Processes of decision-making with adaptive combinations of wiki and chat tools

Kate Thompson and Nick Kelly, CoCo Research Centre, The University of Sydney, Australia
Email: kate.thompson@sydney.edu.au, nick.kelly.mail@gmail.com

Abstract: Debate in the field of event-based analysis of collaborative process data has focused on how best to account for the temporal nature of processes. We reanalyzed data from an online learning environment that included multiple tool use (chat and wiki) in synchronous and asynchronous collaborative settings. In this paper we modeled decision-making, and used time to identify four patterns of tool use to make decisions: single chat, multiple chat, chat/s and follow-up wiki, and integrated. Each pattern of tool use was associated with a unique process model, which indicated the affordances of each tool for that particular activity. The examination of the integrated use of tools, and the integrated decision-making process in a collaborative online setting has implications for designers, instructors and researchers.

Introduction

The field of event-based analysis of processes of collaboration has opened up since Reimann's seminal paper (2009). New methods of analysis (Kapur, 2011; Reimann, 2009; Wise & Chiu, 2011a, 2011b) have been suggested, and critiqued (Goggins, Laffey, & Amelung, 2011). In this paper, we suggest that prior to making a decision about the method of analysis, the role that time plays in the composition of the dataset is important to consider. We build on Reimann, Frerejean & Thompson's (2009b) research by taking their dataset (chat data for a month-long collaborative exercise) and incorporating data from the wiki that was used by the groups. Rather than assuming that linear, or even cyclic processes, that apply to the entire dataset, we identified a number of threads, related to combinations of tool use, and examined the processes involved in decision making for each of these. Unlike the original study, we use first-order Markov models to derive the process models. This study has implications for the design and management of collaborative learning environments, as well as making a contribution to the development of methods in this emerging field of research.

Background

With the ability to collect more complex datasets (for example, log file data, time sequenced video and conversations) has come methodological advances. Research into the processes of learning have included perspectives such as argumentation (Weinberger & Fischer, 2006), knowledge building (Aalst, 2009), decision making (Reimann, 2009) or knowledge construction (Wise & Chiu, 2011a). Research has addressed both synchronous (Ding, 2009) and asynchronous (Weinberger & Fischer, 2006) collaboration, but few address both (Thompson & Kelly, accepted). The units of analysis differ across many of the studies, depending on the research question. As yet, there is not a clear framework in which to place further research. While both synchronous and asynchronous online learning spaces have been researched, few have studied the combination of tools, and even fewer in association with the processes of collaboration. Zenios and Holmes (2010) focused on the social affordances of tools such as chats and wikis. They described a process of collaboration in which students used the online chat tool to reflect on, discuss and modify what they had written in the wiki. They discussed the importance of dialogue before and after the development of collaborative wikis and suggested that chat was needed in order to produce the ideas, whereas the wiki acted as the organisational memory for the group. They concluded that the combination of the chat and the wiki were necessary for epistemic tasks to take place, and group cognition and new knowledge to be developed.

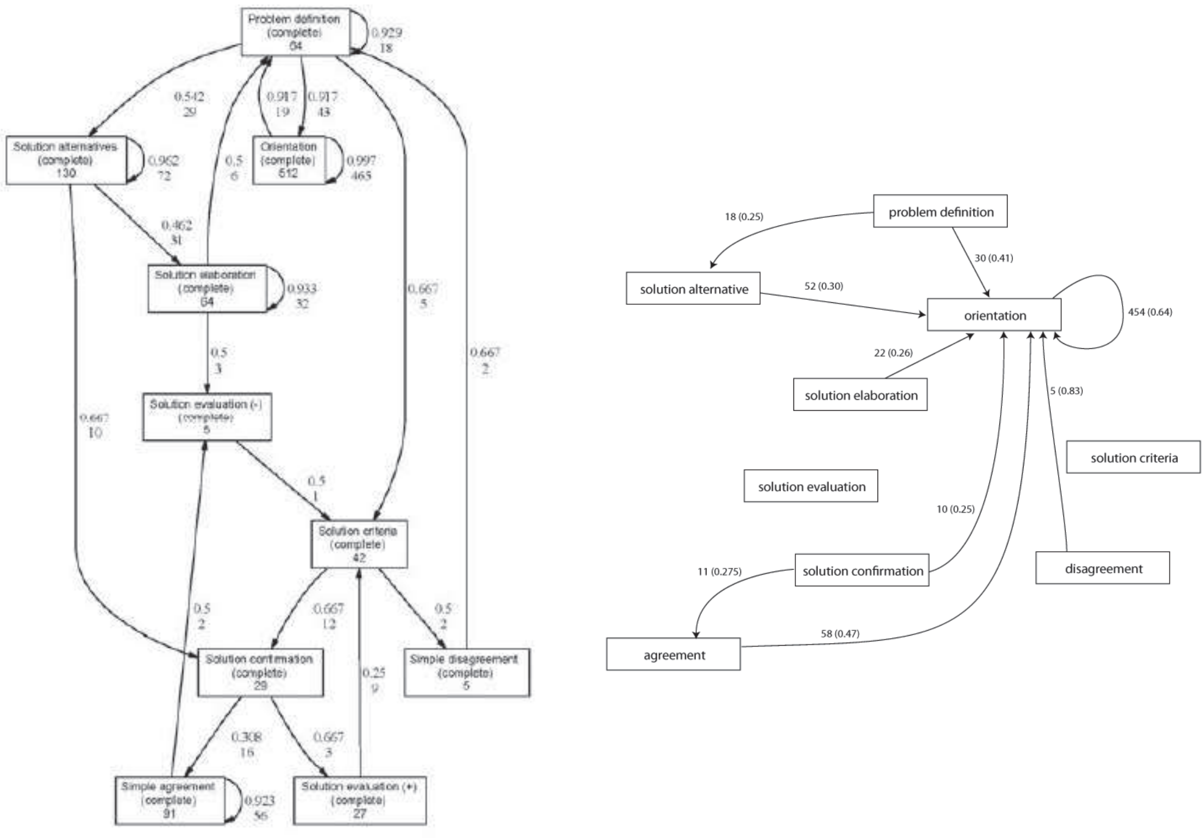
Making use of the large amounts of data produced during online activities, such as that produced with chat and wiki, is something that has received significant attention in recent years (Reimann, Frerejean, & Thompson, 2009a; van der Aalst & Weijters, 2005). Given a set of states (for example, learning phases, decision stages, steps interacting with a computer model) that changes over time for a group of students, what can be mined from this data? Techniques that have been applied successfully for identifying patterns within state transitions include heuristics mining for dependency relationships (Reimann, et al., 2009a; Weijters, Van der Aalst, & Medeiros, 2006), first-order Markov models for state transitions where symbols in the dataset and states of the target are isomorphic (Eddy, 1998; M. S. Poole & Roth, 1989a), and hidden Markov models for determining hidden states from a set of symbols (Jeong, Biswas, Johnson, & Howard, 2010; Shih, Koedinger, & Scheines, 2010; Southavilay, Yacef, & Calvo, 2010).

Methods and analysis

The data collection method and the sample have been described in other papers (Reimann, et al., 2009a). Briefly, data were collected using a tool called Snooker (Ullman, Peters, & Reimann, 2005) from a group of graduate students who worked on a design task of adding instructional design features to an existing system dynamics model, without meeting face to face (Reimann, Thompson, & Weinel, 2007a). Students were expected to coordinate their own work for this task, which required frequent decision-making about the task and managing the group work. We focus on just one of the groups in this paper, composed of three female students and one male student. The learning environment combined synchronous and asynchronous communication components. Students had access to both a chat tool and a wiki.

This paper builds on previously reported studies (Reimann, et al., 2007a; Reimann, Thompson, & Weinel, 2007b). In particular, the analysis of the processes of decision-making were reported on in Reimann et al. (2009a). Process modelling and mining were used to produce a generalised model group of decision-making in a chat. Process was viewed as a sequence of events, which placed it somewhere between an atomistic and holistic granularity of process, with an event-oriented unit of analysis, rather than a variable-oriented view. The DFCS has seven main categories: 1) problem definition; 2) orientation; 3) solution development; 4) non-task; 5) simple agreement; 6) simple disagreement; and 7) implementation. Category Solution Development (3) has five subcategories: 3a solution analysis, 3b solution suggestions, 3c solution elaboration, 3d solution evaluation, 3e solution confirmation. Poole & Roth's (1989a, 1989b) model was used, stating that groups work on multiple threads, or decisions at the same time, and that the decisions are mixed together in observable behaviour. The DFCS was selected as it includes problem definition, orientation and solution development. In other work (Kennedy-Clark, Thompson, & Richards, 2011) a modified version of the scheme was adopted, based on the availability of data that showed implementation of decisions, and we have included this additional code in our analysis.

The 2009 study concluded that the decision process was unstructured, complex, and cyclic, but that within that, differences could still be identified in the two groups of students. In this paper we focus on just one of the groups, which had a more pronounced cyclic pattern to their decision-making. We have included the wiki data in our analysis in order to examine the way in which learners combined their use of the tools.



(a) Heuristics Miner (Reimann et al., 2009) (b) First order Markov model

Figure 1: Process models based on original data

Part (a) of Figure 1 shows the original model presented in 2009, based on an analysis using a heuristics miner. As described in Reimann et al. (2009), this showed that the group began with a period of problem definition, followed either by discussion of solution alternatives, group orientation, and discussion of solution criteria. The dependencies scores indicated that problem definition statements were most often followed by extensive periods of group orientation. The group then discussed either solution analysis, or solution alternatives, generally followed by decision confirmation, although sometimes it was followed by solution elaboration and evaluation. If there was disagreement after deciding on criteria, the group often started from the top again. The confirmation of solutions was followed by positive evaluation or agreement. In some cases the positive evaluation statements triggered discussion of the criteria again.

The second model (b), is the process model derived from the application of a first-order Markov model. The dataset can be considered as a sequence of states (the eleven listed earlier) ordered in time. The set of states $\{S_1, ..., S_k\}$ occur as the series $\{x_1, \ldots, x_n\}$ where k = 11, the number of states, and n = 1828, the number of coded entries. A first-order Markov chain, is created by calculating a table of probabilities for each state transition, expressed by Equation 1. In Figure 1(b) all state transitions with a probability greater than or equal to 0.25 have been represented. For example, the line from *solution confirmation* to *agreement* is labeled 11 (to indicate the number of state transitions) and 0.275, the probability from within this dataset that agreement will follow solution confirmation.

$$\sum_{i=1}^{k}\sum_{j=1}^{k} Pr(X_t = S_i | X_{t-1} = S_j) \qquad (1)$$

It is useful to compare this Markov chain to the heuristic miner representation. The heuristic miner representation conveys dependency relationships, which are heavily influenced by the directional relationships between states – if there is an equal number of transitions from state S_i->S_j as from S_j->S_i then this will not show up as a strong relationship using the heuristic miner. The Markov chain simply shows the probability of each state transition and is appropriate for this work looking at a single group of students and this is adopted for the remainder of the paper.

Data from wiki entries and chats were assigned to the decisions that were made as part of the collaboration (Reimann, et al., 2009a), and coded using a modified version of the Decision Function Coding Scheme. Of the data that was coded by the authors, 2517 lines of chat data were assigned, and 69% agreement was achieved. Further discussion resulted in 100% agreement. Similar coding of wiki data resulted in initial agreement of 98%. In total, 28 decisions were made during four weeks, and were assigned to cases according to the area about which they were deciding (choosing a model to alter, adding to the model, implementing the changes, and general planning). The decisions related to general planning were disregarded not analysed for this paper. 1649 lines of chat data were used, and 179 wiki revisions. A selection of the 28 cases, the distribution of the decisions, and the tools that were used in this process, are displayed in Figure 2.

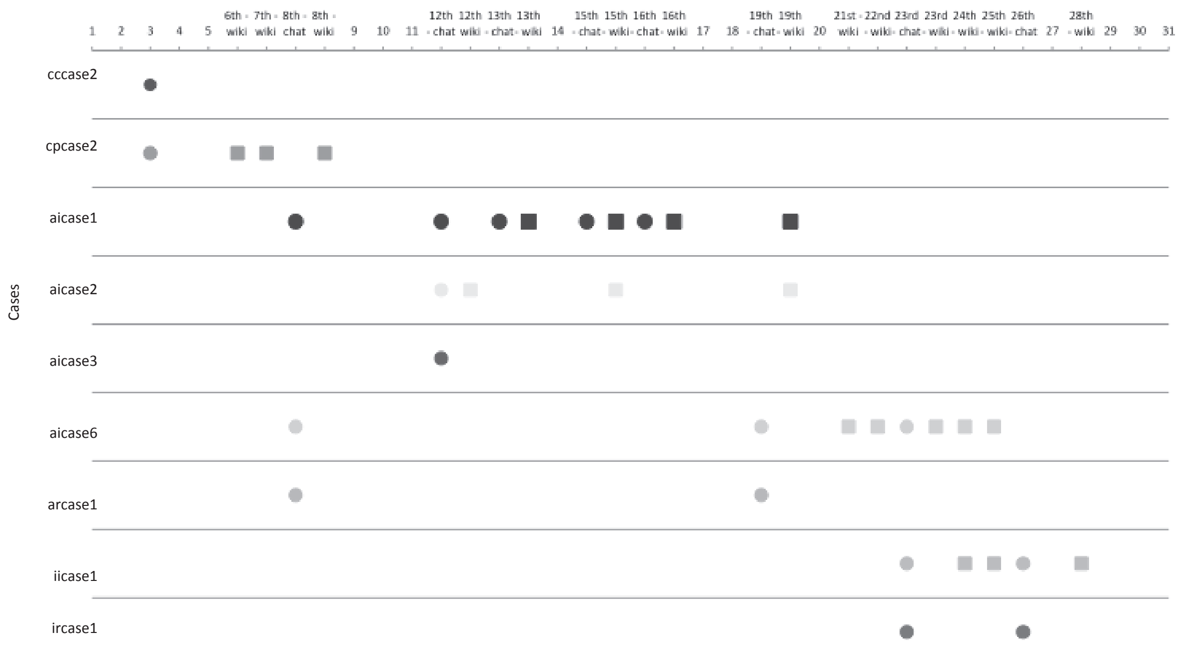


Figure 2. Distribution of a selection of cases over time, by tool use

Figure 2 shows the temporal distribution of the *types of decisions* that were being made, as well as the tool use that was involved. Those decisions concerned with choosing the model were made at the beginning of the collaborative process. Some crossover (time-wise) occurred with the decisions that focused on additions to the model. This was repeated for the transition between deciding on the additions to the model and making decisions regarding the implementation of the additions. The squares and circles represent chat and wiki tool use respectively. Patterns of tool use were identified: single chat (for example cccase2), multiple chat (for example ircase1), chat/s with follow-up wiki (for example cpcase2), and integrated (for example aicase1). This data was also examined at the level of the DFCS (see Figure 3 below).

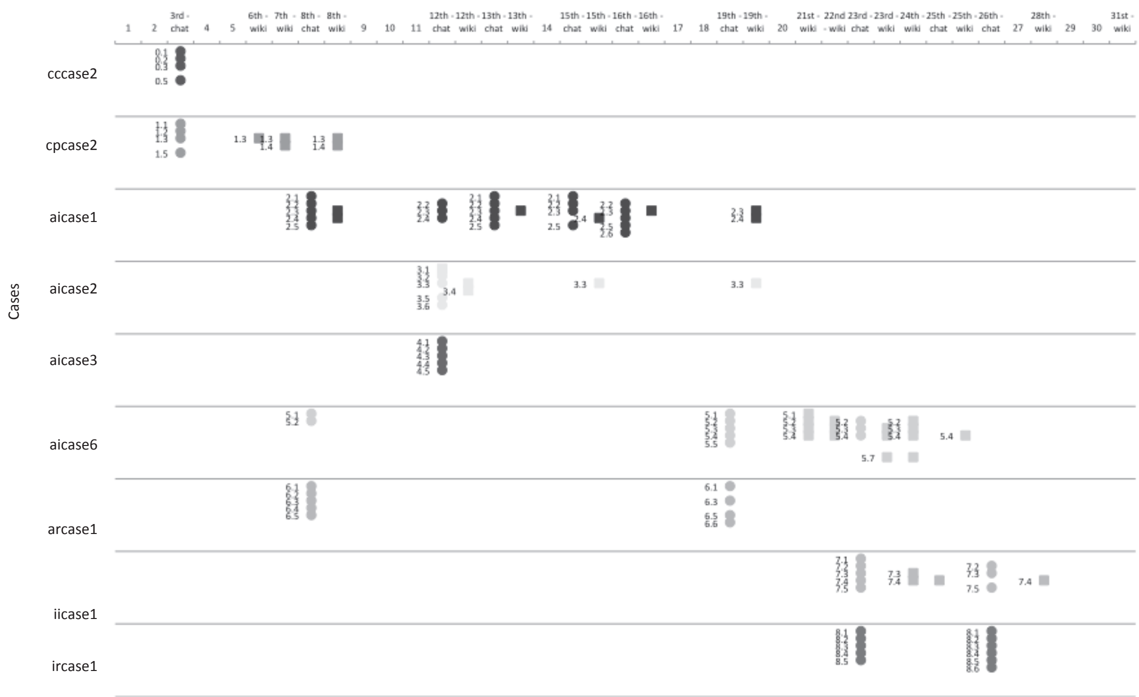


Figure 3. DFCS for selected cases over time, by tool use

Figure 3 shows the different patterns of tool use (circles for chats and squares for wikis) in the states of decision-making (e.g. 0.2 is cccase2, code 2 – orientation). It can be seen that most elements of the decision-making process were observed during chats (e.g. aicase1, in the first chat, codes 1-5 were used: problem definition, orientation, solution development, off-task, and simple agreement), and wikis were used for solution development (e.g. cpcase2) and non-task related postings (e.g. cpcase2, mostly formatting). In the cases of integrated tool use, the wiki was also used for orientation (e.g. aicase6).

We used the patterns identified above and reallocated the cases according to the role that the tools played (not the topic of the decision). The new groups, as outlined above, were single chat (6 cases), multiple chat (4 cases), multiple wikis (1 case), chat/s and follow-up wiki (5 cases), and integrated (4 cases). The multiple wiki group was excluded from any further analysis. This re-classification of the *type of case* (e.g. implementation) into the *role of tool* produced four different 'threads' through the data. Each thread was then mined for processes to compare the state transitions in decision making with a high probability of occurrence.

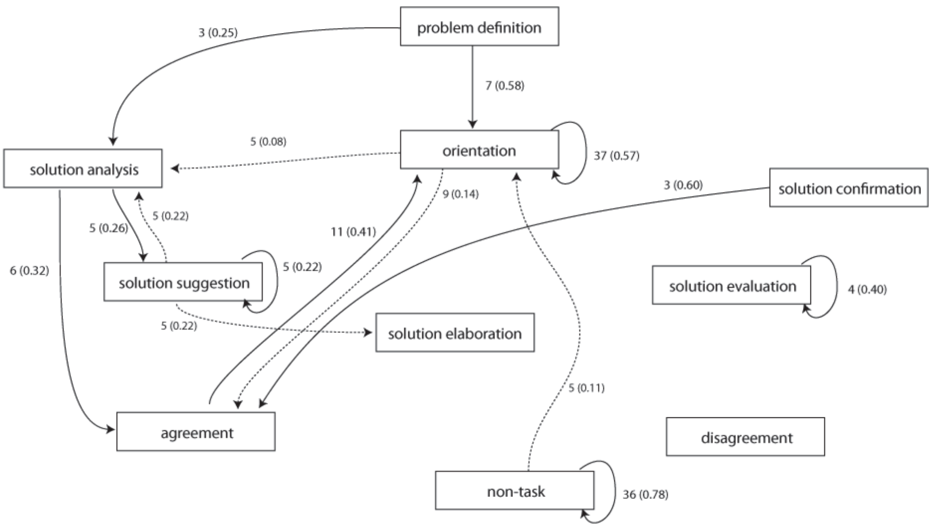


Figure 4. State transitions for decision making in single chats

In the single chats, students began with problem definition, and then moved on to either solution analysis or orientation. Once engaged in orientation there was a high probability that they would continue there, and either agree, or move on to solution analysis. During these single chats, students would move through from solution analysis to solution suggestions, or agreement. Once at solution suggestions, they had a high probability of remaining here (brainstorming) and then either returning to solution analysis or to solution elaboration, at which point students would most probably continue to do this. Links to all other elements of the decision-making process from solution elaboration were fairly evenly distributed. Also worth noting is the high probability that solution confirmation would be followed by agreement.

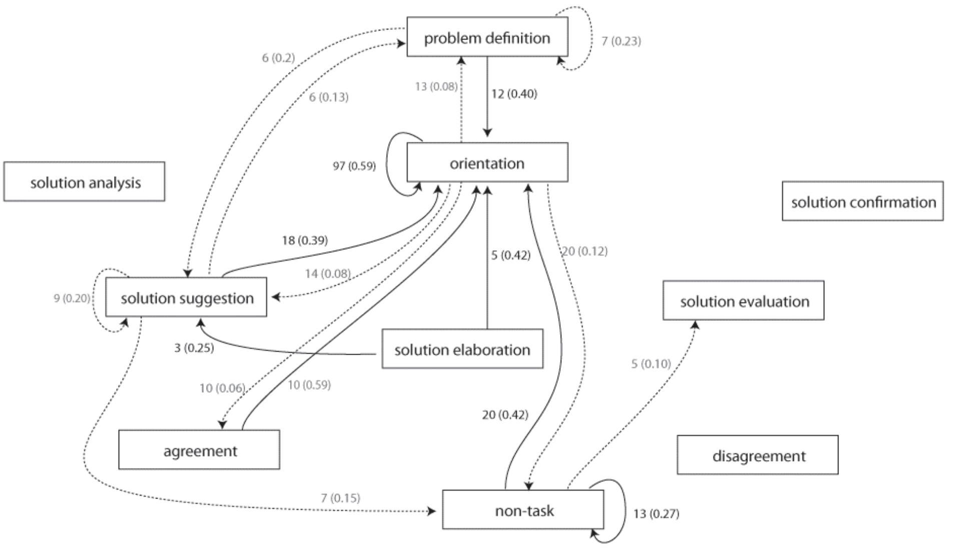


Figure 5. State transitions for decision making in multiple chats

Students who used multiple chats began with problem definition and either continued defining the problem, moved to orientation, or straight to solution suggestions. At orientation, students were most likely to continue with orientation, or begin non-task discussion, however once on non-task they were most likely to return to orientation. Other than these, students either returned to define the problem, moved to solution suggestion, or agreement. Solution analysis was not connected to any other elements, with the probability that students move from here to other aspects of solution development relatively equal. Once students began solution suggestions, they were most likely to return to orientation, or to continue in a brainstorming-like activity. In some instances they returned to problem definition or began non-task discussions. When students engaged in solution elaboration, they either returned to orientation or to solution suggestions. Solution evaluations and confirmations were not connected to other elements, with low numbers of instances and options spread between all other elements of solution development and orientation.

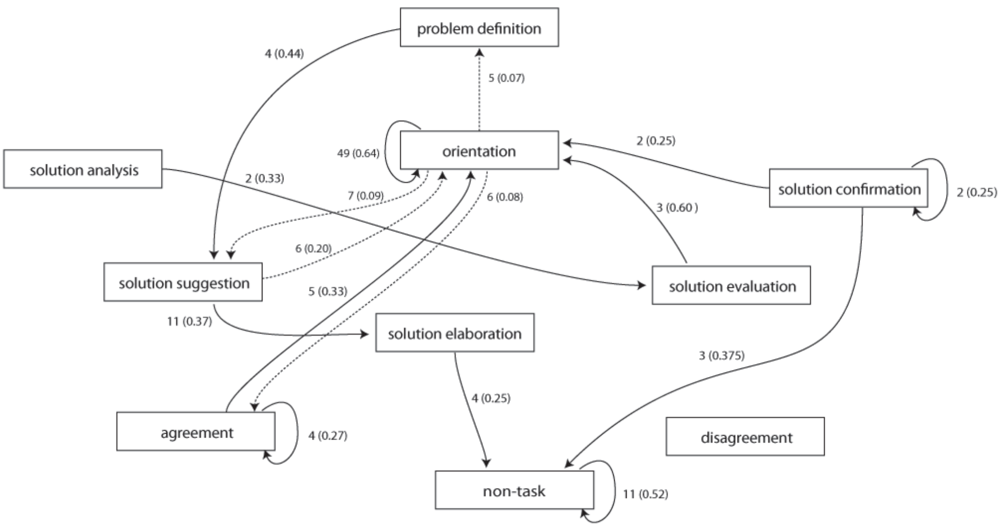


Figure 6. State transitions for decision making in chat/s with follow-up wiki

Students using chats with follow up wiki began with problem definition and were most likely to follow with solution suggestions. If they began their solution analysis, they were most likely to follow with solution evaluation, although this was not common. When they began solution suggestions, they were most likely to follow up with solution elaboration or orientation, they did not continue with solution suggestions. Students engaged in orientation, mostly stayed there, returned to problem definition, agreed, or returned to solution suggestions. When they went to solution elaboration, they most like followed this with non-task (which in almost all cases occurred on the wiki, and is indicative of formatting changes to the text), where they tended to stay. Solution evaluation was likely to be followed by orientation, and evaluation was either followed with non-task, orientation or further confirmation.

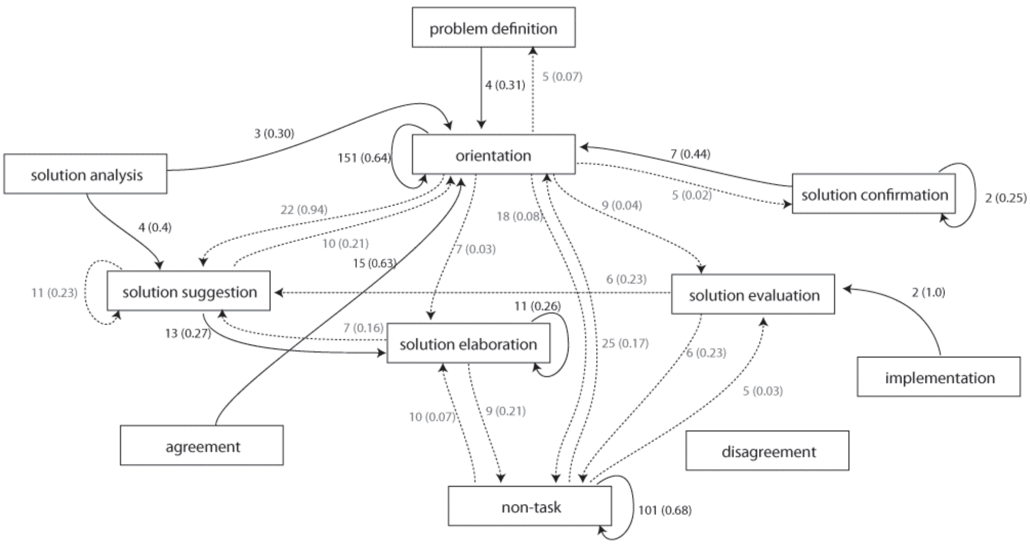


Figure 7. State transitions for decision making in integrated tool use

When students used an integrated approach to their tool use, they began with problem definition, followed by orientation. Once engaged in orientation, they were most likely to stay there. Small probabilities were recorded that flowed back to problem definition, solution suggestion, agreement, solution elaboration, non-task, solution evaluation, and confirmation. When engaged in solution analysis, students either returned to orientation, or proceeded to solution suggestion. At solution suggestion, students would engage in brainstorming, elaboration, or return to orientation. Solution elaboration was most probably followed by further elaboration, orientation, non-task discussion, or a return to solution suggestions. Once students engaged in evaluating solutions, they either moved to non-task discussions (almost all of which were associated with wiki use, and so represented minor formatting changes applied to the text) or returned to solution suggestions. When students engaged in solution confirmation, they tended to remain there, or return to orientation. Evaluation statements followed the two instances of implementation.

Discussion

Processes of decision-making in four patterns of tool use were identified using first-order Markov models. The patterns identified were single chat, multiple chat, chat/s with wiki follow-up and integrated. Relatively efficient decision-making processes characterized decisions made using a single chat. Alternatives were reconsidered early in the process, and once elaborated, were most likely confirmed and agreed upon. Specific decisions that students used this strategy to address were deciding on the subject area on which to focus (cccase2), and adding instructions to the software (aicase3). A different pattern was observed in those decisions that were made over multiple chats. This decision-making process was characterized by a consistent return to defining the problem and identifying features important to making the decision. Topics such as the resources to add to the model (arcase1), and what to do about ongoing software problems (ircase1) were addressed over multiple chats. They needed ongoing discussion, but were relatively straightforward. Nevertheless, the model indicates a careful process of considering options, over several sessions.

The next two patterns of tool use were more complex, as the wiki was incorporated into the decision-making process. From examination of the chat/s with follow-up wiki process, along with the figures presented earlier identifying the temporal distribution of states of the decision-making process, it appears that the chat was used to identify initial solutions and agree on these, and the wiki for further elaboration, asynchronously. Students addressed the process of choosing a model (cpcase2) and the instructional approach adopted (aicase2) in this way. For these decisions, the group did not return to the chat to evaluate the elaborated solutions, or to confirm them. Further consideration needs to be given as to whether these instances could be classified as *implementation* instead. Finally, the integrated approach to tool use in the decision-making process was examined. Questions such as what to add (aicase1) and the issues around the addition (aicase6) and implementation (iicase1) of the storytelling function in the software were addressed in this way. This type of decision-making was clearly separated into stages of understanding the problem, identifying solutions, using the wiki to elaborate on these (often during the chat session, but not always), and then evaluation and confirmation took place in the chat, both of which did, at times, result in further solution development. This is a complex, cyclical process, with truly integrated tool use. This group took advantage of the affordances of the tools identified in other studies.

By dividing the cases by patterns of tool use, we were able to use first order Markov chains to differentiate patterns of decision-making, that fit the different strategies devised by students. Further analysis of the types of decisions included needs to occur, however it would seem at first glance, that time was important in identifying combinations of tool use and then patterns of decision-making within that. The Markov chains were useful in the analysis for showing with clarity the transitions between states that had a high probability of occurring. They lend themselves to a straightforward graphical comparison of different groups. The limitations of this technique of analysis are that the representations only show sequential state transitions and do not reveal hidden states – they assume that all states are known and are present within the dataset. Higher-order Markov models need to be used in cases where non-sequential state transitions are relevant, and hidden Markov models in cases where hidden states may be revealed.

Conclusions

Time is important in investigating processes of learning. In this paper we used time to identify four patterns of tool use and then within these, examined the processes of learning. First-order Markov models were a valuable addition to our analysis. We examined a complex system of collaboration, students made decisions over weeks, and within any one chat session discussed multiple topics. They, seamlessly it seems, used the wiki as both an asynchronous and synchronous collaboration tool. Systematic investigation of the processes involved provides valuable information for designers, manager/instructors and researchers. Generally, we examine patterns of collaboration in order to aid our understanding of the processes of collaboration, to identify indicators of 'good' collaboration in order to aid in classroom management, or to aid in the design of collaborative learning environments. In this study we were able to visualize the temporal nature of the decision-making process, and use this to identify patterns of tool use. The repetition of complexity in both the decision-making process and tool use, suggests that tool use may be a useful indicator of complexity in other elements of the collaborative process (see (Thompson & Kelly, accepted) for more details on this). This has implications for automated feedback to managers of collaborative learning environments, which will be followed up in future research. In addition, the observation of the adaptive nature of tool use in this group would indicate that collaborative learning environments should allow students to have this flexibility to adapt their strategies to meet their needs. Finally, this study contributes to the discussion around the analysis of processes in collaborative learning environments, and recommends consideration of more than time, but of as many elements of the complex process as are possible.

References

Aalst, J. (2009). Distinguishing knowledge-sharing, knowledge-construction, and knowledge-creation discourses. *International Journal of Computer-Supported Collaborative Learning, 4*, 259-287.

Ding, N. (2009). Visualizing the sequential process of knowledge elaboration in computer-supported collaborative problem solving. *Computers & Education, 52*, 509-519.

Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics, 14*(9), 755-763.

Goggins, S. P., Laffey, J., & Amelung, C. (2011). *Context aware CSCL: Moving toward contextualised analysis*. Paper presented at the Connecting computer-supported collaborative learning to policy and practice.

Jeong, H., Biswas, G., Johnson, J., & Howard, L. (2010). *Analysis of productive learning behaviors in a structured inquiry cycle using hidden markov models*. Paper presented at the Proceedings of the 3rd International Conference on Educational Data Mining.

Kapur, M. (2011). Temporality matters: Advancing a method for analyzing problem-solving processes in a computer-supported collaborative environment. *International Journal of Computer Supported Collaborative Learning, 6*(1), 39-56.

Kennedy-Clark, S., Thompson, K., & Richards, D. (2011). *Collaborative Problem Solving Processes in a Scenario-Based Multi-User Environment*. Paper presented at the CSCL2011: Connecting computer supported collaborative learning to policy and practice.

Poole, M. S., & Holmes, M. E. (1995). Decision Development in Computer/Assisted Group Decision Making. *Human Communication Research, 22*(1), 90-127.

Poole, M. S., & Roth, J. (1989a). Decision Development in Small Groups IV: A Typology of Group Decision Paths. *Human Communication Research, 15*(3), 323-356.

Poole, M. S., & Roth, J. (1989b). Decision Development in Small Groups V: Test of a Contingency Model. *Human Communication Research, 15*(4), 549-589.

Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning, 4*, 239-257.

Reimann, P., Frerejean, J., & Thompson, K. (2009a). Using process mining to identify models of group decision making in chat data. 98-107.

Reimann, P., Frerejean, J., & Thompson, K. (2009b). *Using process mining to identify models of group decision making in chat data*. Paper presented at the Computer Supported Collaborative Learning Practices, CSCL 2009.

Reimann, P., Thompson, K., & Weinel, M. (2007a). *Collaborative learning by modelling: Observations in an on-line setting.* Paper presented at the ICT: Providing choices for learners and learning., ASCILITE, Singapore.

Reimann, P., Thompson, K., & Weinel, M. (2007b). *Decision making patterns in virtual teams.* Paper presented at the Supporting learning flow through integrative technologies, Amsterdam.

Shih, B., Koedinger, K. R., & Scheines, R. (2010). *Discovery of Learning Tactics using Hidden Markov Model Clustering*. Paper presented at the Proceedings of the 3rd International Conference on Educational Data Mining, 2010.

Southavilay, V., Yacef, K., & Calvo, R. A. (2010). *Analysis of Collaborative Writing Processes Using Hidden Markov Models and Semantic Heuristics*. Paper presented at the Proceedings of the Third International Workshop on Semantic Aspects in Data Mining 2010, in conjunction with the 10th IEEE International Conference on Data Mining (ICDM 2010).

Thompson, K., & Kelly, N. (accepted). *Combining collaboration spaces: Identifying patterns of tool use for decision-making in a networked learning environment*. Paper presented at the Eighth International Conference on Networked Learning.

Ullman, A., Peters, D., & Reimann, P. (2005). *Developing a research supportive web-based learning System.* Paper presented at the ODLAA Conference 2005, Adelaide.

van der Aalst, W. M. P., & Weijters, A. J. M. M. (2005). Process Mining *Process-Aware Information Systems*: John Wiley & Sons, Inc.

Weijters, A. J. M. M., Van der Aalst, W. M. P., & Medeiros, A. K. A. d. (2006). Process Mining with Heuristics Miner Algorithm, *BETA Working Paper Series*. Eindhoven University of Technology, Eindhoven.

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning.

Wise, A., & Chiu, M. M. (2011a). Analyzing temporal patterns of knowledge construction in a role-based online discussion. *International Journal of Computer-Supported Collaborative Learning, 6*(3), 445-470.

Wise, A., & Chiu, M. M. (2011b). *Knowledge construction patterns in online conversation: A statistical discourse analysis of a role-based discussion forum*. Paper presented at the Connecting computer-supported collaborative learning to policy and practice.

Zenios, M., & Holmes, B. (2010). *Knowledge creation in networked learning: combined tools and affordances*. Paper presented at the 7th International Conference on Networked Learning 2010.

Acknowledgements

This work was funded by the Australian Research Council grant FL100100203.