Bratitsis T. & Dimitracopoulou A. (2006). Monitoring and Analysing Group Interactions in asynchronous discussions with the DIAS system. In (Eds) Y. Dimitriadis, I. Zigurs & E. Gomez-Sanchez.*CRIWG 2006, 12th International Workshop on Groupware, GRIWG2006, Groupware: Design, Implementation and Use,* 17-21 Sept. 2006, Medina Del Campo, Spain Ed. Springer, LNCS 4154, pp. 54-61

Monitoring and Analyzing Group Interactions in Asynchronous Discussions with the DIAS system

Tharrenos Bratitsis and Angelique Dimitracopoulou

LTEE Laboratory, University of the Aegean, 1, Democratias Str,
85100 Rhodes, Greece
bratitsis@aegean.gr, adimitr@rhodes.aegean.gr

Abstract. DIAS is an Asynchronous Discussion Forum Software, mainly developed in order to offer extended monitoring and interaction analysis support, by providing a wide range of indicators jointly used in various situations, to all discussion forae users (individual user/students, groups, moderators/teachers or even researchers/observers), appropriate for their various roles in different activities. In this paper we describe some of the integrated Interaction Analysis (IA) features and provide information concerning case studies, some of which are in progress.

Keywords: Interaction Analysis, Asynchronous discussions, CSCL, CSCW

1 Introduction

Computer Mediated Communication (CMC) tools, allowing communication among users by means of networked computers, for the purpose of discussing topics of mutual interest [3],[7], are actually used in educational, working, or every day life contexts. Nowadays CMC tools, and in particular asynchronous discussion forae are widely used in formal or informal educational contexts, applying principles of constructivism, emphasizing in social interaction during learning activities [3]. The past quinquennium research is focusing towards finding methods for the evolvement and support of critical thinking through interactions, taking place within asynchronous discussions, in order to achieve high quality learning [13]. Such a goal requires tools, frameworks and methods for the facilitation of monitoring, and/or self-reflection and therefore selfregulation, that could be supported by the automated analysis of the complex interactions that occur. Computer based Interaction Analysis (IA) is an emerging field of research within the academic community, focusing in analyzing interactions among users, borrowing elements from the CSCW, CSCL and AI research fields.

2 Interaction Analysis

Computer based IA provides mainly information directly to technology based activities' participants, in order to self assess their activity [5]. The IA results are presented to the participants in an appropriate format (graphical, numerical, literal), interpretable by them. The corresponding information provide an insight of their own current or previous activity,

allowing them to reflect on a cognitive or metacognitive level, and thus act in order to self-regulate their activities. Computer based IA provides also related information to the activity observers that are interesting to analyse the complex cognitive and social phenomena that may occur.

This approach can produce flexible IA tools, which in an educational context, support directly the learning activities' participants (e.g. students, teachers, moderators) or even the observers (e.g. teachers, administrators, researchers) of these activities. The need for such tools derives from the complexity of interactions occurring within computer based learning environments (as described in many CSCL approaches). It would be legitimate to say that the IA research field has partially emerged from the application of methods, frameworks and techniques developed originally within the CSCW field and especially awareness (workspace awareness in particular) information provision, in combination with corresponding elements from the AIED (Artificial Intelligence In Education) field.

Regardless of the origin, the IA research field aims at providing methods and tools that support the participants of learning activities in three major levels: awareness, metacognitive and evaluation level [8]. The expected outcome is the optimization of the activity through:

- Better activity design, regulation, coordination and evaluation by the forum moderator
- Refined participation and learning outcome for the students through reflection, self-assessment and self-regulation

The IA process consists in recording, filtering and processing data regarding system usage and user activity variables, in order to produce the analysis indicators. These indicators may concern: a) the mode or the process or the 'quality' of the considered 'cognitive system' learning activity; b) the features or the quality of the interaction product; or c) the mode, the process or the quality of the collaboration, when acting in the frame of a social context forming via the technology based learning environment. [5]

Our main concern in this paper is IA tools concerning asynchronous discussions.

3 Relative Work

While examining Forum and Forum type software, we find that commercial products such as WebCT and BlackBoard, or open source products such as WebWiz and PhpBB provide minimum analysis information. Most of them present simple usage indicators, such as: Session Information (number of sessions, session length, mean session time), Activity Information (number of messages posted and read) and a few statistical indicators (most and least busy day, etc), online users, number of messages per day, number of unread messages, etc. We consider this minimal information, which supports forum usage only as a subsidiary tool of a learning system [2].

Several new and promising approaches that implement graphical representations of asynchronous discussions' features and parameters can be found while reviewing recent literature. For example, the i-Bee system is a visualization software that represents relationships between users and keywords in online messages, in real time. It also provides snapshots of past discussions and animations. Keywords appear as flowers and users as bees. The distance between flowers and bees, their status (e.g. flying/sleeping bee, blossomed/closed flower) and their orientation depend on discussion parameters, such as frequency of keyword usage and recent user activity [10].

Another example of the use of powerful visualisations via metaphors is the i-Tree system that visualises the discussion status on mobile phones using a tree representation. The tree

corresponds to a single user, whose activities designate the tree's appearance. Thereby the tree's log and branches are relevant to the number of messages, the leaves' range and colour are relevant to message reading, the fruits are relevant to the answers the user has received and the appearance of the sky is designated by the whole discussion status [11].

Mailgroup system is a Forum Type tool with integrated analysis tools emerging from the Social Network Analysis field. Additionally it uses an alternative way of representing the message sequence in an asynchronous discussion, taking into account both chronic and logical constituents [12].

Other approaches also exist, integrating Fuzzy Logic techniques in order to assess and evaluate the collaboration level in a discussion based on several parameters (Degree system) [1] or providing a variety of visualised statistical information (add-on for the AulaNet platform) in order to help the teacher coordinate discussions and obviate undesirable situations or progress of the discussion activity [6].

The aforementioned approaches constitute a representative specimen of asynchronous discussion software, used for learning purposes. All of them provide tools and functionalities for supporting and facilitating user activity in various levels. Nevertheless a closer examination of these systems leads us to the conclusion that they can only be used under specific usage settings. Some of their disadvantages are described in Table 1.

Table 1. Discussion Forum software characteristics

Software	Functionalities	Disadvantages
WebCT, phpBB, WebWiz	Simple statistical awareness information	No real IA indicators
i-Bee	Visualized representation of user – keyword relation	No empirical research about learning utilization of this feature
i-Tree	Visualized representation of user activity on mobile phones	Takes into account very few activity characteristics. Seems to encourage message reading, but not writing
MailGroup	SNA indicators	Indicators are addressed only to the moderator. The system need adequate number of messages to produce meaningful results
Degree	Various indicators and advising mechanisms, regarding collaboration quality	Closed system which is not easy to customize, with non-transparent indicator calculation.
AulaNet add-on	Visualized statistical information drawn from log files	Various diagrams, addressed only to the moderator

4. The DIAS system

The DIAS system (Discussion Interaction Analysis System) has been developed by the LTEE laboratory of the University of the Aegean. It is a fully functional discussion forum platform, with an underlying database management system for data recording and several implemented functionalities in order to facilitate user participation as well as the moderators' alternative discussion strategy planning. Additionally about sixty five (65) visualized indicators are produced (including all possible variations of the indicators), varying from simple statistical awareness information to complex cognitive and metacognitive indicators.

Different sets are addressed to the teacher or moderator and the students - users, along with the corresponding interpretation schema for various discussion strategies or usage scenarios.

Our main goal is to offer direct assistance to user, supporting them in the level of awareness of their actions, as well as those of their collaborators, in order to activate their metacognitive processes, thus allowing them to self-regulate their activities. In parallel, we aim in supporting the discussion moderators (eg teachers) in order to 'identify' problematic situations and difficulties that require regulative interventions. The design of the system is based on three central design principles [2]:

1. Take into account the totality of the users that are involved in a 'learning activity', as well as the cognitive systems they may form, students as individuals (in various roles), but also as members of one or more groups or even communities, teachers in different roles according the category of learning activity, etc.
2. Provide a rich range of IA indicators for the various user profiles and points of view of the activity process, its quality, as well as its product.
3. Create an independent, flexible, customizable and interoperable system. Forae are tools that can be used in a variety of contexts and activity categories. Furthermore forum participants take various roles and have different needs according to their discussion subjects, the available time etc. Thus, customization and flexibility are crucial characteristics.

This lead us to the selection of open source web based technology, making it easy to share with the academic community. More information about the system's architecture and functionality can be reviewed in [2].

5. DIAS Interaction Analysis indicators

By combining some of the indicators produced by the DIAS system and applying the appropriate interpretation schema, several interesting conclusions can be drawn. Let's examine a set of indicators addressed to the teacher, which may help him/her evaluate the quality of a certain student's participation (from now addressed as User X). These indicators are: Classification Indicator, SNA Answers, SNA Reads, User - Tree Structure and several statistical Bar Charts.

User Classification Indicator (Fig 1a): It is a XY scattered chart with the X-Axis representing the amount of contribution and the Y-Axis representing the amount of Interaction by the users. The two Axes are scaled from Low to High. The X-coordinate is calculated by the contributions (messages written) of the user as a percentage of the total contributions, thus placing the lowest number at the left end of the Axis (Low) and the highest number at the right end (High). The Y coordinate is calculated as the percentage of the available messages read by a user (excluding the ones written by him). By inspecting this indicator, the moderator may see how active User X is, in comparison with the other users and the mean values o activity. Activity has two constituents: writing and reading messages. So the first conclusion is whether User X has extreme or balanced behaviour (Arrogant: writes many messages but doesn't read other users' messages. Passive: reads many messages, but doesn't write enough). The second conclusion is whether User X's performance is far ahead from the mean values in any of the two constituents.

SNA Answers Indicator (Fig 1e): The system can produce social matrices according to Ucinet DL format and Agna matrix format for further processing. For N users, the Answers social matrix is a NxN matrix where every user is corresponds to one line and one column. The number placed in the cell designated by line A and column B shows the number of

messages written by user A as an answer to messages of user B. Using the matrix, we may construct the corresponding SNA diagram. By quickly inspecting such a diagram, the moderator can see whether User X is isolated or holds a central position within the discussion. Furthermore, if User X seems active from the previous indicator (writes messages), this diagram can show if he/she exchanges information with other users or not, by writing answers to them. Additionally the number of other users who have posted answers to User X can be detected, revealing interesting information. For example a student who appears to be very active in the Classification Indicator may be isolated in this diagram. Consequently he/she is energetic but doesn't seem to contribute to the quality of the discussion and the overall collaboration, since no one is posting answers to him/her. This could indicate low argumentative value of this user's messages, off topic writing, arrogant behaviour or lack of knowledge regarding the topic. In any of these cases, the moderator may diagnose a problematic situation and act accordingly.

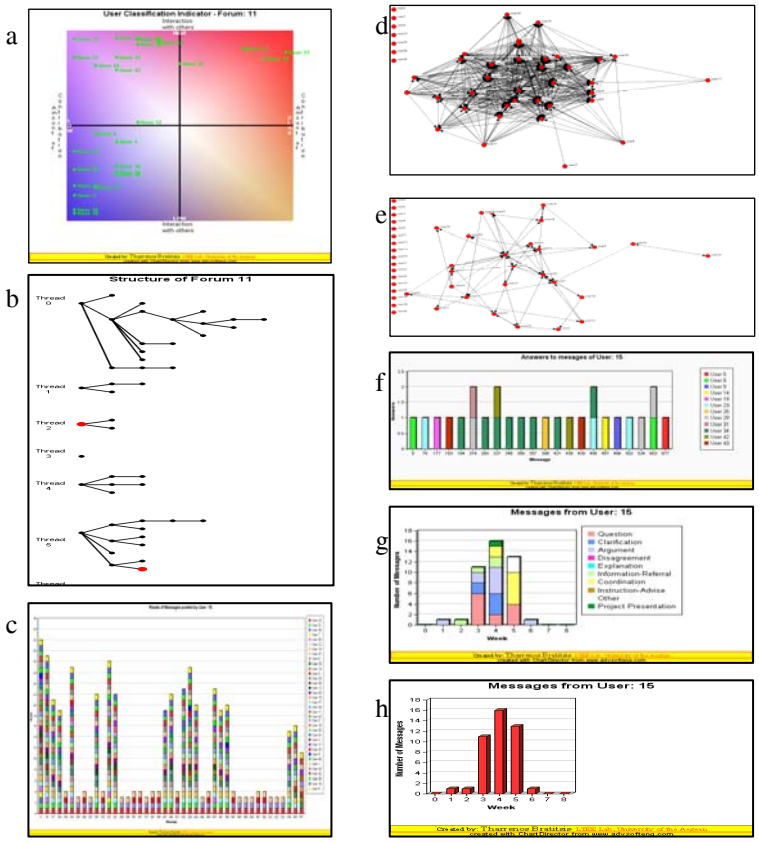

Fig. 1. Interaction Analysis Indicators by the DIAS system

SNA Reads Indicator (Fig 1d): This diagram is similar to the previous. In the social matrix, the numeric value in a cell designated the number of messages that user A has read and have been written by user B. This diagram indicates the amount of other students whose

messages User X reads and consequently the amount of his/her involvement in the collaborative discussion activity. While the Classification Indicator shows the amount of messages read, this diagram additionally shows the dissemination of these messages to the according amount of other students.

In combination with the Answers SNA diagrams, the moderator can see whether User X is participating in a closed user group interacting heavily inter se and very lightly with the rest of the users. This may designate undesired behaviour regarding the collaborative activity.

Furthermore, this diagram reveals the amount of users who have read messages posted by User X. If he/she holds a relatively good position within this diagram but appears to be isolated or obscure in the Answers SNA diagram, then he/she writes messages which are read by many other users, but not answered to. Consequently User X could be a coordinator of the discussion or is possibly facing a participation problem that needs further attention by the discussion moderator.

*User – Tree Structure Indicator (*Fig 1c): This is a visualised representation of the discussion evolution in a tree-like format. Messages appear as dots, where a line segments designates that the message on the right part is an answer to the one on the left part of the segment. Moreover, the messages posted by User X are marked with red dots. By quickly inspecting this diagram the moderator can see whether User X is active mostly in earlier or later phases of the discussion activity. In combination with the previous indicators, interesting conclusions may arise. For example an active user (Classification Indicator) who writes many but receives few answers (SNA Answers) and appears to write messages in later phases of the discussion may possibly have low performance in the activity. This could be the case of a user who simply agrees or disagrees with other users' arguments but doesn't contribute with new information and ideas, which can be revealed by inspecting his/her messages.

Bar Chart Indicators: Besides the aforementioned indicators, many simple awareness, statistical ones regarding User X can be produced in a Bar Chart format *(*Fig 1c, 1f, 1g, 1h show examples). These indicate for example the number of various types of messages (questions, answers, arguments etc) per day. By further examining such information, the teacher may acquire a more concrete reflection of the quality of User X's activity.

The aforementioned indicator set constitutes an example of indicator information utilization. Many combinations may be formed with various indicators and interpretative schemas. Furthermore the information that can be extracted from a single indicator can have different meaning for different kind of user roles or interpretative schemas (combinations with information extracted from other indicators.

6. Case Studies

Our main goal is to assess the indicator's usage, while using asynchronous discussions within learning activities. In particular our aim is to:

- Assess the correctness and clarity of the produced indicators and the proposed interpretative schemas
- Detect the effect of the information provided by the indicators in the users' self-regulation processes.
- Evaluate the contribution of the indicators to the facilitation of the moderator's, coordinator's and observer's work. We intend to provide easy ways coordination and assessment, bypassing the need of thoroughly reading all the messages or using time-consuming methods, such as content analysis.

- Assess the potentiality of a qualitative evaluation of discussion without applying content analysis methods.
- Designate the appropriate set of indicators for each role and phase of a discussion learning activity.

Several case studies have been designed for that matter, one of which is complete and three are still in progress.

Case Study 1: In the first case study forty (40) postgraduate students where involved in a non-restrained discussion activity for six (6) weeks. Their first contact with the system was made through a three hour seminar. The discussion topics were relevant to the course syllabus and the general topics of assignments they had to prepare for the end of the semester. A total of 553 messages were posted, while trying to exchange ideas, information and arguments. Studying the effect of the indicators in their activity behaviour was our intension. The results revealed that the indicators increased the motivation of the student to get involved with the activity (70% increase of messages). Students showed increased interest in observing the indicators, especially the ones providing comparative activity information with the rest of the students (for example the aforementioned Classification Indicator). They were very curious to examine the impression and reception of their messages by other students through indicators containing information about reading of messages and posting of answers. Some additional results could be extracted, such as the fact that the students' criterion for the acceptance of their messages by others was initially the number of answers they received and gradually altered to the number of users reading their messages. Significant part of this alteration was due to the information presented in some of the more complex indicators.

Case Study 2: The second case study is in progress. Thirteen postgraduate students are divided into two equivalent groups. The first group may examine the indicators, whereas the second does not have any indicators at their disposal. They participate in restrained discussion activities for seven (7) weeks, where the teacher has a more active role in coordinating the discussion evolution, according to certain usage scenarios. Our intension is to compare the behaviour of the two groups. Additionally we want to examine the facilitation provided by the system to the teacher, as he may examine the indicators only regarding the first group of students.

Case Study 3: The third case study is also in progress. The settings are the same as in case study 2, following an alternative usage scenario (discussion activity plan).

Case Study 4: The fourth case study is also in progress. Eighty (80) undergraduate students are divided into two equivalent groups. The settings are similar to the ones in the previous case studies. The main variation (apart from the group sizes) is the usage scenario (discussion plan) followed throughout the activity.

In all the case studies, semi-constructed interviews with the users will take place after the conclusion of the activities. The interview activity is already completed for case study 1.

7. Conclusions – Future Work

The main conclusion derived from the first testing of the DIAS system in real settings is that IA indicators influence the discussion activity evolvement, by acting as an additional motive for user's partipation. It can be considered as an additional tool in any distance learning platform and activity, providing means for increased interaction between the students. This effect of the visualized representation of interaction information seems to comply with the results presented by other researchers [2],[9],[10],[11].

Of course it relies upon the teacher to manage this tool to his/her benefit, by proper interpretation of the presented information, as well as by providing an appropriate set of indicators to forum participants so as to selfregulate their own activity. In the first case study attempts were made by some students to 'manipulate" the system and improve their position in the produced diagrams, without significantly contributing to the discussion activity. In one case a student wrote more than 1/3 of her messages the last two days of the activity, in order to appear as one of the most active users. Of course by combining indicator information, as described in the aforementioned interpretative schema, the moderator can designate such possible abnormalities easily, without thorough examination of the messages' content. Currently we have constructed several such schemas, but each moderator may design activities where he/she decides which sets of indicators are appropriate for the designated activity and the participating users.

Our future plans include the completion of the case studies in progress and evaluation of the results. Additional case studies are also under consideration for the near future, mostly addressing questions regarding the moderator's facilitation. Furthermore, we explore the needs of moderators, in asynchronous discussion forae other than for learning purposes (for example, in scientific networks, in open-audience discussions forae within corporative networks and other collaborative discussion activities within the CSCW spectrum). A complementary, overall goal is to try to associate activities and identifiable user action patterns, easily inspected through the visualized IA indicators.

References

1. Barros, B., Verdejo, F.: Analysing student interaction processes in order to improve collaboration. The DEGREE approach. Int Journal of Artificial Intelligence in Education, 11, p. 221-241. (2000).
2. Bratitsis, T., Dimitracopoulou, A.: Data Reording and Usage Interaction Analysis in Asynchronous Discussions: The D.I.A.S. System. Workshop on Usage Analysis in Learning Systems, The 12th International Conference on Artificial Intelligence in Education AIED, Amsterdam. (2005)
3. Collins, M., Berge, Z.: Resources for moderators and facilitators of online discussion. (2001). Available online at: http://ww.emoderators.com/moderators.html
4. Corich S., Kinshuk, Hunt L.: Assessing Discussion Forum Participation: In Search of Quality. International Journal of Instructional Technology and Distance Learning. TEIR Center, Duquesne University, Pittsburgh. (2004).
5. Dimitracopoulou, A et al.: State of the art of interaction analysis for Metacognitive Support & Diagnosis. IA JEIRP Deliverable D.31.1.1. Kaleidoscope network of excellence, December 2005.
6. Gerosa, M.A., Pimentel, G.P., Fuks, H., Lucena, C.: No need to read messages right now: helping mediators to steer educational forums using statistical and visual information. CSCL conference Taiwan 2005
7. Gunawardena, C., Lowe, C. & Anderson, T.: Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. Educational Computing Research, 17(4), 397-431. (1997).
8. Jerman P., Soller A. & Muhlenbrock M.: From Mirroring to Guiding: A Review of State of the Art Technology for Supporting Collaborative Learning. In: P. Dillenbourg, A. Eurelings, & K. Hakkarainen (eds): Proceedings of EuroCSCL, Maastricht, NL, pp.324-331. (2001).
9. Mazza, R., Milani, C.: Exploring Usage Analysis in Learning Systems: Gaining Insights from Visualizations. Workshop on Usage Analysis in Learning Systems, The 12th International Conference on Artificial Intelligence in Education AIED, Amsterdam. (2005).
10. Mochizuki, T., Kato, H., Hisamatsu, S., Yaegashi, K., Fujitani, S., Nagata, T., Nakahara, J., Nishimori, T. & Suzuki, M.: Promotion of Self-Assessment for Learners in Online Discussion Using the Visualization Software. CSCL conference Taiwan 2005

11. Nakahara, J., Kazaru, Y., Shinichi, H., Yamauchi, Y.: iTree: Does the mobile phone encourage learners to be more involved in collaborative learning? CSCL conference Taiwan 2005
12. Reyes, P., & Tchounikine, P.: Mining learning groups' activities in Forum-type tools. CSCL conference Taiwan 2005
13. Stahl, G.: Group Cognition: Computer Support for Building Collaborative Knowledge. Acting with Technology Series, MIT Press. (2006).