Sequentially Analyzing and Mapping the Interactional Processes of Knowledge Construction in Online Learning

Jeong, Allan (2008). Sequentially analyzing and mapping the interactional processes of knowledge construction in online learning. Paper presented at the 2008 American Educational Research Association conference in New York, NY

Abstract

This paper describes how sequential analysis (including specific software tools and techniques) can be used to analyze and map message-response sequences to study the interactional processes of knowledge construction in online learning. Step-by-step instructions are presented to illustrate: a) how sequential analysis can be used to determine to what extent messages elicit responses based on what is said in conjunction with *when*, *how*, *who*, and *why* messages are posted; and b) how it has been used in previous studies to determine how latent variables (message function, response latency, communication style) and exogenous variables (gender, discourse rules, context) affect how likely messages elicit responses, the types of responses elicited, and whether the elicited response sequences (e.g., claim-challenge-explain) support/inhibit knowledge construction.

Introduction

Current research in computer-mediated communication (CMC) is in need of alternative theories, methods, and software tools to achieve a deeper and more thorough understanding of CMC and its effects on group interaction, group performance, and learning. At this time, content analysis is one of the current methods used to identify message categories and message frequencies. This approach generates largely descriptive rather than prescriptive findings, reporting for example the frequencies of arguments, challenges and explanations observed in a discussion. However, message frequencies provide little information to explain or predict how participants respond to given types of messages (e.g. argument - challenge vs. argument - simple agreement), how response patterns are influenced by latent variables (e.g., message function, content, communication style, response latency) and exogenous variables (e.g., gender, personality traits, discussion protocols, type of task), and how particular response patterns help to improve group performance to achieve desired outcomes. Therefore, new approaches are needed to determine to what extent messages elicit responses based on what is said in conjunction with *when*, *how*, *who*, and *why* messages are presented, and whether or not the elicited responses help produce sequences of speech acts that support critical discourse (e.g., claim- challenge- explain) and group performance in decision-making, problem-solving, and learning.

Sequential analysis has been used in studies on inter-personal communication conducted over the last 30 years to examine conversational patterns between married couples, children at play, mother infant play, and studies on human-computer interaction. This method has been claimed by some to be the ‘missing factor’ (King & Roblyer, 1984; England, 1985) in research on the effects of computer-mediated environments and computer-based instruction. As a result, this paper presents seven steps (including software tools and techniques) to using sequential analysis to study the interactional processes of knowledge construction developed in my previous studies. The paper begins with a proposed theoretical framework used to identify the appropriate metrics for measuring group interaction, followed by the presentation of specific methods and software tools to support sequential analysis, and research designs used to investigate factors that influence group interaction.

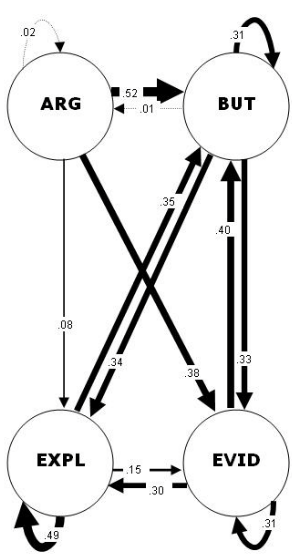
Theoretical framework

The dialogic theory (Bakhtin, 1981) provides a theoretical framework for reconceptualizing and operationalizing group interaction in collaborative learning (Koschmann, 1999). In this theory, the two main assumptions are that a) conflict is produced not by ideas presented in one message alone, such as an argument or claim, but by the juxtaposition of opposing ideas presented in a message and responses to the message, and b) conflicts produced in exchanges help to trigger subsequent responses that can serve to verify (e.g. argument-challenge-evidence) and justify (e.g., argument – challenge - explain) stated arguments and claims. These assumptions imply that we should be focusing on analyzing the frequency of specific message-response pairs (e.g., argument - challenge, challenge - explain) and not the frequency of messages alone (e.g., arguments, challenges, explanations).

Step 1 - Choose a metric for measuring and comparing group interaction patterns

The two metrics described in this paper are *transitional probabilities* and *mean response scores*. Transitional probabilities are computed by tallying the frequency and relative frequency of a particular response posted in reply to a particular message type and by reporting the results in a frequency matrix (Tables 1 & 2). To determine if a particular transitional probability is significantly higher or lower than expected and to determine whether a pattern exists in the way participants respond to certain messages, z-scores are computed and reported in a *z*-score matrix (Table 3). The z-scores takes into account not only the observed total number of responses to a particular message category, but also the marginal totals of each response type observed across all message types. The transitional probabilities can then be examined in the form of state diagrams (Figure 1) to provide a Gestalt view of the group processes and a means to visually identify response patterns and predict event sequences most likely to occur. For example, the diagram can be used to determine or predict how often arguments will elicit challenges versus counter-arguments, and in turn, predict how often challenges will elicit explanations versus counter-challenges.

Figure 1. Transitional state diagram



For example, 52% of all replies to ARGuments were challenges (BUT), and 34% of all responses to challenges were EXPLanations posted to defend the argument.

The mean number of specific responses elicited per message category (mean response scores) determines how many times a given type of message is able to elicit a particular type of response. This metric *describes* the overall level of performance by measuring, for example, the mean number of challenges elicited per argument and the mean number of explanations elicited per challenge, which is similar to measuring the percentage of arguments left unchallenged and the percentage of challenges left unresolved. As a result, this particular metric can be used to determine at what level participants are critically analyzing arguments (e.g., argument-challenge-explain), or to what extent participants engage in processes (e.g., argument-counterargument, argument-no response) that block critical discourse. By using mean scores, statistical methods like *t*-tests and analysis of variance can be used to test for differences in response patterns between experimental conditions, and effect sizes can be computed to determine to what extent the observed differences are meaningful differences.

Use transitional probabilities to explain observed differences in mean response scores. For example, one group might exhibit a tendency to respond to arguments with more challenges than with supporting evidence, whereas another group might exhibit an opposite tendency to respond to arguments with more supporting evidence but fewer challenges. If a significant difference is found in the mean number of challenges elicited per argument between groups, the differences in interaction patterns would suggest that the second group posted fewer challenges in response to arguments *because* more time and resources were allocated by the group to developing evidence to support arguments leaving less time and resources to challenge arguments.

Table 1. Frequency matrix of responses to messages across message categories

	ARG	BUT	EVID	EXPL	Replies	No Replies	Givens	% Targets	% Givens
ARG	3	**101**	**73**	**16**	193	35	112	.25	.30
BUT	3	**82**	88	91	264	24	149	.35	.40
EVID	0	64	50	48	162	22	35	.21	.09
EXPL	0	51	**22**	**71**	144	55	74	.19	.20
	14	307	233	229	763	136	370		

For example, 101 challenges (BUT) were posted in response to arguments (ARG). This frequency was higher than the expected frequency based on its z-score value of 3.96 at p < .01.

Table 2. Transitional probability matrix

	ARG	BUT	EVID	EXPL	Replies	No Replies	Givens	Reply Rate
ARG	.02	**.52**	**.38**	**.08**	193	35	112	.69
BUT	.01	**.31**	.33	.34	264	24	149	.84
EVID	.00	.40	.31	.30	162	22	35	.37
EXPL	.00	.35	**.15**	**.49**	144	55	74	.26
	14	307	233	229	763	136	370	.52

For example, 52% of all responses to arguments (ARG) were challenges (BUT).

Table 3. Z-score matrix

	ARG	BUT	EVID	EXPL
ARG	-0.34	**3.96**	**2.54**	**-7.62**
BUT	-1.05	**-3.76**	1.22	1.95
EVID	-1.96	-0.21	0.10	-0.12
EXPL	-1.82	-1.31	**-4.41**	**5.61**

Z-scores < -2.32 reveal probabilities that were significantly lower than expected. Z-scores above 2.32 reveal probabilities that were significantly higher than expected.

Step 2 – Specify *a priori* tests for specific message-response pairs

The specific message-response pairs examined in your study should be defined *a priori* because the total number of possible event pairs grows exponentially with the addition of each message category to the coding scheme. For example, a coding scheme consisting of four categories (e.g., argument, challenge, explain, evidence) produces a 4 x 4 matrix resulting in 16 possible event pairs (e.g., argument-challenge, challenge-argument, challenge-explain, explain-challenge, and etc.). Testing all 16 event pairs for differences in mean response scores would be too large a number of contrasts to adequately control for Type I error (finding significant differences when the differences are actually the result of random chance alone). Power can be increased by testing only a select number of event pairs – particularly those that are believed to support group performance (e.g., argument-challenge, challenge-explain). To identify the most important sequences to examine in your study, review existing literature and research that present specific models for achieving specific tasks.

Step 3 – Collect discussions and messages parsed and classified by speech act

The next step is to parse the discussion transcripts into discrete units of analysis classified by function (dialog move) based on your coding scheme. One way to facilitate message coding is to instruct students to classify, label, and post messages to address one and only one function at a time (Figure 2). Message labeling has been implemented in a number of computer-supported collaborative argumentation (CSCA) systems to scaffold argumentation and problem solving (Carr & Anderson, 2001; Cho & Jonassen, 2002; McAlister, 2003; Veerman, Andriessen, & Kanselaar, 1999) and to enable participants to see the overall structure of their arguments (Figure 3).

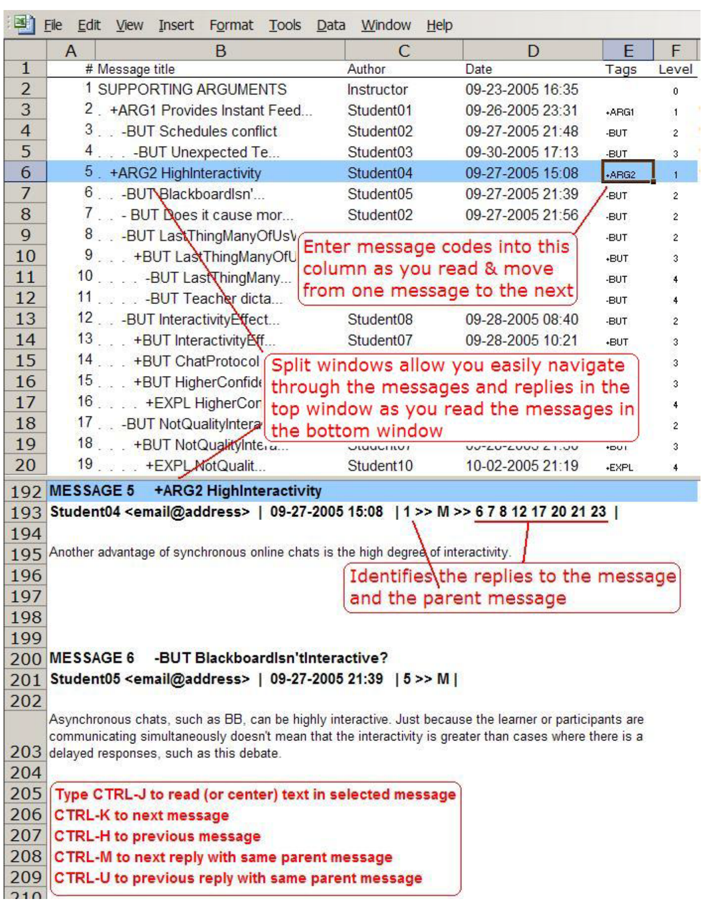
Step 4 – Download messages with message threads intact

Among the software programs that support message downloads, messages are stored into flat files where the explicit links between multi-threaded messages are *not* recorded. Even with existing content analysis tools, such as Atlas-ti and NUDIST, and tools like GSEQ for performing sequential analysis (Bakeman & Quera, 1995), the multi-threaded nature of

Figure 2. Example instructions on how to label messages when posting to an online debate

Symbol	Description of symbol
+	Identifies a message posted by a student assigned to the team supporting the given claim/statement
-	Identifies a message posted by a student assigned to the team opposing the given claim/statement
ARG#	Identifies a message that presents one and only one argument or reason for using or not using chats (instead of threaded discussion forums). Number each posted argument by counting the number of arguments already presented by your team. Sub-arguments need not be numbered. ARG = "argument".
EXPL	Identifies a reply/message that provides additional support, explanation, clarification, elaboration of an argument or challenge.
BUT	Identifies a reply/message that questions or challenges the merits, logic, relevancy, validity, accuracy or plausibility of a presented argument (ARG) or challenge (BUT).
EVID	Identifies a reply/message that provides proof or evidence to establish the validity of an argument or challenge.

Figure 3. Example of online debate with labeled messages from a Blackboard discussion forum downloaded into ForumManager

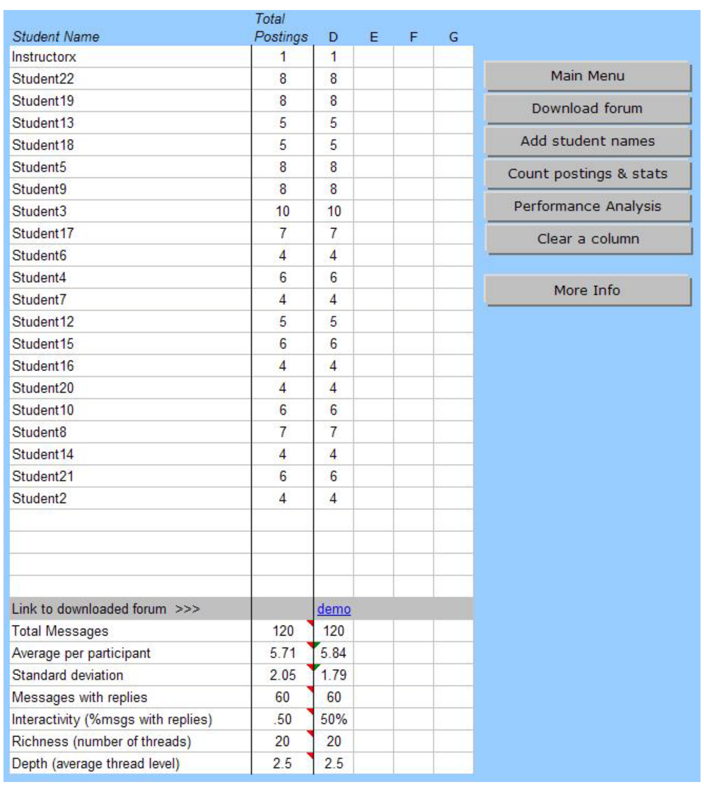


discussions are difficult to record and analyze. However, ForumManager (Jeong, 2004) has been developed and used in previous studies to harvest messages from Blackboard, a course management system (Figure 4a) into Microsoft Excel. Once in Excel, message headers and full texts are archived and the message threads are structurally maintained to enable the user to read and analyze message threads.

Step 5 – Prepare data for analysis according to questions under examination

Use ForumManager to: a) code the messages by *manually* enter codes (Figure 3 column E); b) automatically code messages based on the presence of target keywords (Figure 4b & 4c); or c) automatically pull out the students' labels from message headers (Figure 3 column E) into an Excel worksheet. Note that the code sequences are also extracted by ForumManager and explicitly mapped using a numerical system based on the thread level of each message (Figure 3 column F). Next, modify the codes to identify the data from your experimental groups and enter the codes (along with the thread level data) into the Discussion Analysis Tool (Jeong, 2005) or DAT (Figure 5 column 1 & 2) to identify group interaction patterns based on the variables you have chosen to examine in your study. This presentation will describe ways to manipulate the coded data and use DAT to examine how various factors (function of the message, characteristics of the messenger/responder, of the message text, the response lags) affect and change the response patterns. Specifically, this paper will present findings from my previous studies to illustrate how sequential analysis can be used to understand observed interactions between students based on why, how, who, and when messages and responses are posted in online discourse.

<u>Figure 4a</u>. Screenshot of ForumManager with downloaded discussions



Student Name	Total Postings	D	E	F	G
Instructorx	1	1			
Student22	8	8			
Student19	8	8			
Student13	5	5			
Student18	5	5			
Student5	8	8			
Student9	8	8			
Student3	10	10			
Student17	7	7			
Student6	4	4			
Student4	6	6			
Student7	4	4			
Student12	5	5			
Student15	6	6			
Student16	4	4			
Student20	4	4			
Student10	6	6			
Student8	7	7			
Student14	4	4			
Student21	6	6			
Student2	4	4			
Link to downloaded forum >>>		demo			
Total Messages	120	120			
Average per participant	5.71	5.84			
Standard deviation	2.05	1.79			
Messages with replies	60	60			
Interactivity (%msgs with replies)	.50	50%			
Richness (number of threads)	20	20			
Depth (average thread level)	2.5	2.5			

Figure 4b. Screenshot of ForumManager page used to generate reports on the performance of individual students

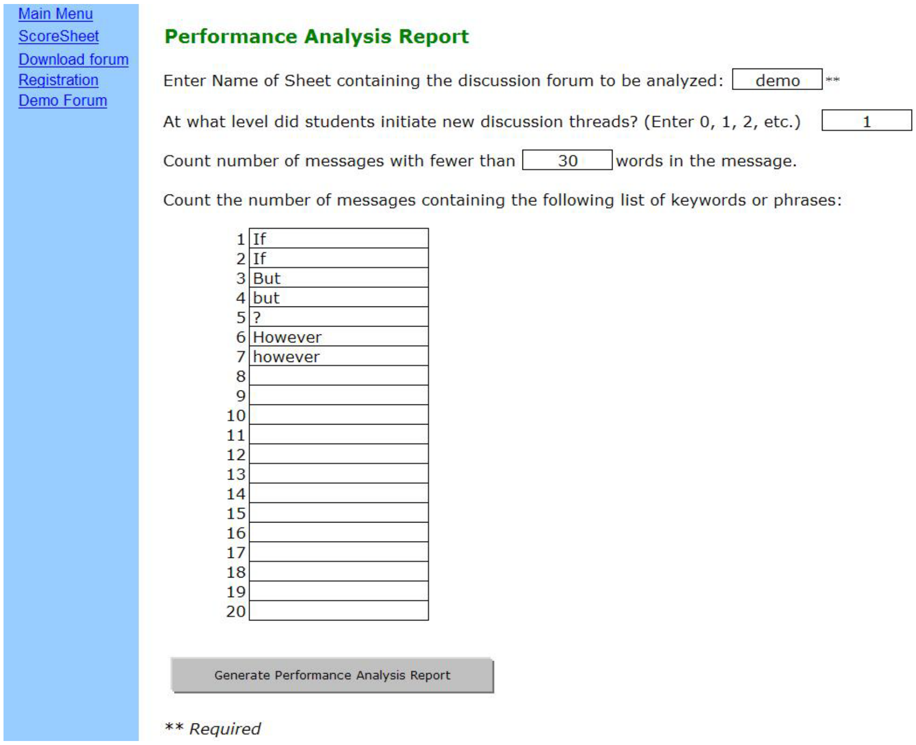


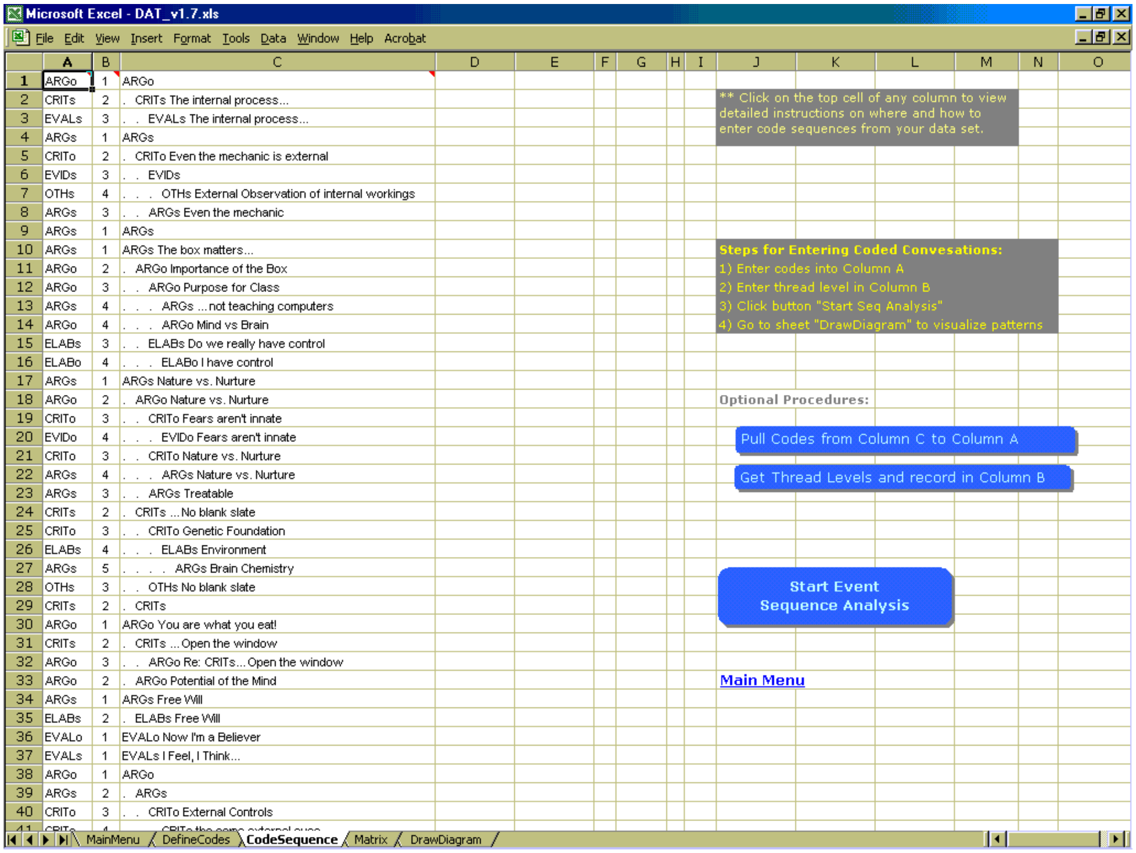
Figure 4c. Example of performance report generated by ForumManager

PERFORMANCE ANALYSIS REPORT: Discussion forum 'demo'

Students	Postings	# of Days	Replies	Replies Elicited	Recip Replies	Total Score	%Posts < 30 Words	Ave# Word/Msg	# Target Keywords	%Msgs w/ Keywords	Forum #
Instructorx	2	1	0	2	0	5	.00	189.0	0	.00	demo
Student22	8	3	6	7	1	25	.13	46.0	4	.25	demo
Student19	8	4	5	5	2	24	.13	84.0	11	.75	demo
Student13	5	2	4	1	0	12	.20	50.4	10	.80	demo
Student18	5	2	4	3	1	15	.20	48.2	3	.40	demo
Student5	8	2	5	5	2	22	.25	41.6	3	.38	demo
Student9	8	3	8	5	2	26	.25	41.1	5	.38	demo
Student3	10	3	8	5	2	28	.30	36.9	12	.40	demo
Student17	7	2	7	3	0	19	.00	102.3	8	.57	demo
Student6	4	1	3	3	0	11	.25	46.0	4	.50	demo
Student4	6	1	6	0	0	13	.50	30.3	4	.50	demo
Student7	4	2	2	3	0	11	.25	62.8	4	.50	demo
Student12	5	1	5	2	0	13	.20	46.6	2	.20	demo
Student15	6	2	4	2	0	14	.00	64.0	4	.33	demo
Student16	4	1	4	1	0	10	.00	41.8	3	.75	demo
Student20	4	1	3	0	0	8	.00	56.5	1	.25	demo
Student10	6	1	5	4	0	16	1.00	20.0	0	.00	demo
Student8	7	2	6	5	0	20	.14	45.7	5	.57	demo
Student14	4	1	4	3	0	12	.75	24.0	2	.25	demo
Student21	6	1	6	1	0	14	.17	56.8	8	.67	demo
Student2	22	5	19	8	3	57	.09	54.7	15	.36	demo

Keywords = (If , If , But, but, ?, However, however)

Figure 5. Screen shot of DAT for processing and analyzing message sequences



Step 6 – Compute transitional probabilities, z-scores & state diagrams

Use the DAT software to compute the frequency, transitional probability, and z-score for each message-response pair. The frequency of event pairs for up to six categories can then be selected to produce state diagrams like those presented in Figures 1 and 8. In addition, DAT supports the analysis of mean response scores by outputting the raw scores (Figure 9) used to compute and test mean response scores in statistical programs to conduct *t*-tests (Figure 10), analysis of variance, regression analysis, multi-dimensional scaling, and other tests that might prove useful in gaining further insights into factors that affect group interaction patterns.

Step 7 – Interpret the transitional probabilities for interaction patterns

Meaningful interpretation of the observed interaction patterns can best be achieved by focusing on only the event sequences that exemplify the processes believed to improve group performance and specified in *a priori* hypotheses. When a particular pattern of interaction is revealed in a z-score matrix, be sure to check that the finding is supported by sufficient cell frequencies for the given message-response pair.

Figure 6. Transitional probability matrix of event sequences produced by DAT

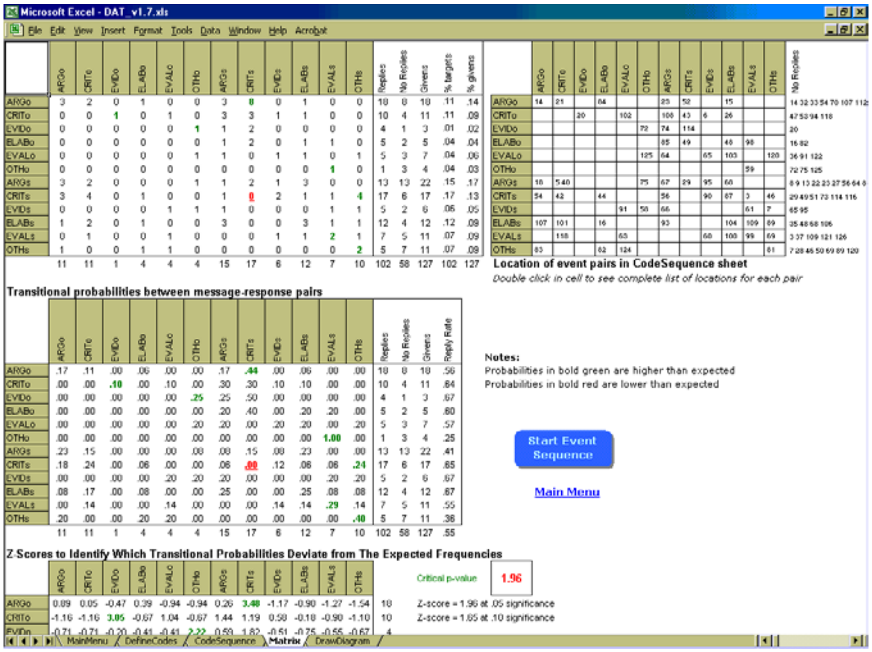
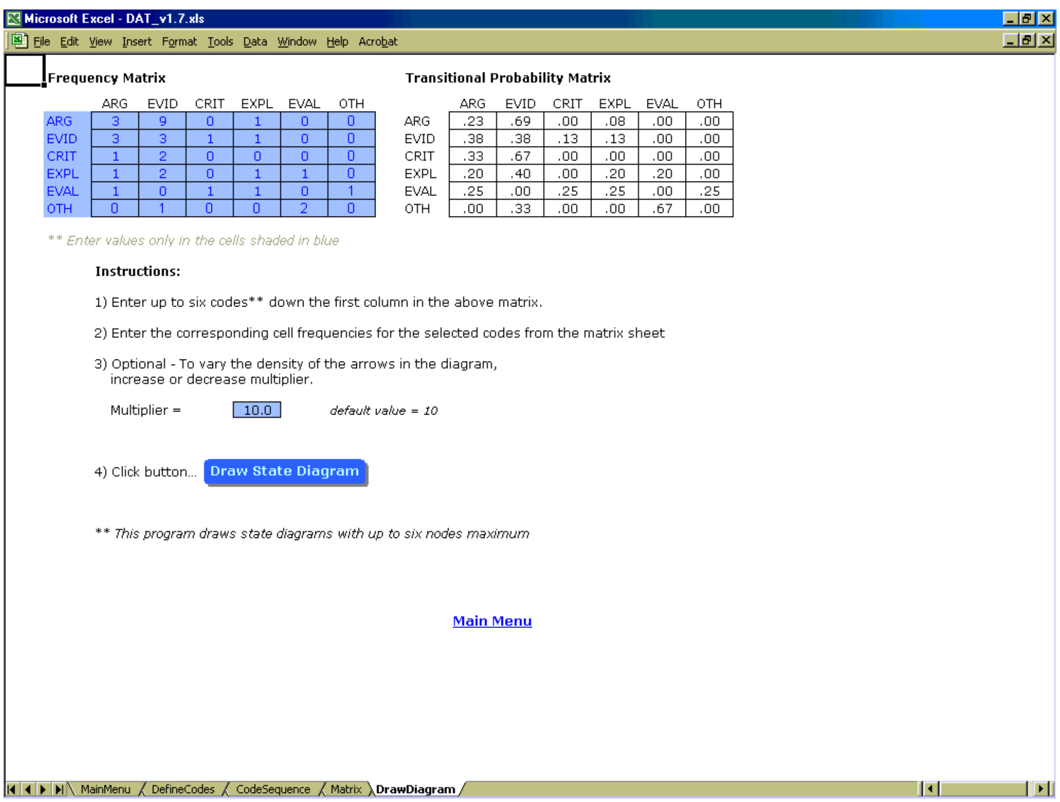


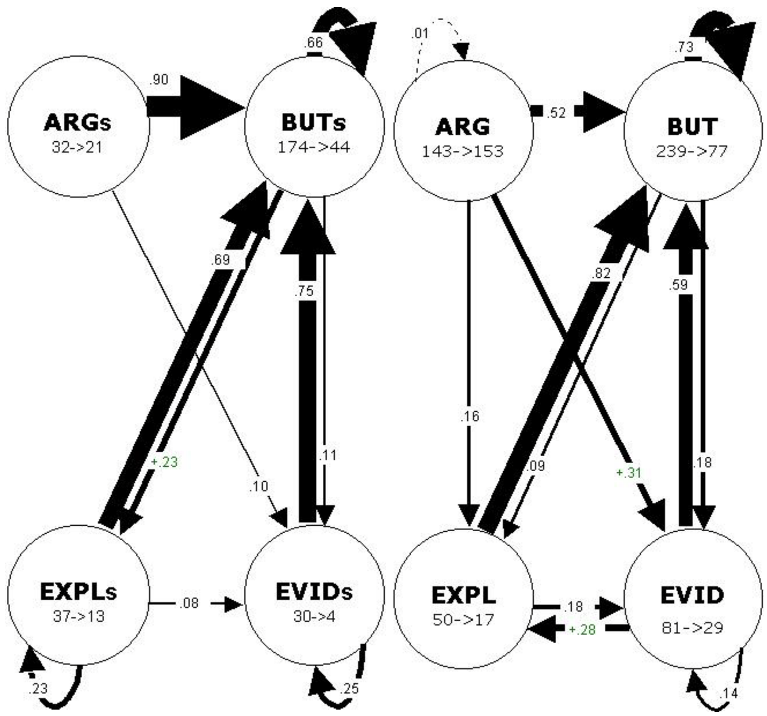
Figure 7. Screen from DAT used to generate a transitional state diagram using frequencies reported in the frequency matrix



Frequency Matrix	ARG	EVID	CRIT	EXPL	EVAL	OTH
ARG	3	9	0	1	0	0
EVID	3	3	1	1	0	0
CRIT	1	2	0	0	0	0
EXPL	1	2	0	1	1	0
EVAL	1	0	1	1	0	1
OTH	0	1	0	0	2	0

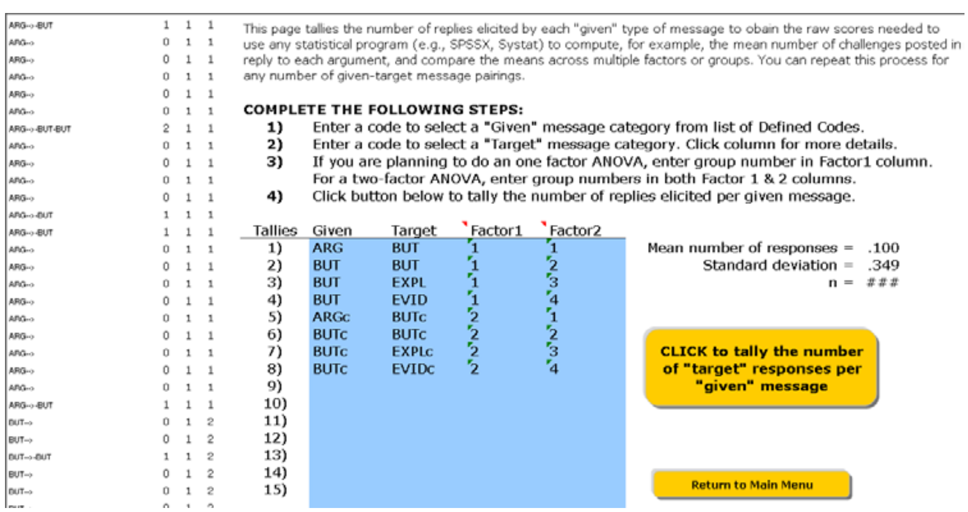
Transitional Probability Matrix	ARG	EVID	CRIT	EXPL	EVAL	OTH
ARG	.23	.69	.00	.08	.00	.00
EVID	.38	.38	.13	.13	.00	.00
CRIT	.33	.67	.00	.00	.00	.00
EXPL	.20	.40	.00	.20	.20	.00
EVAL	.25	.00	.25	.25	.00	.25
OTH	.00	.33	.00	.00	.67	.00

Figure 8. Two transitional state diagrams produced from a previous study comparing interactions produced by messages presented with versus without conversational language



For example: The 32 arguments that were presented using a conversational style (e.g., greetings, emoticons, closing signatures, addressing messages) elicited 21 total responses, where 90% of these responses were challenges. Probabilities presented with "+" indicate those that were significantly higher than the expected probability with z-scores > 2.32 at p < .01.

Figure 9. Screen in Discussion Analysis Tool used to generate raw scores to compute mean response scores.



Note: This page generates the raw scores (in column 2) that identify the number of BUT responses posted in reply to each ARG message (interaction type 1) in group 1 (exchanges produced with conversational language), as well as the number of BUTc replies posted in reply to ARGc messages (interaction type 1) in group 2 (exchanged produced with conversational language).

Figure 10. Mean response score table reporting the mean number of target responses elicited per given message type presented with vs. without a given indicator of conversational language

	M	*n**	*STD*	*t*-test	*df*	*p*	%Diff	Effect size	
ARG --> BUT									
with signature	1.60	25	1.73	3.16	166	.00	.86	.38	**
no indicators	.86	143	.92						
ARG --> EVID									
with signature	.12	25	.33	-2.25	166	.03	-2.55	-.41	**
no indicators	.43	143	.67						
BUT --> BUT									
with reference	.40	10	.70	-.09	247	.927	-.05	-.02	
no indicators	.42	239	.62						
with signature	.36	25	.57	-.46	262	.650	-.17	-.07	
no indicators	.42	239	.61						
with question	.38	44	.54	-.32	281	.747	-.11	-.05	
no indicators	.42	239	.62						
with I agree	.75	20	.71	2.28	257	.023	.79	.35	**
no indicators	.42	239	.61						
BUT --> EXPL									
with reference	.00	10	.00	-.72	247	.470	---	-.23	
no indicators	.05	239	.22						
with signature	.08	25	.28	.63	262	.529	.60	.08	
no indicators	.05	239	.22						
with question	.14	44	.41	2.04	281	.040	1.72	.18	**
no indicators	.05	239	.22						
with I agree	.00	28	.00	-1.02	257	.307	---	-.23	
no indicators	.05	239	.22						

* Number of given messages with the given indicator (and that indicator alone) versus number of messages with no indicators, ** Significant at $p < .05$, 'ARG' = argument, 'BUT' = challenge, 'EVID' = supporting evidence, 'EXPL' = explanation.

Conclusions & Implications

The methods and tools described in this paper provides a road map on how to study the *processes* of collaborative knowledge construction in online learning environments and how factors affect discourse processes in CMC. This approach to studying online interaction will produce the research and findings needed to develop collaborative learning strategies that produce or elicits dialog moves sequences that have or will be proven to maximize collaborative discourse and improve group performance. In other words, the sequential analysis tools and methods discussed in this paper can be used to better understand and improve the processes of collaborative knowledge construction in online learning environments.

In terms of the long-range applications, the proposed methods will provide a starting point for building computational models to explain, predict, and perhaps simulate group discussions in computer-mediated environments. Computational models of group processes combined with the use of techniques like message labeling may serve as the mechanism for building intelligent discourse environments that can diagnose and maximize collaborative knowledge construction. Furthermore, the methods and tools presented here can be used to model interaction patterns in any social exchange, including exchanges between instructor and student, coach and athlete, counselor and patient, and computers and humans in both online and face-to-face environments, and in both group and individual learning tasks. These methods and tools, hopefully, will provide researchers with an effective and alternative approach to studying the processes of human learning and performance.

References

Bakeman, R. & Gottman, J. (1997). *Observing interaction: An introduction to sequential analysis*, Cambridge, University Press.

Bakhtin, M. (1981). *The dialogic imagination*. (Ed. M. Holquist) Austin: University of Texas Press.

Carr, C., & Anderson, A. (2001). Computer-supported collaborative argumentation: Supporting problem-based learning in legal education. Paper presented at the *Computer Support for Collaborative Learning (CSCL) 2001 Conference*. Retrieved October 30, 2003: http://www.mmi.unimaas.nl/euro-cscl/Papers/25.pdf.

Cho, K., & Jonassen, D. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development, 50* (3), 5-22. Retrieved March 3, 2004 from http://tiger.coe.missouri.edu/~jonassen/Argumentation.pdf

England, E. (1985). Interactional analysis: The missing factor in computer-aided learning design and evaluation. *Educational Technology*, 25(9), p24-28.

Jeong, A. (2004). *ForumManager*. Retrieved on July 6, 2005 from http://garnet.fsu.edu/~ajeong.

Jeong, A. (2005). *Discussion analysis tool (DAT).* Retrieved on April 18, 2005 from http://garnet.fsu.edu/~ajeong/DAT.

King, Fi. & Roblyer, M. (1984). Alternative designs for evaluating computer-based instruction. *Journal of Instructional Development*, 7(3), 23-29.

Koschmann, T. (1999). Toward a dialogic theory of learning: Bakhtin's contribution to understanding learning in settings of collaboration. In *Proceedings of the Computer Support for Collaborative Learning (CSCL) 1999 Conference*.

McAlister, S. (2003). Assessing good argumentation. Retrieved April, 10, 2004 from: http://iet.open.ac.uk/pp/s.r.mcalister/personal/AssessingGEA.htm

Veerman, A., Andriessen, J & Kanselaar, G. (1999). Collaborative learning through computer-mediated argumentation. In C. M. Hoadley and J. Roschelle (Eds.), *Proceedings of the Computer Support for Collaborative Learning (CSCL) 1999 Conference* (pp. 640-650). Palo Alto, CA: Stanford University [Available from Lawrence Erlbaum Associates, Mahwah, NJ].