

Spilberg & Pilkington, 2009

Spilberg, R. & Pilkington, R. (2009).

Interaction on a timeline: visualising spurts and lulls

Submitted for the CSCL 2009 Pre-Conference Workshop on: Interaction Analysis and Visualization for Asynchronous Communication: Analysis Methods, Tools, and Research Questions.

Rose Spilberg
Faculty of Media, Humanities & Technology,
University of Lincoln, Brayford Pool, Lincoln,
LN6 7TS, UK.

Rachel M. Pilkington
School of Education, The University of
Birmingham, B15 2TT, Birmingham, UK.

Interaction on a timeline: visualising spurts and lulls

Rose Spilberg and Rachel Pilkington

Abstract

Asynchronous Computer Mediated Communication (CMC) technologies have a recognised pedagogic value in both conventional and distance learning courses. These tools can facilitate learning and foster collaboration, critical thinking, reflection and the development of communities of practice. However, CMC has also often failed to deliver these desired educational benefits. In part this is believed to be because dialogue through these tools can be difficult to support, monitor and analyse. There is considerable interest in finding ways to use graphical tools to help identify and represent patterns of interaction thought to be pedagogically effective, and to distinguish these patterns from those that might be considered less effective. There are two key potential uses of such visual representations: the first is as a research and evaluation tool to compare and contrast collaborative and discursive activity in order to identify effective interaction and its correlates; the second use is as an aid to the participant (tutor or discussant) when monitoring and reflecting on the quality of discussion in progress, and with a view to helping students and tutor steer dialogue toward more effective collaborative outcomes. This paper focuses on one particular kind of graphical representation – a time-based visualisation model – showing how it is applied to asynchronous CMC interactions enabling interaction patterns to be more easily perceived, compared and analysed. The paper discusses the analysis which the model affords and describes the data requirements and method of constructing the model. Future potential development of software tools for automated construction of the model to support dynamic online facilitation or reflective processes are briefly explored.

Introduction

In Higher Education it is generally held that collaborative learning and discussion can help adult learners learn to think more critically and to engage more deeply with subject content through interaction in communities of learning and of practice (Roschelle, 1992; Koschmann, 2002; Stahl, Koschmann & Suthers, 2005). However, there is considerable time pressure upon opportunities for students to develop collaborative discussion skills. Moreover, for some students, even when group-size is small, the prospect of engaging in critical debate with tutor and fellow students can be daunting (Pilkington & Kuminek, 2004). Helping students gain confidence and motivation to develop critical thinking through debate can be difficult. Asynchronous Computer Mediated Communication (CMC) technologies have a recognised pedagogic value in both conventional and distance learning courses as they can help to foster collaboration, critical thinking, reflection and the development of communities of practice (Warschauer, 1999; Garrison & Anderson, 2003; Guldborg & Pilkington, 2006). However, learning to participate in or facilitate dialogue through CMC also presents problems. Previous work suggests that although students may be adept users of communication technologies (e.g. using online public chat rooms to manage social networks), setting up a chat room or discussion board to support a course does not guarantee students will use these tools to meet the course aims (Walker & Pilkington, 2000; Jones et al. 2001). Tutors have to know when to take a strong, supportive and facilitating role when groups are struggling, yet disappear to avoid inhibiting discussion when not (Kuminek & Pilkington, 2001; Guldborg & Pilkington, 2007). Creating a

healthy discourse community depends on tutors *orchestrating* healthy interaction including establishing, reinforcing and maintaining appropriate ground-rules for discussion to enable these processes whilst not *over-scripting* learners but engaging learners in monitoring and reflecting on their own contributions and roles to make collaboration effective (Crook, 1997; Dillenbourg, 2002).

Visualisation and Interaction Analysis

It has been argued above that one reason why CMC has often failed to deliver the desired educational benefits which we know are possible through this medium is that discussion using CMC can be difficult to support, monitor and analyse. There is considerable interest in finding ways to use graphical tools to identify and effectively represent those patterns of interaction thought to be pedagogically effective and to distinguish these from those considered less robust or educationally effective (Erickson, 2003). There are two key uses of such representations:

1. The first use of these visualisations is for research and evaluation, to compare and contrast collaborative and discursive interaction patterns as an aid to analysis, particularly in triangulation with other qualitative and quantitative data (such as content analysis or task performance outcomes). In this way it is hoped to gain greater insight into those patterns of interaction which seem to signal or accompany more effective collaboration and might even be antecedent to them.

2. The second potential use of such visualisations is as an aid to the participant (whether the discussant is tutor or student). Here the aim is to monitor and reflect on the quality of the ongoing discussion whilst it is in progress and in ways that can help participants steer the dialogue toward more effective solutions or more productive conclusions.

To some extent it is recognised that the latter (use 2 above) is likely to become more effective as knowledge from the former use (use 1 above) becomes more robust. Thus, one possibility is that a graphic showing an interaction pattern known typically to be a less effective pattern could provide a prompt to the tutor to look more closely at the textual content to see if they need to make a timely intervention. Equally, such a picture could point discussants to where more attention to the discussion content or their role is needed. It is also possible that we may learn something of the potential of such graphical tools in helping to monitor and support tasks from simply making such tools available to learners and tutors and observing how they make use of them.

A number of studies have suggested alternative graphical views onto data to help participants or tutors identify patterns of interaction (Kuminek & Pilkington, 2001; Erickson, 2003; Reimann et al. 2005). Erickson (2003) argues for social visualisation models that are built up out of many smaller components and which can be viewed at macro and micro resolutions, enabling the user to zoom in on characteristics of the interaction through a range of alternative visualisations. For example, the social proxy tool in Babble (Erickson et al. 2002) indicates how active a participant currently is by showing them as a colour dot inside a circle representing, relatively, how more or less at the centre of the discussion they are. The Loops timeline proxy represents the last month of online activity on a timeline on which each participant has a row (Erickson et al., 2006).

These tools differ in their aims to show aggregated information concerning the group, how members interact with each other within them and individual performance data relative to other members. Similar but alternative tools are suggested in (Reiman & Zumbach, 2003; Reimann et al., 2005; Kay et al. 2006). The former study used tracking

data from asynchronous conferences to generate visual aids to provide an automatic feedback mechanism for participants. The study went on to evaluate how learners might benefit from this feedback in terms of successful outcomes to problem-solving. The Kay et al. report includes a timeline described as a wattle tree which differs in some respects to that described in the implementation of Loops by Erickson et al. 2006. The wattle tree represents a number of different activities of group members as leaves on a branch, where each branch represents an individual. The picture thus shows group members' activity on a number of different tasks over a period of time, but is not explicitly concerned with communication within the group. The Loops timeline functions as a social proxy for members within the Loops group; in other words, it exists to enhance Loop users' experience of the online environment as a space shared with others, enhancing the sense of 'social presence' (Connell et al, 2001) over a period of time.

It is this timeline aspect of the visualisation of data that the study described here is also first hoping to explore. It is noted that the goals of this study differ somewhat from those of collaborative group problem-solving or the Computer Supported Collaborative Learning/Working (CSCL/W) environments described in the studies above. The task in the case discussed here is one of discussion or debate, in a learning context, as an educational activity in its own right. This is expected to affect the kinds of visualisation tools likely to give insight into effectiveness and/or be most useful to participants in reflecting on performance. The visualisation discussed here is very simple, and clear, and we hope, therefore, accessible to participants. The objectives are first and foremost pedagogical i.e. concerned with the most effective design of learning activities to meet learning objectives. In particular, we hope the representations, together with the discussion transcripts from which they are created, will be able to help address the following research questions:

1. Whether the frequency of interaction (number of posts) over time (i.e. the intensity of interaction) in asynchronous discussion is a reliable indicator of more or less effective interaction, where effectiveness is concerned not with interaction *per se*, but with the achievement of the educational aims of the discussion activity defined in qualitative terms.
2. Whether relatively short periods of intense and reciprocal exchanges - *spurts* (as opposed to the opposite - *lulls*) are a better indicator of highly effective periods of discussion.
3. Whether and how spurts relate to the typical length of post and whether the length of post (message word-count), appears to influence the effectiveness of the discussion (as defined at 1 above).
4. How to optimise the instructional design, in online discussion tasks, to foster the interaction patterns which appear to be most fruitful, either by structuring the task (for example the overall discussion period), or by varying the instructions given to participants (for example limiting the length of post).

This paper focuses on one particular kind of graphical representation – a time-based visualisation model – showing how it is applied to asynchronous CMC interactions enabling interaction patterns to be more easily perceived, compared and analysed. The paper discusses the analysis which the visualisation model affords, and describes the data requirements and method of construction of the visualisation. Future potential for development of a software tool for automated construction of the visualisation model and its potential to support more dynamic online facilitation are then briefly explored.

The Study

This position paper reports work in progress based on actual data from a case study undertaken by Rose Spilberg as tutor-researcher. The use of graphical representations of interaction reported here is therefore, primarily as a research and evaluation tool which can be triangulated with other analyses to provide insight into what might be associated with effective periods of interaction or the factors that might trigger them. However, it is hoped that in the future insights gained from these analyses might suggest ways the tools could be used more dynamically to help students and tutor facilitate or steer discussion more effectively.

The data come from three cohorts of adult students. Students attended an English university, studying a third-year unit concerned with technology and society and the social aspects of computing. Students were required to participate in on-line discussions as part of a collaborative group assessment. The students were divided into small groups, (4-6 people) and each group undertook two cycles of online discussion, each taking place over a period of three to five weeks.

Over the three cohorts of students, over three academic years (2005-06, 2006-07, 2007-08), altogether 13 collaborative discussion groups, have undertaken a total of 26 separate on-line discussions. The full transcripts of these discussions form the raw data for the current on-going research, which seeks to evaluate applications of computer-mediated communication (CMC) used for educational purposes, to inform instructional design and pedagogical practice. This paper reports only on the initial quantitative analysis of the data drawn from one cohort: a total of four groups, yielding eight discussion transcripts, hereinafter referred to as discussion instances. Each individual contribution to a discussion is referred to as a post.

Quantitative analysis

Basic quantitative data from each discussion were manually entered into a spreadsheet application (Microsoft® Office Excel 2003) to facilitate quantitative analysis and comparison. A template was created for the entry of data from each discussion instance, enabling the following data to be entered for each post:

Participant	Date	Time	Word-count	Thread
-------------	------	------	------------	--------

The data were obtained by downloading the discussion transcripts, and importing the files into a word-processing application. The date and time data is that provided by the on-line discussion system. A protocol was developed to systematise word-count, using the word-count tool in the word processor, in order to ensure consistency and minimise inaccuracy.

This basic spreadsheet data yields simple aggregate and mean information (total word-counts, mean number of posts per group or per person), and affords some interesting insights and comparisons between groups and between discussion instances (see Figure 1, below).

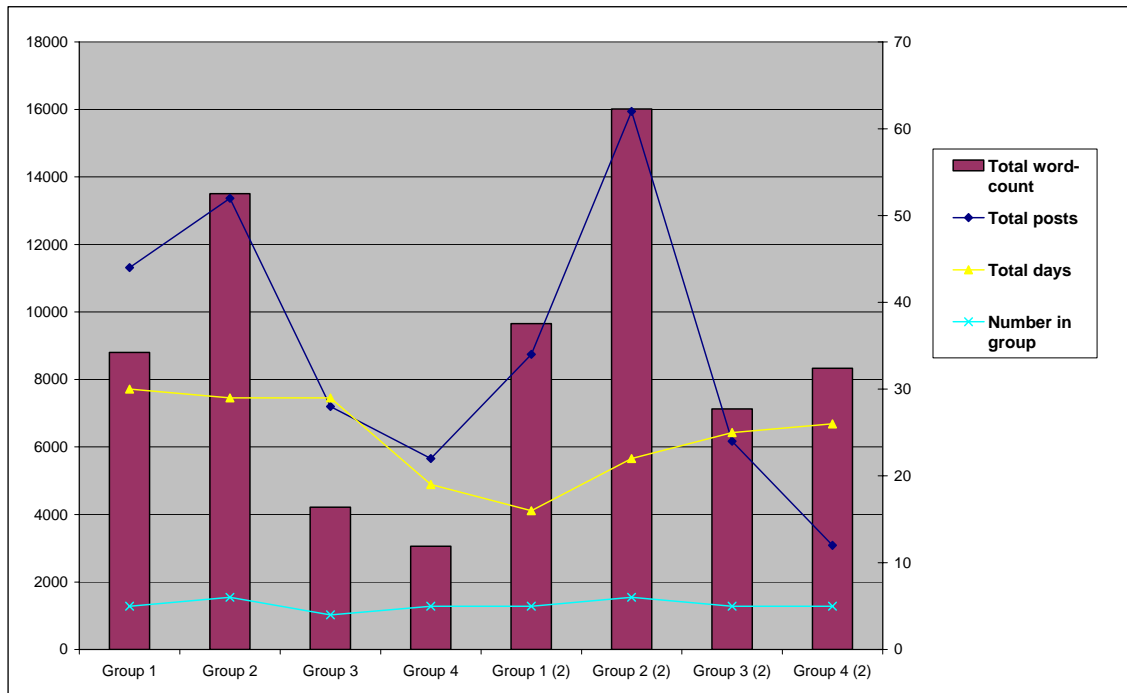


Figure 1. Aggregate data: four groups, each having two discussions

It is interesting to note that, for each of the four groups in Figure 1, the second discussion instance yielded higher contributions than the first, in terms of word-count and, for most groups also in terms of posts made, and that this overall 'size' of discussion does not appear to be related to either the number of participants in the group, or to the amount of time (in terms of days) given to the task. As can be clearly seen, the groups vary considerably in the total size of the discussion, as measured in word count.

Analysis is on-going to yield information for detailed comparison of quantitative data between individual participants. The quantitative analysis is being followed by, and is informing, qualitative analysis of the discussions to analyse and evaluate the evidence for the desired learning outcomes for the activity. Here, the quality of discussion will be defined by: evidence of critical thinking, analysis, argumentation, use of evidence and structuring an argument.

The total number of discussion instances available, at 26, provides a large volume of data to be analysed. One of the functions of the quantitative analysis, in addition to yielding the quantitative data, which is intrinsically of interest, is to afford a basis for selection of those discussion instances which may be most fruitful for further qualitative study. It was felt that additional productive and interesting information could be revealed by the quantitative data, if a way could be found to provide a richer picture of aspects of each discussion instance, particularly of the patterns of participant interaction in the discussion. Conversely, to avoid confirmatory bias, textual content analysis using qualitative indicators of effective debate will also be undertaken on a sample of discussion and used to suggest periods of discussion that are most productive. Having identified these productive periods they will also be visually modelled to check if they are associated with predicted indicators (e.g. spurts) or are otherwise distinctive. Taken together, these analyses might yield a greater

understanding of the ways students engage and interact in CMC and also inform how such activities may be made most engaging and effective.

The timeline visualisation

A number of ways were explored to attempt to provide a graphical representation of the above discussion interactions. The process of interaction over time is believed to be critical to the robustness and success of asynchronous CMC, so a time-based representation was sought. It quickly became clear that a graphic which showed the interaction, post by post, on a timeline, would meet the requirement. Such a graph, while simple in principle, was complex to produce through commonly available software charting facilities, but after considerable manipulation, an appropriate chart template has been designed.

The template is based on the available XY-scatter graph template in Excel, customised to provide the desired representation. This visualisation uses a time-line axis, using the date and time data recorded by the discussion system to plot each post. The time-stamp data, recorded as text in the discussion files, was converted and formatted to the required date/time numeric format in Excel, and then plotted onto the time-line.

In Figure 2, below, the X-axis is the timeline, with each division showing one day; the Y-axis values represent the word-count of each post, yielding an immediate and rich visualisation of the progress of a discussion in terms of the frequency of participants' engagement with the CMC system and whether it is both frequent and reciprocal over time (a spurt) during the discussion period.

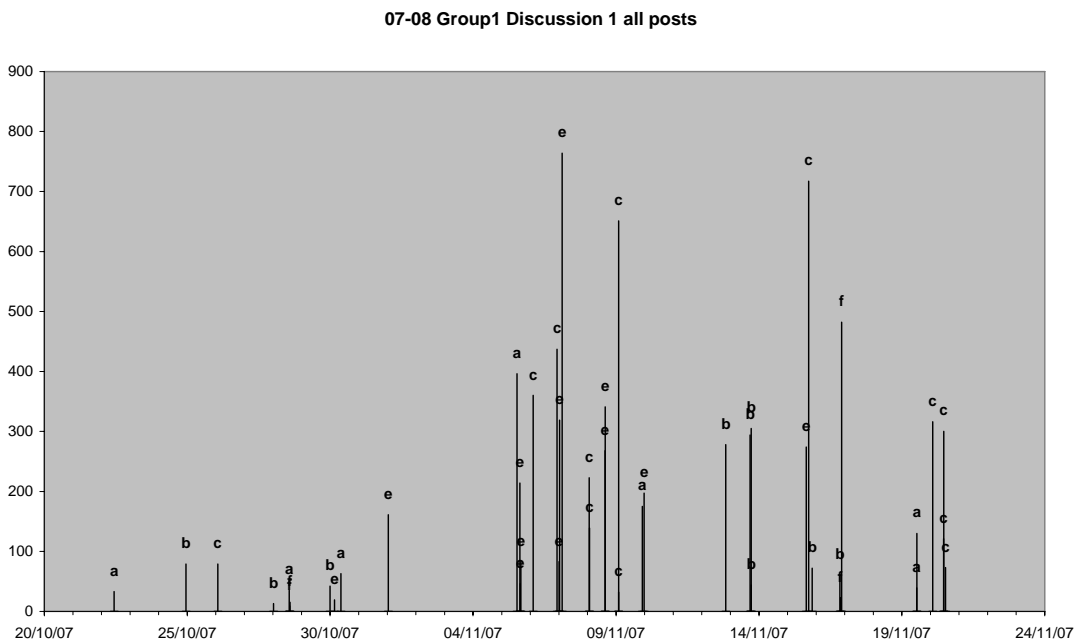


Figure 2. Interaction on a timeline: Group 1's first discussion

In addition, inferences can be made about the kind of interaction which is taking place, based on the size of the Y-value, the word-count. For example, an initial check of the content of contributions seems to show that very short posts are likely to involve social interaction, or management and organisation of the communication, rather than engagement on the task proper. The letters labelling each post serve to indicate different participants, and also to differentiate posts where they are clustered very

closely on the timeline. A key strength of this approach is the ease with which the graphical representation can be linked back to the raw data - actual transcripts - enabling textual analysis of the interaction at particular points on the time-line, so that inferences can be queried and verified.

The picture of interaction, as represented by spurts and lulls on the time-line, serves to alert and suggest elements of the discussion which may be particularly fruitful for qualitative analysis. For the purposes of this analysis *spurts* may be defined as posts from two or more participants at short intervals over a few hours or days, constituting a fairly intense interaction, while *lulls* are periods, of days, of intermittent or desultory posting by different individuals, with little apparent reciprocity with each other. *Nulls* are periods of time over 24-48 hours or more, in which no activity on the discussion occurs. In the picture above, for example, the five days from the 5th to the 9th of November constitute a spurt, in that they suggest an intense interchange of postings, of which most are contributions of several hundred words. Study of the actual transcripts for those dates reveals a period of active on-task contribution by most or all group participants, during which a substantial proportion of the group task is undertaken. Two further spurts (12-13 and 15-16 November) contain more content-driven discussion and summarising, preparatory to the completion of the final task – a group summary and position paper, which is circulated for final comment 19-20 November. The early days of this discussion are characterised by occasional and quite brief posts, before starting the group task itself.

Interaction patterns of different groups, or of the same group at different points in time, are clearly presented with this visualisation. Figure 3, below, shows the interaction pattern of a different group over the same discussion time-period. There are similarities in the interaction, which again is characterised by spurts of activity, and is broadly interactive. Here some very long posts lie beyond the y-axis range of the graph, but the axis scales have been retained to better afford the visual comparison of the overall picture.

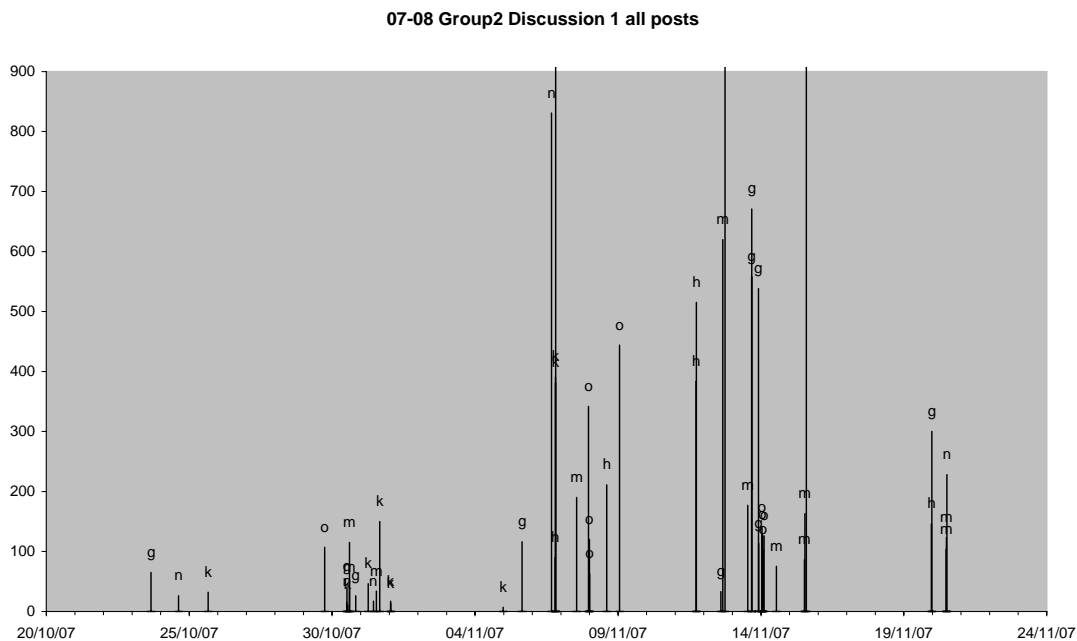


Figure 3. Interaction on a timeline: Group 2's first discussion

In Figure 4, below, a picture of a much sparser interaction can be seen. This group discussion was active over a shorter period of time altogether – and shows one apparent spurt. However, careful study reveals that all but one of the posts in the cluster around 28-29 October had the same author (**u**), who in fact contributed many more posts than any other participant in the discussion. The discussion transcript reveals that, while participants were engaged on the task, there was much less engagement and debate with each other than in Group 1, for example. The pattern presented here, of occasional posts along the timeline, is neither spurt nor lull. Where the pattern is of repeated posts by the same individual, the interaction type is termed *solos*.

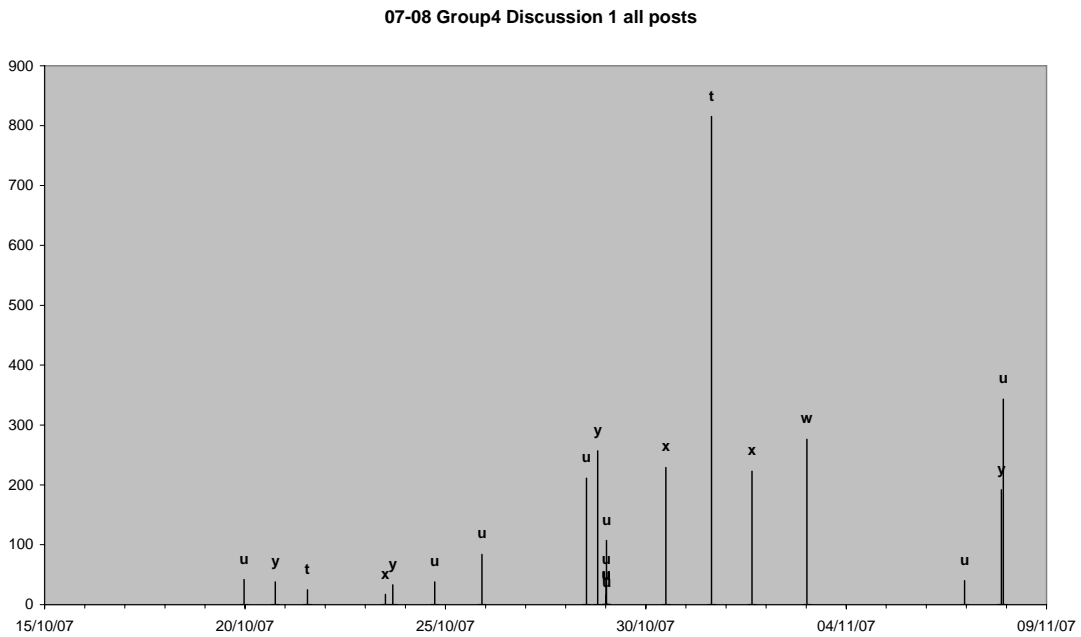


Figure 4. Interaction on a timeline: Group 4's first discussion

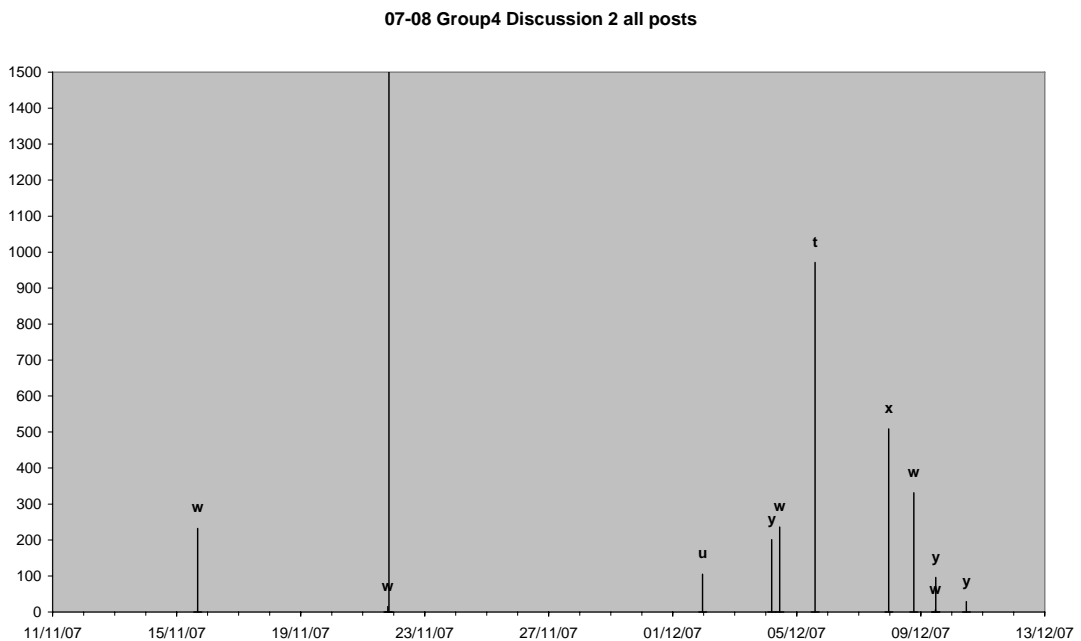


Figure 5. Interaction on a timeline: Group 4's second discussion

Figure 5 shows the second discussion of Group 4, whose first discussion is given above. Again this group's interaction is very sparse, but there is one very long post – actually over 5,000 words.

Flexibility of the model

This chart template can be used to provide a visualisation of any discussion interaction where the date/time information is available. It can show the interaction with the system of any individual or subset of the whole discussion group, and the time-line scale can be expanded to show more detail, for example over a short but intense period of a few days.

Additional richness

Attempts are being made to add richness to the interaction picture, by adding further timing information, in the form of background time-based light and dark shading, to give a sense of day/night posting. Adding days of the week to the X-axis is also under consideration. Both of these additions would enhance the visual representation of the discussion interaction over time day-by-day and week-by-week.

Automating the tool

As currently shown, the model requires manual preparation of the chart data series, and manual creation of the graph using the pre-set template. The automation of both of these elements is an objective. Transcripts from text-based CMC systems, if available as text files, should be capable of being imported into an application such as Excel, although some CMC systems deal with threading by storing text in tables which may be nested, creating complications for data transfer. The manipulation of the text, once exported, to provide time-date data to create the graphical output is straightforward, but a flexible interface enabling the creation of customised charts as required would be more demanding.

The first technology to explore in this regard might be Visual Basic. Excel functions can be programmed in VBA, and the development of an interface which would allow entry of the required parameters for the chart creation, and then automatically produce the chart formatted as requested, might be possible. Interfacing with external and open source systems, using, for instance, XML tools, is an alternative approach being considered for the future.

Conclusion

The work reported here is at an early stage of development but aims to use commonly available off-the-shelf tools such as Excel, interfacing with macros, templates and functions, for example using VBA, to provide insightful representations of discussion activity. Management and processing of the transcript data into a form that can be imported into and manipulated within a spreadsheet application, and automating the graphical chart production within the spreadsheet, is the general problem, and a number of possible solutions remain to be explored.

It is hoped that in the future the graphical application should help identify hotspots of lively activity, or patterns of interaction that are potentially a cause for concern, alerting the tutor to monitor and if needed intervene to facilitate and support group work. Initial work suggests the following may be such indicators:

- 1 Spurt: indicator of effective and fruitful interaction;
- 2 Lull, 3 Solo, and 4 Null: all indicators of possible problems in the interaction.

However, having identified 1, 2 or 3 the tutor needs to check the quality. All spurts so far studied have evidenced productive CMC activity, although their quality varies. Some lulls and solos appear to be part of the process of getting started or building up to effective discussion, while others may be indicative of failure to do so. It is hoped that by triangulating these indicators with qualitative analysis of argument content, structure and form we can further determine the utility of these indicators for intervention and in particular which factors might play a more direct or causal role in creating spurts and lulls and how these relate to local discussion outcomes. Undertaking this analysis with a further two cohorts of students, covering all 26 transcripts, will provide a foundation upon which to validate such claims. We may then be able to utilise these insights in optimising the instructional design of the on-line activity towards more fruitful interaction.

References

- Connell, J.B., Mendelsohn, G. A. Robins, R. W. & Canny, J. (2001) "Effects of Communication Medium on Interpersonal Perceptions: Don't Hang Up on the Telephone Yet!" *Proceedings of ACM Group Conference 2001*, (pp. 117-124), Boulder, CO. ACM Press.
- Crook, C. (1997) Children as computer users: the case of collaborative learning. *Computers & Education*, 30, 237-247.
- 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?* (pp. 61-91). Heerlen: Open University of the Netherlands.
- Erickson, T. (2003) Designing visualisations of social activity: six claims. *Proceedings of CHI 2003*, (pp.846-847), Florida, ACM Press.
- Erickson, T., Halverson, C. A., Kellogg, W. A., Laff, M. & Wolf, T. V. (2002) Social translucence: designing social infrastructures that make collective activity visible. *Communications of the ACM*, 45, 41-44.
- Erickson, T., Kellogg, W. A., Laff, M., Sussman, J., Wolf, T. V., Halverson, C. A. & Edwards, D. (2006) A persistent chat space for work groups: the design, evaluation and deployment of loops. *Proceedings of the 6th ACM Conference on Designing Interactive Systems*, (pp. 331-340), University Park, PA, ACM Press .
- Garrison, D. R. & Anderson, T. (2003) *E-Learning in the 21st Century*, London, RoutledgeFalmer.
- Guldberg, K. R. & Pilkington, R. M. (2006) A community of practice approach to the development of non-traditional learners through networked learning. *Journal of computer Assisted Learning*, 22, 159-172.
- Guldberg, K. & Pilkington, R. (2007) Tutor roles in facilitating reflection on practice through online discussion. *Educational Technology and Society*, 10, 61-72.
- Jones, C., Asensio, M., Goodyear, G., Hodgson, V. & Steeples, C. (2001) *Networked Learning in Higher Education Project: Final Report on the Field Studies*. JISC/CALT, Lancaster, Lancaster University.
- Kay, J., Maisonneuve, N., Yacef, K. & Reimann, P. (2006) *The Big Five and Visualisation of Team Work Activity*, Technical Report 581, University of Sydney.
- Kuminek, P. A. & Pilkington, R. M. (2001) Helping the tutor facilitate debate to improve literacy using CMC. IN Okamoto, T., Hartley, R., Kinshuk & Klus, J. P. (Eds.) *Proceedings of the IEEE International Conference on Advanced Learning Technologies, ICALT 2001*, (pp. 261-263). Madison, IEEE Computer Society.
- Koschmann, T. D. (2002) Dewey's contribution to the foundations of CSCL. IN Stahl, G. (Ed.) *Proceedings of CSCL, 2002*, (pp.17-21). Boulder, Colorado, LEA.
- Pilkington, R. M. & Kuminek, P. A. (2004) Using a role-play activity with synchronous CMC to encourage critical reflection on peer debate. IN Monteith, M. (Ed.) *ICT for Curriculum Enhancement*, (pp. 69-85). Bristol, Intellect.

- Reimann, P. & Zumbach, J. (2003) Supporting virtual learning teams with dynamic feedback. *Paper presented at the International Conference on Computers in Education, ICCE 2003, Hong Kong.*
- Reimann, P., Kay, J., Yacef, K. & Goodyear, P. (2005) Adaptive visualisation support for self-managed learning groups. *Paper presented at the AIED 2005 Workshop on representing and analysing collaborative interactions, Amsterdam.*
- Roschelle, J. (1992) Learning by collaborating: convergent conceptual change. *Journal of the Learning Sciences, 2, 235-276.*
- Stahl, G., Koschmann, T. D. & Suthers, D. D. (2005) Computer Supported Collaborative Learning: an historical perspective. IN Sawyer, R. K. (Ed.) *The Cambridge Handbook of the Learning Sciences*. New York, Cambridge University Press.
- Walker, S. A. & Pilkington, R. M. (2000) Networked Communication and the Collaborative Development of Written Expression at Key Stage Three. IN Asensio, M., Foster, J., Hodgson, V. & McConnell, M. (Eds.) *Proceedings of the Second International Conference on Networked Learning 2000*, (pp. 354-361). Lancaster, Lancaster University and Sheffield University.
- Warschauer, M. (1999) *Electronic Literacies*, Mahwah, New Jersey, Erlbaum.

Acknowledgements

The authors would like to thank Ian Davison, School of Education, Birmingham, U.K. for his earlier assistance with aspects of using Excel.