Monoxide and wind values. We describe the requirement for, and the collaborative and evolutionary development process of, a visual method through drawing overview activity maps. Following this we discuss some issues arising and future challenges for overview activity maps.

## Introduction

Despite an increasing emphasis on technology-mediated scientific enquiry, there is a marked need for analytical tools for studying how technology can be best integrated in the design of effective learning experiences. Several authors have stressed the need for an established framework for analysing technology-supported collaborative interactions (e.g. Lehtinen, 2003; Littleton & Light, 1999; Pea, 2004). The challenge for researchers is to develop a framework that will situate the learning experience in the social and cultural context. Moreover, such a framework must allow for the analysis of the collaborative experience as a whole that is more than the sum of individual actions. Ideally it must also be sufficiently flexible to allow for the changing needs of accepted practice in interaction analysis (Jordan & Henderson, 1995).

We report on the process of developing a tool for analysing technology-supported collaborative interactions towards building a framework of this kind. The context for this work is two related projects that investigated the potential of mobile technology to support collaborative scientific enquiry. Both projects involved small groups of students (3-4 in each) investigating pollution levels in their local area. Students were aged between 14 and 16 years. They used the same hand-held data-logging technologies in 30 minute, exploratory sessions to develop collaborative ideas about the behaviour of Carbon Monoxide in different locations and wind conditions. The projects differed, however, in the level of structure provided to students (asked to develop hypotheses versus asked to explore using the equipment) and the facilitator role (degree of support through questioning to encourage participant to verbalise their ideas). The data analysis we report focused on these data-collection sessions in the field. For further detail on the projects see Underwood et al (2004) and Stanton Fraser et al (in press)

Our initial analysis was driven by broad research questions such as:

- How did the students engage in the data-collection process
- how did they engage with multiple devices

- what sorts of behaviours/interactions were enabled
- and how the social and physical context supported or hindered the learning experience

We collaboratively created a series of overview activity maps to analyse the video-ed sessions. The activity maps enabled our analysis to become more focussed as interesting patterns of talk and behaviour were identified across all sessions. The process of developing the maps was thus data-driven, based on observation of emerging patterns and enabled us to progressively refine our research questions.

**Methodology**

The video data available from each session was supported by a collection of multimedia, including: students' handwritten wind data, automatically captured CO data, Global Positioning System (GPS) location data and photographs taken by the facilitator. A large portion of this data collection occurred in parallel and so it was a combination of all the data available that provided the most complete record of what happened and when. We found we needed a method of analysis that allowed for incomplete data in some sessions. This occurred for a variety of reason, including wind or noise on camera, being too distant from speakers, recordings not always being continuous video, students forgetting or otherwise to focus on devices or speaking people.

The initial activity map (v1, see figure 1) was created to provide a timeline of the actions that occurred within a single group during a session. Version 2 (figure 2) was created to allow the capture of some of the variables that were not constant across sessions (e.g. group size, hypotheses verbalised, planning where to take readings) for an early categorisation of the types of actions observed. At this stage we also agreed the variety of types of location visited by all groups. These locations included: open field, bust stop, tree area, vents on buildings and machinery. Also, we identified social interactions that influenced actions, such as facilitator input, distractions and changes to the initial plan that were made on the fly. These initial maps provided the means to extract patterns from the data but also to identify important information that was not adequately represented (e.g. physical location, whether standing still or moving, malfunctioning of equipment) that led to refinement of the maps. We identified a further need to define the observed behaviours to ensure each researcher was categorising in the same way.

Developing the maps further was a collaborative process during which we categorised talk and actions (communicating readings, commenting on readings, hypothesising, discussing, distracting); structured interactions in terms of devices and people; and defined a framework for capturing the context data that was relevant to our study.

A key aspect of the development of the activity maps was finding ways of representing the information visually that would enable us to 'scan' the data and rapidly identify clips/points of interest: e.g. spotting changes in behaviour when a person switched to using a different device in the same session, knowing where to target different kinds of interaction / behaviour / comments. This involved minimising text and substituting for symbols, colour-coding, arrows to show links (see figures 3 and 4).

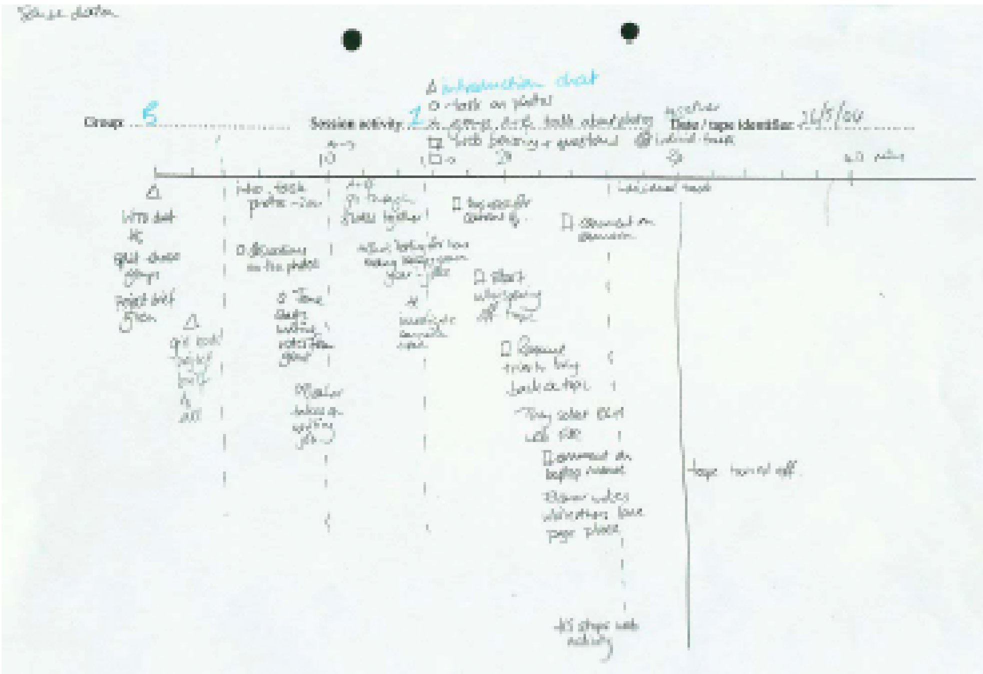



Figure 1 - activity map version 1

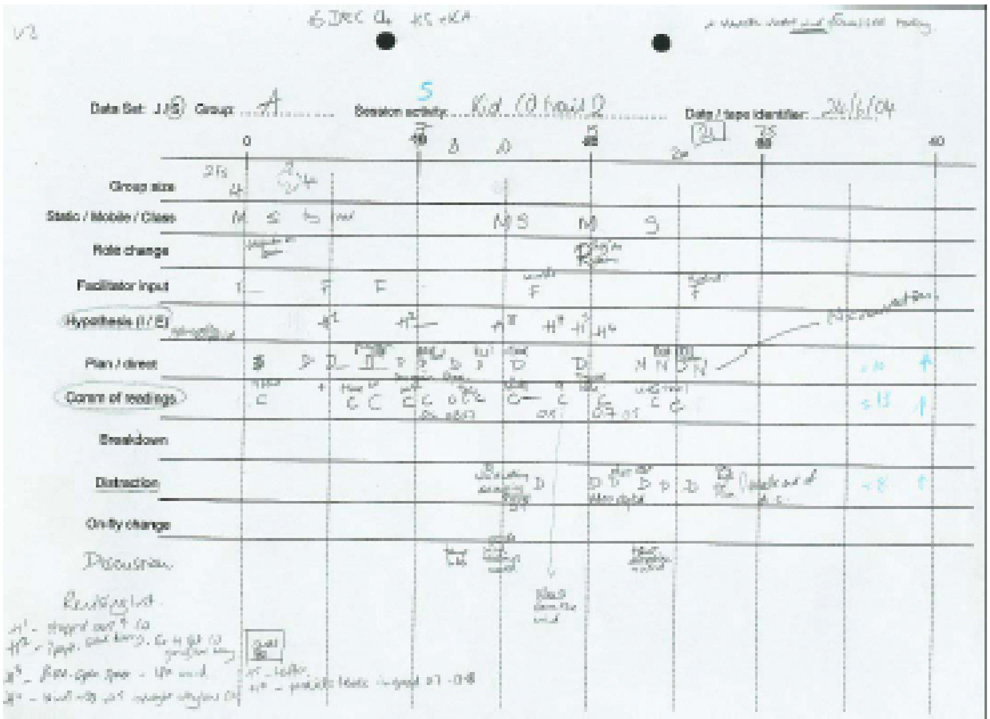



Figure 2 - activity map version 2, where letters on map refer to category of behaviour on that line e.g. S = static, M = mobile, F = Facilitator comment, H1 = hypothesis 1

The final version of the maps (figure 4) allowed us to combine an individual and group perspective on the data. More specifically, on an individual level we were able to match actions and talk to people and the equipment they were using at the time, analyse individuals' use of technology, and detect who requested to swap devices with whom during the session. On a group level we were able to analyse individual contributions to group activity, identify how the context influenced group activity, and identify how the technology enabled or hindered collaboration and group activity (ease/difficulty of sharing information/creation of device roles).

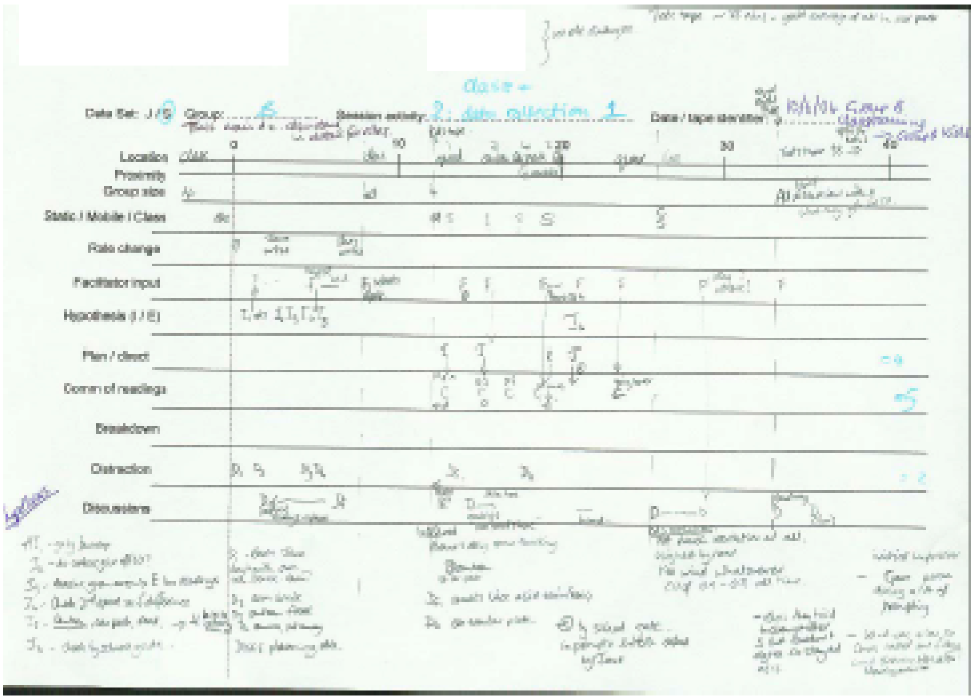



Figure 3 - activity map version 3, letters on map refer to method used in figure 2

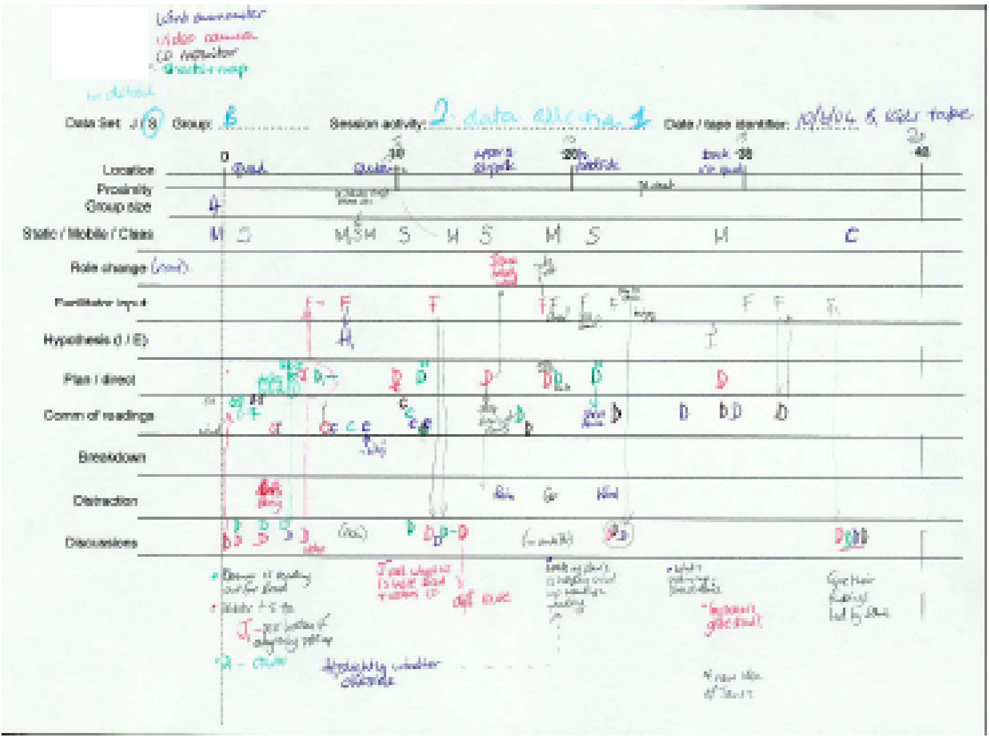



Figure 4 - activity map version 4, letters on map refer to method used in Figure 2

## Discussion

The iterative development of the activity maps enabled events within the data to be flagged as interesting and for us to use significant behaviour patterns to drive the focus of the analysis. The open nature of version 1 (almost a blank sheet) led to the use of categories to focus attention on particular types of interaction and behaviour and enabled researchers to work through data separately but in parallel. We fortified our individual analysis through joint working sessions to develop our ideas about the ways we defined particular behaviours. We were also able to share ideas about how best to arrange the activity maps e.g. to change the way we 'chunked' the maps. In version 2 we had provided vertical lines at set time intervals. These were found to be less meaningful than the idea of chunking activity by location of the group.

The collaborative nature of researchers being fully involved in video analysis in parallel with similar data meant the development of our categories for analysis focus was much reduced than it would be for one researcher working alone on each data set in turn. Our foci were subtle in their differences: by the very nature of our multidisciplinary backgrounds (of user centred design, psychology, education, computer science).
The visualisation of data in this way provided a very rapid way to scan through tapes in an initial pass, and identify which would be most useful for more in-depth analysis. It meant that transcription effort could be targeted efficiently, rather than across the board. Furthermore we found we were able to easily analyse from an individual or whole group perspective, or isolate a sub-group or particular device combination when investigating the roles students took on with each device.

In developing activity maps we propose to tie in various supporting data where available, for example by overlaying a graph of the collected wind and Carbon Monoxide readings, or location readings. This will help us identify reasons behind flurries of activity we have noticed e.g. in identifying the triggers to discussions that were student led rather than facilitator led.

## Acknowledgements

We would like to thank our project collaborators Joshua Underwood, Danae Stanton Fraser, Sara Price, Ella Tallyn, the SENSE project collaborators at Nottingham, session participants including David Daniels, Steve Rogers, teachers and pupils at Varndean and Hove Park Schools, Portslade Community College. The e-Science project was part of the Equator IRC (www.equator.ac.uk). The SENSE project was funded by the JISC.

*References*

Jordan, B. and Henderson, A. (1995). "Interaction Analysis: Foundations and Practice." The Journal of the Learning Sciences, 4(1): 39-103

Lehtinen, E. (2003). Computer-Supported Collaborative Learning: An Approach to Powerful Learning Environments. In E. De Corte, L. Verschaffel & N. E. A. Entwistle (Eds.), Powerful Learning Environments: Unravelling Basic Components and Dimensions (pp. 35-53). Amsterdam: Elsevier.

Littleton, K., & Light, P. (Eds.). (1999). Learning with Computers: Analysing productive interaction. New York: Routledge.

Pea, R., D. (2004). The Social and Technological Dimensions of Scaffolding and Related Theoretical Concepts for Learning, Education, and Human Activity. The Journal of the Learning Sciences, 13(3), 423-451.

Stanton Fraser, D., Smith, H., Tallyn, E., Kirk, D., Benford, S., Rowland, D., Paxton, M., Price, S. and Fitzpatrick, G. (in press) The SENSE project: a context-inclusive approach to studying environmental science within and across schools. Accepted for publication for CSCL'05

Underwood, J., Luckin, R., Fitzpatrick, G., Steed, A., Spinello, S., Greenhalgh, C., Egglestone, S. and Hampshire, A. (2004) From e-Science for Children to e-Services for Educators. In Proceedings of Grid Learning Services (GLS'04) workshop at ITS2004

# Extending an Online Individual ScientificProblem-Solving Environment to Support and Mediate Collaborative Learning

Alessandra GIORDANI[1], Luca GEROSA[1], Amy SOLLER[2], Ron STEVENS[2]
*1 Dipartimento di Informatica e Telecomunicazioni, Università di Trento*
*Via Sommarive 14, 38050 Povo di Trento, Italy*
*{agiordani, lgerosa}@dit.unitn.it*
*2 IMMEX Project, UCLA, 5601 W. Slauson Ave, #255, Culver City, CA. 90230*
*immex_ron@hotmail.com, asoller@ida.org*

**Abstract**: Students sometimes develop unproductive problem solving approaches that can persist for months if not detected and addressed. Previous work has shown that such impasses can be detected through probabilistic modelling of students' strategic approaches. Unproductive approaches might be avoided and/or modulated for some students by placing them in collaborative groups of 2-4 students. Although the specific collaborative activities responsible for these strategic changes are unclear, often such group interactions result in effective problem solving. We developed a web-based synchronous collaborative environment into which online problem solving activities can easily be embedded, allowing an analysis of not only the problem-solving process, but also the collaborative activities and events as teams perform simulations. A structured collaborative interface allows segmentation of the collaborative learning event log, facilitating the establishment of linkages between the collaborative interactions and the problem solving efficiency and effectiveness. In this paper, we present our preliminary analysis explaining how this structured collaborative environment may improve the performance of individuals, and how collaborative learning might be assessed by comparing the models of students' strategic approaches to their interaction.

## 1. Introduction

Documenting students' problem solving performance and progress in real-world simulation environments is difficult given the complexity of the tasks. A recurring challenge for instructors is determining which students are applying the knowledge gained in introductory courses to think critically, and which ones require interventional supports. We have been addressing this challenge by constructing online problem solving systems, and layered analytical machine learning tools, collectively called IMMEX, that support the development, implementation and analysis of online problem solving [1], [2]. The singleuser version has been widely used in science classes across middle and high schools, universities, and medical schools and has logged over 400,000 scientific student problem solving episodes over the past 3 years.

From these logged problem-solving datasets, we have developed performance and progress models of student problem solving using artificial neural network clustering and Hidden Markov Modeling [3]. These analyses have enabled us to identify when students develop and stabilize with unproductive problem solving approaches. Because these strategies sometimes appear and persist for months, there is a need to provide feedback and/or interventions that might encourage students to alter their problem solving approaches [4], [5].

Collaborative learning is one potential interventional approach, as studies have shown that groups may learn faster, make fewer errors, recall better and make better decisions than individuals working on their own [6]. When we initially observed students in small collaborative groups, we noted that fewer students stabilized with inefficient and ineffective strategies [7]; however, problem solving was not effective for all student groupings. These results suggested that the group composition or interactions may contain important and as of yet unknown variables influencing strategic performance and progress. We began documenting such collaborative events and components by designing an online collaborative software environment into which the IMMEX problem solving simulations could be embedded, and through which the student interactions could be tracked and modeled [8].

The study described in this paper validates the design of the IMMEX Collaborative interface and shows that, at the problem solving level, the group actions, approaches and outcomes share similarities with those that occur in face-to-face groups. The preliminary results suggest this may be a useful approach for linking collaborative interaction analysis with problem solving outcomes. In the next section, we provide an overview of the IMMEX Collaborative environment and describe our research theses. We then discuss the preliminary experimental results and put forward some hypothesis for further work.

## *2. The Problem Solving Environment*

In designing IMMEX Collaborative, we felt that it should interpret actions in a shared workspace as acts of communication, as if the students were seated around a table engaging in problem solving. We also felt that it would be important to construct a structured environment that realistically reflected the nature of the problem solving. For this we drew on earlier verbal protocol research showing how students propose hypotheses, run physical and chemical tests, and reflect on the results of those tests in a repetitive fashion as they solve IMMEX problems [9]. The final requirement was that the environment needed to be structured to facilitate the automated modelling of the group interactions in a way that would accommodate the thousands of current and future IMMEX users.

The IMMEX Collaborative client interface (Figure 1) is divided into three portions. The main window is a shared workspace dedicated to the collaborative navigation of the IMMEX multimedia web pages. Actions taken by students in this frame are automatically reflected on the other group members' screens. The vertical frame on the left side shows the structured chat interface with a three tabbed panel. The horizontal frame along the bottom shows a graphical representation of the service and synchronization facilities, which are used to manage the flow of actions in the collaborative space. The mouse image moves over the name of the student who has control of the workspace, as if the members were seated in front of the same monitor, passing the mouse among each other. The client runs in a browser, and is managed through Java applets that communicate with the IMMEX Collaboration Server. The Collaboration Server is an HTTP server acting as a proxy, that filters, edits, and synchronizes the IMMEX HTML pages through JavaScript, and sends them to the clients.

IMMEX problems require students to frame problems from descriptive scenarios, judge what information is relevant, plan a search strategy, gather information, and eventually reach a decision that demonstrates understanding. One problem set researched extensively is *Hazmat*, which provides evidence of students' ability to conduct qualitative chemical analyses. A multimedia presentation is shown to the students, explaining that an earthquake caused a chemical spill in the stockroom and their challenge is to identify the unknown chemical by gathering information using a 22 item menu containing a Library of terms, a Stockroom Inventory, and a number of different Physical or Chemical Tests (e.g. litmus test, precipitate test). This problem set contains 38 cases that can be performed in class, assigned as homework, or used as quizzes. As students perform multiple cases that vary in difficulty, student ability can be estimated by Item Response Theory analysis by relating characteristics of items and individuals to the probability of solving a given case.

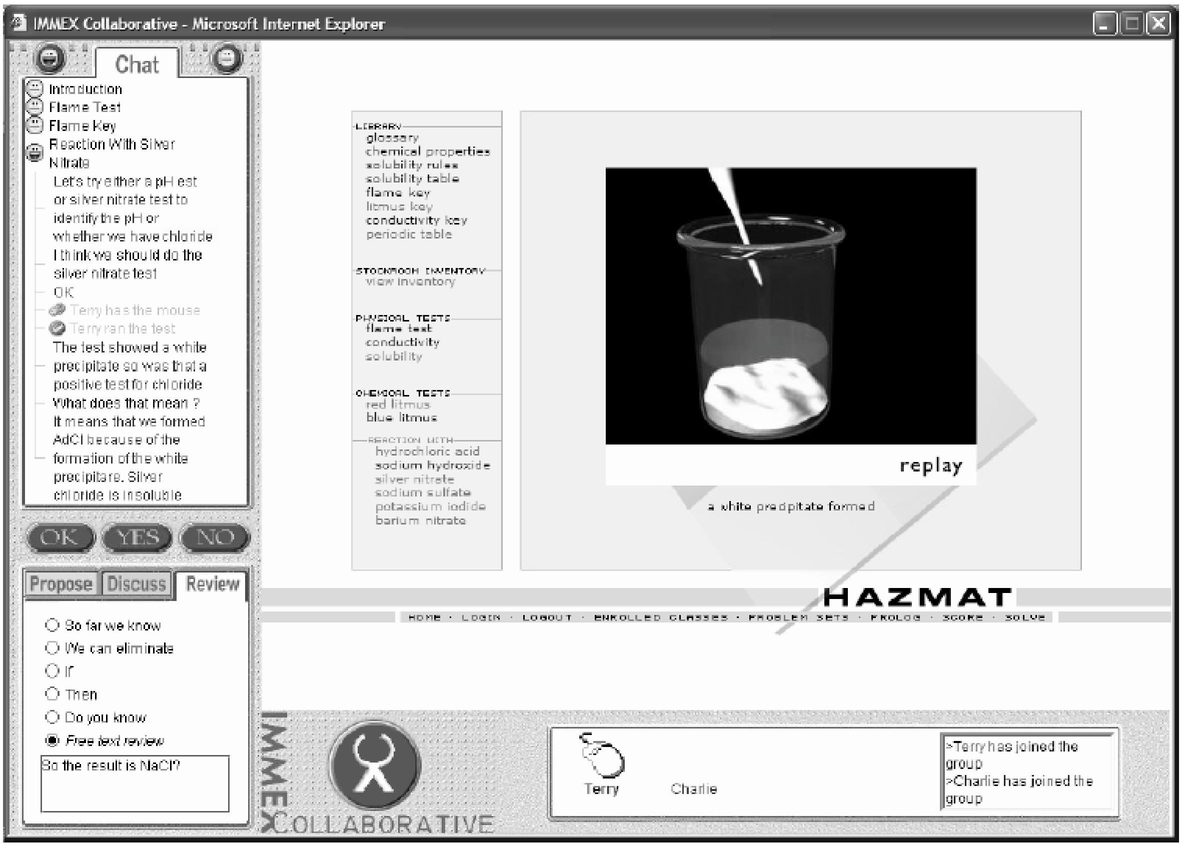



**Figure 1**. The main frame shows the IMMEX problem solving environment embedded within the *IMMEX Collaborative* environment which allows groups of students to use chat, sentence openers (far left) and shared mouse control (bottom) to solve problems.

An analysis of the face-to-face interaction between pairs of students learning with the IMMEX environment showed that their discourse often followed a predictable pattern. First, they tended to discuss which chemical or physical test to run next (the *proposal* 3 *episode*); second the students ran the test (the *event*); and third, the students discussed the results of the test (the *discussion episode*). A set of sentence openers for each episode was then developed based our manual analysis, and taking into account earlier work on effective peer dialogue [10]. Table 1 lists these eighteen openers, also located in the three tabbed panel on the lower-left hand corner of the interface (see Figure 1).

**Table 1**. The eighteen sentence openers are distributed across three problem solving phases: propose, discuss, and review. Each represents a different cognitive process related to the problem solving phase.

| Propose | Discuss | Review |
|---|---|---|
| 1. Let's try… | 7. The test showed … | 13. So far we know … |
| 2. Why …? | 8. What does that mean … ? | 14. We can eliminate … |
| 3. We should … | 9. Can you explain … ? | 15. If … |
| 4. What do you think …? | 10. It means … | 16. Then … |
| 5. Because … | 11. Do you think …? | 17. Do you know …? |
| 6. I think (*Free text proposal*) | 12. I think (*Free text discussion*) | 18. I think (*Free text review*) |

The sentence openers are simple enough for users to find and select those that are most appropriate, but we also allow for the use of free text by providing the opener “I think” in each of the three proposal, discussion, and review categories. IMMEX Collaborative manages the conversation through the use of topic threads (based on the context). These topic threads attempt to structure the student discussions to reflect the structure of their decision processes in selecting and explaining the results of the various physical and chemical tests. The topic threads are also used to automatically segment the sub-dialogues into episodes. Conversations that are segmented at different levels of granularity may be useful in the future for modeling structured collaborative interactions (for example, using quantitative indices, Hidden Markov Modeling [10] and neural networks [12]). In this way, the IMMEX Collaborative environment supports learners through the various phases of problem solving, facilitating an extended, in-depth, on-topic discussion, and providing a coherent view of the argument [11].

The structured interaction model is obtained by segmenting the chat log according to the flow of conversational contributions and students' actions. Every proposal segment ends with an EventType = *TestItem* (e.g. *View Inventory*), and may or may not be followed by a segment in which the students discuss the results of the test. If a discussion phase begins after a test is ordered, it will continue until another proposal opener (Stems 1 through 6) is used to propose a new test (see Table 2). In other words, discussion and subsequent proposal episodes always follow events in which students order tests. Special episodes for review can occur at any point of the chat and are identified from the third range of stems. The three "quick" buttons ("OK", "Yes", and "No") can be used to indicate agreement. Their type is inferred from the episodes to which they belong.

We have defined several quantitative indices that could be useful indicators for understanding the coherency of problem solving in this environment: *episodic alignment, episodic balance and dynamic structure*. Episodic alignment measures the linkage of the proposal/test/discussion segments with regard to each other as they occur repeatedly during the problem solving event. IMMEX verbal protocols [9] suggest that discussions should contain linked aligned sequences of proposals, tests and discussion. Episodic balance refers to the ratio of contributions in the proposal segments to those in the discussion segments of each episode. Many proposals without consensus may indicate a less effective teamwork. Dynamic structure is a coarse-grained coherency measure from a problem solving perspective: more proposals would be expected during the early framing stages of problem solving, and as the students converge upon a solution, we should see proportionally more discussion. 4

**Table 2**. Proposal segments (grey) begin when someone proposes a new test, and end when the test is ordered

| Time | User | Sentence Opener | Text | Event Type | |
|---|---|---|---|---|---|
| 2:21 | Charlie | 6 | Let's view the inventory | Chat | PROPOSAL |
| 2:22 | Terry | 3 | We should click on the view the inventory sheet to see the message | Chat | |
| 2:22 | Charlie | 6 | Do you want to spend these points? | Chat | |
| 2:23 | Terry | 20 | Yes | Chat | |
| **2:23** | **Charlie** | **0** | **View Inventory** | **Test Item** | |
| 2:25 | Terry | 7 | The test showed we have a sodium cation so we have carbonate, chloride, hydroxide, nitrate, or sulfate anions. | Chat | DISCUSSION |
| 2:25 | Charlie | 19 | OK | Chat | |
| 2:26 | Terry | 7 | Let's try either a pH test or silver nitrate test to identify the pH or whether we have a chloride. | Chat | |

In our analysis we also examine the symmetry of the contributions from the different individuals in the group [13] assuming that more symmetrical collaboration should include near equal participation by individuals in the proposal and discussion sessions, near-equal responsibility of test ordering (as evidenced by mouse sharing), and symmetry across the framing and closure sections of the problem solving session.

## 3. Performance Models and Analysis of Collaborative Events

The first experiment involved eight freshman university chemistry students working in pairs at a distance through the IMMEX collaborative environment. Each group performed four randomly presented cases of *Hazmat,* which varied in difficulty. Fifteen problem solving sessions were logged and analyzed for both the collaboration events (Table 3).

Table 3 lists the average percentages and total numbers (in the boxes) of chat interactions (proposals, discussion and review), actions (ordered items and background information) and mouse control (handling rate) for each individual and group. The chat percentages are to be interpreted referring to the averages of total number of groups' contributions as the mouse control's percentages relate to the rates of total ordered and background's items selected by each members. Even where the work was not shared (e.g. low mouse transfer rate), all of the chats included near equal participation of individuals proposing, discussing and reviewing. Analysis of the values in the table shows the following:

• Group 1 selected more tests and spent less time reviewing than the other groups.
• Group 2 chatted the most, and had the greatest percentage of proposals, but ran the fewest tests. This group was one-sided in that they never passed the mouse, indicating that one individual may have dominated the problem solving session.
• Group 3 chatted little, but its members shared their workload better than the other groups.
• Group 4 spent a lot of time reviewing information, and shared problem solving responsibilities as evidenced by the passing of the mouse

**Table 3**. Communication, test ordering, and mouse control actions (data averages)

| | Chat | | | | Item | | Mouse |
|---|---|---|---|---|---|---|---|
| | Propose | Discuss | Review | Total | Ordered | Background | Control |
| User 1 | 22% | 27% | 8% | 57% | 9 | 9 | 78% |
| User 2 | 20% | 20% | 4% | 43% | 2 | 1 | 22% |
| **Group 1** | **41%** | **47%** | **12%** | ***35*** | **11** | **10** | ***21*** |
| User 3 | 32% | 17% | 7% | 57% | 6 | 5 | 100% |
| User 4 | 24% | 13% | 7% | 43% | 0 | 0 | 0% |
| **Group 2** | **56%** | **30%** | **14%** | ***50*** | **6** | **5** | ***11*** |
| User 5 | 25% | 17% | 7% | 49% | 5 | 6 | 64% |
| User 6 | 23% | 19% | 9% | 51% | 3 | 3 | 36% |
| **Group 3** | **48%** | **36%** | **17%** | ***20*** | **8** | **9** | ***17*** |
| User 7 | 15% | 16% | 14% | 49% | 3 | 3 | 53% |
| User 8 | 16% | 23% | 16% | 55% | 5 | 7 | 47% |
| **Group 4** | **31%** | **39%** | **30%** | ***33*** | **8** | **10** | ***18*** |

We also studied the coherency of collaborative events using our interaction models. In general, groups' conversations were episodically balanced and aligned: the number of subdialogues (pairs of episodes) was proportional to the number of problem solving events, reflecting the fact that the members usually make some hypotheses and *proposals* before ordering tests or background information, and then they *discuss* the results obtained. The data showed that these episodes were paired together. Moreover we found that there is a strength coherency between the trends of the interactions and the problem solving framing stages (dynamic structure). As expected, more proposals occur during the early framing stages of problem solving (88% cases) and, as the students converged upon a solution, there was proportionally more discussion (69% cases). In 94% of the chat logs, the amount of discussion increased (from 25% to 64%) in the second half of the performances, as the proposal rates decreased. The interaction analysis gives us some information about the students' interaction, but it does not tell us much about the effectiveness of the group learning, and how the interaction models match with more or less efficient problem solving approaches.

To follow students' performance and progress, we have automated layered analytic models of how strategies are constructed, modified and retained as students learn to solve online problems like *Hazmat*. These strategies are modeled first by self-organizing artificial neural network analysis, using the tests that students choose to solve the problems as the classifying inputs. This generates a 6 x 6 matrix of strategies detailing the qualitative and quantitative differences among problem solving approaches. In Figure 2 the highlighted boxes in each neural network map indicate which strategies are most frequently associated with each state. Then, progress models are developed across sequences of performances (defined by the neural network nodal classifications) by HMMs, which stochastically describe problem solving progress with regard to different strategic stages in the learning 6 process. These analytic layers operate as background processes and can generate performance measures in real-time (see [3] for more detail).

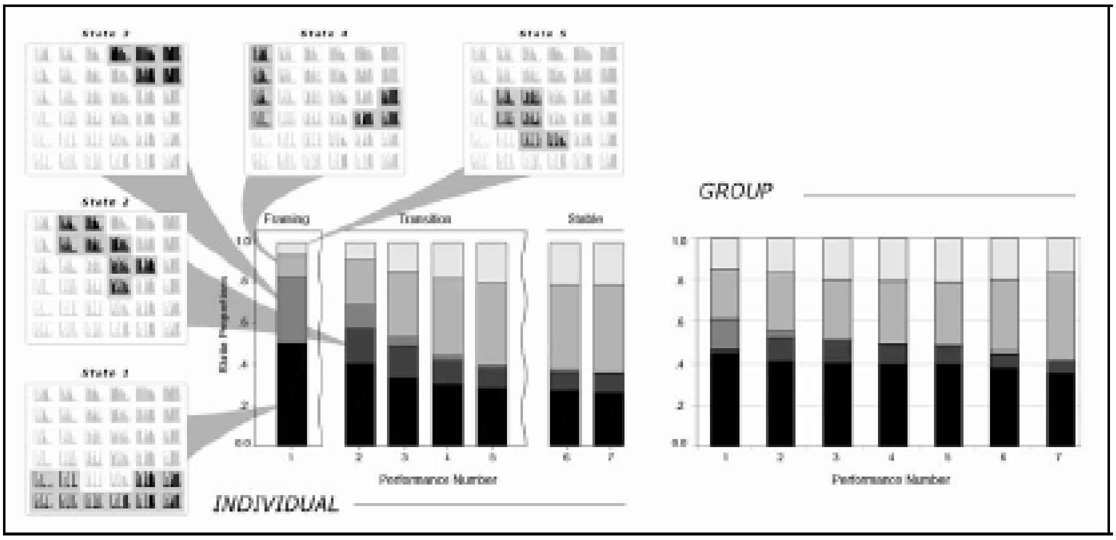



**Figure 2.** The learning trajectories for individuals (A) and groups (B) suggest that once students adopt a strategy in stable state (1-4-5), they are likely to continue to use them. In contrast, students adopting State 2 and 3 strategies are less likely to persist with those approaches and more likely to adopt other strategies.

An example of student strategy development is shown in Figure 2A, illustrating the distribution of HMM state usage as students solved 7 different *Hazmat* cases. On the first case, when students are framing the problem space, the two most frequent states represent either a limited number of test selections (State 1), or an extensive number of test selections (State 3). As students develop experience, they transit through states, as shown by the state transitions in Figure 2A. Students transit from State 3 (and to some extent State 1), pass through State 2 and into States 4 and 5. By the fifth performance the distribution of approaches appears to have stabilized, showing that without intervention, individuals learning alone generally tend not to switch their strategies, even when their strategies were ineffective.

Also shown in this figure is a similar learning trajectory for 5452 Hazmat performances which were collected from students who worked collaboratively in face-to-face groups of 2 or 3. Consistent with the literature [6] [10], having students work in collaborative groups significantly increased their solution frequency. More importantly, ANN and HMM modeling of these performances showed (Figure 2B) that the collaborative learners stabilized their strategies more rapidly than individuals, used fewer of the transitional States 2 and 3 and more State 1 strategies (limited and/or guessing approaches). This suggests that the group's interaction helped students see multiple perspectives and reconcile different viewpoints, events that seem to be associated with the transitional states. The collaboration may have replaced the explicit need for actions that are required to overcome impasses, naturally resulting in more efficient problem solving.

Let's now examine the four groups' fifteen performances as they learned collaboratively. First, from Table 4A ,the solve rates on the first and second attempts combined was 68% (average of columns "1st" and "2nd" combined in Table 4A) which was very similar to the performance of face-to-face groups and significantly higher than individuals (~53%) [3]. Similarly, from Table 4B the progress states obtained by mapping of the ANN performance nodes of each group performance to the associated states derived from HMM showed enrichment for States 1 and 4, much like face-to-face group performances (refer to Figure 2B). A surprising finding was that most of the groups stabilized their approaches quite quickly as evidenced by the use of consistent strategies from case to case, even when the cases were of different difficulties and whether they 7 solved the case or not (see Table 4B). However, it appeared that the collaborative interface was flexible and allowed the groups to solve the cases in multiple ways. Group 4 was the only one that changed its stabilized strategy passing from State 5 (the more effective, see solve rates in Table 4A) to State 3 (the less effective). This is not surprising because this case, Iron III Nitrate, is the third most difficult Hazmat case and deserves more attention. It is interesting that the flexibility shown by this group enabled them to solve this case, demonstrating that chats and collaboration events are as sensitive to problem difficulty as they are when individuals perform IMMEX cases.

**Table 4A:** Solve rates of HMM stases.
**Table 4B:** Strategy trajectories: solving tries, ANN nodes and progress states

| A | Not solved | 1st try | 2nd try |
|---|---|---|---|
| **State 1** | 30.4% | 51.7% | 17.9% |
| **State 2** | 35.2% | 43.1% | 21.7% |
| **State 3** | 46.3% | 32.4% | 21.3% |
| **State 4** | 36.1% | 44.9% | 19.0% |
| **State 5** | 23.5% | 56.5% | 20.0% |

| B | Group 1 | | | | Group 2 | | | Group3 | | | | Group 4 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Solved** | 1st | 1st | 1st | 1st | 1st | 1st | no | 2nd | no | 1st | no | 2nd | 1st | 1st | no |
| **Node** | 18 | 1 | 7 | 23 | 26 | 33 | 33 | 16 | 16 | 2 | 16 | 20 | 27 | 21 | 11 |
| **State** | 4 | 4 | 4 | 4 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 5 | 5 | 5 | 3 |

Note: *1st try/ 2nd try* means solved on the first or second attempt

The next step is to document correlations between interaction models and strategic problem solving approaches. One starting point may be to use the HMM state differences we have seen between individual and face-to-face groupings. Face-to-face groups use fewer State 2 strategies; collaboration seems to help these students transit through this state faster. In the data presented here, Group 3 stabilized with State 2 strategies, and this was also revealed in the nature of the interaction. The analysis in Table 3 shows that they had fewer discussions, fewer interaction overall, and no mouse sharing. While this could happen because of an incorrect use of the tool or because of the tool itself, this group also had the lowest solution frequency and ordered the fewest number of tests suggesting a lower quality problem solving and collaboration experience overall.

## *4. Conclusion and Future Work*

The goal of these studies was to begin validating the usefulness of a synchronous, symmetrical approach for relating the dynamics of online collaboration with the effectiveness / efficiency of concurrent problem solving. The preliminary results are encouraging in that the solution frequency, time on task, and interaction statistics were similar to what is observed in face-to-face collaboration, suggesting that the interface neither significantly changed the nature of interaction and problem solving in this environment, nor interfered with the overall problem solving. At the strategic level, this was further supported by greater than expected usage of HMM States 1 and 4 by the groups, also mirroring that found with face-to-face collaboration in *Hazmat*.

Perhaps the most unexpected finding was that most groups rapidly develop a rapport that resulted in the negation of a strategy that was repeatedly used across previous tasks. To our knowledge this is not a well documented phenomenon although, given our results with individuals, perhaps not overly surprising . Analysis of a second chemistry problem set with 19 groups is providing similar results. The most disappointing aspect of these studies is that while the chat sessions were episodically aligned and balanced, the students did not always use the tabs and sentence stems as designed and intended, and instead often used the free-text box. While manual coding by a series of raters was possible for this small performance sample, future scale-up and automation efforts will require student training on the use of these features (perhaps in the context of formal instruction on problem-solving and critical thinking) and/or restriction of the free-form text interface.

While these results are based on only a limited number of groups and performances they suggest an approach around which to integrate the prior strategic models with models of the collaborative interaction. To establish a mapping between the nature of the collaborative event and the strategic problem solving approaches and validate other emerging hypotheses, we plan to continue gathering data and running dynamic analyses so that we might better understand how students' strategies might be improved through strategic collaborative learning situations.

Supported in part by grants from the National Science Foundation (ROLE 0231995, DUE 0126050, HRD-0429156). The authors would like to thank Dr. Melanie Cooper and Charles Cox at Clemson University, and Dr. Marcia Sprang (Placentia-Yorba Linda School District) and Tricia Um (MIT),for participating in pilot design and validation studies, and Dr. Marco Ronchetti at the University of Trento for guidance and support.

## References

1. Underdahl, J., Palacio-Cayetano, J., & Stevens, R., (2001). Practice Makes Perfect: Assessing and Enhancing Knowledge and Problem-Solving Skills with IMMEX Software. Learning and Leading with Technology. 28: 26-31
2. Stevens, R. H., & Palacio-Cayetano, (2003). Design performance frameworks for constructing problem-solving simulations. Cell Biology Education, 2, 162-179.
3. Stevens, R., Soller, A., Cooper, M., and Sprang, M. (2004) Modeling the Development of Problem Solving Skills in Chemistry with a Web-Based Tutor. Intelligent Tutoring Systems. Lester, Vicari, & Paraguaca (Eds). Springer-Verlag Berlin Heidelberg, Germany. 7th International Conference Proceedings (pp. 580-591).
4. Stevens, R., Wang, P., Lopo, A. (1996). Artificial Neural Networks Can Distinguish Novice and Expert Strategies During Complex Problem-Solving. Journal of the American Medical Informatics Association vol. 3 Number 2, pp. 131-138.
5. Simon, H., Kotovsky, K, & Hayes, J. (1985). Why are some problems hard? Evidence from the Tower of Hanoi. Cognitive Psychology, 17, pp. 248-294.
6. Webb, N., & Palincsar, A. (1996). Group processes in the classroom. In D. Berlmer & R. Calfee (Eds.), Handbook of Educational Psychology (pp. 841-873). New York: Simon & Schuster Macmillan.
7. Case E., 2004. The effects of collaborative grouping on student problem solving in first year chemistry. Unpublished thesis, Clemson University.
8. Giordani, A., & Soller, A. (2004). Strategic Collaboration Support in a Web-based Scientific Inquiry Environment. European Conference on Artificial Intelligence, "Workshop on Artificial Intelligence in Computer Supported Collaborative Learning", Valencia, Spain.
9. Chung, G.K.W.K, deVries, L.F., Cheak, A.M., Stevens, R.H., & Bewley, W.L. (2002). Cognitive Process Validation of an Online Problem Solving Assessment. Computers in Human Behavior, 18: 669-684.
10. Soller, A., & Lesgold, A. (2003). A Computational Approach to Analyzing Online Knowledge Sharing Interaction. Proceedings of Artificial Intelligence in Education, 2003. Australia, 253-260

11. McAlister, S., Ravenscroft, A., & Scanlon, E. (2003). Combining Interaction and Context Design to Support Collaborative Argumentation in Education. Institute of Educational Technology (IET), Open University, UK, Learning Technology Research Institute, London Metropolitan University, UK.
CALRG Report No. 204 (http://iet.open.ac.uk/research/calrg/)
12. Goodman, B., Linton, F., Zarrella, G., & Gaimari, R. (2004). Using Machine Learning to Predict Trouble During Collaborative Learning. ITS 2004 Workshop on Computational Models of Collaborative Learning, Maceio-Alagoas, Brazil. pp. 25-30.
13. Jermann, P. (2004). Computer support for interaction regulation in collaborative problem solving. Unpublished Dissertation. University of Geneva.

# TagHelper: An application of text classification technology to automatic and semi-automatic modeling of collaborative learning interactions

Carolyn ROSÉ & Pinar DONMEZ
*Carnegie Mellon University,*
*5000 Forbes Avenue, Pittsburgh PA, 15213*

**Abstract.** In this paper we introduce the objectives of the TagHelper project, which is an applied research project exploring the application of text classification technology to support corpus coding for behavioral research, including collaborative learning research. We present an overview of this problem and introduce some of our work in this area. Despite the availability of text classification technology that would be usable for this purpose from the Computational Linguistics community, none of the tools that are commonly used in the CSCL community and other behavioral research communities make use of it to support coding. We discuss a series of corpus based experiments where we explored alternative approaches to automating a multi-dimensional process analysis for characterizing collaborative learning interactions. Our results show great promise that text classification technology can have a tremendous impact on the cost of data analysis within this community. We discuss our work in the context of other work on multi-class text classification for highly skewed data sets.

## 1. Introduction

In this paper we introduce the objectives of the TagHelper project, an applied research project exploring the application of text classification technology to support corpus coding for behavioral research. A wide range of behavioral researchers including social scientists, psychologists, learning scientists, and education researchers collect, code, and analyze large quantities of corpus data as an important part of their research. For example, corpus analysis is important in educational research related to human tutoring and collaborative learning. Within this sphere, research in Computer Supported Collaborative Learning often depends upon quantitative process analyses through multi-dimensional coding schemes such as those described in (Weinberger, 2003). Other behavioral researchers, such as social psychologists studying communication in Computer Supported Cooperative Work settings, painstakingly code by hand and analyze large quantities of natural language corpus data to explore how alternative forms of computer-mediated communication affect patterns of interaction. The goal of our research is to develop technology to address concerns specific to classifying spans of text using coding schemes developed for behavioral research. The goal is to reduce the cost of research involving time intensive corpus analyses as well as potentially increasing the quality of that work.

The corpus coding process can be supported by making automatic predictions about which codes should be assigned to spans of text. Automatic text classification technology for supporting analysis of corpus data shows great potential benefit to CSCL and CSCW researchers and practitioners. Applying a categorical coding scheme can be thought of as a text classification problem where a computer decides which code to assign to a text based on a model built from examining "training examples" that were coded by hand. When the technology is able to make these judgments with an acceptable level of reliability and validity, no human intervention is required, and there is a 100% savings of human effort. Furthermore, making the judgments in an automatic way ensures total consistency of coding in a way that is not possible with human coding.

For types of judgments where the technology is not able to perform at a reliable enough level, the predictions can be checked and corrected by a human. In our previous work we have developed a coding interface that displays predictions and allows an analyst to quickly make necessary corrections through a simple menu selection (Gweon et al., submitted). This stands in contrast to typical coding support provided to analysts. Despite the availability of text classification technology that would be usable for this purpose from the Computational Linguistics community, none of the tools that are commonly used in the CSCL community and other behavioral research communities make use of it to support coding. For example, specialized software for supporting text, audio, and video data analysis such as HyperResearch, MacShapa, and Nvivo offer well designed interfaces for hand annotation and cataloguing of codes. Yet, none of them go as far as to support automatic assignment of codes. Similarly, database programs in common use for cataloguing categorized data, such as Excel or MS Access, also do not have automatic or even semi-automatic coding of textual data. Not even more flexible, free form databases, such as AskSam, support this either. In this paper, we investigate the possibility of incorporating automatic classification of text into data analysis applications widely used in CSCL research and practice.

Our evaluation of our prediction support interface demonstrates that even with a modest level of prediction accuracy (anything above 50%), analysts save time by using the interface that displays predictions that can be corrected than using an otherwise equivalent interface that does not display predictions (Gweon et al., submitted). Furthermore, even with a modest 50% prediction accuracy, on average coders using the interface produce results that agree better with each other and with a gold standard set of codes. Nevertheless, our goal is to develop technology to achieve the highest accuracy possible with the types of coding schemes that are important for behavioral research, including collaborative learning research. And our ultimate goal is to strive for total automation rather than requiring a human to check and correct the assigned codes.

In this paper we discuss a series of corpus based experiments where we explored alternative approaches to automating a multi-dimensional process analysis for characterizing collaborative learning interactions. Our results show great promise that text classification technology can have a tremendous impact on the cost of data analysis within this community.

## 2. Motivation and Research Context

Research in CSCL often depends upon quantitative process analysis through multi-dimensional coding schemes such as those described in (Fischer et al., 2002). Often only detailed process analyses reveal plausible interpretations of the effects of technological variations in computer supported collaboration environments (Weinberger, 2003). Thus, we begin by contextualizing our technological explorations within a high profile CSCL project (Weinberger, 2003).

The purpose of the multi-dimensionsal coding scheme from (Weinberger, 2003) is to model the process of argumentative knowledge construction. Argumentative knowledge construction is based on the perspective of cognitive elaboration, the idea that learners acquire knowledge through argumentation with one or more learning partners. Computer-supported collaboration scripts apply on specific dimensions of argumentative knowledge construction, e.g., a script for argument construction could support learners to ground and warrant their claims or a social cooperation script can support productive conflict resolution strategies (Weinberger, 2003). The complete process analysis from a series of studies in which these and other types of scripts were varied experimentally comprises about 200 discussions of about 600 participants with altogether more than 12,000 coded text segments. Students in all experimental conditions had to work together in applying theoretical concepts to three case problems and jointly prepare an analysis for each case by communicating via web-based discussion boards. They were asked to discuss the three cases against the background of attribution theory and to jointly compose at least one final analysis for each case, i.e. they usually drafted initial analyses, discussed them, and wrote a final analysis. The discussion boards provided a main page with an overview of all message headers, which were graphically represented in a discussion thread structure. Learners could read the full text of all messages, reply to the messages, or compose and post new messages. All of the messages that comprised the interactions between students were recorded and analyzed using the multi-dimensional coding scheme. Three groups of about six coders each were trained to apply the coding scheme to the corpus data. Between the training and the coding itself, one quarter of the total duration of the research project was used for the coding of collaborative processes.

Very little work has been done on automating this type of corpus analysis. Soller & Lesgold (2000) and Goodman et al. (to appear) present work on automatically modeling the process of collaborative learning by detecting sequences of speech acts that indicate either success or failure in the collaborative process. What is different about our approach is that we start with the raw text and detect features within the text itself that are indicative of different local aspects of the collaboration. Thus, rather than presenting a competing approach, we present an approach that is complementary.

## 3. Corpus and Coding Scheme

Many sentence classification tasks have been addressed successfully by a range of text classification approaches. For example, classifying spoken utterances into dialogue acts or speech acts has been a common way of characterizing utterance function since the 1960s, and many automatic approaches to this type of analysis have been developed (Serafin & Di Eugenio, 2004). Other applications of sentence classification technology include identifying rhetorical relations in legal documentation (Hachey & Grover, 2005) or distinguishing subjective versus objective statements (Wiebe & Riloff, 2005). Recent approaches focus on the problem of assigning sentences to classes that represent an idea that might occur within an essay (Rosé et al., 2003).

While these alternative problems vary from one another in terms of the primary type of linguistic distinction that separates the alternative classes, what these problems share is that all of the classes for a single problem are distinguished from one another by a common type of linguistic distinction. For example, in (Rosé et al., 2003), classes are distinguished by propositional content whereas in (Hachey & Grover, 2005), the classes are distinguished by rhetorical function. The advantage of this commonality is that cues that are specific to the type of linguistic distinction can be identified through careful corpus analysis and exploited by the classification algorithms. For example, in (Rosé et al., 2003), features from a deep syntactic functional analysis of the sentence were used in order to accommodate the causal nature of the propositional content. In (Hachey & Grover, 2005) sentence placement information was used as a cue for rhetorical function.

Argumentative knowledge construction must be evaluated on multiple process dimensions (Weinberger, 2003). The coding scheme we use in our experiments has 7 dimensions (See Figure 1). These dimensions are derived from different theoretical approaches and focus on different conceptualizations of argumentative knowledge construction. The main concepts are (1) epistemic activity, formal quality of argumentation, (2) microlevel argumentation, (3) macrolevel argumentation, and (4) social modes of interaction, (5) reaction to a previous contribution. In accordance with the theoretical approach, the number of categories differs between dimensions from 2 (e.g., quoted) to 35 (e.g., epistemic). For experimental reasons, there is also a (6) dimension on which student responses to script prompts are coded for appropriateness and a (7) quoted dimension, which distinguishes between new contributions and quoted contributions. In our text classification experiments, the Epistemic and Social dimensions were the most difficult with which to achieve acceptable performance, thus we examine these dimensions in greater depth in the next sections. Each sentence receives a classification on each of the seven dimensions, each of which represents a different aspect of that sentence's role in the collaborative discourse.

| Dimension | Description | Number of Classes |
|---|---|---|
| Epistemic (EPI) | How learners work on the learning task, e.g., what content they are referring to or applying | 35 |
| Micro (ATOL) | How an individual argument consists of a claim which can be supported by a ground with warrant and/or specified by a qualifier | 4 |
| Macro (ALEI) | How argumentation sequences are examined with respect to how learners connect single arguments and create an argumentation pattern together | 6 |
| Social Modes (SOC) | To what degree or in what ways learners refer to the contributions of their learning partners | 21 |
| Reaction (REA) | Coarser grained version of SOC | 3 |
| Appropriatenes s (PRO) | How learners make use of prompts, i.e., whether learners could use the prompts in the intended manner, e.g. write an argument when they are asked to write a counterargument | 4 |
| Quoted (QUO) | Distinguishes between new contributions and quoted contributions | 2 |

Figure 1 This table describes the 7 dimensions that are part of the coding scheme used in the text classification experiments reported in this paper. Notice that the dimensions vary both in terms of the type of knowledge that is brought to bear in making a judgment as well as the number of alternative categories that may be assigned on each dimension.

### 3.1 High Level Overview of Coding Scheme

When it comes to annotating corpus data for behavioral research, the coding schemes developed for it are different from those typically used in automatic sentence classification tasks within the computational linguistics community. Coding schemes used in behavioral research tend to be motivated by top-down, theoretical concerns (typically socially or cognitively motivated categories) related to the experiments providing the context for the data collection, whereas sentence classification research related to computational linguistics concerns, such as dialogue act tagging, is motivated by bottom-up, linguistically motivated observations from the collected corpus. Thus, while alternative classes within behavioral coding schemes may all be ontologically similar on a theoretical level, the types of linguistic distinctions that distinguish one class from another might vary widely between pairs of classes.

Another characteristic of many coding schemes that are motivated top-down by theoretical concerns is that they tend to be very highly skewed, containing many infrequently occurring but nevertheless important classes. Hence, automatic classification approaches often achieve high percent agreements (i.e., in the high 90s) and low or even negative kappa values (as a measure of reliability of coding). Thus, addressing concerns related to text classification performance on highly skewed data sets is an important part of our research agenda.

In the remainder of this section we discuss challenges specifically related to aspects of the coding scheme that we worked with for the experiments we present in our evaluation below. As mentioned, we found that the Epistemic and Social dimensions of the Weinberger coding scheme were the most challenging.

### 3.2 The Epistemic Dimension

On the epistemic dimension, argumentative knowledge construction processes are analyzed with respect to questions of how learners work on the learning task, e.g., what content they are referring to or applying. As part of the analysis on this dimension, each text span was assigned one of 35 separate categories, each of which indicated something about the content communicated. 18 of these categories occurred very infrequently in the corpus, where by infrequent we mean having 10 or fewer instances in the corpus. One important distinction on the epistemic process dimension is to what extent learners stay on topic or digress off task. In order to solve a problem, learners may need to talk about parts of the case study, concepts from the theory, and connections between these two sets of ideas. With an exploration of the case study itself, learners are to acquire an understanding of the problem they are supposed to work on. As they relate parts of the case study to the theory, they demonstrate their understanding of the theory. Learners connect individual theoretical concepts or distinguish them from another. 30 of the classes in the Epistemic dimension represented specific content connecting evidence from case studies with concepts from the theory students were applying to their analysis. Thus, the problem of distinguishing between these categories takes on a topic detection flavor similar to that presented in (Rosé et al., 2003). However, the remaining categories were very differently construed. For example, one category was for off-topic conversation, and another was for arguments that make use of reasoning not specifically related to the theory. These classes were defined by what they were not rather than what they were and did not refer to a specific topic or idea. From a linguistic standpoint, these categories are quite different from the other 30. Nevertheless, ontologically they all belong on the epistemic dimension.

### 3.3 The Social Dimension

The social modes dimension indicates to what degree or in what ways learners refer to the contributions of their learning partners. In this dimension there are 21 separate categories, seven of which have 10 or fewer instances in the corpus. Learners may explicate their knowledge, e.g., by contributing a new analysis of a problem case. Externalizations are discourse moves that do neither refer to preceding contributions of peers nor aim to elicit information from the learning partners. Learners may use the learning partner as resource and seek information (elicitation) in discourse from the learning partners in order to solve a problem case. Learners need to build at least a minimum consensus regarding the learning task in a process of negotiation in order to improve collaboration. There are different styles of reaching consensus, however. Quick consensus building means that learners accept the contributions of their learning partners not in terms of taking over his or her perspective, but in order to be able to continue discourse.

Recent approaches towards collaborative learning stress that collaborative learners may eventually establish and maintain shared conceptions of a subject matter (integration-oriented consensus building). Learners approximate and integrate each others perspective, synthesize their ideas and jointly try to make sense of a task. Conflict-oriented consensus building has been considered an important component in the socio-cognitive perspective upon collaborative learning.

## 4. Evaluation

Coding schemes developed for behavioral research tend to be multi-class coding schemes, sometimes with multiple dimensions with multiple classes on each, similar to that used in our experiments reported in this paper. We consider a multi-class classification problem to be one where a training set consists of data points, each of which is assigned to one of k different classes. The goal is to predict the correct class label for a given new data point. There are two separate, and yet related general schemes for multi-class classification: one-against-all approaches and pairwise approaches. In both cases, the multi-class problem is broken down into multiple binary classification problems, and the solutions are then combined. What distinguish the two is how the problem is divided into sub-problems, and how these sub-solutions are combined later. The two new approaches that we evaluate in this section are variations on the on-against-all approach.

We used as a test-bed the Minorthird text-learning toolkit (Cohen et al, 2004), which contains a large collection of configurable machine learning algorithms that can be applied to text classification tasks, as a framework in which to conduct our research. Because Minorthird includes a wide range of text classification algorithms that all operate over text coded in the same format, it is a convenient test environment for experimentation. We measure our success in terms of agreement with the hand-coded gold standard corpus in terms of the Kappa statistic since it is accepted by our target user population (i.e., behavioral researchers) as a standard for coding reliability.

We compared a wide range of basic classifiers in order to establish a baseline with which to compare the performance of our new techniques. We selected baseline classifiers known to perform well with highly skewed data sets (Yang & Liu, 1999). It's not necessary for the reader to understand the technical details of what distinguishes these different baseline classification algorithms. Specifically, we tested the K-NN method proposed by Yang and Chute (1994) and Voted Perceptron (Freund & Schapire, 1998) using a One-versus-All strategy (Fuerncratz, 2002; Allwein et al., 2000), in which a separate binary classifier is learned for each category, and the final prediction is made by picking the most confidently-predicted class. The class of a test item is based on a weighted vote of the 30 closest neighbors in this space. An advantage of this learning method is that it can efficiently handle many classes. We tested this approach using 10-fold cross-validation, which is a standard practice in machine learning research that allows us to use our data efficiently, while keeping separate what is used for training and what is used for testing. We went through this process separately for each of the 7 dimensions. The results are presented in Figure 2. The K-NN classifier only achieved an acceptable level of agreement with the gold standard in the case of REA, achieving a Kappa of .81. Note that the voted perceptron using a one-versus-all strategy achieved an acceptable level of agreement with the gold standard on 4 out of the 7 dimensions.

Next we evaluated two techniques alone and in combination for improving the performance of the standard voted perceptron algorithm. The first new approach is what we call Confidence Restricted Cascaded Binary Classifiers (CR-CBC), which operate within single dimensions and yield acceptable kappa measures (.7 or above) over 1 of the remaining 3 dimensions in our coding scheme. 88% of the sentences or more are assigned a code within each dimension based on a confidence rating. Furthermore, it achieves a boarderline acceptability in an additional dimension (kappa of .68) over 50% of the sentences, selected according to a confidence rating. In combination with the first approach, the second technique, which we refer to as Cross-Dimensional Constraint Classifiers (CD-CC), achieves an acceptable level of agreement (kappa of .7) over 80% of the sentences the final remaining dimension and maintains a near acceptable kappa (kappa of .69) over 67% of the sentences along an additional dimension. We measure our success in terms of the Kappa statistic since it is the accepted standard measure for coding reliability among behavioral researchers.

We evaluated the performance of our first proposed approach, namely Confidence Restricted Cascaded Binary Classifiers (CR-CBC). For each dimension, we trained a one-versus-all classifier for each class and then rank ordered the binary classifiers for the categories within each dimension according to their accuracy (in terms of kappa) over a validation set. We then applied them in rank order to the test set, selecting as an assigned code the first binary classifier that indicated a positive match for an example text. We compared our results with the baseline performance using the nonbinary version of voted perceptron algorithm, which is known to perform well with highly skewed data sets (Yang & Liu, 1999). See Figure 2 below.

In all cases except SOC, EPI and ALEI our approach achieved a substantially higher accuracy than the baseline technique in terms of kappa statistic. In fact, our data has multiple interrelated dimensions along which classifications must be made. Thus, correlations between classes might be used to improve predictions for these challenging dimensions. Based on this hypothesis, using a form of cross-training (Sarawagi et al., 2003) we add the predicted labels (obtained from cascaded binary classification (CR-CBC) by means of cross validation) as features into our training and testing corpus, and then built and evaluated the new classifiers on them. We will call this technique Cascaded Binary Cross-dimensional Classification (CB-CDC). We used a two level form of cross-validation evaluation in order to avoid testing on training data. The results are shown in Figure 2 where we achieved a kappa of .72 within 80% of the data receiving non-null classification for EPI, and a kappa of .69 within 63% of data for SOC and a kappa of .85 within 96% of data for ALEI. While it was necessary to restrict the portion of the corpus that we committed a code two on two dimensions in order to achieve an acceptable level of reliability for the coding that we did, this would still result in a very substantial savings of coder time. First, coders would only have to be trained on 2 out of 7 dimensions, which would be a savings of 71% of training time. Furtheremore, only 20% of one dimension and 37% of another dimension would need to be checked and corrected. Thus, at least 83% of coding effort could be saved.

| Dimension | NonBinary kNN | NonBinary VotedPerc | CR-CBC | CD-CC | CR-CBC + CD-CC |
|---|---|---|---|---|---|
| EPI | .51 | .55 | .49/.53 (43%) | .63 (80%) | .72 (80%) |
| ATOL | .54 | .6 | .76/.83 (92%) | n/a | n/a |
| ALEI | .54 | .7 | .67/.7 (88%) | n/a | .83/.85 (96%) |
| SOC | .35 | .5 | .55/.68 (50%) | .59 (80%) | .69 (63%) |
| REA | .81 | .84 | n/a | n/a | n/a |
| QUOTE | .63 | .91 | .98 | n/a | n/a |
| PRO | 0 | .7 | .73 | n/a | n/a |

Figure 2 Summary of results on experiments in terms of agreement with gold standard, measured with kappa. Numbers in parentheses indicate percentage of sentences in corpus that received a non-null classification. n/a indicates that we did not yet evaluate the associated approach on the associated dimension.

## 5. Conclusions and Current Directions

In this paper we describe the goals and early work of the TagHelper project, an applied research project in which we explore the application of automatic text classification technology to supporting the analysis of corpus data. Our early results show great promise that such technology could be used to substantially reduce the cost of behavioral research involving corpus analysis.


## Acknowledgements

This work was funded through NSF grant SBE0354420.


## References


Allwein, E., Schapire, R. & Singer, Y. (2000). Reducing multiclass to binary: A unifying approach for margin classifiers. Journal of Machine Learning Research, 1:113-141, 2000.

Cohen, W. (2004). Minorthird: Methods for Identifying Names and Ontological Relations in Text using Heuristics for Inducing Regularities from Data, http://minorthird.sourceforge.net.

Fischer, F., Bruhn, J., Gräsel, C., & Mandl, H. (2002). Fostering collaborative knowledge construction with visualization tools. Learning and Instruction, 12, 213-232.

Freund, Y. & Schapire, R. (1998). Large Margin Classification Using the Perceptron Algorithm, Computational Learning Theory, pp209-217 http://citeseer.nj.nec.com/freund98large.html .

Fuernkranz, J. (2002). Round Robin Classification, Journal of Machine Learning Research 2, pp 721-747.

Goodman, B., Linton, F., Gaimari, R., Hitzeman, J., Ross, H., & Zarrella, J. (to appear). Using Dialogue Features to Predict Trouble During Collaborative Learning, to appear in the Journal of User Modeling and User Adapted Interaction.

Gweon, G., Rosé, C. P., Wittwer, J., & Neuckles, M. (submitted). Supporting Efficient and Reliable Content Analysis using Automatic Text Processing Technology, submitted to INTERACT '05.

Hachey, B. & Grover, C. (2005). Sequence Modeling for Sentence Classification in a Legal Summarization System, Proceedings of the 2005 ACM Symposium on Applied Computing.

Rosé, C., Roque, A., Bhembe, D., VanLehn, K. (2003), A Hybrid Text Classification Approach for Analysis of Student Essays , Proceedings of the HLT-NAACL 03 Workshop on Educational Applications of NLP

Sarawagi, S., Chakrabarti, S. & Godbole, S. (2003). Cross training: learning probabilistic mappings between topics. In Proc. of the Nineth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD-2003), Washington D.C., USA, August 2003.

Serafin, R. & Di Eugenio, B. (2004). FLSA: Extending Latent Semantic Analysis with Features for Dialogue Act Classification, Proceedings of the Association for Computational Linguistics

Soller, A. & Lesgold, A. (2000). Modeling the Process of Collaborative Learning, Proceedings of the International Workshop on New Technologies in Collaborative Learning, Awaiji-Yumebutai, Japan.

Weinberger, A. (2003). Scripts for Computer-Supported Collaborative Learning Effects of social and epistemic cooperation scripts on collaborative knowledge construction, from http://edoc.ub.uni-muenchen.de/archive/00001120/01/Weinberger_Armin.pdf

Wiebe, J. & Riloff, E. (2005). Creating Subjective and Objective Sentence Classifiers from Unnanotated Texts, Proceedings of CICling 2005.

Yang, Y. and Liu, X. (1999). A re-examination of text categorization methods. Proceedings of ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'99, pp 42--49).

Yang, Y. & Chute, C. (1994). An example-based mapping method for text classification and retrieval, ACM Transactions on Information Systems 12(3), 1994.

# Characteristics of Asynchronous Online Mathematics Help Environments: Do They Provide Conditions for Learning?

Dragana MARTINOVIC
*Ontario Institute for Studies in Education of the University of Toronto*
*252 Bloor Street West, Toronto, ON M5S 1V6, Canada*

**Abstract.** This qualitative and quantitative study intended to define and explore the characteristics of asynchronous online mathematics help environments and look for some evidence that they provide conditions for learning. Data for the study were collected from online contributions on three purposely selected, public mathematics help sites; written reflections of expert tutors; and interviews with them. The study looked into the tutorial discourse with the intent to characterize communicative goals, form, cognitive level, hedging and degree of specification of students' questions in mathematics online help; and to distinguish what and how tutors in mathematics online help teach, do they hedge, and how helpful are their answers. The study showed that users of this service have specific goals; that peer tutors and expert tutors teach differently; and that carefully designed discursive environment provides conditions for learning mathematics.

## Introduction

The analysis of online help opens avenues for deeper investigation of online tutorial discourse and the educational gains it offers. The objective of this study was to analyze public mathematics online help environments with text-based, asynchronous communication, and open access to archived questions and answers. The reasons for focusing on such sites were: (a) Anonymity provides conditions for honest communication that is free of stereotyping; (b) bulletin board system provides the fast and cheap access; (c) asynchronous communication enables reflection, vicarious learning, and more time for a tutor to respond; and (d) public sites (with free access) are more likely to attract a diverse population of visitors.

## 2. Research Question and Methods

The main research question for the present study was: What are the characteristics of asynchronous online mathematics help environments and do they provide conditions for learning? Following Strauss and Corbin [23] recommendation to apply qualitative methods for exploring a little known phenomenon, gaining new perspectives on known phenomena, or gaining more in-depth information that may be difficult to express quantitatively, the mixture of qualitative and quantitative research methods were used here. More quantitative approach was used to analyze random samples of 200 threads from each of the three purposely selected Web sites, where: Site A offered both expert and peer tutoring; site B only peer, and; site C only expert tutoring. In particular, on site A there was only one expert tutor who also monitored correspondence between peers, while archives on site C contained only questions recommended by the tutors. More qualitative approach was applied on interviews with five expert tutors from two expert tutoring sites and tutoring logs obtained from four of the tutors. Analysis of the messages posted on these Web sites was based on the taxonomies of tutorial discourse developed by the researchers in Intelligent Tutoring Systems: Graesser, Person and Huber [9]; Brandle and Evens [1]; Shah et al. [22]; and refinements of some specific categories used in Hume et al. [11] and Kim [14]. Categories that were relevant for the present research problem were selected, further extended to fit the data, and some new were added. The site with both expert and peer tutoring was initially used in order to exercise the methodology of the research, finalize the Taxonomy of Online Tutorial Discourse [16], achieve the reliability of coding done by two independent coders, and to establish the comparisons between sites. Chi's Verbal Data Analysis [2] was applied in order to quantify qualitative data collected from the Web sites. In the quantitative part of the study, online communication was segmented after reviewing the whole question (answer) and eliminating non-content verbalizations, where the segments were groups of sentences and/or concepts directly related to the cognitive level, communicative goal, delivery mode, hedging, acknowledgement, and degree of specification. Each code contained the statement number (the order in the sample), the category (categories), and the level on ordinal scale (if applicable) measured by the researcher and one independent rater. Most of the measures used here were the relative frequencies of the observed contextual values and the chi-square measure of the dependency between variables. While the observation of communication on three Web sites provided uncontrolled data, the rest of the data were collected through contacts with those who have the most hands-on experience with mathematics online help – online tutors. Interviews with five expert tutors and 20 logs from four of them were used in the qualitative part of the study. Such process of data and method triangulation provided for a fuller understanding of mathematics online help.

## 3. General Results

Being based on asynchronous computer-mediated communication (CMC), mathematics online help shares its benefits and its problems. For example, Davie and Inskip [5] reported that CMC suffers from being text-based, missing visual clues, temporal delay in communication and taking too long for the exchange to finish. Although asynchronous online communication is much less efficient than face-toface communication, analysis of the time it took for the first tutor to respond showed that, in the majority of cases, tutors reacted quickly. The first response was usually given the same day the question was posted, or the day after. On the peer tutoring site B, 75% (149, n = 200) of the first messages in a thread were answered the same day and 23% (45) were answered the next day; while on the expert tutoring site C, 60% (119, n = 200) of the first messages in a thread received a response the same day and 29% (58) the next day. Also, in most threads it took one to two days for communication to finish. Specifically, exchange among peers lasted on average shorter than one with expert tutors (1.73 days on site B, compared to 2.22 days on site C). Almost half (96, 48%, n = 200) of the threads on the peer tutoring site contained more than one answer and more than one question, while on the expert tutoring site there were less than one-third of such (56, 28%, n = 200). Therefore, the transactional distance as psychological space between the learner and the teacher [13] was smaller on the peer tutoring site B. Also, a sense of community was stronger on this site which was obvious through greetings like: "Hi All," "Greetings All"; through inquiries like: "Can somebody help?" and "Anybody, please?" and through more informal speech where there were much more social topics than on the expert tutoring site. The first person pronouns (i.e. "I" and "we"), which "indicate the author's personal involvement with the activity portrayed in the text" [17, p. 5] were almost four times more frequent in expert than in peer tutors' answers (5.56 on site C vs. 1.42 per answer on site B). Experts used "we" much more often than peers (3.09 vs. 0.45 per answer), which probably suggests that they were speaking with the authority of the mathematical community [17]. Both experts and peer tutors used the second person pronoun more frequently than any first person pronoun, which "may indicate a claim to relatively close relationship between author and reader or between reader and subject matter" [17, p. 6].

### 3.1 Online Medium for Providing Help

Users of mathematics online help sites need to know how to use a browser and how to search a Web site. All three analyzed Web sites provided mathematics help in the form of a bulletin board with options to search through the archives or post a new question. The sites also had some additional resources that visitors could access by selecting a hyperlink. Although these features seem easy to use, there were students who could access the bulletin board, but did not know how to post a question. It was obvious that recommendations given by Tolhurst [24] and Reushle [20] regarding the design of hypermedia resources were not fully met, since some students were left to discover the features of the sites without enough guidance for them as novices. For example, one student posted this message as part of the existing thread: "can anyone help me [sic] is anyone even out there [sic] i [sic] don't really know how to use this program" (site A). On sites A and B, some students posted questions as answers because it was not clear to them how to start a new thread. All three online help sites were of the exploratory type [12], designed so that users can explore and find their own way around. A user needed to know what to look for in order to find it, which can be a disadvantage when the user does not know the exact name of the mathematical discipline, or which terms to use as search words. Either because visitors were left unguided; or some were not interested enough to explore; or some could not make the transition between the information provided in the resources and their questions, there were a lot of similar questions asked on each of three online help sites (as evidenced from the archived questions on sites A and B, and interviews with expert tutors from site C).

### 3.2 Online Help as Environment for Communicating Mathematics

The difficulty to effectively communicate mathematics online was both pointed out by the tutors in their interviews and observed in communication on all Web sites. Text-based communication has little means for presenting graphs, diagrams, and tables. Both tutors and students suffered from an inability to use proper mathematical symbols and sometimes had to put in extra effort to use text editing capabilities for visual presentations. One of the tutors explained how an otherwise simple task becomes tedious and long in an online environment: It's very difficult to explain mathematical ideas without actually drawing them and because it is only using the characters on the keyboard, it's also very hard to show, for example long division. Trying to show long division to somebody by pressing the keys on the keyboard is a little challenging. It also takes about four times the amount of space as if you did it with a piece of paper and pencil in front of you. (Beth, interview) This was certainly one of the reasons that some posted questions (5%, n = 251, Mixed site A; 3%, n = 249, Peer site B) were incomplete or unclear (Fragments). Expert site C did not contain Fragments, probably because tutors there chose which questions to archive.

### 3.3 Mathematics Language and Social Dimensions on Online Help Sites

Many factors in face-to-face communication that can cause passivity in students [8, 28] do not exist in an online help environment. Partial anonymity and taking charge of their own learning [21] provide a secure setting for the students. However, evidence from the Web sites (see Table 1) shows that many students still do not engage in a fruitful mathematics communication, but provide telegraphic messages without any context. As a result, almost every tenth (9%) question on the expert tutoring site was Implicit (consisting just of a text of mathematical problem) and almost every third (31%) was a Low Specification question (not helpful enough for the tutor), while these percentages were even higher on the other two sites.

**Table 1.** *Percentages of Implicit and Low Specification Questions on Sites A-C*

| Questions | Mixed Site (A, n = 251) | Peer Site (B, n = 249) | Expert Site (C, n = 264) |
|---|---|---|---|
| Implicit | 29% | 28% | 9% |
| Low Specification | 63% | 61% | 31% |

Peer tutors too, on both sites A and B, could have been more helpful in one quarter to one-third of their answers respectively (see Table 2). Experts on site C were most helpful as less than one-tenth of their answers were Low Specification, thus providing information without any explanation, rationale, or example, which may be difficult for students to follow.

**Table 2.** *Percentages of Low Specification Answers on Sites A-C*

| Answers | Mixed Site (ET) (A, n = 53) | Mixed Site (PT) (A, n = 212) | Peer Site (B, n = 341) | Expert Site (C, n = 232) |
|---|---|---|---|---|
| Low Specification | 23% | 25% | 34% | 9% |

As noted by English and Yazdani [6], students can feel lost in the absence of the person of trust. That was obvious on the peer tutoring site where the following message was posted:

Sorry [*tutor*] tried that and its [*sic*] wrong!! (probably because its [*sic*] e^something^something) Can ne [*sic*] one else help??? (Student, site B)

Two related problems were observed here: (a) Answered question might leave the student unsatisfied, but the thread can be skipped by other tutors who may consider the job to be done, and (b) answers may be wrong. Since public communication adds to the credibility of provided assistance [27], answers in online help seem to have communal approval. Many cases when wrong answers get corrected very quickly can give the false impression that all answers that are not corrected are accurate, which is not true. Even on the expert tutoring site some answers were not corrected until a long time (even years) after they were first posted.

### 3.4 Mathematics Discourse on Mathematics Online Help Sites

Students often do not see mathematics as a subject where discussion is welcome [18]. In school they get used to a common practice of "path-following" [15] and trying to reproduce [4] the "right way" of solving the problem. This was addressed by the expert tutors:

I think watching somebody else solve the problem is not a good way at all for doing math, and yet that's what students are expecting. (Morgan, interview) [Online students] do not want two pages to go into to figure what they want to find out, unfortunately…. They kind of want answers right away and that's not necessarily helping them. (Beth, interview)

All interviewed tutors followed "we are not doing your homework" policy and used hinting in their answers. They would provide the necessary guidelines for the students to get them started and then they would wait to see if students required more help. Hinting is certainly the most efficient delivery mode for online tutors, since it enables them to answer more questions in the same amount of time–they do not have to do the mathematics problem themselves and do not have to bother with typing in long messages. Crossreference of data regarding hinting on Web sites A-C (see Table 3) shows that experts used hinting more than peer tutors. A relatively low percentage of hints in expert tutor answers on site A (Mixed Site, ET) probably indicates that the expert tried to be more elaborate in order to educate not only the students, but the peer tutors as well.

**Table 3.** *Percentages of Hints in Answers on Sites A-C*

| Answers | Mixed Site (ET) (A, n = 53) | Mixed Site (PT) (A, n = 212) | Peer Site (B, n = 341) | Expert Site (C, n = 232) |
|---|---|---|---|---|
| Hinting | 38% | 29% | 41% | 72% |

Although researchers [18, 19, 25] point to the value of a debate in mathematics class, there are cases when learning is most effective when relevant and direct instruction is provided [10, 15]. Especially when learners are novices in the field, direct answers or trails that they can follow can help them to build their confidence. Also, students who reach a certain mastery level in the field may require only direct instruction [7]. These results show that even straight answers or answers with a low level of specification can be useful for some students. However, not knowing enough about the student or the origin/appropriateness of the question, tutors sometimes make an extra effort to ensure that their answers are really helpful. They do it by providing alternative (Multiple) answers. In mathematics online help, alternative answers in one message make the message longer and for the inexperienced or unmotivated student more difficult to comprehend. On the other hand, they make communication more efficient, since the tutor does not have to wait (often in vain) for the student to improve the question but tries to anticipate possible problems in advance and answers in alternative ways. This shows students that there exists a variety of optional paths that they can take in solving some problem and presents mathematics as a discursive subject. As can be seen from Table 4, peer tutors used Multiple Answers much less than expert tutors.

**Table 4.** *Percentages of Multiple Answers on Sites A-C*

| Answers | Mixed Site (ET) (A, n = 53) | Mixed Site (PT) (A, n = 212) | Peer Site (B, n = 341) | Expert Site (C, n = 232) |
|---|---|---|---|---|
| Multiple | 11% | 3% | 2% | 7% |

A relatively high percentage of expert tutor's Multiple Answers on Mixed site A can be explained through the expert tutor's need to educate and serve as an example to peer tutors. At the same time, about 41% of threads on the peer site contained more than one answer with about 30% of answers coming from the second, third, and following tutors. As different tutors were likely to answer the same question in different ways, the final effect for the students would be similar as to that on the expert site. Chi [3] cautioned that poor learners think they understand most of the time when in fact they do not, implying that they do not detect any conflicts between their understanding and what the text says. Chi further noted that self-repairing requires the successful detection of comprehension failure, or some conflict between what the student thinks is going on versus what the text is presenting. Otherwise, when students do explain examples to themselves, they seem not to be aware of their ignorance, so their knowledge gaps persist and cause further errors [26]. In the present study, taking the student's acknowledgement in the context of the whole message revealed if the student's feedback was just general (a "thank you" note) or it also provided some evidence that the student accepted an answer, understood it (Comprehended) and was able to put it to some use (Fully Comprehended). A comprehended type of acknowledgement would be: "Thank you for your reply, you it [*sic*] explained it so I could understand and I am grateful" (Site A). A Fully Comprehended type contains more references to the answer, as in: "Thank you. What I had backwards was the order of composition. I was thinking you had to do F(x) first and plug it into G(X), not vice versa. Doing G(X), then F(X), set me straight" (Site A). Acknowledgements in which students expressed satisfaction with an answer and how such answers affected them, show that the correspondence on the Web sites resulted in actual learning gains for some (see Table 5). However, the problem is that not too many students engage in a longer communication with a tutor and rarely do they provide elaborate feedback to the tutor.

**Table 5.** *Students' Comprehension across the Sites A-C*

| Questions | Mixed Site (A, n = 251) | Peer Site (B, n = 249) | Expert Site (C, n = 264) |
|---|---|---|---|
| Comprehended or Fully Comprehended | 8 (3%) | 5 (2%) | 27 (10%) |

While the low percentage of Comprehended questions on the peer tutoring site can be the consequence of the culture of the site – students simply might not be used to providing feedback to peer tutors; when one compares the data from all three sites the morelikely cause is probably somewhere else. Absence of the person of trust [6] may leave thestudents uncertain if the help was successful and if they really resolved the matter that caused them to ask for help in the first place.

## 4. Conclusions and Recommendations

The objective of this study was to analyze public, text-based, asynchronous, mathematics online help environments and find if they provide conditions for learning. The present nature of such communication is that:
(a) It is more complex than faceto- face communication since a simple two-step dialogue process, becomes a multi-step process in a written form;
(b) It takes patience and skills to write a formula, use mathematical symbols, or sketch a diagram using only symbols on a keyboard;
(c) The whole exchange may take too long to finish, and in the meantime students can lose interest or hope and can turn to other topics or resources;
(d) Tutors' answers have certain credibility in public communication, so errors in tutors' answers that do not get repaired can cause students to lose trust in this service; and
(e) Not all the questions get answered.
All analyzed sites were of the exploratory nature, thus letting students find their way around, which evidently not all of them did. Also, students' poor writing skills emerged as an important deficiency in the text-based communication, which was obvious from the number of spelling and grammatical errors; as well as the number of incomplete and unclear questions.
Mathematics online help is more efficient on sites with peer than with expert tutoring, as peer tutors need less time than experts to respond to a question and to finish communication with a student. There is a stronger sense of community and a smaller transitional distance among peers. On the other hand, based on the students' feedback, the educational gains are more evident on the sites that involve experts. Overall, peer tutoring sites may be used differently than expert tutoring sites. If one needs a quick answer to a specific, typical question, peer tutoring sites are certainly more beneficial. If, one needs advice, reinforcement of ideas, or an answer to some esoteric question, expert tutoring sites are more useful. Potential benefits to users of mathematics online help sites are:

1. The sole act of coming up with a question can be the result of self diagnostics in terms of locating knowledge gaps or contradictions in reasoning.
2. One can learn from answers by adopting the procedure and a "way of thinking."
3. Since not all visitors ask questions, but they browse (the majority of sites advise their visitors to try to find the answer to a similar question, before they propose one), this provides conditions for vicarious learning. Witnessing how others think and communicate mathematics is an important factor of learning mathematics language, learning how to ask questions, and how to answer them.
4. Different approaches in answering similar mathematical questions enrich the student's process of self-explaining and prove that there is not only one way of solving problems.
5. Sharing anxiety with others provides comfort and ease.

The previous analysis also pointed to some problems with mathematics online help. The following improvements could be incorporated into this service: (a) More guidelines for the users (how to use the site, how to start a thread, etc.); (b) Better search options – as a student submits the question, the search engine can look for the similar resources on the site and make a list of suggestions; (c) Option for students to express their satisfaction with the answer – a flag can alert the tutors that this thread is still open for the discussion; (d) Tools for editing mathematics text and making drawings; and (e) Syntax analysis of entered mathematics text can point to errors and omissions, and increase efficiency of communication on the sites. With the spread and further development of online communication technologies, more students and tutors will become used to them. If properly used, mathematics online help can assist students to reflect, self-explain, and build their confidence. It can give them the opportunity to also provide assistance to others. As such, online help has value as a learner support feature and should warrant more recognition from institutions at all levels of schooling.

**References**

[1] Brandle, S., & Evens, M. (1997). Acknowledgments in tutorial dialogue. In E. Santos (Ed.), *Eighth Midwest AI and Cognitive Science Conference* (pp. 13-17), Dayton, OH, 1997, Menlo Park, CA: AAAI Press.
[2] Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences*, *6*(3), 271-315. Lawrence Erlbaum Associates, Inc.
[3] Chi, M. T. H. (1998). Self-explaining: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp.161-237). Mahwah, NJ: Erlbaum. [4] Crawford, K. P., Gordon, S., Nicholas, J., & Proser, M. (1993). Learning mathematics at university level. In W. Atweh (Ed.), *Contexts in mathematics*. *The Proceedings of the Mathematics Education Research Group of Australasia* (pp. 209-214). Annual conference. Brisbane, July 1993.
[5] Davie, L. E., & Inskip, R. (1992). Fantasy and structure in computer-mediated course. *Journal of Distance Education*, *Fall 1992*, *VII* (2), 31-50.
[6] English, S., & Yazdani, M. (1999). Computer-supported cooperative learning in a virtual university. *Journal of Computer Assisted Learning*, *15*, 2-13.
[7] Flanders, N. (1970). *Analysing teacher behaviour*. London: Addison-Wesley.
[8] Good, T. L., Slavings, R. L., Hobson-Harel, K., & Emerson, H. (1987). Student passivity–A study of question asking in K-12 classrooms. *Sociology of Education*, *60*, 181-195.
[9] Graesser, A. C., Person, N. K., & Huber, J. (1992). Mechanisms that generate questions. In T. Lauer, E. Peacock, & A. Graesser (Eds.), *Questions and information systems* (pp. 167-187). Hillsdale, NJ: Erlbaum.
[10] Hammond, N. (1993). Learning with hypertext: Problems, principles and prospects. In C. McKnight, A. Dillon & J. Richardson (Eds.), *Hypertext: A psychological perspective* (pp. 51-69). Chichester, West Sussex: Horwood.
[11] Hume, G. D., Michael, J., Rovick, A., & Evens, M. W. (1996). Hinting as a tactic in one-on-one tutoring. Learning Sciences, *5* (1), 23-47.
[12] Jonassen, D. H. (1993). Effects of semantically structured hypertext knowledge bases on users' knowledge structures. In C. McKnight, A. Dillon & J. Richardson (Eds.), *Hypertext: A psychological perspective* (pp. 153-168). Chichester, West Sussex: Horwood.
[13] Jones, D. (1996). Computing by distance education: Problems and solutions. In G. Davies (Ed.), *Proceedings of the First Conference on Integrating Technology into Computer Science Education* (pp 139- 146). Barcelona, Spain: ACM Press.
[14] Kim, J. H. (2000). Natural language analysis and generation for tutorial dialogue. Doctoral dissertation, Graduate College of the Illinois Institute of Technology, Illinois,

USA. Retrieved from http://www.csam.iit.edu/~circsim/documents/kjhdiss.pdf
[15] Nemirovsky, R., & Monk, S. (2000). "If You Look At It the Other Way…" An Exploration Into the Nature of Symbolizing. In P. Cobb, E. Yackel & K. McClain (Eds.), *Symbolizing and communicating in mathematics classrooms: Perspectives on discourse, tools, and instructional design* (pp. 177-221). Mahwah, NJ: Erlbaum.
[16] Martinovic, D. (2004). Communicating Mathematics Online: The Case of Online Help. Doctoral dissertation. University of Toronto, Toronto, Canada.
[17] Morgan, C. (1996). The language of mathematics. Towards a critical analysis of mathematical texts. *For the Learning of Mathematics*, *16* (3), 2-10.
[18] Pimm, D. (1987). *Speaking mathematically*: *Communication in mathematics classrooms*. London: Routledge & Kegan Paul, Ltd.
[19] Resnick, L. B. (1988). Treating mathematics as an ill-structured discipline. In R. I. Charles & E. A. Silver (Eds.), *The teaching and assessing of mathematical problem solving* (pp. 32-60). Hillsdale, NJ: Erlbaum.
[20] Reushle, S. (1995*).* Design considerations and features in the development of hypermedia courseware. *Distance Education*, *16* (1), 141-156.
[21] Santoro, G. (1995). What is computer-mediated communication? In Z. Berge & M. Collins (Eds.), *Computer-mediated communication and the online classroom: Overview and perspectives* (pp. 11-27). Cresskill, New Jersey: Hampton Press.
[22] Shah, F., Evens, M., Michael, J., & Rovick, A. (2002). Classifying student initiatives and tutor responses in human keyboard-to-keyboard tutoring sessions. *Discourse Processes*, *33* (1), 23-52.
[23] Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage Publications, Inc.
[24] Tolhurst, D. (1992). A checklist for evaluating context-based hypertext computer software. *Educational Technology*, 32 (3), 17-21.
[25] Usiskin, Z. (1996). Mathematics as language. In P. C. Elliott & M. J. Kenney (Eds.), *Communication in mathematics, K-12 and beyond, (1996 Year Book*) (pp. 231-243). Reston, VA: Author.
[26] VanLehn, K., & Jones, R. M. (1993) What mediates the self-explanation effect? Knowledge gaps, schemas or analogies? In M. Polson (Ed.), Proceedings of the 15th Annual Conference of the Cognitive Science Society (pp. 1034-1039). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
[27] Wells, C. (1994). Communicating mathematics: Useful ideas from computer science. Retrieved from http://www.cwru.edu/artsci/math/wells/pub/pdf/commath.pdf.
[28] Willson, J. (1999). *High and low achievers' classroom interaction patterns in an upper primary classroom*. AARE, Melbourne, Australia, November/December, 1999. Retrieved from http://www.aare.edu.au/99pap/wil99741.htm.

# Analysing Learner Interaction in an Adaptive Communication Tool

Agoritsa GOGOULOU, Evangelia GOULI and Maria GRIGORIADOU Department of Informatics & Telecommunications,
University of Athens,
Panepistimiopolis,
Ilissia,
Athens 15784,
Greece
rgog@di.uoa.gr, lilag@di.uoa.gr, gregor@di.uoa.gr

**Abstract**. This paper presents an approach to analyse learner interaction, both quantitatively and qualitatively, in the context of a synchronous adaptive communication tool, named ACT. The quantitative analysis aims to provide information about learners' contributions to the dialogue and the results are presented in a graphical form. _he qualitative analysis aims to exploit various attributes regarding learners' contributions and investigate their collaboration behavior in terms of collaboration indicators. The analysis focuses on (i) the cognitive skills that learner develops with respect to the learning outcomes addressed by the activity, and (ii) the behavior that learner exhibits in promoting the collaboration by initiating/stimulating and advancing the discussion. The teacher has the possibility to personalize the analysis process by defining weights to the attributes analysed denoting this way their importance with respect to the underlying activity. Indicative cases from the formative evaluation of the ACT tool illustrate and explain the proposed collaboration indicators

## Introduction

Towards the direction of supporting learners during their collaboration, research efforts attempt to either structure the collaboration or regulate the collaboration or both [3]. In this context, the automatic analysis of learners' interaction is at the forefront of research in the field of Computer Supported Collaborative Learning (CSCL) and concerns learners' dialogue if synchronous or asynchronous text-based communication is supported and/or learners' actions if shared group workspaces are available. The analysis process includes (i) data selection regarding learners' contributions (messages and/or actions) and (ii) application of processing methods in order to aggregate the selected data and produce one or more *indicators* that indicate the 'quality' of the individual activity, the 'quality' of the collaboration, and/or the quality of the collaborative product [8].

Barros & Verdejo in their asynchronous newsgroup-style system called DEGREE [2], use the structured form of the dialogue in terms of asking learners to denote their contribution type from a predefined set, with respect to the underlying conversational structure. The system records all the actions performed by the learner and supports both a quantitative and a qualitative analysis resulting in to output ratings for attributes such as attitude and promote discussion. The ratings are exploited by an advisor agent offering tips on improving learners' interaction. Synergo [1] builds on the Object-oriented Collaboration Analysis Framework (OCAF) where learner interaction and workspace actions are analysed from the shared objects' point of view. The objects that learners manipulate independently compile statistics on their use, and contribute to the definition of indicators describing their owners' collaboration behavior. The collaboration factor (CF) is proposed, which provides a degree of collaboration of the group and is graphically displayed on the time axis, facilitating the analysis of collaboration over a set time period. Padilha et al [9] propose performance reports based on a set of quantitative (e.g. interaction numbers in chat tool) and qualitative (e.g. degree of explanations sent) indicators. In the EPSILON system [10], learners collaboratively solve object-oriented design problems while they communicate through sentence openers. The system codes learners' communication and actions and determines whether or not they effectively share new knowledge and what sort of guidance might be helpful by applying the Hidden Markov Model technique.

Our research efforts extend previous work in interaction analysis in synchronous text-based communication tools. More specifically, we attempt (i) to analyse learners' interaction (i.e. dialogue) with respect to the learning outcomes addressed by the collaborative activity or the model of collaboration followed, and (ii) to offer teachers the possibility to personalize the analysis process by defining weights to the attributes analysed denoting this way their importance with respect to the underlying activity. To this end, in the context of the ACT (Adaptive Communication Tool) tool, we follow a quantitative and a qualitative approach aiming to analyse and investigate learner collaboration behavior in terms of Learner Indicators and group collaboration behavior in terms of Group Indicators. The work presented in this paper focuses on Learner Indicators.

The rest of the paper is structured as follows. In Section 1, a brief overview of the ACT tool is given while in Section 2 we present the modeling of the scaffolding sentence templates on which the dialogue is based. Section 3 is devoted to the description of learner indicators illustrating and explaining them through specific case studies. The paper ends with the main points of our work and our near future plans.

## The ACT Tool

ACT (Adaptive Communication Tool) supports the synchronous communication of learners in groups of up to four persons in the context of a collaborative activity. The activity may address cognitive skills that are classified to one of the four levels: *Comprehension level* (Remember + Understand), *Application level* (Apply), *Checking-Critiquing level* (Evaluate) and *Creation level* (Analyse + Create) [4]. Moreover, a specific model of collaboration is followed; the group members may collaborate either having the same duties or undertaking different roles. In any case, one of the group members plays the role of the moderator, being responsible for the coordination of the group process (e.g. proceed to the next question), the summarization of the debate and the submission of the answer.

The automatic analysis of the dialogue is quite difficult, in cases where the free dialogue is supported. In order to facilitate the tracing of the dialogue states and enable the automatic interpretation of learners' interaction [6], the structuring of the dialogue through sentence openers or communication acts is usually supported. Towards this direction, the ACT tool follows the structured form of the dialogue, aiming to (i) guide learners towards the underlying learning outcomes of the activity or the duties and responsibilities implied by the model of collaboration, and (ii) enable the automatic analysis and interpretation of learners' interaction. In ACT, the structured form of the dialogue is supported utilizing both sentence openers and communication acts. The provided *Scaffolding Sentence Templates*

(SST) (i.e. sentence openers and communication acts) are adapted on the basis of (i) the level of the learning outcomes (i.e. cognitive skills) addressed by the activity, and (ii) the specific roles that learners undertake in the context of a specific model of collaboration [4]. More specifically, the sentence openers are aligned with the Comprehension, Application and Checking-Critiquing level of the cognitive skills, while the communication acts are aligned with the Creation level and the role that each learner undertakes. Also, the communication acts are used in case learning activities do not explicitly address one out of the four above mentioned levels of cognitive skills, but they rather aim to enable learners to discuss/exchange ideas on a specific topic or on the subject/solution of the activity.

All the group members have at their disposal the same set of SST if they collaborate having the same duties; the moderator of the group has available additional SST compatible to the additional duties. In case a model of collaboration with roles is followed, the provided SST are different for the group members supporting their roles appropriately [4]. Besides the predetermined sets of SST, a learner may determine his/her own SST in case the available ones do not cover his/her needs. The learner's determined SST are part of his/her learner model and are available each time s/he uses the ACT tool.

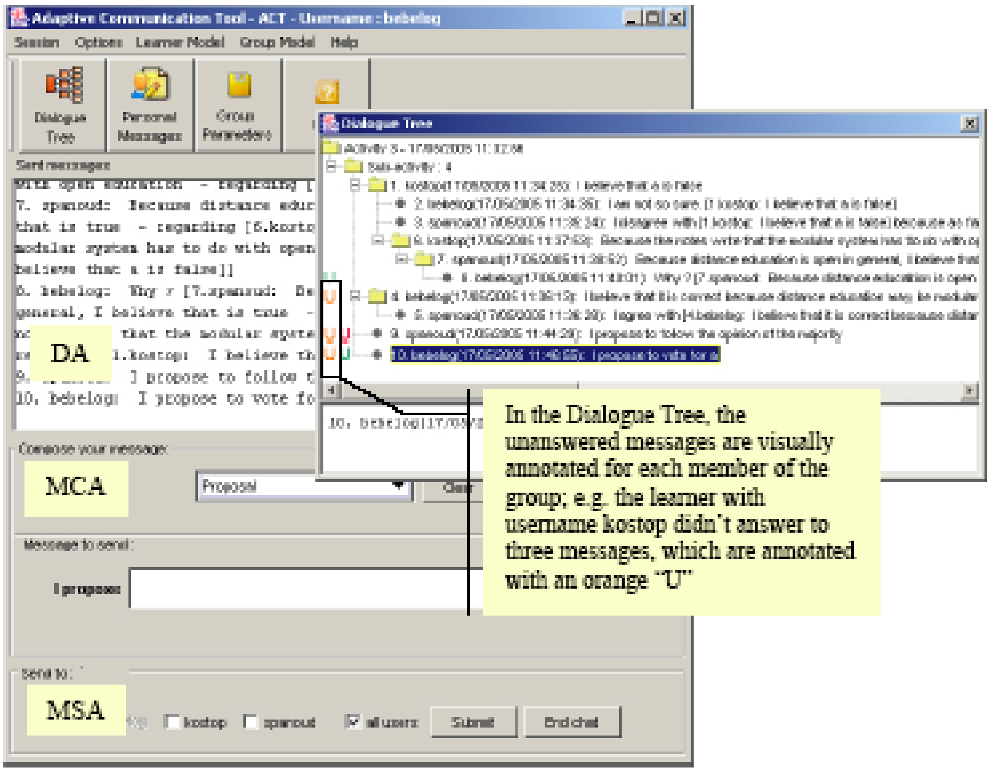



**Figure 1**. A screenshot of the ACT tool with the "Dialogue Tree" window. The Dialogue Area (DA) presents the messages in a chronological sent order; the Message Composition Area (MCA) enables learners to select the desired SST and compose their message; the Message Submission Area (MSA) enables learners to submit their message to all or to selected members of the group.

During the collaboration, learners can have access to their model as well as to the group model in order to have an insight to their own contributions and to their collocutors' contributions in a graphical form. Figures 2 and 3 present the contributions of two groups respectively; the left (blue) column corresponds to the group's contributions followed by a coloured column for each member of the group. This facility acts as a mirroring tool and supports the regulation process. The messages are grouped according to the message they are referring to and are visually represented in a tree structure through the "Dialogue Tree" window. In particular, ACT supports a facility for the automatic construction and update of the Dialogue Tree as learners submit their messages. The learners have access to the Dialogue Tree at any time during the communication. Such a graphical representation of the dialogue enables learners to trace the sequence of the dialogue more easily, to have a clear view of the dialogue progress and to receive feedback, in a visual form, about their contributions (e.g. in the Dialogue Tree presented in Figure 1, following the analysis of the dialogue, the unanswered messages are notified for each member of the group with a different colour in correspondence to the colours used in the graphical representation of their contributions). Therefore, the Dialogue Tree can stimulate learners to reflect on their dialogue and improve their participation.

## 2. Modeling Scaffolding Sentence Templates

The predetermined SST as well as learner's defined SST are categorized to one or more of the following *discourse categories*: Proposal (P), Opinion (O), Question (Q), Reasoning (R), Clarification (C), Agreement (A), Disagreement (D), Inference (I), Motivation (M), Need (N) and Social Comments (S). The predetermined set of the SST includes:
(i) a subset dedicated to the development/cultivation of cognitive skills aligned with the addressed learning outcomes,
(ii) a subset facilitating the communication, and
(iii) a subset available only to the moderator of the group.
Each SST is defined as a set of the following attributes:
• *SSTT* (SST Type): the type of the SST may be either a Sentence Opener or a Communication Act.
• *ST* (Skill Type): the type of the skills that the SST mainly concerns, may be either cognitive, with respect to the activity under consideration, or communication.
• *OL* (Outcome Level): the outcome level that the SST is aligned with.
• *DC* (Discourse Category): the discourse category of the SST denoting the intention of learner's contribution.
• *SR* (Supporting Roles): the roles that the SST serves in case the model of collaboration implies specific roles to the group members with specific duties and responsibilities.
• *T* (Text): the text forming the SST, which may be composed of one or two parts depending on the number of arguments.
• *FA* (Filling Actor): if the SST consists of one or more arguments, then the argument(s) may be filled either by the learner (in case of text field) or the tool (in case of a reference to an already sent message) or both.
• *UI* (User Input): in case that the argument of the SST is a text field, the user input may be optional or obligatory.
• *W* (Weight): the degree of the SST denoting the value of the underlying contribution (w[0, 100]).
To clarify the above attributes, let's consider the following examples of SST:
{SO, CS, C, P, null, "I propose", U, Ob, 100}: The Sentence Opener "I propose" concerns Cognitive Skills, is aligned with the Comprehension level of learning outcomes, denotes learner's intention to contribute to the dialogue through a Proposal, the accompanying text field has to be filled in by the User and is considered Obligatory, and the weight of the SST is 100.
{CA, CS, null, A, Assessor-Driver, "Agreement", SU, Op, 80}: The Communication Act

“Agreement” concerns Cognitive Skills, denotes learner’s intention to make an Agreement, is made available to those learners undertaking the role of Assessor in the context of the “Brainstormer-Assessor” collaboration model or the role of Driver in the context of the “Driver-Observer” collaboration model, the accompanying arguments include (i) a reference to an already sent message; the already sent messages are made available by the System, and (ii) a text field, which has to be filled in by the User and is considered Optional, and the weight of the SST is 80. In case of learner’s defined SST, the values of ST and W are inferred from the learner’s defined DC and the correspondence between the DC, ST and W values as assigned by the teacher. The weights assigned by the teacher reflect the degree of importance of the provided SST with respect to the learning outcomes addressed by the activity. This way, the teacher has the possibility to personalize the analysis process of learners’ interaction. For example, in case the teacher wishes to analyse and have an evidence of learners’ ability to make proposals and provide explanations or arguments, s/he may set higher weights to SST belonging to the discourse categories of Proposal and Reasoning.

## 3. Collaboration Indicators

The analysis of learners’ interaction follows both a quantitative and a qualitative approach. The quantitative analysis aims to provide information about learners’ contributions to the dialogue and the results are presented in a graphical form, as described in Section 1, while the qualitative analysis aims to exploit various characteristics regarding learners’ contributions and investigate (i) learner collaboration behavior in terms of Learner Indicators and (ii) group collaboration behavior in terms of Group Indicators. The design of the analysis process and subsequently of the supported indicators is based on the design principles of the ACT tool as well as on the literature research [5], [7], regarding the skills that contribute to a creative and “good” collaboration. In the context of the current work, we describe the Learner Indicators and elaborate on them through specific empirical data.

### *3.1 Learner Indicators*

As one of the main objectives of the ACT tool is to guide learners towards the development of cognitive skills in line to the addressed learning outcomes, it is considered necessary to analyse learners’ contribution in view of this perspective. As stated above, the teacher has the possibility to assign the desired weights to the provided SST taking into account the learning outcomes and the discourse category of the SST. In this sense, the *Cognitive Skills Indicator for the ith Learner (CSI(Li))* is defined as the sum of the weights of the SST used by the learner during the communication (Formula F1). Depending on the weights assigned to the SST, this indicator denotes the degree that the learner contributes to the dialogue using SST aligned with the cognitive skills addressed by the learning activity; and subsequently, the degree of cultivating such skills.

$$\mathrm{CSI\ (Li)} = \sum_{\text{for all SST used by Li}} W_{SST} \quad \text{(F1)}$$

During the collaboration, the learner may exhibit initiatives in promoting the dialogue/collaboration by (i) making proposals or expressing an opinion and in this way initiating/stimulating the discussion, (ii) answering to the contributions of other group members (whether it is required or not) or elaborating further on his/her own contributions by reasoning for his/her point of view, and (iii) elaborating on a point of view although it is

not considered necessary. More specifically:
• The *Initiating the Discussion Indicator for the ith Learner (IDI(Li))* concerns the first abovementioned point and it is measured as the sum of all the messages sent, characterized as proposals or opinions in the total of all the sent messages; Formula (F2) quantifies the degree of learner's initiative to stimulate the discussion:

$$IDI(Li) = \frac{\sum_{\substack{\text{for all SST used by Li} \\ \text{where DC = P or O}}} SSTi}{\sum_{\text{for all SST used by Li}} SSTi} \quad (F2)$$

• The *Advancing the Discussion Indicator for the ith Learner (ADI(Li))* refers to the second abovementioned point and reflects learner's behavior in advancing the discussion taking into account a number of factors:
o the messages that the learner Li answered and s/he had to do so; that is, a collocutor expresses an opinion/proposal or a question referring to one of the Li's previous sent messages and expects an answer.
o the messages that the learner Li answered although s/he had not to do so; for example, in case a collocutor disagrees to Li's contribution or makes an inference and Li attempts to elaborate further on the collocutor's contribution by posing a question or expressing his/her agreement/disagreement, etc.
o the elaborations that the learner Li made on his/her own messages in order to give a clarification or a justification despite s/he was not asked to do so. For each of the above factors, a weight is assigned by the teacher denoting the importance of each factor in the context of the underlying activity.
• The *Further Elaboration on a collocutor's view Indicator for the ith Learner (FEI(Li))* reflects that the learner Li not only acknowledges his/her collocutors' point of view but also wants to stress and elaborate further on the point under discussion and therefore promotes the discussion. The FEI(Li) is measured according to the formula (F3) which shows the mean number of the messages that learner Li has further elaborated on by filling in the "optional" text field.

$$FEI(Li) = \frac{\sum_{\substack{\text{for all SST used by Li} \\ \text{where UI = Op \& filled by Li}}} SSTi}{\sum_{\substack{\text{for all SST used by Li} \\ \text{where UI = Op}}} SSTi} \quad (F3)$$

The above three indicators contribute to the quantification of the *Promotion of Discussion Indicator for the ith Learner (PDI(Li)),* which shows learner's collaboration behavior in participating in a creative discussion. Each one of these indicators is partially contributing to PDI(Li) with respect to the corresponding weights assigned by the teacher and reflecting the degree of importance of each one in the context of a specific learning activity. More specifically, PDI(Li) is measured as in Formula (F4):

$$PDI(Li) = IDI\ (Li) * W_{IDI} + ADI\ (Li) * W_{ADI} + FEI\ (Li) * W_{FEI}\ (F4),$$

*where* $W_{IDI}$ is the weight denoting the contribution of IDI to PDI,
$W_{ADI}$ is the weight denoting the contribution of ADI to PDI,
$W_{FEI}$ is the weight denoting the contribution of FEI to PDI,

The above defined indicators have a complementary value in interpreting the interaction for each learner separately and making comparisons between the collocutors' contributions. They are estimated on the fly revealing the evolution of the cultivation of the desired skills and of learners' collaboration behavior in the context of the activity.

### 3.2 A Case Study

In the context of the formative evaluation of the ACT tool [4], an analysis of the learners' dialogues was carried out, in terms of the above defined learner indicators. In the following, we elaborate on two indicative cases in order to illustrate and explain the indicators. Figures 2 and 3 present raw data of two groups' contributions, aggregated for each learner separately as well as for the whole group, with respect to the corresponding discourse categories. Group G1 consists of two learners (usernames: tsourak and kourt, moderator: tsourak) while group G2 consists of three learners (usernames: kostop, bebelog and spanoud, moderator bebelog). The collaborative learning activity under consideration addressed learning outcomes of the Comprehension level, asking learners to identify as true or false five statements related to the subject matter of "Distance Education" and justify their answers. In this context, the teacher assigned high weights to the discourse categories of Proposal and Opinion, following the categories of Question, Reasoning and Clarification and afterwards the categories of Agreement and Disagreement. Low weights were given to the discourse categories of Inference, Motivation, Need and Social Comments.

As far as group G1 is concerned, in an attempt to interpret the learner indicators, as these are depicted in Table 1, we notice the following:

- The value of CSI indicator is greater for learner kourt than learner tsourak, meaning that learner kourt developed at a greater degree those skills addressed by the collaborative learning activity. This seems to be true as learner kourt made more proposals and opinions than learner tsourak (4 and 3 respectively), more agreements (3 and 2 respectively) and gave more reasons (2 and 1 respectively) (Figure 2).
- The above point is consistent with the value of IDI indicator, which denotes the mean number of the proposals and opinions made by a specific learner (for learner kourt, IDI indicator has greater value than for learner tsourak).
- Both learners seem to have equally attempted to advance the discussion by elaborating and contributing to their collocutor's point of view (the corresponding values of ADI indicator are very close). Having a close look at their dialogue, we observed that learner tsourak answered to more collocutor's messages than learner kourt (3 out of 5 and 2 out 4 respectively) while learner kourt elaborated further on one of his proposals although he had not to do so.
- Learner tsourak reasoned for his agreements (value of FEI is 0,5; 1 out of 2) while learner kourt didn't justify his agreements at all (value of FEI is 0; 0 out of 3).
- Both learners contributed to a creative and productive collaboration as it results from the PDI indicator. The difference of the corresponding values is due to the difference in the values of ADI and FEI indicators.

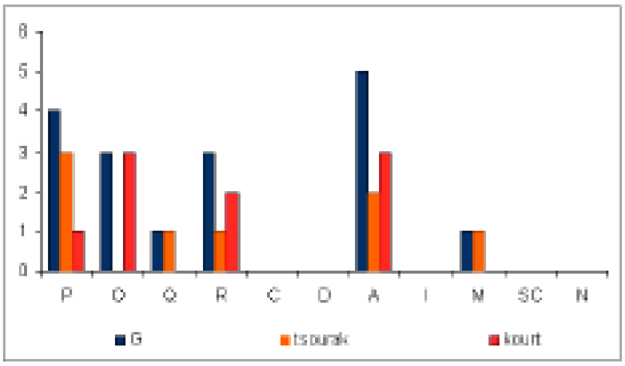



**Figure 2.** The contributions of each member of group G1 as they are presented in ACT

| Group G1 | CSI | IDI | ADI | FEI | PDI |
|---|---|---|---|---|---|
| tsourak | 0,39 | 0,43 | 0,48 | 0,5 | 0,46 |
| kourt | 0,76 | 0,44 | 0,43 | 0 | 0,39 |

**Table 1.** The values of indicators for each member of the group G1

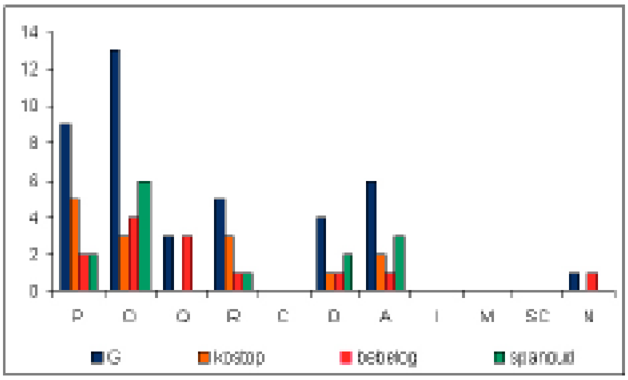



**Figure 3.** The contributions of each member of group G2 as they are presented in ACT

| Group G2 | CSI | IDI | ADI | FEI | PDI |
|---|---|---|---|---|---|
| kostop | 0,68 | 0,57 | 0,07 | 1 | 0,41 |
| bebelog | 0,32 | 0,5 | 0,06 | 0 | 0,27 |
| spanoud | 0,45 | 0,57 | 0,19 | 0,6 | 0,42 |

**Table 2.** The values of indicators for each member of the group G2

As far as group G2 is concerned, we can reach to analogous conclusions in accordance to group G1. However, in the case of group G2, it is worthwhile noting the value of the indicator ADI with respect to the contributions of each member. Although, each member of the group G2 has exhibited skills in making proposals and expressing opinions (the value of IDI is quite high for all members), they seemed to be reluctant to negotiate, questioning their collocutors and reach an agreement after a creative debate (the values of ADI are very low). Examining their dialogue, we observed that they could not agree on what answer to give to a specific question, and each one of them attempted to propose a solution without elaborating on his/her collocutor's point of view. In general, we can say, that all of them tended to express their opinion by making a new proposal instead of making an agreement to a collocutor's opinion even though they agreed with him/her.

## 4. Conclusions

In this paper, we present an approach to analysing learner interaction in the ACT tool, both quantitatively and qualitatively. The quantitative analysis aims to provide information about learners' contributions to the dialogue and the results are presented in a graphical form. The qualitative analysis aims to exploit various attributes regarding learners' contributions and investigate learner collaboration behavior in terms of (i) the cognitive skills that the learner develops with respect to the learning outcomes addressed by the activity, and (ii) promoting the collaboration by initiating/stimulating and advancing the discussion. A discriminative characteristic of the approach is that the teacher has the possibility to personalize the analysis process by defining weights to the attributes analysed denoting this way their importance with respect to the underlying activity. The illustration of the defined indicators with specific empirical data reveals that they give a valid evidence of learners' behavior during the collaboration process. However, the investigation of the validity of the proposed indicators with more groups of learners is considered necessary. Also, we plan to exploit these indicators in the development of guiding mechanisms, which will be adapted to learners' collaboration behavior.

**References**


[1] Avouris, N., Margaritis, M., & Komis, V. (2004). Modelling interaction during small-groups synchronous problem-solving activities: The Synergo approach. *Proceedings of ITS 2004 Workshop on Designing Computational Models of Collaborative Learning Interaction*, (August 31, 2004, Maceió, Brazil), 13-18.
[2] Barros, B. & Verdejo, F. (2000). Analysing student interaction processes in order to improve collaboration. The DEGREE approach. *International Journal of Artificial Intelligence in Education*, 11, 221-241.
[3] Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. *In P. A. Kirschner (Ed.) Three worlds of CSCL. Can we support CSCL*, (Heerlen, Open Universiteit Nederlands, 2002), 61-91.
[4] Gogoulou, A., Gouli, E., Grigoriadou, M., & Samarakou, M. (2005). ACT: A Web-based Adaptive Communication Tool. *Proceedings of CSCL 2005 (to appear)*.
[5] Harwood, D. (1995). The Pedagogy of the World Studies 8-131 Project: the influence of the presence/absence of the teacher upon primary children's collaborative group work. *British Educational Research Journal*, 21(5), 587-609.
[6] Jermann, J., Soller, A., & Lesgold, A. (2004). Computer Software Support for CSCL. *In J. W. Strijbos, P. A. Kirschner & R. L. Martens (_ds.), What we know about CSCL and Implementing it in Higher Education*. 2004 Kluwer Academic Publisher, 141-166.
[7] Johnson, D. & Johnson, R. (1999). Learning Together and Alone, Englewood Cliffs, NJ: Prentice Hall.
[8] Kaleidoscope NoE, ICALTS - JEIRP (2004). State of the Art on Interaction Analysis: "Interaction Analysis Indicators". Retrieved from http://www.rhodes.aegean.gr/LTEE/KALEIDOSCOPE-ICALTS/
[9] Padilha, T., Almeira, L., & Alves, J. (2004). Mining Techniques for Models of Collaborative Learning. *Proceedings of ITS 2004 Workshop on Designing Computational Models of Collaborative Learning Interaction*, (August 31, 2004, Maceió, Brazil), 89-94.
[10] Soller, A. (2004). Understanding Knowledge Sharing Breakdowns: A Meeting of the Quantitative and Qualitative Minds. *Journal of Computer Assisted Learning*, 20, 212-223.

# Knowledge Development through Collaborative Work Supported by Interactive Technologies

**Jorge Ferreira FRANCO**[1,2], **Roseli de Deus LOPES**[1]
*1-Laboratório de Sistemas Integráveis (LSI), Departamento de Engenharia de Sistemas Eletrônicos*
*Escola Politécnica, Universidade de São Paulo, São Paulo, SP, Brasil*
*2-Secretaria Municipal de Educação de São Paulo (SME), SP, Brasil*
*Av. Prof. Luciano João Gualberto, trav.3 – 158*
*CEP: 05508-900 – São Paulo*
*Phone: +55-XX11-30915661 – FAX: +55XX11-3091-5665*
*{jfranco, roseli}@lsi.usp.br*

**Abstract.** This paper addresses experiences based on collaborative work aiming to improve basic education domain's teaching/learning practices, as well how the experiences have enhanced individuals' knowledge through using interactive technologies. The strategies for reaching such goals have evolved the synergy of learning theories and practices, high and basic education cooperative work, individuals' direct manipulating web accessible standard languages, including using virtual reality techniques and multimedia tools as well files.

The investigation and application of such convergence has brought about school community's intellectual, social and technical enhancements within a dynamic, active and creative mood. Such enhancements have influenced individuals' cooperative/collaborative learning/teaching attitudes, stimulated more horizontal relationship and respect among educators and children, as well contributed for increasing individuals' traditional and digital literacy competences and skills, including inspirited ones' lifelong learning autonomy.

## 1. Introduction

The development of technological artifacts has reached a point, in which they are much more accessible for people, even the individuals who are under economic disadvantage. On the other hand, there are many of these individuals who are not prepared for using such artifacts in way they can benefit from the artifacts far beyond entertaining moments. Conversely, the industry have used technological artifacts and explored interactive technologies with success in diverse fields, such as Internet applications, Education, Research, Entertainment, Marketing, Scientific Visualization via Virtual and Augmented Realities techniques, Business, Information and Communication. Such exploration have brought about the necessity of preparing a skilled and competent workforce able to synthesizing and understanding knowledge, as well employing such technologies and knowledge for building high quality content within the age of interactive media [1-10].

Then, it is relevant using better the technological artifacts within social environments, such as at school domain. This action can provide people's critical uses and reflection about such technologies contribution for improving individuals' traditional and digital literacy, learning autonomy, developing self-esteem and mental models [11-16]. In conjunction with school's individuals we have used, adapted and developed learning models as well strategies related to the application of interactive technologies as support for enhancing and transforming school's curriculum as well teaching/learning experiences in something more practical and interactive.

On the other hand, we have faced educators' lack of technical skills for manipulating technological artifacts, which makes difficult educators using them effectively [4, 17]. Unfortunately, according to Valente [18] there are not enough skilled individuals for training educators' contingent who need assistance for improving their multimedia and interactive technologies' skills and competences. Educators from public basic education have been under economic disadvantages related to low salaries, implicating in educators' double work journey and less possibilities for doing high-level courses in order to improve their technical skills and competences. For many educators as well for children the school environment is still a unique opportunity for accessing, investigating, reflecting, training and being aware about how to use interactive technologies further than entertaining.

Hence, historically there has been much more access to economic resources and high technology for educators on high education than the ones in elementary and secondary institutions. For instance, "(…) public education faces government cutbacks that naturally force reform of the school system. Educators are faced with the dilemma of providing a high-quality education with less funding, with more students in our classes (…)" Vrieze in [19]. Conversely, Valente [18] on his investigation related to learning using virtual environments for training teachers recommends 20 students per educator to be feasible the several interactions and keep educational standard high quality. Then, again, based on Vrieze ideas [19] if educators work within a more adequate educational environment, educators can face the challenge from business world demanding that its future employees, our students, be educated to keep up with the rapidly changing technological advancements to be competitive in a global economy.

Due to the broad range of technologies, which have evolved such changes and advancements, including educators' lack of technological skills and knowledge, at least within the context of Brazilian' s public basic education, a possible and realistic solution for decreasing such problems is developing collaborative/cooperative knowledge's networking. So, in that direction collaborative educational actions among individuals from Higher and Basic Education are fundamental. Fortunately, the projects and experiences we have researched and carried out, confirm the importance of collaborative/cooperative teaching/learning actions.

## 2. Related Work

Technological advances and the constant decreasing on devices' prices have enabled the use of innovative learning tools in several experiences related to the Education domain, bringing about better support for collaborative work in teaching/learning actions.

**Kaufmann** [16] explored the state-of-the-art of technological advances using collaborative Augmented Reality (AR) within education in the context of immersive virtual learning environments through the experiences made during the development of a collaborative AR application, specifically designed for mathematics and geometry education called Construct3D, which is based on the mobile collaborative AR system "Studierstube" [20]. In **Simpson** [21] work "The Virtual Designer – The Application of VRML in Collaborative Design" is looked at the process of collaborative virtual design, undertaken to improve online architectural communication. The author argues that through analyzing the needs of designers and addressing their responses to collaborative design scenarios a clear process of approach to optimizing virtual communication can be developed. **Liarokapis** et al [22] inspirited in Billinghurst [23] as well in some ARToolkit functionalities [24] developed a multimedia augmented reality interface for e-learning (MARIE) focused on enhancing teaching/learning process through a low cost AR collaborative system able to provide interactions between immersed user (using Head Mounted Displays – HMD) and non-immersed one through PC monitor and voice. **Roussos** et al [25] NICE (Narrative Immersive Constructionist /Collaborative Virtual Environment) project focuses on informal and formal education, social content domains, embracing a constructivist approach, collaboration, plus narrative development. For that, the project explores virtual reality main power: a combination of immersion, tele-presence, immediate visual feedback, and interactivity. Software development is based on open standard languages such as JAVA, VRML and C++. The virtual reality environment is designed for both multi-projection CAVE™ and PC systems.

## 3. Pedagogical Concepts

The convergence of learning theories and methodologies such as Piaget's constructivism and Papert's construcionism can enhance individuals' understanding about how people learn and grow, providing better support for designing teaching and learning environments [25, 26]. Thoughts which come from researchers such as Paulo Freire and Ivan Illich refer to the necessity of making a revolution in the curriculum content and the pedagogy of the present-day school, transforming them in more practical and inclusive based on horizontal relationship between educator and pupil, love, humility, hope, faith, confidence and respect for the freedom of expression [27]. For reaching such harmonic state among individuals in an educational environment it seems necessary to have excellent communication among people. It is thought that the word that can define successful communication in human relationship is *empathy,* as suggested by Peters [28]. The teaching/learning actions we experienced showed that using the same technological tools which have entertained students can be a strong support for achievening *empathy* as well carrying out collaborative/cooperative work situations [4, 34].

The spirit of human and technical interactions based on Freire's ideas related to more horizontal and dialogic communication during teaching/learning experiences can come from using the convergence of information and communication technologies innovations, arts and culture on curriculum development. This cultural and technological application can be the base for individuals' reflection about such synergy contribution for social development plus the construction and maintenance of meaningful learning environments [29]. Other relevant supports and advantages of using interactive media and multimedia devices for enhancing individuals' teaching/learning practices come from Gardner's theory related to multiple types of intelligence, as well Projects' pedagogy related to John Dewey

[28, 30]. Supports and advantages are associated to the flexibility of working through projects bringing about exploring individuals diverse learning interests. Using interactive technologies in teaching/learning actions have attracted more students toward stimulating their diverse perceptual and sensorial channels such as auditory and visual.

The convergence of learning theories and advances on technology has brought about space for the growth of innovative educational designing and intervention, which is called *Emergent Design*. "It is an approach used for educational intervention; the claim is a more general one, however, in that the strategy is appropriate in settings for technologically enabled paradigmatic change (...)" [14]. Through Emergent Design experiences it is possible to meet a balance between digital technology and the approach to management of organization and of organizational change that has come in the wake of the technology. A distinction must be made because the temptation to use either of them alone has led to failure. It is the combination that offers an optimistic vision for the future of learning—the combination of these two products of the digital age along with a theoretical framework based on the work of pre-digital-age thinkers who knew what to do but did not have the means to do it. Among these the most central is Paulo Freire, but also represented are John Dewey and Jean Piaget, although he did not focus on education per se [14].

## 4. Work Development and Strategies

We have developed work based on collaborative, co-operative, constructivist and construcionist approaches for educators' training, bringing about human sustainable improvements. It includes applying concepts from the experiential learning model in synergy with the spiral model for enhancing individuals' technology skills and competences. Such competences are capabilities of analysis, interpretation and selection of information, facts or situations; skills to manage group activities, as well awareness of social surrounds being able to transform it [31- 33].

### 4.1 Background – Collaborative and Cooperative work

Since 2002, we have carried out work for improving individuals' knowledge through using the convergence of interactive technologies, arts, culture and communication at Ernani Silva Bruno Public School, which is situated in the surrounds of north zone of São Paulo city [9, 34]. In general, citizens living around the school are under economic disadvantage. We met on Cavallo's reflections a picture about our current concerns. "The growing "digital divide"—concerns about the potential of a widening gap between rich and poor in the new, knowledge-based global economy due to a lack of modern, technological skills among people in lower social-economic strata, and a growing concern about the potential of educational systems to ameliorate this situation— all point to a serious problem becoming seemingly permanently intractable (...)" That is the case of the named municipal district school in which author 1 works. On the other hand, "(...) The same technology that can be a primary factor in widening the divide may be the best hope for eliminating the divide" [14].

After three years of collaborative/cooperative work within basic and higher education through small-scale projects it seems that such experiences have become useful tools for inspiriting teaching/learning practices not only in experimental situations, but also with real students, as expressed in [16], within regular basis as classroom problem solving instruments. Such actions meet support on Papert's ideas in [35] recommending; to add age-appropriate knowledge of computer programming to "the basics", to bring back the arts as factors that drive learning in a high-tech environment, to integrate students into the process of deciding what shall be studied and especially how it shall be assessed at all levels, to begin a program of increasing the portion of time spent on project-based learning using the new content to give intellectual depth, rigor and toughness to project work.

Experimental situations have been very important to Ernani's educational environment. Such situations have improved individuals technical skills, self-confidence and social relationship through learning and using Web standards such as Hypertext Markup Language – HTML and Virtual Reality Modeling Language - VRML within educational projects. Such experience led school community to participate in two major projects developed at high education aiming to stimulate collaborative work and research at basic education "FEBRACE" and "The City We Want".

**FEBRACE** - Feira Brasileira de Ciências e Engenharia is a science and engineering fair coordinated by author[2]. The project brought about in 2003, 2004 and 2005 opportunity for educators and students from basic education developing collaborative projects based on their cultural and educational context, after that, sharing knowledge with other individuals during a week fair. At the end of the week some students' projects are chosen for representing Brazil in an international fair called ISEF [36]. Due to Ernani's small-scale projects developed from 2002 to 2004 using accessible Web standards, author 2 invited educators and children from school for being protagonists of three days workshop researching and direct manipulating interactive media, arts and culture [9, 34]. **The City We Want** is a project similar to FEBRACE in terms of educational goals, but it is designed for attending educators and children from public municipal schools through a partnership between the Municipal Education Secretary of Sao Paulo city and LSI-USP. The project infrastructure has been improved since 2002 and in 2004 was possible to support several municipal schools in diverse regions from the city [37].

Due to Ernani School participation at FEBRACE 2004, individuals were invited to participate on "The City We Want" project in August 2004. After that, there was a three months period for sharing knowledge with other teachers and students in internal and external interactive experiences, including the development of a project's proposal aiming to solve a city problem at school's region. There were four major meetings in diverse educational city points where individuals presented and shared initial projects' implementations, in middle October. Educators and children from Ernani's school showed animations related to social problems that have influenced local life's quality. Such problems are violence, lack of leisure activities, as well difficulties for accessing arts and cultural programs.

In November within a regional Education Information Systems fair called "II Educainfo", designed for educators and students from district schools showing and sharing their work carried out during the year, individuals could feel how important was the implementation of "The City We Want" project in several schools. It brought about students and educators a great synergy for improving individuals' researching, collaborative learning attitudes, traditional and digital literacy, including communication skills and competences [31]. For instance, it was evident for author 1, Ernani School community's engagement on "II Educainfo". Educators and children from Ernani were invited for carrying out an animation workshop and sharing knowledge with other scholars on the event. They prepared themselves very well for participating on the workshop. The animation models used at "The

City We Want" project were revised and enriched for the demonstrations. Ernani School's principal, a pedagogical coordinator, fifteen children, as well a Math and Arts' teachers who were on FEBRACE workshop and coordinated the animation project were in the event. There, it was carried out two hours animation workshop, in which children from Ernani's School taught and shared knowledge with more 30 children from events' school. Such Ernani's community engagement brought about individuals' self-esteem enrichment and better pedagogical support for outlining and running more structured projects in 2005.

### 4.1 Education and Interactive technologies Project

This project is a further implementation of the author 1 investigations on how to introduce and use Web standards to support non-technical educators and children teaching/learning practices. From, initially working with 6 students informally on Fridays' morning during 2003/2004, author 2 in a collaborative work with the other school computers lab's coordinator have investigated since February 2005 interactive technologies with 20 youth children from 8th grade direct manipulating accessible Web standards (browsers and languages), multimedia instruments and files [9, 34]. Through supervised, unsupervised and reinforcement learning, including learning from observation we are using such tools and techniques as a platform for stimulating individuals' learning to learn, developing collaborative attitudes, reflecting about the importance of carrying out traditional and digital literacy skills and competences since basic education [38, 39]. Although we have run the project based on Web technologies with 8TH grade students, the parallel educational processes carried out at dynamic school environment have converged in several moments. Such convergence led us to use VRML with 7th grade students within a project evolving 30 students from 7th A, B, C and D classes.

Due to students' difficulties for understanding cartographic, measure and scale concepts in Math, Geometry and Geography subjects, from March 2005 students have been measuring the school and designing a 2D blue print using Paint™ program. Because of the collaborative work between the computers lab's coordinators, exposition of other educators and students to virtual worlds and Virtual Reality (RV) techniques, it was suggest to students from 7th grade modeling the school in 3D using interactive and universal Internet tools. So, after presenting author 1's VRML examples, the other school lab's coordinator in conjunction with students started building a school's 3D model for presenting that to a Ernani's community during the 10 years celebration of School existence on 7th May. Students accepted the challenge and in conjunction technical and non-technical individuals start developing the model (figure- 1).

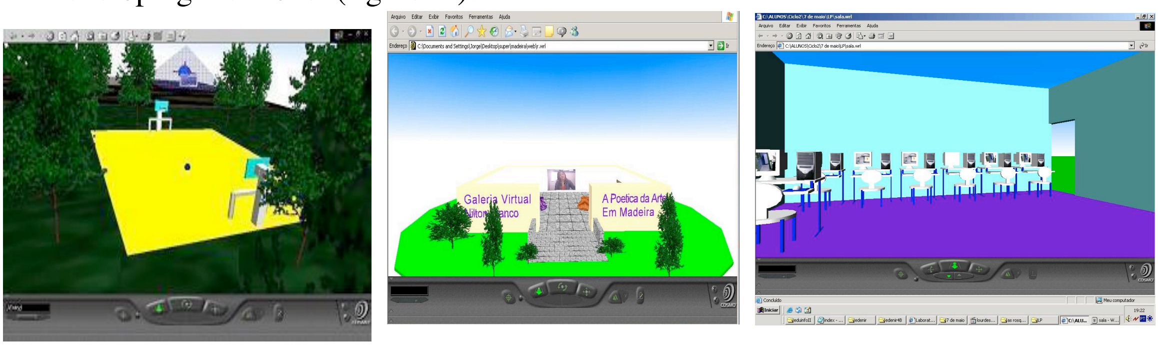



**Figure-1** – Left and middle images are hybrid interfaces built using VRML audio, video, animation features. Right image is School Computers lab's model designed by 7th students.

**5. Infrastructure and Tools**

Management, actions and technical designing are based on HTML and VRML languages as well as their similarities which brought about better comprehension on how to use them even for non-technical individuals [40-42]. **Hardware**: school computers' lab facilities are; Internet access, intranet with 21 computers, Pentium III™, 128 RAM, webcam, microphones and datashow. **Software**: Paint™ and Gimp™ for producing and editing images; Word pad™ and Notetab light™ for textual programming. **Browsers**: Internet Explorer™, Cosmo player 2.1 ™ and Cortona™.

**6. Discussion**

Although, we do not have an AR structure for running high-tech AR experiments as in [16], we have applied similar learning techniques and concepts explored in such experiments through desktop VR using VRML as an instrument for supporting children's understating cartographic, measure and scale concepts, developing spatial cognition and solving simple and complex problems [12]. Experiences with children developing content using interactive technologies have showed that for being feasible and of high quality the interactions among educator, students and content, working within small groups, about ten per educator, is recommended [46]. Conversely, we work with 30 children at once in the Lab, but we divided them in two groups. Two computers lab's educators guided a group of 15 pupils developing several parts of school blueprint on Paint™ and the other students constructing school's 3D model using VRML. Such situation suggests that it is necessary to implement a policy for improving schools environmental conditions, such as smaller quantity of children per class, increasing teacher's technical training opportunities.

For starting up students' learning curve the other lab coordinator with little technical skills in VRML guided the students using supervised learning technique through a tutorial [42]. After developing a small-scale room for understanding how to work language syntax, the visualizations and browser manipulation on the screen, students collaboratively exchanged code via computers lab's intranet, reusing and adapting VRML code for building desks, tables and other objects, in actions similar to the work in [21, 22]. The construction of WEB based educational applications in synergy with the pedagogical concepts discussed in session 3 and in [25] supported face-to-face interactions, meaning in our learning context cooperative/collaborative improvements on the social relationship among educators and children, enhancing traditional teaching/learning methods, increasing individuals' self-esteem, as well knowledge and technical skills.

It is thought that through such actions it will be constructed a cultural and resonance knowledge network able to contribute with a society that needs adaptable and creative entrepreneurs, as well workers capable of taking advantage of the latest technologies to develop niche products and markets. Due to such society needs, it is relevant using information and communication technologies – ICTs during curriculum development. It can allow individuals' ICTs social appropriation and redefinition according to their needs [2, 43]. So, people will need to be powerful learners imbued with powerful knowledge. This will come from individuals' awareness about how much important is to foster toughness of mind in our students [44]. Then, from that, toward using the convergence of electronic systems, traditional and new digital media we can encourage in our children three important skills, which Harel [45] calls the "**three Xs:** eXploring, eXpressing and eXchanging ideas". And one of the best tools for learning these skills is the Internet as well its tools.

In terms of lifelong learning autonomy, students have got technical skills and developed perseverance attitudes for reading tutorials, doing the suggested exercises and sharing knowledge with their classmates. Some of them who left Ernani's have shared knowledge with their neighbors and back to our school for acquiring more information and also training new students' generation.

## 7. Conclusion and Future Work

We presented collaborative teaching/learning experiences and strategies for introducing interactive technologies to individuals at basic school domain. The experiences showed that educators who have dominated multiple languages such as arts, culture and improved their technical skills can dramatically empowering their communication with students, offering to and developing with them high quality, dynamic, creative, active and collaborative teaching/learning environment, contributing for decreasing digital divide.

It is relevant giving the CIC sciences (computation, information, complexity) as well as elements of engineering at least the same status as the traditional sciences (biology, physics, math, geometry etc...). This will take time to phase in: but worth it because of the several possibilities they can offer for supporting transdisciplinary projects [35]. Using electronic systems in convergence with Web based applications can bring about a flexible learning environment able to collaboratively improve technical and non-technical individuals' knowledge based on scientific rigor, but simultaneously attending people's diversity and necessity of effective lifelong teaching and learning practices [29, 44].

5

For the future, the idea is investigating in depth tools such as ARToolkit and Magic Book adapting them to our learning context. Although we have already downloaded such software, due to lack of programming skills an time we could not exploring the tools potential, but we think that the collaborative work we have carried out will bring about the necessary support for implementing such experience.

6 8. Acknowledgements

Our thanks to Ernani School and LSI's educators, researchers and children support. God bless you.

## 9. References

[1] PREECE, J. Human Computer Interaction, Great Britain: Addison-Wesley, 1999, 29-51p.
[2] CASTELLS, M. A Galáxia da Internet: reflexões sobre a Internet, os negócios e a sociedade, (The Internet Galaxy: Reflections on the Internet, Business and Society – Oxford - 2001) tradução Maria Luiza X. de A, Borges, Rio de Janeiro: editor JZE, 2003.
[3] SIMS, E. M. Web3D Technology in learning, Education, and Training: The 2 Sigma Challenge, in Proceedings of LET-WEB3D, First International Workshop on Web3D Technologies in Learning, Training and Education, September 30 – October 1, 2004, Udine, Italy. http://hcilab.uniud.it/let-web3d/proceedings.pdf Acessed: 05 March 2005.
[4] FRANCO, J., F., Teaching and Learning Using Interactive 3D Virtual Environments and Multimedia tools. ICEL 2000, The University of Auckland, New Zeland. (www.el.uct.ac.za/icel/)
[5] BOWMAN, D., A., et al, 3D User Inferfaces: Theory and Practice, USA: Pearson Education, 2005, 311-418 p.
[6] GRADECKI, J., The virtual Reality: programmer's kit, John Wiley & Sons, Inc. Canada, 1994.
[7] Billinghurst, M., Baldis, S., Matheson, L., Philips, M., 3D Palette: A Virtual Reality Content Creation Tool http://www.hitl.washington.edu/publications/index.html
[8] KOTLER, P., Marketing Management: Millennium Edition, Tenth Edition, Translation Bazán Tecnologia e Linguística, São Paulo: Prentice Hall, 2003.
[9] FRANCO, J., F., e Lopes, R., D., Improving and Sharing Knowledge Using 2D/3D Web and

Interactive Technologies in Education, Web3D Symposium, University of Wales, Bangor & Technium CAST 28th March to 1st April,
http://www.hpv.informatics.bangor.ac.uk/s2005/FinalProgramme.pdf accessed April 04 2005.
[10] SOULES Marshall. Virtual Reality / Hyper-Reality, 2003
http://www.mala.bc.ca/~soules.index.htm access Dec 28 2004
[11] DAVIDSON, M., J., Dove, L., and Weltz, Mental Models and Usability, Depaul University, Cognitive Psychology 404, November 15, 1999
http://www.lauradove.info/reports/mental%20models.htm accessed in February 26, 2005.
[12] OSBERG, K. Spatial Cognition in the Virtual Environment (Technical Report R-97-18). Seattle: Human Interface Technology Lab, College of Education University of Washington, 1997. www.hilt.washington.edu/publications/r-97-18
[14] CAVALLO, D., Emergent Designing and learning environments: Building on indigenous knowledge, IBM Systems Journal, Vol. 39, NOS 3&4, 2000.
http://www.research.ibm.com/journal/sj/393/part2/cavallo.html
[15] VARIS, T., Approaches to Media Literacy and e-learning
http://www.elearningeuropa.info/extras/pdf/media_literacy.pdf
[16] KAUFMANN, Hannes, Collaborative Augmented Reality in Education, Institute of Software Technology and Interactive Systems, Vienna University of Technology.
[17] FRANCO, J., F., e Lopes, R., D., Novas Tecnologias em Ambientes de Aprendizagem; Estimulando o Aprender a Aprender, Transformando o Currículo e Ações, III Ciclo de Palestras sobre Novas Tecnologias na Educação, http://www.cinted.ufrgs.br/ciclo3/IIICicloPrograma.htm Accessed in 03 April 2005.
[18] VALENTE, J., A, Formação de Educadores para o Uso da Informática na Escola (Teachers' preparation for Using Informatics at school), Brasil: Nied, 2003.
[19] RANDALL, N., Teach Yourself the Internet in a week second edition, USA: SAMS NET, 1995, 490-491 p.
[20] SZALAVÁRI, Z., et al. Augmented Reality Enabled Collaborative Work in "Studierstube", Institute of Computer Graphics, Vienna Universty of Technology, Austria.
http://www.cg.tuwien.ac.at/research/vr/studierstube/
[21] SIMPSON, M., S., The Virtual Designer – The Application of VRML in Collaborative Design, Information Environments Program, School of ITEE, University of Queensland, Australia.
[22] LIAROKAPIS, F.; Petridis, P.; Lister, F., P.; White, M. Multimedia augmented reality interface for e-learning (MARIE), The University of Suxess, Falmer, UK, Center for VLSI and Computer Graphics, School of Engineering and Information Technology.
[23] BILLINGHURST, M.; Kato, H.; Poupyrev, I., The magic Book: A traditional AR Interface. C and Graphics, 25, 2003, 745-753 p.
[24]Augmented and Mixed Reality, ARToolkit,
http://www.hitl.washington.edu/people/poup/research/ar.htm#artoolkit.
[25] ROUSSOS et al, Learning and Building Together in an Immersive Virtual World. In Presence vol. 8, n. 3, June 1999, pp. 247 - 263. (www.ice.eecs.uic.edu/~nice )
[26] ACKERMANN, Edith, Piaget' s Cosntructivism, Papert's Constructionism: What is the difference? Massachusetts Institute of Technology, (www.meida.mit.edu ).
[27] GADOTTI, M., Reading Paulo Freire, his life and work, Translated by John Milton, USA: University of New York Press, 1994.
[28] PETERS, Otto. Didática do Ensino a Distância, Brasil: Editora Unisinos, 2001.
[29] FREIRE, P., Pedagogia da Autonomia: Saberes necessários à prática educativa. (Pedagogy of Autonomy: Necessary knowledge to educational practices) Brasil: Paz e Terra, 2004.
[30] NOGUEIRA, N. Ribeiro, Pedagogia dos Projetos, Uma jornada interdisciplinar rumo ao desenvolvimento das Múltiplas Inteligências, Brasil: ed. Érica, 2001.
[31] PERRENOUD, P., 10 Novas Competências para Ensinar (10 New Competences for Teaching), Brasil: Artmed, 2000.
[32] ICEL, International Consortium for Experiential Learning
http://www.el.uct.ac.za/icel/index.html accessed in 2004
[33] BOEHN, B., Egeyed, A., Port, D., Shah, A., Kwan, J., and Madachy, R., A Stake holder win-win approach to software engineering education, Annals of Software Engineering 6 (1998) 295-321
http://sunset.usc.edu/research/WINWIN/winwinspiral.html A. in 2004.
[34] FRANCO, J., F., and Lopes, R., D., Novas Tecnologias em Ambientes de Aprendizagem; Estimulando o Aprender a Aprender, Transformando o Currículo e Ações, III Ciclo de Palestras

sobre Novas Tecnologias na Educação, http://www.cinted.ufrgs.br/ciclo3/IIICicloPrograma.htm accessed in February 19, 2005.
[35] PAPERT, S., THE LEARNING STATE + Tough Thinking about Tough Learning, Notes For Conference on Education, Technology and The Future of Maine's Economy. http://ase.tufts.edu/cogstud/papers/papert.htm .
[36] FEBRACE (www.lsi.usp.br/febrace ).
[37] http://guarani.lsi.usp.br/cidade Acessed em 2005
[38] NORVING, P., and Russell. S., Artificial Intelligence: A Modern Approach, USA: Prentice Hall, 2003, 649-785 p.
[39] PERRENOUD, P., Construir as Competências desde a escola, (Building competences since school), Brasil: ARTMED, 1999.
[40] AMES, A., L.; Nadeu, D., R.; Moreland, J., L., VRML 2.0 Sourcebook second edition, USA: Wiley and Sons, 1997. On-line book examples in: www.wiley.com/compbooks/)
[41] ZAKOUR, J., Foust, J., and Kerven, D., HTML How-To, USA: WAITE GROUP PRESS™, 1997.
[42] MANSSOUR, H., I., VRML tutorial in Portuguese http://www.inf.pucrs.br/~manssour/VRML/index.html Accessed 15.2.2005.
[43] CASTELLS, M., A Sociedade em Rede (The rise of the network society) Brasil: PAZ e TERRA, 2005, 65 – 118 p.
[44] PRESSMAN, R., Engenharia de Software (Software Engineering) Brasil: Mc Graw Hill, 2002, 751 – 779 p. [45] HARE, I., To succeed in this brave new information-driven world, what are the skills your child will need to thrive? The three Xs. And where can your child hone these skills? The Net, of course! http://www.mamamedia.com/areas/grownups/new/21_learning/main.html
[46] RICCIARDI, L, A Glimpse of the School of Tomorrow http://www.papert.org/articles/MaineYouthCenterArticle.html

# Short Papers

# An annotation scheme to analyse discussions in online classrooms

Aiman BADRI, Floriana GRASSO, and Paul LENG
*Department of Computer Science*
*University of Liverpool*
*{A.Badri,F.Grasso,P.H.Leng}@csc.liv.ac.uk*

## 1_ Abstract

Classroom discussion is an important element in many forms of online collaborative learning. We here discuss the design of systems to support tutors in the tasks of moderating such discussions and evaluating the contributions of participants. We aim to support the online instructor by providing a theoretical account of the discussion meta-structure and practical tools to assist the instructor in the moderating activity. Methodologically, we base our research on the analysis of a corpus of classroom interactions from an on-line Master's degree in Information Technology. In this paper we present a framework for the analysis of classroom discussions, intended as the basis of a system to support the role of the instructor. We also illustrate our experience in implementing the framework, and some preliminary results of evaluation of our process are also presented. The results demonstrate a high level of agreement among instructors in using the framework to annotate discussion contributions. We expect further results of instructor satisfaction will play an important role in shaping our final conclusions.

## 2_ Introduction

Research into systems of e-Learning has primarily focussed on the modelling of students, as the recipients of the education materials. Examples range from Q&A type [1] to user modelling [2] to intelligent tutoring systems [3]. Very few researchers have emphasised the importance of the role of the moderator of the classroom discussion [4]. Here we take the teacher's perspective, rather than the student's, as we seek to produce both a theoretical account of how discussion in a virtual classroom takes place, and practical tools to assist the teacher in the moderating activity. This classroom discussion has a central role in the teaching and learning paradigm used: it demands the active engagement of students in the learning process, and promotes collaborative learning and sharing of experiences among the class. Much research, however (e.g: [5]), while underlining the value of this kind of interaction, has noted that without direction from a moderator, a proportion of students will not participate in the discussion. It is the responsibility of the instructor both to guide the discussion and to evaluate the contributions of students for assessment.

The context for the work we describe is an MSc degree programme that is being delivered entirely online. Within this programme, described more fully in [6], students are taught in classes of no more than 20, interacting asynchronously with each other and their instructor using the SoftArc FirstClassTM system [7]. Modules are subdivided into seminars, each taking place over a single week, during which a topic of study introduced by the instructor is explored by the class. This includes a discussion of aspects of the topic being examined, carried out through contributions to an open folder in the module virtual classroom.

Our objective in this paper is to investigate the possibility of developing a framework that can objectively support the online instructor in moderating and evaluating these discussions. We focus on the discussion meta-structure aspect as we believe this can be made essentially topic-independent, a great advantage for portability over several fields of discussions. A more detailed description of our framework can be found in [8].

## 3 Annotation of Class Discussions

We operate on “contributions”, i.e. textual messages posted by the various students to the virtual classroom. The systematic analysis and evaluation of these contributions will provide the instructor with a summary of the virtual class discussions and activities. While the content of the single contribution is crucial to the analysis, our aim is primarily to model the discussion meta-structure, so one can use shallower natural language processing techniques to retrieve the contribution content [9], and focus on how the student’s contribution relates to others in the discussion [10]. We wish to capture the intensity of the discussion, the chain of responses, discussion turns and responses to responses created in the conversation [11]. The notion of a dialogue move, as an abstraction which captures the participation of an “agent” in the discussion, helps in this task [12]. Dialogue moves are defined on the basis of factors such as cognitive plausibility, ease of coding, reliability and computational tractability [10]. Here we concentrate on a technique for manual annotation of such moves, and the evaluation of this technique to inform the creation of a system that could at least partially automate some of these tasks.

First we must decide what are the “basic units” of the contribution: for example, single clauses, prosodic units, dialogue turns, sentences or intentionally defined discourse segments [13]. It has been argued that errors in discourse segmentation greatly influence the quality of the interpretation [14]; larger basic units, while computationally better, may leave significant rhetorical information out [15]. In our analysis we adopted a simple approach: we took simple end of sentence markers to segment paragraphs into sentences (such as “.” “?” and “,”). Secondly, it must be decided which types of *rhetorical relation* can exist among textual units [16]. There is debate on the number and the nature of the relations that should be used [17]. We are interested, however, not much in identifying what the contribution is about, but especially in understanding how it relates to the rest of the discussion, and to create a dialogue meta-structure in which all the contributions are represented.

To produce an annotation scheme for the analysis of class discussion, we selected “real” cases from a real classroom and conducted two consecutive studies, segmenting and annotating 3 weeks discussions. We started the first analysis with a list of 32 tags derived from a preliminary investigation. For the second study, we applied a revised annotation, reducing the list to 17 tags that included examples such as “Agree”, “Own Opinion” and “Source of Information”; other examples relate to the “Owner”, “Date” and “Subject”. From our experience in annotating over 10,000 sentences, we devised a grammar to guide our reconstruction process. In this grammar, each contribution is described as follows:

Contribution: <Context>.<Content><br>
Context:<Module_Name>.<Source>.<Subject>.<Time>.<br>
<Addressee>|<Aattached>|<Signature><br>
Content: <Block>*<br>
Block: <Main>|<Components>**<br>
Main:<Agree>|<Disagree>|<Own_Opinion>|<Own Experience>|<Question>

Components:<Repeat_DQ>|<Introduce_a_response>|<Statement_or_fact>| <Textbook>|<Contrast>|<Provide details>|<Source for information>|<Repeat sentence>|<Image>

*Note: The "*" means at least one occurrence, the "**" means 0 or more occurrence, the "." means "must exist together" and the "|" means or.*

A contribution's tag can therefore refer either to the *context* of the discussion thread (who the sender is, who the addressee is, what time and which thread of discussion it is part of, etc.), or to the *content* of the contribution (what the author wants to say to the hearer). A *content* annotation tag can refer to the *main* moves related to the discussion ("Agree", "Disagree", etc.) or to those that compose them. The *component* tags refer to additional elements that the contribution's author provides. In order to set up an experiment with annotators, we developed a manual that explains in simple words the meaning of each tag and how an annotator can recognise it.

## 4 Evaluation

We are evaluating the annotation scheme in two ways. The first experiment involves the annotation manual, and aims to test whether this leads to consistent annotation by different annotators. This experiment was intended to test the agreement among annotators using the manual for a pre-selected group of tags (the Main tags) and to verify the ease of use and usefulness of the manual. This involved 25 annotators from various backgrounds annotating 30 sentences. To investigate the reliability of the annotation scheme, we need to measure the level of agreement among annotators in assigning the tags to each sentence. Although there are several measures of agreement, such as Kendall coefficient of agreement $u$ [19], our research clearly identified the Kappa coefficient as the most commonly used tool applied to measure the agreement among annotators in a discourse environment [18, 20 among many others]. In fact the Kappa coefficient $k$ [19], as a conservative statistical measure of agreement, attempts to exclude any chance, in the agreement, of k raters in assigning one of m categories to N objects. In our case, k = 25; m = 5; and N = 30. The value of $k$ should range from 1 (total agreement) to 0 (no agreement). The $k$ value alone is however not sufficient: further analysis of the *significance* of $k$ is needed (z). The value of z should be greater than 2.32 (for _ = 0.01) for us to be able to say that the raters exhibited significant agreement on their ratings [19 page 290]. In our experiment, the result reported $k = 0.53$ for the total group, while the value of the significance was $z = 8.3$. This indicates a very strong agreement among our annotators. We also tested $k$ for 2 groups of our annotators; those who have some knowledge about our topic of discussion ($k = 0.54$) and those who do not have any knowledge about the topic ($k = 0.52$), showing that there is no significant difference between the two groups. This supports our claim that the annotation scheme can be used without any knowledge about the topic, an important indicator of the portability of the scheme to various topics. Analysis to each category of our main tags reported strong respondent agreement on the correct answer 93% for the Agree tag, 53% for the Disagree tag, 71% for the Own Opinion tag, 62% for the Own Experience tag and 89% Question tag.

In our second experiment, which is still on-going, we target the online instructors, with the aim of reviewing their satisfaction/opinion on the usefulness of our work. The experiment involves 2 groups of online instructors, using 2 full weeks worth of discussions.

## Conclusions

We presented in this work our preliminary investigation in the development of a framework to analyse discussion in online classrooms. Our objective is to support the online instructor by reflecting two views of the classroom: an activity view drawn from statistical data and a discussion view drawn from analysis of the corpus of the discussion. We have introduced an annotation scheme for the analysis of classroom discussions, and presented our experience in conducting extensive annotation processes. We have described our evaluative experiments and reported our preliminary findings that supported to a great extent our claims and objectives. The results clearly illustrate that it is possible to analyse a discourse without any knowledge base with the aid of a simple manual. We plan to investigate the usefulness of our framework to the online instructor, as an aid in their mentoring activities.

## References


1) John Caron: Applying LSA to Online Customer Support: A Trial Study (2000)

2) Conati, C., Gertner, A. S., VanLehn, K., Druzdzel, M. J.: Online Student Modeling for Coached Problem Solving Using Bayesian Networks. Proceedings of the Sixth International Conference, UM97 (1997)

3) Jerinic, L., Devedzic, V., Radovic, D.: The Component Based Model Of The Control Knowledge For Intelligent Tutoring.

4) Salmon, G.: E-Moderating: The Key to Teaching and Learning Online. Kogan Page (2000)

5) Klemm, W., Snell, J.: Enriching computer-mediated group learning by coupling constructivism with collaborative learning. Journal of Instructional Science and Technology 1 (1996)

6) Grasso, F., Leng, P.: Issues of quality in on-line degree programmes. In Ghaoui, C., ed.: Usability Evaluation of Online Learning Programs. (2003) 198–217

7) Persico, D., Manca, S.: Use of firstclass as a collaborative learning environment. Innovations in Education and Training International 37 (2000) 34–41

8) Badri, A., Grasso, F., Leng, P.: Evaluation of Discussions in Online Classrooms. In 7th International Conference on Knowledge Based Intelligent Information and Engineering Systems KES2003, Oxford, UK. (2003)

9) Riloff, E., Lehnert, W.: Information exteraction as a basis for high-precision text classification. ACM Transactions on Information Systems 12 (1994) 296–333

10) Cooper, R., Larsson, S., Matheson, C., Poesio, M., Traum, D.: Coding instructional dialogue for information states. Technical Report Deliverable D1.1, Task Oriented Instructional Dialogue (TRINDI) project (LE4-8314) (1999)

11) Smith, M.: Tools for navigating large social cyberspaces. Communications of the ACM 45 (2002) 51ss

12) Traum, D., Bohlin, P., Bos, J., Ericson, S., Larsson, S., Lewin, I., Matheson, C., Milward, D.: Dialogue dynamics and levels of interaction. Technical Report Deliverable D3.1, Task Oriented Instructional Dialogue (TRINDI) project (LE4-8314) (2000)

13) Carlson, L., Marcu, D., Okurowski, M.E.: Building a discourse-tagged corpus in the framework of rhetorical structure theory. In van Kuppevelt, J., Smith, R., eds.: Current Directions in Discourse and Dialogue. Kluwer Academic Publishers (to appear)

14) Marcu, D.: A decision-based approach to rhetorical parsing. In: Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL'99). (1999) 365–372

15) Marcu, D., Romera, M., Amorrortu, E.: Experiments in constructing a corpus of discourse trees: Problems, annotation choices, issues. In: Workshop on Levels of Representation in Discourse, Edinburgh, Scotland. (1999) 71–78

16) Marcu, D.: The rhetorical parsing of natural language text. In: Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics (ACL'97). (1997)

17) Hovy, E., Maier, E.: Parsimonious or profligate: How many and which discourse structure relations? available at http://www.isi.edu/naturallanguage/people/hovy/papers/93discproc.pdf (1995)

18) Marcu, D., Romera, M., Amorrortu, E.: Experiments in constructing a corpus of discourse trees. In: The ACL'99 Workshop on Standards and Tools for Discourse Tagging, Maryland. (1999) 48–57

19) Sidney Siegel and N.J. Castellan. 1988. *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill, second edition.

20) PARADISE: A Framework for Evaluating Spoken Dialogue Agents Marilyn A. Walker, Diane J. Litman, Candace A. Kamm and Alicia Abella AT&T Labs—Research 180 Park Avenue Florham Park, NJ 07932-0971 USA walker, diane,cak,abella@research.att.com

# Iterative Development of a Methodology for Analysing Collaborative Interactions.

Lucinda KERAWALLA, Nicola YUILL, Darren PEARCE, Rosemary LUCKIN and Amanda HARRIS..
*University of Sussex, Falmer, Brighton, BN1 9RH, UK.*

***Abstract***: Single-user interfaces can be detrimental to the collaborative process, even when pairs of children are given a mouse each. We describe a novel user interface, Separate Control of Shared Space (SCOSS), and present a study that explores its potential as a tool to resource collaborative interactions between children doing an estimation task. We then discuss the iterative development of two methodologies for analysing how the interface worked to support desirable collaborative behaviours. This included the generation of a coding scheme and timeline graphs that depict behavioural events. The advantages and disadvantages of these are discussed and used to inform the second iterative methodology which focused on more global measures.

## 1. Background

The Riddles project is concerned with increasing children's language awareness to foster their reading comprehension (e.g.[1]). We are developing software designed for children to work collaboratively in pairs on language tasks. We have had to develop an alternative to the usual single-user interface because such interfaces can be detrimental to the collaborative process, even when each child has their own mouse (e.g. [2] and [3]). We argue that the findings from both of these studies can be explained by the fact that the single user interfaces of the software used, allowed only one child to have access to each single feature at any one time, thus promoting turn-taking rather than concurrent task activity.

This paper will first briefly describe our own interface: Separate Control Of Shared Space (SCOSS), and a study that explored its use. This will be followed by a discussion of the iterative development of both our research questions and the analysis of the collaboration between pairs of children using SCOSS.

## 1.2 Separate Control of Shared Space (SCOSS): Features

When children are working collaboratively at a computer, the whole screen can be conceptualised as a shared space; ideally it mediates joint decision-making but the limitations discussed above mean that this does not always occur. In an attempt to overcome such problems, the design of SCOSS focuses on promoting the following outcomes: 1) joint understanding, 2) joint agreement, 3) working towards a shared goal, 4) an equal contribution to the task process, 5) minimal domination, 6) no deletion of each other's work, and 7) equitable input.
The SCOSS interface includes the following features and hence should enable the intended collaborative behaviours:

- The screen is divided in half so each of the pair has their own space in which to simultaneously carry out an identical task. Each space can be used both to mediate and to represent each child's current understanding of the task
- Each child can control elements only within their own space, therefore they cannot delete their partner's work. It also enables each individual to contribute to every part of the task process. The need for turn-taking is eliminated and domination is made less likely.
- Agreement and disagreement are made visually explicit: agreed-upon elements are coloured green whereas elements that are not agreed upon are not green. The children can then use the similarities and differences between the representations on the screen as a resource to mediate discussions about coming to an agreement.
- Users have to agree by pressing a 'happy key' before they can proceed to the next task or part of the task. This should encourage reasoning and explanation as part

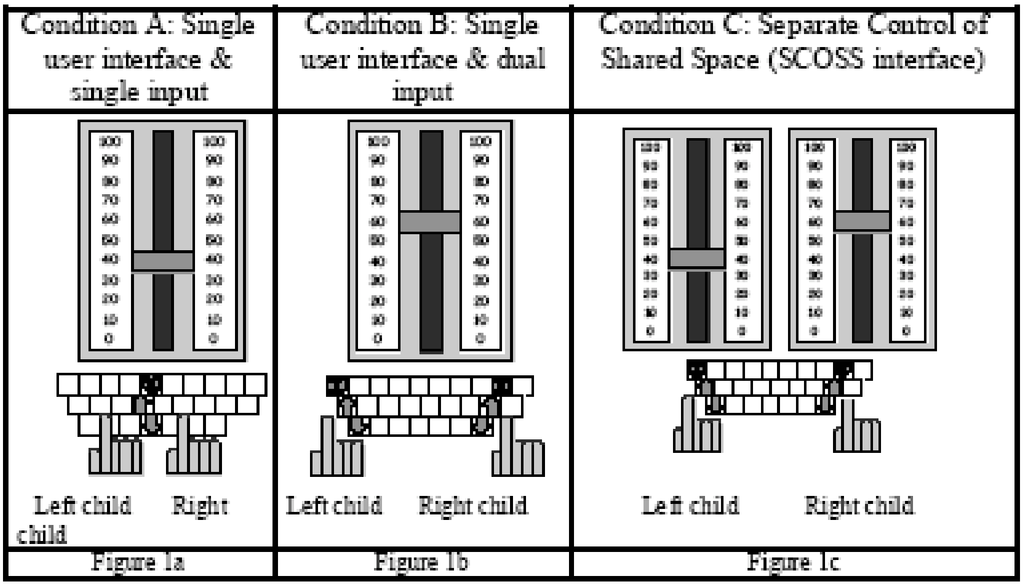
of the collaborative process.

## 2. Exploring SCOSS in an estimation task

### *2.1 Design and methodology*

This study compared the utility of the SCOSS interface with single control of a singleuser interface, and dual control of a single-user interface to determine whether children could use SCOSS as a tool to mediate their collaborative progress through the task. Thirty six pairs of 8-9 year olds estimated the number of sweets in eight ‘real’ containers.

In condition A the children shared a single-user interface (figure 1a), in condition B the children had dual input into a single-user interface (figure 1b) and in Condition C the pairs had dual input into the SCOSS interface (figure 1c).




### *6.1.1 2.2 Analysis and findings: first iteration of a coding scheme*

Our initial research question was *‘what are the children doing that is good collaborative practice and does the interface mediate this?’* We defined good collaborative practice in the term of the seven outcomes outlined above and then identified the finer behaviours that constituted each of these. For example, asking for their partner’s opinion (which meets outcomes number 1 to 4), explaining and justifying opinions by comparing containers with each other (1-3), the nature of the comparisons (outcomes 1-3), not pressing their partner’s keys ( 4-7) and asking their partner before pressing the ‘happy’ key (1-5).

We consequently developed an extensive coding scheme identifying a total of fifty-four discreet behaviours and utterances. These codes were applied by watching the video footage and completing paper-based tables, indicating the code and timestamp.

It was possible to quantify the number of times various behaviours occurred in each condition. For example, we found that there were significantly more comparisons between containers made in condition A than in B ($H=7.6$ $p<0.01$), and in A compared to C ($H=5.2$ $p<0.05$). However, such findings could not reveal anything about the quality of the events. We therefore decided not to pursue this path of enquiry further.

We also built upon previous work by Luckin [4] to explore whether it was possible to use this data to illustrate visually the events that resource the process of reaching a joint agreement. We designed a database and used it to produce timeline graphs (e.g. figure 2). These depicted where each behaviour (represented individually as symbols on the Y axis) occurred in time (represented on the X axis) for each child of the pair for each individual container. Long symbols represent events that lasted for several seconds. Logs of keyboard activity were overlaid (on the secondary Y axis) and these recorded movement up and down the estimation scales for each child, which are represented here as solid lines.

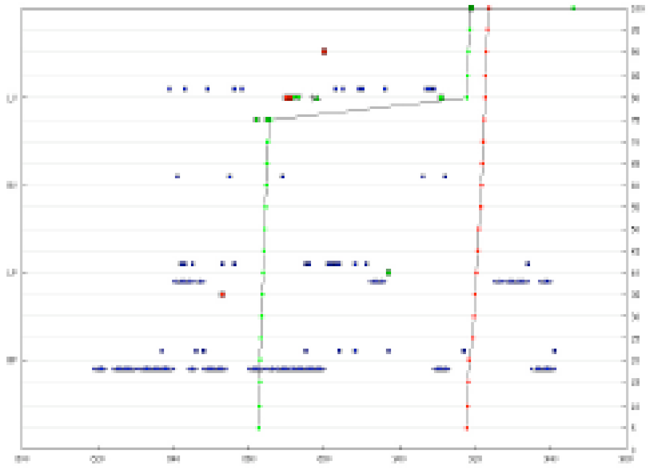


Figure 2: An example of a graphical representation of the joint decision-making process.

We envisaged that it would be possible to identify various patterns across the graphs, thus enabling us to systematically categorise and represent different ways in which joint decisions were reached. Some types did emerge, such as those where one child made only a minimal contribution, but overall no useful types were identified. Significantly, we also concluded that the nature of the coding scheme, in treating behaviours as discrete events, was detrimental to understanding how the children were working together.

The graphs also did not depict how the various interface conditions were having a direct role in mediating specific behavioural events. They were, however, a useful tool in aiding our exploration of the data and for informing our reiteration of the research question. The next section describes our second iteration.

***2.3 Analysis and findings: second iteration of a coding scheme***

The research questions for the second iteration of the analysis asked more specific questions about precisely how the SCOSS interface afforded the seven outcomes listed above. In contrast to the above coding scheme, in this analysis we made global judgements about whether the interaction showed evidence of outcomes 1-7.

Analysis revealed that there was no difference in the quality of the collaboration, across the three conditions which indicated that some children are not spontaneously good at collaboration and suggests that SCOSS alone cannot mediate the quality of the discussion surrounding decisions. However, the following transcript is evidence of the potential for the SCOSS interface to mediate joint decision-making (actions are in *italics* and the pair are estimating jar JJ):

> R: that one *[picks up box and holds it next to JJ and compares their size]* it's the same height *[stacks one on top of the other and compares width]*
> L: not the same. It's 75 at the moment *[uses own keys]*
> R: I think 80, 85
> L: 85.You do 85 and I'll do 75 at the moment
> R: *[uses own keys]* yeah but our answers have to be the same
> L: I know we need to see *[picks up box labelled 40]* 40 and then that'll be two of them

These children not only use the screen to recognise that they have changed their mind but also made good use of the possibility to represent transient disagreement on their scales. There is clear evidence that they both understand that they need to agree eventually on the final answer, which they do finally achieve.

This finding indicates that our analysis needs to be even more finely tuned to exploring the exact moments when the interface is having a direct role in encouraging an awareness of agreement and disagreement and the collaborative benefits of both. Our current work on iteration three will hopefully address these issues for the future.

## 3. Conclusions

We have found that an iterative approach towards the analysis of collaborative interactions has been useful in terms of understanding the rich and complicated nature of collaboration mediated by technology and for enabling us to be clearer about our own focus and research questions. This paper does not mark the end of our iterative cycle but serves as a useful example of the benefits of iteration as a methodology in itself. We are currently refining our analysis methodology to assess the role that SCOSS plays in the mediation of desirable collaborative interactions, and will be using it to analyse a further set of data from pairs of children working on a language task.


[1] Yuill, N. and Oakhill, J. (1991), *Children's problems in text comprehension*, UK: Cambridge University Press.

[2] Scott S., Shoemaker G. and Inkpen K. (2000), Towards seamless support of natural collaborative interactions, in *Proceedings of Graphics Interface*, Montreal; Canada May. pp103-110.

[3] Benford, S., Bederson, B., Akesson, K., Bayon, B., Druin, A., Hansson, P., Hourcade, P., Ingram, R., Neale, H., O'Malley, C., Simsarian, K., Stanton, D., Sundblad, Y. & Taxen, G (2000), Designing Storytelling technologies to encourage collaboration between young children. In *Proceedings of CHI 2000*, The Hague, Netherlands, April1-6.

[4] Luckin, R. (2003), Between the Lines: documenting the multiple dimensions of computer supported collaborations. *Computers and Education* (Special issue on Documenting Collaborative Interactions). Volume: 41, Issue : 4. pp 379-396.

# Towards adaptable interaction analysis tools in CSCL

José Antonio Marcos[1] and Alejandra Martinez[1] and Yannis Dimitriadis[2]
[1]*Department of Computer Science*
[2]*Deparment of Signal Theory,Communications and Telematics Engineering*
*University of Valladolid*
*Camino del Cementerio, s/n. 47011, Valladolid, Spain*
*jamarcos@ulises.tel.uva.es, amartine@infor.uva.es, yannis@yllera.tel.uva.es*

**Abstract**. Interaction analysis has become a basic function in the field of collaborative learning as a means for supporting *evaluation* processes. These processes can benefit from the use of automatic or semi-automatic interaction analysis tools. If these tools considered the different *roles* implied in the analysis processes, this could permit to exploit the results of the analysis in function of who is the user and what is his/her purpose. The experience of awareness systems in CSCW that use roles to decide the type and amount of information that they show suggest that this can be an appropriate approach. However, a review of the concept and classification of roles in the CSCL literature has shown a great diversity of classifications and a lack of common vocabulary to describe roles that also ignore the dynamic aspects of real situations. These aspects demand a new dimension for the classification of roles capturing dynamic aspects, such as the evolution of roles in an activity. Moreover, they demand a common vocabulary for defining and describing roles in learning scenarios. This would allow to automatically adapt the functionalities of interaction analysis tool to the evolving needs of the roles. This paper elaborates two proposals that help to detect the changes of roles produced during the collaborative activity and identify the needs established for these roles.

## 1. Introduction

Interaction analysis supports different functionalities in Computer Supported Collaborative Learning (CSCL), such as the *evaluation* of collaborative learning processes. Currently, the evaluation of CSCL systems, and of the learning promoted with them is a priority of research in the area. For this purpose, the researchers propose the elaboration of powerful tools and methods for interaction analysis in the study of collaboration [1].

Our group has been working in the evaluation and the analysis of interactions for the last years. A main research effort has been the development of a system for supporting formative evaluation in CSCL settings. In order to meet this aim, we proposed the *Mixed Evaluation Method* [2]. It defines a general approach oriented to support the formative evaluation of participatory aspects of collaborative learning in real classrooms.

This method is partially supported by an interaction analysis tool called *SAMSA* that builds social networks and computes a set of indexes that are shown to the users for its later analysis. Although this tool was designed to be used by teachers and researchers, its use has shown that it might be useful to support self-evaluation by the students [3]. However, these types of users (teachers, researchers, students) have different needs. This has lead us to work on a more general problem, consisting on the study of these users' characteristics and how the interaction analysis tool can be dynamically adaptable to their different needs.

From the experience of awareness systems that use role-based proposals in the CSCW field [4],[5] we can state that the collaboration-support tools would benefit from considering this adaptation to the user, in order to improve the collaborative processes supported by them. A main problem in these systems is the quantity and type of information that they have to display in a moment, as well as how it is to be shown, depending on the different roles that participants take during collaboration. From this perspective it is possible to think on interaction analysis tools meeting the needs of different types of users, providing them with different functionalities. For example, the data obtained by the teacher in an intermediate evaluation with *SAMSA* could be used to support the students' self-regulation, but it does not seem reasonable to show the same information to the teacher than to a K-12 or to a university student, neither to use the same format for all the cases. This requires considering the needs of teachers and learners (and other potential participants) in every moment.

It is therefore necessary to identify the *roles* that can appear in the collaborative process, and the requirements they pose to interaction analysis. Also, it will be necessary to detect dynamically the changes of roles during the development of the collaborative activity to adapt the functionalities of interaction analysis tools to their evolving needs. We aim at supporting these processes automatically or semi-automatically.

This paper elaborates on these issues, and proposes a *structured description* for roles, and a new *dimensions* for the classification of roles that permits to capture dynamic aspects, such as the evolution of roles in an activity. This results in a two-way relationship between roles and analysis of interactions. First, analysis of interactions helps to identify roles, and then, these roles (i.e., the people representing them) will be supported by interaction analysis functions. According to the general goal presented beforehand, this support will be adapted to the needs of a particular role.

The rest of the paper is structured as follows: The next section presents a brief summary of the different role definitions and classifications found in the literature. Next, section 3 presents our proposal, which includes the new dimensions for the classification of roles based on dynamic aspects, and the structured description of roles in CSCL. The paper concludes presenting the open research questions and an overview of our future research plans related to these topics.

## 2. A Review of Roles in Learning Systems

A review of the existing proposals related to roles, based on works from the CSCL, e-learning, CSCW, group dynamics and classroom-based research shows a great diversity. There is a lack of common vocabulary to describe roles, multiple definitions and very different classifications of roles, many of them, domain-dependent.

In this review we have detected a rather high consensus with respect to the generic roles (to which we will refer to as actors) that can be identified in a learning scenario, such as the teacher, the student or the designer [9,13,11,12]. On the contrary, the teachers' and students' roles (to which we will refer to as functions) depend very much on the approach and on the context of each work, and that there is no such consensus between the different authors [7,18,16,17].

We have detected different functionalities for the same role. For example, about the Teacher-facilitator role, [7] states that "must create learning situations and improve the motivation of learners", but [10] considers that "they monitor the collaboration activities within a group, detect problems and intervene". But also we have identified different roles whit the same functionality (e.g. [16] and [13] describe in the same terms the roles of the teacher-guide and the tutor).

Moreover, in these descriptions, roles are presented as static entities, in the sense that membership to a given role is established early and rarely changed. Nevertheless, potential role membership can vary from moment to moment during the lifetime of a collaborative session. These roles ignore situational dynamics ("the real world"), and they do not provide the flexibility needed in many situations [6]. For example, during the evolution of a certain task, a teacher should be able to shift from the initial teacher-guide function, when students need more help, to the teacher-observer function, when the students have reached some autonomy [21].

These aspects demand a common vocabulary for defining and describing roles in learning scenarios, and a new dimension for the classification of roles capturing dynamic aspects, such as the evolution of roles in an activity. With these elements, it would be possible to define a description of roles able to be managed computationally, and detect the changes of roles produced during the activity. This would allow to automatically adapt the functionalities of interaction analysis tools to the evolving needs of the dynamic roles. We present our proposals regarding this issue in the next section.

## 3. A new proposal for describing and classifying roles in CSCL systems

We have identified two dimensions that dynamically helps to identify the roles established before the beginning of the activity, and to detect the changes of roles produced during the activity.

Regarding *the moment of their appearance* we define **pre-established** and **emergent roles**. *Pre-established roles* are those that are assigned before the beginning of the collaborative activity (e.g. role defined by a type of task). *Emergent roles* are those that are not assigned in advance, but that appear spontaneously during the development of the activity [22].

According to their *variability* we define **static** and **dynamic roles**. *Static roles* are those that remain invariable from the moment of their appearance until the collaborative activity finishes. *Dynamic roles* are those that vary during the development of the collaborative activity (e.g. due to a rotation of roles among the members of a group).

To detect emergent and dynamic roles, it will be necessary to define a set of indicators and the values that identify the transitions between the different roles. These indicators will be a component into the formal description of a role in a common vocabulary. Then, the interaction analysis tool will be able to identify a change of role by means of these indicators, and it will be able to adapt to their new needs, which will be to specify into the formal description of this role.

Then, the problem would consist on making a common framework for defining and describing roles, their functions and needs for a generic context. We have elaborated a proposal of a structured definition and description of roles in CSCL composed by four dimensions that aim to solve the lack of common vocabulary detected to define and characterize roles in learning scenarios. These four dimensions are: actor, function, needs and indicators.

In this context, an ***actor*** represents a generic role, that is, a human, an agent or any combination of them [9] (i.e., the *teacher* and the *student* roles have been pre-established roles in a traditional classroom [17]).

A ***function*** is a characterization of an actor. With a function we could specify its activities, duties and responsibilities (i.e., as a *facilitator*, "a teacher perform a minimal pedagogical intervention in order to redirect the group work in a productive direction or monitor which members are left out of the interaction" [10]).

A ***need*** is a requirement for each pair role-function. These requirements relate to the necessary information (quantity and type) and the functionalities of tools, and they are influenced by diverse parameters related to the context, such as the environment (synchronous or asynchronous), the educational level of the students (University, K-12, etc), the goal pursued with the interaction analysis tool (regulation, formative evaluation) or the specific activity (i.e. collaborative edition).

An ***indicator*** is a parameter that helps to identify the transitions between the different roles. Each indicator is composed by a name and the values that identify a possible change of role. The values can be different depending of the context. A tool that recognized changes of roles would use the limit values described by the indicators.

This structure allows for a static description of a role, in the sense that it can be established at the beginning of the collaborative activity. In addition, we have included the *indicators* dimension for defining the values that permit to detect transitions between roles. The next step will be to implement this approach for its computational management.

These approaches aim to provide for the adaptation of the functionalities of interaction analysis tools to the changing needs of roles.

## 4. Conclusions and Future Work

This paper has presented the need of considering roles when designing tools for the analysis of interactions. This will permit to support the analysis of interaction data collected from a learning experience and to exploit them depending on who is the user and what are his needs. This approach is based on existing proposals from awareness systems in CSCW that adapt to the users' profiles.

An initial review of the concept of role in the literature has shown many different definitions and a great diversity of classifications, many of which are domain-dependent and ignore the dynamic aspects of real situations. Due to this diversity, we have proposed a new dimension based on dynamic aspects for classifying roles, such as the evolution of roles in an activity. Moreover, we have proposed a structured description of roles in a common vocabulary. Then, the interaction analysis tool will be able to identify a change of role, and it will be able to adapt to their new needs, which will be to specify into the structured description of this role. These are previous steps towards building analysis interactions tools capable of adapting their functionalities depending on the needs of roles at a given moment.

We have to go further in the description of the roles that are involved in collaborative learning scenarios, and establish their functional and user-interface needs. These needs will define the type of support that the different roles will need, which must me achieved by the interaction analysis functions. Moreover, it will be necessary to define the set of indicators and the values that identify the transitions between the different roles, and find a formal representation of this information for its computational and automatic management.

**Acknowledgements**

This work has been partially funded by European Commission Project EAC/61/03/GR009, the Kaleidoscope network of excellence FP6-2002-IST-507838, and Spanish Ministry of Science and Technology Project TIC2002-04258-C03-02. The authors would also like to thank the rest of "Intelligent & Cooperative Systems Research Group" at the University of Valladolid for their support and ideas.

**References**

[1] Jermann, P., Soller, A., & Muehlenbrock, M. (2001). "From Mirroring to Guiding: a Review of the State of the Art Technology or Supporting Collaborative Learning". *Proceedings of European Conference of Computer Support Collaborative Learning (EuroCSCL´2001).* Maastricht.
[2] Martinez, A., Dimitriadis, Y., Gomez, E., Rubia, B., & de la Fuente, P. (2003). "Combining Qualitative and Social Network Analysis for the Study of Classroom Social Interactions". *Computers and Education, Special Issue on Documenting Collaborative Interactions Issues and Approaches*, vol.41, no.4, pp.353-68. Kluwer Academic Publishers
[3] Martinez, A. (2003). "Method and Model to computational evaluation support in CSCL" (in Spanish). Department of Computer Science. University of Valladolid. Valladolid, Spain
[4] Drury, J. & Williams, M.G. (2002). "A Framework for Role-Based Specification and Evaluation of Awareness Support in Synchronous Collaborative Applications". *Proceedings of 11th International Workshop on Enabling Technologies for Collaborative Enterprises (WETICE02).* Carnegie Mellon University, Pittsburgh, IEEE Computer Society Press.
[5] Dourish, P. & Belloti, V. (1992). "Awareness and Coordination in Shared Workspaces". *Proceedings of the Computer Support for Collaborative Working (CSCW) Conference*. Toronto, Canada, ACM Press.
[6] Edwards, W.K. (1996). "Policies and Roles in Collaborative Applications". *Proceedings of ACM Conference of Computer Supported Cooperative Work (CSCW'96).* pp. 11-20
[7] Marín, M. & Pérez, G. (1985). "Social Teaching and Sociology of the Education: Roles of Teacher as the Leader of Classroom" (in Spanish). pp. 228-48. Madrid, Spain: UNED Publications.
[9] ISO/IEC JTC1 SC36 N0065. (2001). "Expertise and Role Identification in Learning Environments." *Information Technology for Learning, Education and Training*.
[10] Chen, W. (2004). "Supporting Teachers Intervention in Collaborative Knowledge Building". *Proceedings of 16th European Conference on Artificial Intelligence (ECAI'2004). Workshop on Artificial Intelligence in Computer Supported Collaborative Learning*. Valencia, Spain. E.Gaudioso & L.Talavera (eds). pp. 1-5
[11] Lijnse, P. ( 2000). "Didactics of science; the forgotten dimension in science education research?". *Improving science education: the contribution of research. Millar, R. and Osborne, J. (ed*). Buckingham, UK:Open University Press.
[12] Mizoguchi, R. & Bourdeau, J. (2000). "Using Ontological Engineering to Overcome Common AIED Problems." *International Journal of Artificial Inteligence in Education*. 11, pp.107-121.
[13] Lund, K. (2003). "Human Support in CSCL". Netherlands: Kluwer Academic Publishers.
[16] Shank, P. (2004). "Competences for Online Instructors." *Learning Peaks LLC*.
[17] Tinzmann, M. (1990). "What Is the Collaborative Classroom". *Proceedings of NCREL.* Oak Brook.
[18] Gisbert, M. (2000). "The Teacher of 21th Century" (in Spanish). *New Technologies for Learning Imporve. J. Cabero et al(ed*). pp. 315-330. Sevilla, Spain: Kronos.
[21] Marcos, J.A. , Martinez, A., & Dimitriadis, Y. (2004). "The Role of Roles in the Analysis of Interactions in Collaborative Environments". *Proceedings of 16th European Conference on Artificial Intelligence (ECAI'2004). Workshop on Artificial Intelligence in Computer Supported Collaborative Learning*. Valencia, Spain. pp. 23-25
[22] Strijbos, J.W. (2003). "Functional Versus Spontaneous Roles During Computer-Supported Collaborative Learning". *Proceedings of* 2nd *International Conference on Multimedia and ICTs in Education (m-ICTE 2003).* Badajoz, Spain.

# Insightful Problem Solving Promoted by Collaboration: The Effect of Role Reversals between Trial and Observation

Taro IZAWA[1]
*Graduate School of Interdisciplinary Information Studies, University of Tokyo,*
*7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan*
*Sachiko KIYOKAWA*
*Center for Research of Core Academic Competences, School of Education, University of Tokyo,*
*7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan*
*Kazuhiro UEDA*
*Department of General System Studies, University of Tokyo,*
*3-8-1 Komaba, Meguro-ku, Tokyo, 153-8902, Japan*

**Abstract**. We experimentally examined the effects of role reversals on insight problem solving. Forty-two undergraduate students were randomly assigned to one of two conditions: single and pair. In the single condition (14 solos), the participants worked alone. In the pair condition (14 pairs), each pair of two participants tried to carry out the task by alternating the role of working with the pieces and the role of watching a partner's performance every 20 seconds. The results showed that the rate to reach the solution in the pair group was significantly higher than the rate to reach the solution in the single group. In addition, the solution time turned out to be significantly shorter in the pair condition. Moreover, there existed a significant difference in frequencies of deviation from the constraints between the two groups.

## 1. Background

Previous studies have shown that collaboration can facilitate problem solving (e.g. Miyake, 1986; Okada and Simon, 1997). In these studies, protocol analyses were used in order to examine the processes of collaborative problem solving. Although these post hoc analyses are useful, it is hard to clarify how the facilitative effect emerges only by post hoc analysis because a variety of processes occur during collaboration. In this study, therefore, we investigate the effects of role reversals between task-performing and observation on insight problem solving by controlling the way of interaction between members.

In addition, theoretical framework is needed to analyse the processes of collaborative problem solving. We rely on *dynamic constraint relaxation theory* (Hiraki and Suzuki, 1998).

In the theory, three types of constraints, *object-level*, *relational* and *goal* are hypothesized. The object-level constraint is our natural tendency to encode objects at a basic level, although there are numerous other ways of interpretations. The relational constraint is a tendency to make a choice of specific relations among innumerable alternatives. The word "relation" is defined as the manner in which objects relate to each other and each object has a specific role. The goal constraint gives feedback to the two other constraints mentioned above, by evaluating a match between present and desired states. A desired state and an evaluation function are based on the representation of a goal. Hiraki and Suzuki (1998) suggested that an impasse is constructed by these constraints and the incremental relaxation of the constraints driven by failures brings about qualitative transitions probabilistically.

In this study, we experimentally examine the effects of role reversals between taskperforming and observation on insight problem solving and analyse the processes in the basis of the dynamic constraint relaxation theory.

## 2. Method

### *2.1 Participants*

Participants were 42 undergraduates. They were basically asked to come as a pair, as they were initially asked to bring a friend of the same-sex. They were randomly assigned to one of the two conditions: *single* and *pair*. The independent variable was whether participants played without role reversals (in the single condition) or with role reversals (in the pair conditions).

### *2.2 Task*

T puzzle was used as a material. The puzzle consists of four wooden pieces. A shape of "T" is to be constructed from these pieces (Fig. 1).

### *2.3 Procedures*

In both conditions, the participants were presented a sheet of paper with a 2/3-sized image of a finished "T" in front of them prior to beginning the problem. Then they were asked to construct a shape of T using four wooden pieces. They were not allowed to utter any word. They were notified about the following points prior to the experiment; (1) a bell rang every 20 seconds; (2) the whole time limit was 20 minutes; (3) a sheet of paper was taken away before they start working with the puzzle. The entire course of experiment was video-taped for analysis. In the single condition, the participants were required to engage in the task on their own.

In the pair condition, the participants were asked to work in the setting shown in Figure 2. The setting was designed so that the participants saw their partners not directly but through displays. The reason why each participant watched a partner's performance through a display instead of looking at it directly is to avoid making the participants look at trials from different angles.

At the beginning, a board on which the participants worked with the puzzle was in front of participant A and s/he started working with the puzzle. While participant A was working, participant B watched her/him working through Display B in front of her/him. They were asked to watch the display, thinking how to solve the puzzle. After 20 seconds, a bell rang and participant A passed the board with pieces to participant B without touching the form s/he constructed. Then participant B began the task and participant A watched his/her trial through Display A in front of her/him. They repeated (3) and (4) until they discovered the correct answer.

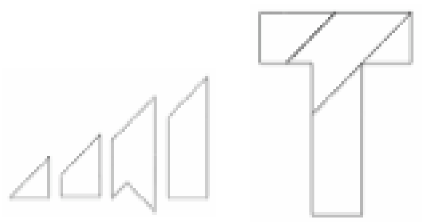


**Figure 1**. T puzzle. Construct "T" as shown on the right putting together four pieces.

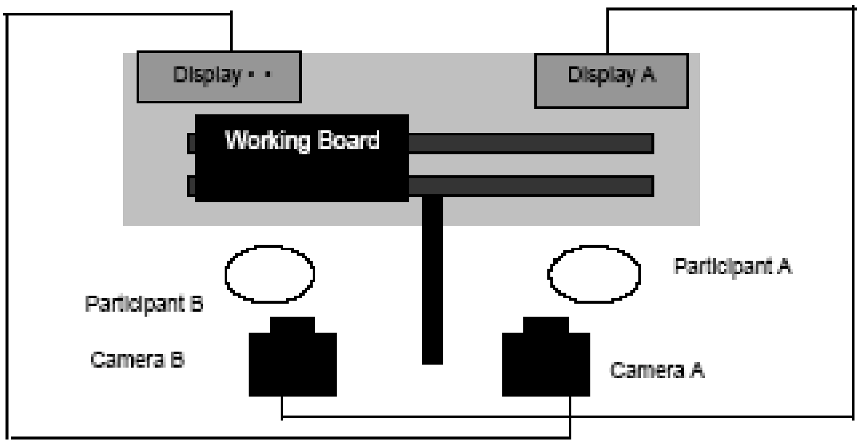


**Figure 2**. The experimental settings in the pair condition.

## 3. Results

*6.1.2 3.1 Performance Analysis*

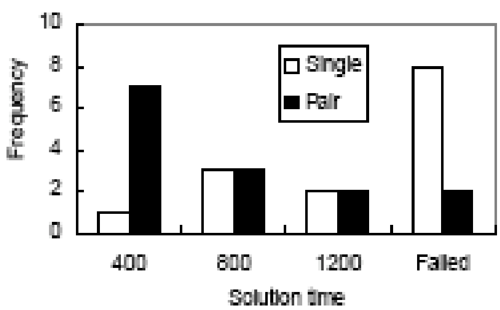
In order to examine the effect of role reversals on performance, the participants who could complete the task were classified into three groups, according to the time taken to solve the puzzle (Fig. 3). In the single condition, many of the participants were classified as "failed", so it can be presumed that the peak solution time in the single condition should be more than 20 minutes. Meanwhile, in the pair condition, the peak of the solution time lies less than 400 seconds and the number of participants who could solve the puzzle drops as solution time becomes longer. This difference in distribution of solution time in each condition is verified by Fisher's exact test (p=0.04). These results show that role reversals have a facilitative effect on solving the puzzle.

*3.2 Analysis of how pieces were placed*

According to the framework of dynamic constraint relaxation theory mentioned above, we examined the effects of role reversals on the relaxation of constraints.

Because people have a strong tendency to place the pentagon piece either vertically or horizontally (Suzuki and Hiraki, 1997), putting the pentagon diagonally was regarded as an index of relaxation of the object-level constraint.

In terms of the relational constraint, people tend to connect pieces in order to fill notches. Therefore, the index of relaxation of the relational constraint was defined as connecting the pentagon with other puzzle pieces at the central part of "T". The central part indicates that the pentagon is placed at the intersection of two bars of "T".



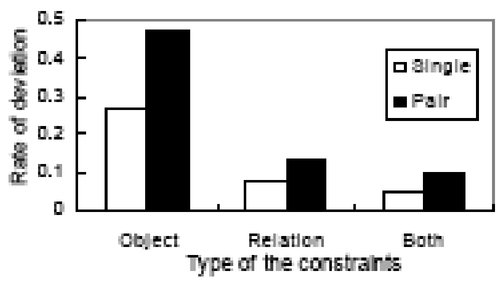



**Figure 3.** The frequency of the participants who accomplished the task within 400, 800 and 1200 seconds, and the number of those who could not solve it within the given time in each condition.

**Figure 4.** Mean rate of deviation from the constraints in each condition.

Furthermore, the number of times which the participants "put the pentagon diagonally (including at the right angle place)" and "connected the pentagon with other piece as the center part" is counted as an index of the strong relaxation of the constraints, since these factors are related not only with object-level and relational constraints but also with a goal constraint.

Figure 4 shows the mean rates of all the indexes of constraint relaxation in each condition. __2-tests show that the participants in the pair condition deviated from constraints more often than those in the single condition (__2 = 30.31, 8.27, 9.70, p<0.01, respectively). These results indicate that, in the pair condition, the participants were more likely to be activated to deviate from the constraints in both object and relational.

### 4. Discussion

The purpose of the experiment was to investigate the effect of role reversals between performing and watching partner's trials in insight problem solving. The result showed that the participants in the pair condition generally took less time to solve the puzzle than those in the single condition. Moreover, the participants in the pair condition placed pieces in ways which deviated from the constraints more frequently than those in the single condition. The difference between the pair condition and the singe condition was whether the role reversals existed or not. Thus, the facilitative effect found in the pair condition could be attributed to the effect of role reversals. It can be considered that role reversal is a type of effective collaboration to solutions of insight problems.

Based on protocol analyses, several previous studies pointed out the importance of role reversals between doing tasks and watching a partner's performance (e.g. Shirouzu, Miyake, and Masukawa, 2002). Their suggestion was tested by a controlled experiment in the present study. In addition, it was revealed that role reversals facilitate relaxation both of object-level and relational constraints. In consequence, it can be said that the method used in the study is useful in clarifying how the facilitative effects of collaboration emerge.

However the method used in the study has some limit. First, because we focus on the effects of the role reversals in the study, we were not able to compare different type of collaboration. As mentioned earlier, there exists a variety of processes during collaborative problem solving in our daily settings. Therefore it is necessary to compare these various processes and to determine relative effectiveness among different types of collaboration.

Second, because we restricted utterance during problem solving, we were not able to examine what kinds of mental activities the participants engaged in when they were observing their partners.

Finally, further researches are needed in order to determine the mechanism that role reversals facilitate constraint relaxation during insight problem solving.

### References

[1] Miyake, N. (1986). The constructive interaction and the iterative process of understanding. *Cognitive Science*, 10, 151-177.

[2] Okada, T., & Simon, H.A. (1997). Collaborative Discovery in a Scientific Domain. *Cognitive Science*, 21, 109-146.

[3] Hiraki, K., & Suzuki, H. (1998). Dynamic constraint relaxation as a theory of insight. *Cognitive*

*Studies: Bulletin of the Japanese Cognitive Science Society*, 5, 69-79.
[4] Hiraki, K., & Suzuki, H. (1997). Constraints and their relaxation in the processes of insight. *ETL Technical Report*, TR-97-13.
[5] Shirouzu, H., Miyake, N., & Masukawa, H. (2002). Cognitively active externalization for situated reflection. *Cognitive Science*, 26, 469-501.

# Adaptive Visualisation Support for Self-managed Learning Groups

Peter REIMANN*, Judy KAY**, Kalina YACEF**,
Peter GOODYEAR*
** Computer-supported Learning & Cognition (CoCo)*
*** School of Information Technologies*
*University of Sydney, Sydney, Australia*

**Abstract:** In authentic, long term projects involving groups, people typically use multiple communication systems and representational notations. This paper describes our work towards a system that will support groups in learning how to work more effectively as groups: a system that exploits the trails of data about group interaction across diverse media and tools. We review our foundational theories and describe our approach. Core to our work is the use of visualisations of learner models to support group members in learning to select suitable media for their collaborations and to improve their understanding of how to be effective group members in long term activities.

## Introduction

In authentic, long term group work, it is the norm that people make use of a rich, diverse collection of communication systems, such as chat, discussion forums, and video conferencing. It is also typical that they make use of a range of tools and representational notations within one medium including, for example, written text and diagrams. We believe it is critical to begin to explore group support systems that can operate in the context of such media richness, exploiting the potentially huge amounts of data that could be available. We are particularly interested in three classes of learning that could occur in such situations:

_ Learning to solve problems in a domain more effectively;
_ Learning about the team, its members, and effective ways of cooperating and collaborating;
_ Learning to use communication media and representational notations that match the
demands of the tasks at hand, including tasks of member and collaboration management.

Our work focuses on *self-managed* learning groups. This is in contrast to forms of group work characterized by a high degree of didactically-imposed structure, with tightly specified roles for group members and where the interaction between group members is scripted in detail. This kind of collaborative learning certainly deserves its place in the didactical repertoire of teachers and other educators, for example in situations where groups are short-lived and/or formed on an ad-hoc basis. In contrast, our focus is on forms of collaboration that takes place in groups that (attempt to) form real social units, work together over longer stretches of time (weeks and months rather than hours), and where there is little or no guidance from outside. Examples of such groups are the small teams of students formed for problem-based learning in universities. One of *our* examples is a group of students learning about eXtreme programming, and various communication and software 2 technologies, in the context of completing a classic capstone group software project. The other is a group of instructional designers learning about ontology engineering and applying it in the context of designing a course. Ideally, these groups are learning at all three levels mentioned above. Such learning groups, being highly self-organized, have to establish norms and processes that are conducive to knowledge building and sharing. Importantly, members of a group will naturally focus on their "production task" and on learning about domain issues. For their parallel learning goals of learning how to facilitate the effectiveness of the group, and about media use, they need to establish and maintain common ground, keep the group stable, and take care of individual members' concerns [1]. The move to machine media interaction among group members has a dual effect. First, it makes it even more important to attend to these group functions. Second, it opens the potential for exploitation of the data trail that can be collected by the machine.

**Our approach.** A number of researchers in the field of Computer-Supported Learning (CSCL) have begun to address this issue of collaboration management. Managing on-line collaboration by means of intelligent support can take a number of forms: mirroring, metacognitive and advice tools [2]. They all require the ability to trace the interaction between the team members at some level of detail. We are building upon this work and intend to extend it into two directions: Firstly, in addition to supporting member interaction directly with feedback and/or advice systems, there is a need for learners to develop skills in choosing the right communication medium and tool for the situation at hand. Approaches to collaboration management that rely on a single communication medium, and/or on strongly restricted notational systems used for communicating [3, 4] need to be extended, because groups typically do not accept such limitations over longer stretches of time [5]. Having the choice among various communication and representation systems, however, adds to the demands groups face: they now have to deal with the additional issues of task-to-media fit [6] and task-to-representation fit [7]. Secondly, we address human-computer interface issues extensively; not only because the management of task and interaction information distributed across various communication media raises serious attention and cognitive load issues, but also because of the social signals that come with using certain media [8] and which have not been reflected sufficiently in research on computer-supported learning. We suggest an approach where the shared interface can be adapted to the needs of the work on the task as well as to the needs of interaction and member management. In the absence of a conclusive research base to derive advice from, our short term goal is to create an environment where such phenomena can be studied under controlled conditions and to experiment with various ways of visualizing information for groups and facilitators/moderators.

### Adaptive Collaboration Visualisation

There has already been some work towards adaptive systems to provide advice on collaborative learning, for example [9]. There has also been recognition of the importance of social parameters, such as participation patterns [10]. We will explore the use of adaptive information presentation using visualisations of the collaboration. These seem particularly promising because they are easier to implement than advice systems and no normative model of collaboration is required.

**What to record.** We are working on finding research-based answers to three questions around the process: (1) What to record about the learners' performance; (2) How to aggregate and then analyse the traced information; (3) What and how to visualize the results from step 2, in a manner that is adapted to the group's needs. With respect to question (1), we propose to capture *all* task- and group-related exchanges available, 3 regardless of whether these involve the whole group, sub-groups, or individual members. Since we expect to be able to motivate the group members to help monitor their own interactions, we will be able to encourage the use of tools that we have set up to capture a rich record of interactions.

**How to aggregate.** An immediate effect of this is that we have to deal with large amounts of information. This must be analysed and summarised. Our approach with respect to question (2) is to collect the full set of available, un-interpreted data and then to perform a series of analyses to create both individual learner models and collective group models. We will use machine learning and data mining techniques (association rules, classification and clustering techniques such as hierarchic clustering, k-means, decision trees and data visualisation in particular) to identify patterns in groups' performance and relate those to outcome measures such as the quality of the groups' decision models and participants' satisfaction with the group process. Data mining and machine learning techniques have been successfully used for user modelling and, to a lesser extent, in education contexts. In particular, mining data based upon learners' interactions with a learning environment is promising [11].

Since a user model captures the system's beliefs about the learner's knowledge, beliefs, preferences and other attributes, it has the potential to play an important role in providing external representations of the individual and group learner models relevant to the group interaction and learning. There has been a growing appreciation of this possibility, with learner models being shared with learners in order to support reflection [11-14] and to help learners work collaboratively [15]. The challenges in this project are to mine the available data sources to support the construction of a student model [16], to provide natural interfaces that enable learners to see and understand the externalised form of that model [17], to explicitly contribute to it and, finally, but most importantly, to improve our understanding of the ways that this externalised user model can support learning and as well as the operation of the group.

**What and how to visualize.** Once relevant information is identified, the challenge remains how to communicate this back to the group (question 3). While the question of information visualisation has been researched before, including our own work [17, 18], research has so far been mainly limited to analysing *individual* displays of task and participation parameters [19]. The overall configuration of information displays – the interface elements that make up the shared work space – has been assumed as being static. We propose to dynamically adapt not only the content of individual information displays, but the *overall configuration* of information displays. For instance, when the group has to work on complex information together, social information should be reduced (in the absence of conflicts or member problems) so that all the cognitive resources can go into task information processing. Similarly, if interaction problems require attention, then the task information should temporarily be reduced and social information should be displayed with greater salience and detail. If both the task representation(s) and the social information representation(s) are properly adapted, then it should be feasible to provide suitable tradeoffs between the cognitive effort for the core task versus that for processing group and member information.

We also propose to differentiate more systematically between 'person awareness' and 'team awareness'. For instance, the video/audio display of a user – as a "rich" medium [6] –primarily provides information about an *individual* group member. It does not depict information about the team as such. The user lists that are part of most chat tools, however, are a rudimentary *team awareness* component – showing who is currently "in" the group activity. Visualisations can, and probably should, play a much stronger role in supporting team awareness. For instance, [20] makes a number of suggestions on how to visualize social configurations of team members in digital spaces such as chat rooms.

Our current prototype collaboration environment comprises various synchronous and asynchronous communication and information representation tools, including one that allows for co-located team work. We are experimenting with a number of computational approaches to aggregate collaboration information and identify psychologically and pedagogically meaningful patterns and trajectories. We are also developing means for visualising information relevant for task-, team-, and person-awareness. Building on these, we will experiment with ways to dynamically modify the respective information displays to make the overall interface adaptive to situational parameters (cognitive load, social conflicts, member problems) and to group members' preferences and individual needs.

## References


[1] McGrath, J.E., *Time, Interaction, and Performance (TIP). A Theory of Groups.* Small Group Research., 1991. **22**(2): p. 147-174.

[2] Jermann, P., A. Soller, and M. Muehlenbrock, *From mirroring to guiding: a review of the state of the art technology for supporting collaborative learning*, in *European Perspectives on computersupported learning*, P. Dillenbourg, A. Eurelings, and K. Hakkarainen, Editors. 2001, University of Maastricht: Maastricht, NL. p. 324-331.

[3] Conklin, E. J. (1993). Capturing organizational memory. In R. M. Baecker (Ed.), *Groupware and computer-supported cooperative work* (pp. 561-565). San Francisco: Morgan Kaufman.

[4] Kuminek, P. A., & Pilkington, R. M. (2001). Helping the tutor facilitate debate to improve literacy using CMC. In T. Okamoto, R. Hartley, Kinshuk, & J. P. Klus (Eds.), *Proceedings of the IEEE International Conference on Advanced Learning Technologies* (ICALT 2001), 6-8th August. (pp. 261-263). Madison, USA.: IEEE Computer Society.

[5] Buckingham Shum, S. (1997). Balancing formality with informality: User-centred requirements for knowledge management technologies, *AAAI Spring Symposium on Artificial Intelligence in Knowledge Management (Mar. 24-26, 1997), Stanford University, Palo Alto, CA. AAAI Press.*

[6] Daft, R.L. and R.H. Lengel, *Information richness: A new approach to managerial behavior and organizational design.* Research in Organisational Behavior, 1984. **6**: p. 191-233.

[7] Suthers, D.D., *Towards a systematic study of representational guidance for collaborative learning discourse.* Journal of Universal Computer Science, 2001. **7**(3): p. 1-23.

[8] Robert, L.P. and A.R. Dennis, *Paradox of richness: a cognitive model of media choice.* IEEE Transactions on Professional Communication, 2005. *48*(1): p. 10-21.

[9] Constantino-Gonzalez, M., D. Suthers, and J. Escamilla, *Coaching web-based collaborative learning based on problem solution differences and participation.* International Journal of Artificial Intelligence in Education, 2003. 13(2-4): p. 263-299. [10] Barros, B. and F. Verdejo, *Analysing student interaction processes in order to improve collaboration. The DEGREE approach.* International Journal of Artificial Intelligence in Education, 2000. 11.

[11] Bull, S., Brna, P., & H Pain, Extending the scope of the student model. *User Modeling User-Adapted Interaction,* 1995 *,* *5*(1), 44-65.

[12] Bull, S., Brna, P., & H Pain, Mr Collins: a collaboratively constructed, inspectable student model for intelligent computer assisted language learning. *Instructional Science,* 1995*, 23*, 65-87.

[13] Crawford, K., & J Kay, Metacognitive processes and learning with intelligent educational systems. In P. Slezak, T. Caelli & R. Clark (Eds.), *Perspectives on Cognitive Science,* 1993, pp. 63 – 77, Ablex

[14] Kay, J. The um toolkit for cooperative user modelling. *User Modeling and User-Adapted Interaction,* 1995*, 4*(3), 149 - 196.
[15] Bull, S., & E Broady, *Spontaneous peer tutoring from sharing student models* Proceedings of the International Conference on Artificial Intelligence in Education, 1997, 143 – 150.
[16] Kay, J., & R C Thomas, Studying long term system use. *Communications of the ACM,* 1995, *38*(7), 131 – 154.
[17] Uther, J., & J Kay, Visualising Large User Models. In P. Brusilovsky, A. Corbett & F. d. Rosis (Eds.), *Proceedings User Modeling,* Johnston (ed) Springer-Verlag in Lecture Notes in Artificial Intelligence, 2003, pp. 198—202.
[18] Zumbach, J. and P. Reimann, *Influence of feedback on distributed problem based learning*, in *Designing for Change in Networked Learning Environments*, B. Wasson, S. Ludvigsen, and U. Hoppe, Editors. 2003, Kluwer: Dordrecht. p. 219-228.
[19] Jerman, P. R. (2004). *Computer support for interaction regulation in collaborative problem-solving.* University of Geneva, Switzerland. [20] Erikson, T. and W.A. Kellog, *Social translucence: An approach to designing systems that mesh with social processes.* ACM transactions on computer-human interaction, 2000. *7*(1).