Journal of Universal Computer Science, vol. 21, no. 1 (2015), 66-92
submitted: 16/5/14, accepted: 1/10/14, appeared: 1/10/15 © J.UCS

Learning Analytics at "Small" Scale: Exploring a Complexity-Grounded Model for Assessment Automation

Sean Goggins
(School of Information Science & Learning Technologies
University of Missouri, Columbia, USA
GogginsS@missouri.edu)

Wanli Xing
(School of Information Science & Learning Technologies
University of Missouri, Columbia, USA
wxdg5@mail.missouri.edu.edu)

Xin Chen
(School of Engineering Education
Purdue University, West Lafayette, USA
chen654@purdue.edu)

Bodong Chen
(Department of Curriculum and Instruction
University of Minnesota, Minneapolis, USA
bodong.chen@gmail.com)

Bob Wadholm
(School of Information Science & Learning Technologies
University of Missouri, Columbia, USA
rrwadholm@mail.missouri.edu)

Abstract: This study proposes a process-oriented, automatic, formative assessment model for small group learning based on complex systems theory using a small dataset from a technology-mediated, synchronous mathematics learning environment. We first conceptualize small group learning as a complex system and explain how group dynamics and interaction can be modeled via theoretically grounded, simple rules. These rules are then operationalized to build temporally-embodied measures, where varying weights are assigned to the same measures according to their significance during different time stages based on the golden ratio concept. This theory-based measure construction method in combination with a correlation-based feature subset selection algorithm reduces data dimensionality, making a complex system more understandable for people. Further, because the discipline of education often generates small datasets, a Tree-Augmented Naïve Bayes classifier was coded to develop an assessment model, which achieves the highest accuracy (95.8%) as compared to baseline models. Finally, we describe a web-based tool that visualizes time-series activities, assesses small group learning automatically, and also offers actionable intelligence for teachers to provide real-time support and intervention to students. The fundamental contribution of this paper is that it makes complex, small group behavior visible to teachers in a learning context quickly. Theoretical and methodological implications for technology mediated small group learning and learning analytics as a whole are then discussed.

Keywords: Learning Analytics, Small Group Learning, Assessment, Complex Systems
Categories: L.0.0, L.1.1, L.3.6, L.6.2, E.0

1 Introduction

One of the important goals of learning analytics is to measure and assess data about learners and their contexts in order to automatically provide actionable knowledge for teachers and students [Simens and Baker, 12]. Assessment is not only important in evaluating learning outcomes: it may also be a motivating factor for students who have a performance goal orientation [Dennen, 08]. While measurement and assessment of learning is a major objective of learning analytics, it is also a demanding experience for many teachers due to the heavy workload and time-consuming nature of the assessment activities, especially when learning takes place in technology mediated group settings [Strijbos 11][Gress et al. 10].

The assessment of small group learning in computer-supported environments is more than the measurement of outcomes**;** the quality of collaborative learning processes [Strijbos, 11] is also salient. However, the automation of process-oriented assessments for small group learning in a socio-technical context is a difficult problem to resolve, as learning in this context takes place through the complex processes and interactions of numerous factors, artifacts, and environments [Barab, 01][Stahl 12]. Due to this complexity, assessment of group learning has been dominated by "after collaboration" measurement [Gress et al., 10], where the performance of each group is measured by the quality of the solutions or products generated. This type of group assessment centers on the intellectual results of the learning process rather than the process itself [Kumar et al., 10], overlooking group dynamics, interaction, and the technology-mediated processes. To address the complex interactions of small groups, some process-oriented assessment endeavors require integration across multiple coding schemes and sources of data [Hmelo-Silver et al., 11]. These assessment efforts often rely heavily on conventional methods such as content analysis, coding of observations, interaction analysis, etc., and are therefore very time-intensive. Moreover, a "coding and counting" approach omits important information about elements such as temporality and group interactions [Reimann, 09][Suthers, 03].

An automatic and process-oriented assessment model for small group learning would be ideal for relieving teachers' burden of assessment and providing insights into the collaborative process, In order to build such a model, we must overcome problems of high data dimensionality (too many factors affecting group learning and performance). For example, a large set of chosen variables can dramatically diminish both statistical and machine learning performance [Vanneschi and Poli, 12]. Feature selection can be employed to reduce data dimensionality, but the automatic processing of data generated by these environments creates a conceptually "blunt instrument" due to feature selection algorithms, statistical models and machine learning that is grounded in mathematical theories rather than theories of human behavior. Thus, it is more suitable to reduce data dimensionality by constructing variables according to established theories [Fancsali, 11] in educational situations where human judgment is key [Siemens and Baker, 2012]. Besides the feature selection approach, variable selection and construction are often based on ad-hoc guesswork, or significant experience in the educational field [Tair and El-Halees, 12]. A principled, theory-based method for constructing variables and contextualizing raw data automatically would provide a rich and likely more useful set of measures that could then be evaluated over many iterations of a particular socio-technical context; in this case a synchronous, small group based math learning environment, Virtual Math Teams with Geogebra (VMTwG).

Machine learning methods that are widely applied in Learning Analytics demand hundreds to thousands of rows of data to produce reliable models [Kantardzic, 11]. However, educational datasets are often too small to directly apply machine learning methods [Hämäläinen and Vinni, 06]. Online learning supports increased class sizes, and it is possible to accumulate several years' worth of data**.** Excluding Massive Open Online Courses (MOOCs), technology mediated small group learning in many settings generates too little data for conventional machine learning methods.

To address the gap between learning analytics aims and "thin data", this study proposes a method for constructing an automated assessment model for small groups as a whole using a small dataset. We use the two-fold meaning of "Small." This work is informed by complexity theory to produce a set of simple rules to characterize the interactional dynamics among group members in an online collaborative environment. We then operationalize these rules to produce a series of temporally-embodied variables (features)**.** This theory-based feature construction method in combination with a feature selection algorithm effectively reduces data dimensionality, which contextualizes these features in a semantic background and in turn reduces the complexity and "noise" of data analysis. This calming of interpretive noise enables teachers to focus on aggregated signals hidden in the noise and to provide concrete feedback to groups. We then build a Tree-Augmented Naïve-Bayes (TAN) classifier was coded to build the assessment model. Finally, a web-based tool was discussed that not only showed the small group learning performance result but also provided actionable information for teachers to support small group learning.

This paper is organized as follows: Section 2 describes related literature. Section 3 provides background on complex systems theory and small group learning. Section 4 shows the research context and data format. Section 5 describes methodology. Section 6 presents experimental results and the web-based tool. Section 7 discusses results. Section 8 summarizes this study, pointing out limitations and future search directions.

2 Background Studies

After reviewing 186 articles, [Gress et al., 10] stated that assessment for group learning remains by and large summative in nature. Typically, evaluation of groups in CSCL is carried out by means of examination of the final product of collaboration. For example, [Zhu, 12] studied student knowledge construction in CSCL from a cross-cultural perspective, scoring each class group according to the quality of the group product for each task, and assigning a final group score based on an average score of all assignments. Similarly, [Kapur and Kinzer, 09] looked at productive failure in collaborative learning by comparing groups solving ill-structured and well-structured problems. The performance of each group was measured by the quality of the solution produced according to a holistic rubric. While ‘real-time,’ ‘process-orientation’ and ‘context’ are critical characteristics of group learning [Reimann, 09], individual, summative evaluations are usually administered after collaboration. There is a mismatch, then, between the pedagogical approach of small group learning, and the assessment of individuals, which ultimately risks undermining the foundation of small group learning approaches.

To address the “small group learning, individual assessment” paradox, both qualitative and quantitative approaches are applied to process-oriented group assessment. To analyze the complexity of collaborations, many researchers use qualitative methods such as interactional analysis, discourse analysis and conversation analysis. For instance, [Safin et al., 10] assess the success in collaboration between two teams in the domain of architectural design. All collaborative activities are video recorded and analyzed using an in-house protocol. [Arnold et al., 01] studies collaboration among foreign language graduate students using wikis. Through content analysis of the dialogue, they evaluate student performance in planning, contributing, seeking input, reflection, and social interaction to determine whether collaborations are successful. These qualitative coding-and-counting methods provide difficult to gather accounts of the complex process of group interaction [Hmelo-Silver et al., 11]. However, these methods are rather time-consuming, and are difficult, if not impossible, for teachers to implement. Moreover, post-collaboration analysis eliminates opportunities for real-time feedback [Kumar et al., 2010].

To overcome limitations in scalability found in deep, qualitative methods, another branch of literature attempts to perform small group assessment using quantitative methods. These methods attempt to assess complex collaborative processes either by building ad-hoc measures or by using intricate coding schemes that quantify categories of actions or utterances (e.g. [Mirriahi et al., 13], [Hmelo-Silver, 03]). Quantitative content analysis has been employed widely to characterize group discussion by coding and counting the frequencies of different aspects of discourse (e.g. [Kapur et al., 11], [Stijobs et al., 06]). Although results are ultimately quantitative, these methods require time-intensive qualitative analysis first, rendering them unsuitable for assessment automation.

Other quantitative approaches for assessing group collaboration include experimental methods (e.g. [Xing et al., 14], [Suthers et al., 03]), social network analysis (e.g. [Goggins et al., 09]) and multilevel analysis (e.g. [Cress, 08]). However, these quantitative measures and methods, including quantitative content analysis, are unable to automatically conceptualize the temporal evolution of group collaboration and problem-solving processes [Chen and Resendes, 14] [Reimann, 09]. Also, the ad-hoc nature of measure selection and the construction of the quantitative analysis skirts systematic measurement of group interaction and collaboration [Xing et al., 14 & 14].

Most qualitatively focused assessment-related research takes place at the scale of 20 – 60 groups and does not incorporate individual interactions between participants (e.g. [Xing et al., 14], [Kapur et al., 11]). In contrast, learning analytics, which incorporates machine learning algorithms, will work with every interaction, often analysing in excess of tens of thousands of data points. Algorithms used in learning analytics research include Bayesian networks, decision trees or fuzzy logic (e.g. [Ferguson and Shum, 11], [Coffrin, 14]). According to computational learning theory, data size in machine learning problems has a significant influence on machine learning performance [Hoyle, 08]. Therefore, without careful selection of appropriate algorithms, learning analytics emerging in the context of "Big Data" may not migrate well to these smaller datasets.

This study aims to provide guidance for algorithm manipulation with respect to educational datasets, by employing systems theory to model the interactions and dynamics within small groups in a synchronous, math learning context, VMTwG**.** Current application of complex systems in the learning sciences is relatively sparse, but gaining momentum [Jacobson and Wilensky, 06]. A significant portion of current studies that frame learning sciences through a complex systems lens focus on curriculum, teaching and knowledge transfer (e.g. [Hmelo-Silver et al., 00], [Goldstone and Wilensky, 08]). Our aim here is to illustrate one way in which complex systems theory may contribute theoretical conceptions and methodologies that could potentially expand the toolkit for learning analytics and of learning sciences as a whole [Kapur, 11]. Specifically, the present study explores the potential to better understand and automatically assess technology mediated small group learning.

3 Small Group Learning as Complex Systems

Small group learning is complex in the sense that each small group has a unique history, motivations, purpose and background, as well as its practices for making sense of how to collaboratively learn through technology. Groups and their members have differing styles and interests. These groups also exist in relation to their environment and are affected by that environment [Mennin, 07], evolving as they optimize their learning and functionality in a system filled with conflicts, constraints and opportunities [Kauffman, 95].

According to [Arrow et al., 00], "small groups are complex systems that interact with smaller systems (group members) embedded within them and the larger systems (organizations, classes, society) within which they are embedded. Groups have fuzzy boundaries that both distinguish them from and connect them to their members and their embedding contexts." From this perspective, each group member is also a complex system with many interactions among diverse agents [Mennin, 07] and similarly is embedded in and influenced by a particular context. Therefore, small groups act as complex systems nested in other complex systems at different levels to form a multidimensional web of interactions [Capra, 96]. These interactions are usually nonlinear and highly dynamic, which gives rise to emergent properties of the system as a whole [Kauffman, 95]. Complex systems research focuses on studying relationships and interactions among agents instead of agents themselves because it is the interactions that are most important for the emergence of learning [Holland, 98].

Since complex systems research emphasizes how complexity as a whole relates to the complexity of its constituent parts, the concept of emergent behavior, or how macro-level behaviors emerge from micro-level interactions of individual agents, is fundamental in gathering an understanding of this relationship [Kapur et al., 11][Bar-

Yam, 97] [Holland, 98] [Kauffman, 95]. In a small group learning context, this emergent behavior is essential to learning [Mennin, 07] [Arrow et al., 00]. To facilitate this understanding, we must first understand two principles of complexity: emergent simplicity and emergent complexity [Kauffman, 95]**.** The research of [Kapur et al., 11] provides an example. Consider the brain as a collection of neurons. Individual neurons are quite complex, yet they exhibit simple binary behavior in their synaptic interactions. This kind of behavior is referred as emergent simplicity, where complexity at an individual level results in simplicity at a collective level. Further, these simple (binary) synaptic interactions among neurons can produce complex brain "behaviors" e.g. learning, memory, nonexistent at the level of individual neurons. This type of emergent behavior is considered to be emergent complexity, where simplicity at the individual level generates complexity at the collective level. Therefore, it is not always necessary to seek complex models and explanations for complex behavior.

Complex collective emergence, operationalized in the current study as group learning, can be modeled from the "bottom up" using simple and minimal information e.g. functions, rules [Kapur et al., 11]. We explore the assessment of small group learning with a focus on simple, theoretically informed, rule-guided and low-level interactions that contribute to higher levels of complexity, which we call learning emergence. This focus on the rule-guided interactions in small group learning reduces the number of data factors we need to consider.

4 Research Context

This study focuses on the interactions and learning behaviors of small groups participating in a structured geometry curriculum within a synchronous math discourse tool. Specifically, we focus on one module of a 2013 course designed for Virtual Math Teams employing Geogebra (VMTwG) software (Figure 1). The module is called "Exploring Triangles", and aims to explore the built-in dependencies and relationships of the different dynamic-math triangles. To assist students in accomplishing this goal, this module is further divided into four sequential parts: "Equilateral," "Relationships," "Where's Waldo?" and "Exploring", each with more granular objectives and instructions to guide study. The full curriculum currently contains a total of 21 modules, and is available on the project website (http://vmt.mathforum.org).

The module analyzed in this paper includes groups consisting of three to five members, with a total of 28 groups involved in the study. Based on content analysis and clustering, groups were rated on performance using a 1-9 scale with 9 as highest performance. We further transformed this into a categorical data set, with 8-9 representing High, 5-7 representing Medium and < 5 representing Low performance. Results indicated 7 groups belonging to the High performance category, 12 to Medium-level performance, and 9 to Low performance.

Figure 1 provides a guide to understanding cognitive learning discourse in VMT. Section A of Figure 1, the VMT replayer bar, reveals the time dimension. Each action within VMTwG is logged with a timestamp. Section B is the chat window, where text is entered into chat. Future analytics in this project will focus on the analysis of the text in chat windows in concert with GeoGebra gestures. Sections C and D are related to Geogebra actions. C is the “Take Control” button, which gives an individual user control of the VMTwG environment. Section D is the GeoGebra window itself. Here, students work to create an equilateral triangle within an equilateral triangle using multiple approaches.

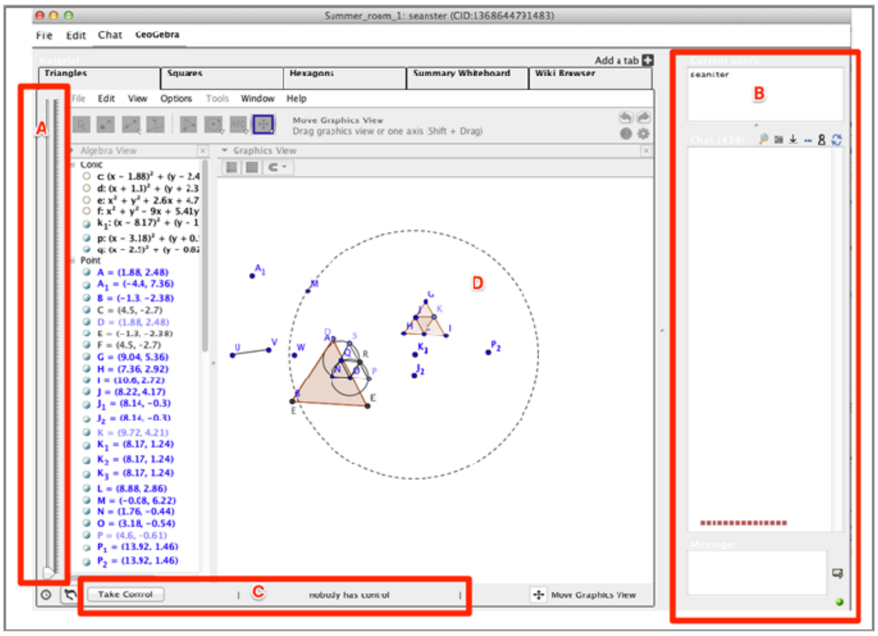

Figure 1: VMTwG, an analytical tool for collaborative math discourse

All log data was collected in .txt format and centers on specific event types VMT: ‘Awareness’, ‘Geogebra’, ‘System’, ‘Chat’, and ‘WhiteBoard’ (Wb). The ‘Chat’ event-type logs all messages in the group. ‘Awareness’ records the actions of erasing chat messages when the chat bar is full. ‘Geogebra’ logs information on how students virtually construct a geometry artifact (e.g. add a point, update a segment). The ‘System’ event-type records information on how VMT is accessed. For example, a student joining a virtual room, or viewing different tabs created by students or teachers. ‘Wb’ logs more specific actions in the whiteboard area such as the resizing of objects. For every event type, we have logs of which actions (Add a point, Send a chat, or Creating a text box, etc.) the student makes under what subjects and topics, as well as the initiator (Source) and receiver (Target) of those messages. In addition, the environment logs information about the time during which an action takes place and the virtual room (group) in which the event occurs. Figure 2 shows a sample of original log data.

Community	Subject	Topic	Room	Source	Target	Time	Finish Time	Event Type	Event
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	Argy123	Argy123	0:00:01	44:46.1	system	2013-02-24 09:44:46.062 - Argy123 -> Now viewing tab Relationships
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	Argy123	0:00:01	44:47.6	Geogebra:I	2013-02-24 09:44:47.562 - mathminion1 -> tool changed to Move
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	mathminion1	0:00:03	44:50.7	Geogebra:I	2013-02-24 09:44:50.681 - mathminion1 -> updated group of objects A,B
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	mathminion1	0:00:03	44:54.5	Geogebra:I	2013-02-24 09:44:54.487 - mathminion1 -> updated Point F
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	mathminion1	0:00:02	44:56.6	Geogebra:I	2013-02-24 09:44:56.619 - mathminion1 -> updated Point F
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	mathminion1	0:00:05	45:01.7	Geogebra:I	2013-02-24 09:45:01.701 - mathminion1 -> updated Point F
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	mathminion1	0:00:08	45:10.0	Geogebra:I	2013-02-24 09:45:10.042 - mathminion1 -> updated group of objects A,B
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	mathminion1	0:00:10	45:21.0	Geogebra:I	2013-02-24 09:45:20.983 - mathminion1 -> updated Point F
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	mathminion1	0:00:28	45:52.0	chat	2013-02-24 09:45:51.999 - mathminion1 -> yea it is definetely not the midpoir
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	mathminion1	0:00:40	46:33.2	chat	2013-02-24 09:46:33.179 - mathminion1 -> i also noticed that only point f was
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	mathminion1	0:0:-11	46:21.8	Geogebra:I	2013-02-24 09:46:21.799 - mathminion1 -> tool changed to Move Graphics Vie
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	mathminion1	0:00:13	46:35.2	Geogebra:I	2013-02-24 09:46:35.154 - mathminion1 -> tool changed to Move
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	mathminion1	0:00:00	46:36.1	Geogebra:I	2013-02-24 09:46:36.053 - mathminion1 -> tool changed to Move Graphics Vie
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	CT.mates	mathminion1	0:00:20	46:59.6	chat	2013-02-24 09:46:59.618 - CT.mates -> yes, i agree, point F is free
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	CT.mates	0:00:21	47:06.5	chat	2013-02-24 09:47:06.495 - mathminion1 -> is that because that those points ar
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	Argy123	mathminion1	0:00:55	47:51.8	awareness	2013-02-24 09:47:51.805 - Argy123 -> [fully erased the chat message]
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	CT.mates	Argy123	0:00:02	47:46.0	awareness	2013-02-24 09:47:46.011 - CT.mates -> [fully erased the chat message]
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	CT.mates	CT.mates	0:00:01	47:48.3	awareness	2013-02-24 09:47:48.295 - CT.mates -> [fully erased the chat message]
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	CT.mates	CT.mates	0:00:12	48:01.2	awareness	2013-02-24 09:48:01.164 - CT.mates -> [fully erased the chat message]
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	Argy123	CT.mates	0:00:04	47:56.9	chat	2013-02-24 09:47:56.902 - Argy123 -> yes exactly
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	Argy123	Argy123	0:00:24	48:24.3	awareness	2013-02-24 09:48:24.289 - Argy123 -> [fully erased the chat message]
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	CT.mates	Argy123	0:00:15	48:17.0	chat	2013-02-24 09:48:17.021 - CT.mates -> Points A and B can change the size of th
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	CT.mates	0:00:14	48:25.9	awareness	2013-02-24 09:48:25.874 - mathminion1 -> [fully erased the chat message]
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	CT.mates	mathminion1	0:00:09	48:27.9	chat	2013-02-24 09:48:27.920 - CT.mates -> points C, E, and D all depend on the size
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	Argy123	CT.mates	0:00:15	48:40.7	chat	2013-02-24 09:48:40.750 - Argy123 -> correct too haha alright ill take control n
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	Argy123	0:00:14	48:40.6	chat	2013-02-24 09:48:40.638 - mathminion1 -> exactly ct....good observation
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	mathminion1	0:00:03	49:06.2	awareness	2013-02-24 09:49:06.203 - mathminion1 -> [fully erased the chat message]
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	Argy123	mathminion1	0:0:-1	49:04.4	Geogebra:I	2013-02-24 09:49:04.388 - Argy123 -> tool changed to Move
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	Argy123	0:00:05	49:12.1	awareness	2013-02-24 09:49:12.139 - mathminion1 -> [fully erased the chat message]
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	mathminion1	mathminion1	0:00:42	49:55.3	chat	2013-02-24 09:49:55.323 - mathminion1 -> we kinda went out of order bc we a
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	Argy123	mathminion1	0:0:-12	49:43.1	Geogebra:I	2013-02-24 09:49:43.134 - Argy123 -> tool changed to Move
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	Argy123	Argy123	0:00:02	49:45.7	Geogebra:I	2013-02-24 09:49:45.740 - Argy123 -> tool changed to Move Graphics View
Spring 2013	Dynamic Geometry	Topic 02	Colon_Group_1	Argy123	Argy123	0:00:24	50:27.6	awareness	2013-02-24 09:50:27.611 - Argy123 -> [fully erased the chat message]

Figure 2: Sample logs from VMT

5 Methodology

5.1 Simple Rules in Small Group Learning

The notion of emergent simplicity hypothesizes that simple rules can govern the low-level interactions in small groups [Kapur, 11]. Though individual group members in are complex with different backgrounds, prior knowledge and experience, the impact of their interaction with other group members is guided by a set of simple rules. In the context of this study, group members are considered to be agents that interact with each other to accomplish the learning goal of exploring relationships between different dynamic-math triangles and constructing an isosceles triangle in VMT. Therefore, local interactions among group members can be conceived as goal-directed analysis performed by operators in problem space [Newell, 72]. As a result, a set of rules naturally emerges. According to [Kapur, 11], each interaction has an impact that:

- Moves the group towards the goal, or
- Moves the group away from the goal, or
- Maintains the status-quo.

That is, some interactions between groups members aid in developing comprehension of the relationships between different dynamic triangles and contribute to the construction of isosceles triangle, some diverge away from the group comprehending the dynamic triangles and its development, and other interactions have little or no effect on the comprehension and triangle development. Further, these simple rules governing local interactions among individual agents give rise to group learning viewed as emergent complexity.

5.2 Rule Operationalization

5.2.1 Data Formation

To assess small group performance, we can measure interactions that cause a group to move closer to or farther from solving a problem, or in neither direction. Measures constructed to reflect these interactions have a rule-semantic background, allowing teachers to understand the rationale behind the assessment model. Specifically, we simplify the data in a manner that is coherent with both theory and our observations about learning in VMTwG, gathered from years of research. For chat log data, we first remove all stop words, leaving a series of keywords and symbols. Then based on previous experience, question marks, exclamation marks, and words in all capitals are also selected as indicators of influence of interaction on group movement. The other four event types ('Awareness', 'Geogebra', 'System', 'Wb') each contain a definite number of actions. Therefore, we put actions of these four event types into a union ($\cup$) set. Because temporal evolution of the group collaborative process is key to developing an understanding of the interaction and collaboration within both complex systems and CSCL [Reimann, 09], we also explore the incorporation of time factors into our measures. Again, the assessment model considers the small group as a unit of interest instead of a single student and all the selected actions or words performed by any student member in the group are incorporated in the feature space, and then used for feature selection and construction and further model building.

Let G_{m} notate any group in this course, $m=1,2,\cdots,M$. In this context, M is 28, indicating 28 groups in total. N represents the number of students in that particular Group G_{m}, $n=1,2,\cdots,N$. T denotes event type, ['Awareness', 'Chat', 'Geogebra', 'System', 'Wb']. A represents tool actions ('Awareness', 'Geogebra', 'System', 'Wb'), key words, or key symbols. A_{ij} denotes the frequency of use of a specific tool/word/symbol j in event type i, where $i \in [1,5]$, $j \in [1,\mathrm{J}]$, $i \in Z^+$, $j \in Z^+$. k indicates the specific position in the sequence of the topic to put a sequential and temporal background behind each entity, where for this specific topic, $k \in [1,4]$. Therefore, A_{ijk} represents the frequency of a tool/word/symbol in one of the four sequential parts of the topic. Particularly, Stemming technique [Runeson, 07] is used for the keyword search. X indicates the constructed variable/measure and x is the value of that measure. The dataset for a single group is a fourfold structure as shown in Figure 3.

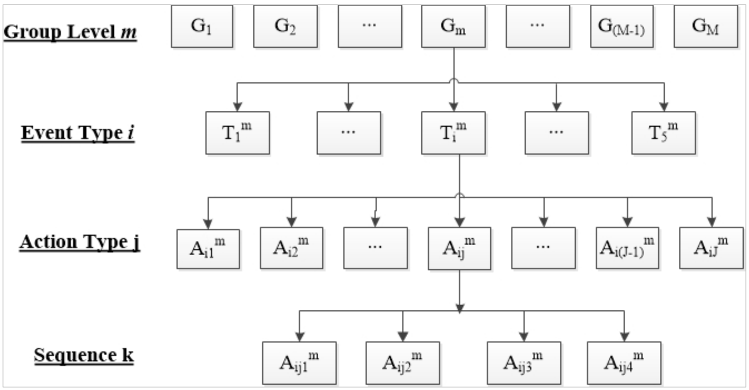

Figure 3: Four-fold data structure

5.2.2 Measure Construction

After processing the chat and tool action data into a bag of words, 30 types of actions and words are selected to measure the influence of interactions as guided by the three rules. Specifically, for interactions that move group learning towards or away from their goal, T_2 and T_3 'Chat' and 'Geogebra' are chosen, as these two action types indicate concrete exploration and the main task of the development of triangles. Further, actions A_{ij} are chosen to represent interactions that contribute to the forward movement based on the curriculum, where $i \in \{2,3\}$, $j \in [1, J]$.

For each group, in order to embody the temporal effect in this model, the frequency of each action A_{ijk} is divided into four consecutive parts. Specifically, the A_{ijk} is represented as a four-dimensional set for group G_m : $\left[A_{ij1}, A_{ij2}, A_{ij3}, A_{ij4}\right]$. However, actions occurring in different parts of these sequential sub-topics have different levels of influence. Thus, we use a weight function to indicate this difference. Specifically, the golden ratio φ, a concept also called division in extreme and mean ratio by Euclid [Smith, 53] is applied in this study. The golden ratio arises from dividing a segment so that the ratio of the whole segment to its larger part is equal to the ratio of the larger part to the smaller part as shown in Figure 4, expressed algebraically, $\frac{1}{b} = \frac{b}{1-b}$, where $(X > 0.5)$. φ is the answer to the equation $b = \frac{(-1+\sqrt{5})}{2} \approx 1.618033...$.

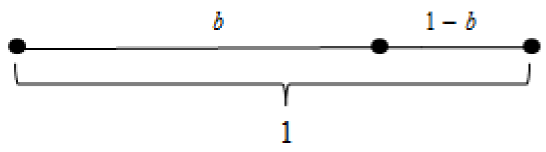

Figure 4: Line Segment in Golden Ratio

The golden ratio is a heuristic used to divide and analyze natural objects and man-made systems [Smith, 53]. In the context of our study, the golden ratio φ is multiplied by the frequency of actions that are required or important to the completion of a particular sequential task, and the sequential parts that are less critical are multiplied by $\varphi-1$. Similarly, among the collection of actions, the golden ratio φ is further multiplied by A_{ijk} because some actions have more prominent interaction influence than others. In sum, each action or keyword in the particular section of a topic is represented as $\varphi^2 A_{ijk}$, $(\varphi-1)^2 A_{ijk}$, or $\varphi(\varphi-1)A_{ijk}$ depending on its degree of influence over the interaction and on group movement over time.

On the other hand, each action or word identified should have one computed value x as the final variable X. Therefore, we multiply the four sequential elements to produce a value calculation. Since we employ *union* ($\cup$) rather than *intersection* ($\cap$), there is a good possibility that some groups may receive a value of 0 for the frequency of the sequential parts simply because the group may never use that tool or symbol. To avoid this phenomenon, e is used as a base for the final calculation of a variable when the instance is 0. Multiplication is applied in this calculation to improve the influence of positive movement during the final assessment of the small group learning performance. Also because the number of group members is different from group to group, an average function is used in our formula to leverage the difference. In sum, each action x for the interaction that moves the group towards the goal is expressed as:

$$X = \begin{cases} \prod_{k=1}^{4} \frac{rA_{ijk}}{N}, \forall A_{ijk} \neq 0,\ r \in \left\{\varphi^2, (\varphi-1)^2, \varphi(\varphi-1)\right\} \\ e^{A_{ijs}} \prod_{k=1}^{s-1} \frac{rA_{ijk}}{N} \prod_{k=s+1}^{4} \frac{rA_{ijk}}{N},\ \exists A_{ijk} = 0,\ s \in \{1,2,3,4\},\ r \in \left\{\varphi^2, (\varphi-1)^2, \varphi(\varphi-1)\right\} \end{cases} \tag{1}$$

Actions in T_2, which move away from problem solving to other directions in that there is less need for these actions to complete the current task, have development procedures similar to those that move the group towards their goal. The modification that has been employed is to use a $\log$ function to dampen the influence of those actions on assessment model construction. Similarly, the 0 frequency situation is also taken into account.

$$X = \begin{cases} \prod_{k=1}^{4} \frac{\log(rA_{ijk})}{N}, \forall A_{ijk} \neq 0,\ r \in \left\{\varphi^2, (\varphi-1)^2, \varphi(\varphi-1)\right\} \\ \log(A_{ijs}+10) \prod_{k=1}^{s-1} \frac{\log(rA_{ijk})}{N} \prod_{k=s+1}^{4} \frac{\log(rA_{ijk})}{N},\ \exists A_{ijk} = 0,\ s \in \{1,2,3,4\}, r \in \left\{\varphi^2, (\varphi-1)^2, \varphi(\varphi-1)\right\} \end{cases} \quad (2)$$

Drawing inferences from key words and key symbols is different than drawing inferences from tool use. When a tool is used, we can make a deterministic judgment about whether that tool is relevant to a particular topic. With words, our determinations of the utility of the words for problem solving and assessment must be based on probabilities. Also, groups may use vocabularies and signs that are widely different from one another. Therefore, we limit the number of key words or symbols in the assessment model, focusing only on those that are most useful and commonly used by the group while completing a task. We do not assign different weights to reduce the influence of those words. A summation function is used instead of multiplication to reduce this influence. As a result, for interactions that move the group forward or backward, variables from the key words are presented as:

$$X = \frac{\sum_{k=1}^{4} A_{ijk}}{N} \quad (3)$$

Besides some keywords which are identified as neural functions, T_1 ('Awareness'), T_4 ('System'), and T_5 ('Wb') are seen as actions that do not directly move the group towards or away from problem-solving. Therefore, the influence of those interactions is the same with definitions of keywords using a simple summation, such as (3). In sum, there is not much difference between which group members perform the action or use the key word because the group is considered as a whole. However, the frequency and amount of time it takes to perform a particular action or use a particular keyword from any member of the group may influence the group performance significantly because different weights are assigned to different actions/keywords performed at different times.

5.3 TAN Model

The aim of the assessment model is based on the previously discussed aim of exposing temporally embodied features to automatically estimate the performance of each team (high, medium, or low). This assignment focuses on process rather than simply evaluating a final solution or product. In this VMT case, and in the discipline of education in general, the feature space contains too many attributes to build an accurate classifier model. Even if machine learning algorithms could be incorporated under these conditions, it is still extremely useful to select only the most influential factors; those best able to distinguish among the different classes. Therefore, a correlation-based feature subset selection method [Hall, 99] is used to reduce the dataset domains to 9 features only: triangle, angle circle, intersect, segment, question mark, drag, viewing tab, and perpendicular line. As a result, each group is represented as a 9 dimensional set. Furthermore, these groups and their associated performance category (High, Medium, Low) constitute the 28 lines of data that are used to train an assessment model, where we use the 9 features to predict the categorical value of the group.

In educational contexts, which often produce less than 100 lines of data, we decided to exclude commonly used techniques such as nearest neighbor classifiers, neural networks and variations of decision trees, as they usually require much more data in order to function accurately. Regression, especially linear or logistic regression, is a more optimal candidate for a "small data" assessment model. The main limitations associated with these regression techniques are their sensitivity to outliers and collinearity. However, educational data almost always contain exceptional students and groups (outliers) who can achieve a good performance with little effort or fail without any sensible reason. Collinearity indicates strong linear dependencies among those variables or features. It is not uncommon that educational factors are more or less related to each other. It is difficult to provide any exact threshold value for correlations to be harmless. Empirical studies [Hämäläinen and Vinni, 06] show that correlation coefficient r > 0.7 affects model performance significantly, while weaker correlations usually do not have any significant impact. Support vector machines (SVM) are another good choice for the classification model of small educational data in that the classification model depends only on some data points [Schuldt et al., 04]. However, SVM is a "black-box" model, which does not show a tangible relationship among variables. This is not ideal from the perspectives of application and comprehension, which require the model to be transparent.

Alternately, a Bayesian network model is an attractive and understandable option within the educational domain, where an element of uncertainty is always involved. However, general Bayesian networks are too complex for small data sets, and as a consequence, the developed model over fits easily. The Naïve Bayes model, due to its simplicity, avoids this problem to some extent. In the present study specifically, Group G is represented by a tuple of 9 selected feature values $\{x_1, x_1, x_1, ..., x_n\}$, where x_i is the value of attribute X_i. Let C denote performance level (high, medium, low) and use c as the value of C. According to Bayes rule, the probability of an example $G = \{x_1, x_2, x_3, ..., x_n\}$ in class c is

$$p(c \mid G) = \frac{p(G \mid c)\, p(c)}{p(G)} \tag{4}$$

Assuming that all attributes are independent for the value of the class variable, then

$$p(G \mid c) = p(x_1, x_2, x_3, ..., x_n) = \prod_{i=1}^{n} p(x_i \mid c) \tag{5}$$

Hence, Naïve Bayes classifies group performance by selecting

$$\arg\max_c \left(\hat{p}(c) \prod_{i=1}^{n} \hat{p}(x_i \mid c) \right) \tag{6}$$

where $\hat{p}(c)$ and $\hat{p}(x_i \mid c)$ are estimations of the probabilities derived from the frequency of their arguments in the training data respectively. This is the basic reasoning behind the Naïve Bayes model. Compared to a Bayesian network, this model only has two layers, the class variable c in the root node, and all the other variables X_i in the leaf nodes, as shown in Figure 5. The Naïve Bayes model assumes that the all leaf nodes are conditionally independent, given the class value. Even though this assumption is often unrealistic, especially in the educational sphere, in practice Naïve Bayes model has worked well and has in some cases outperformed more sophisticated models such as decision trees and general Bayesian networks, especially in small datasets (e.g. [Domingos, 97], [Friedman, 97], [Hämäläinen, 06]).

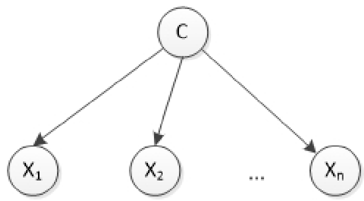

Figure 5: An Example of Naïve Bayes Model

Notwithstanding the fact that some violations of the variable independence assumption for Naïve Bayes model do not matter, many do. Therefore, substantial effort has been invested to remove any violations to variable independence, while maintaining a desirable level of simplicity (e.g. [Friedman et al., 97], [Webb, 01]). Of these techniques, TAN has demonstrated remarkable accuracy [Keogh 99][Friedman, 97]. TAN enlarges the Naïve Bayes model by allowing additional dependencies among features. That is, the TAN model allows every attribute x_i to depend on the class and on another attribute $p(x_i)$, named the parent of x_i as shown in Figure 8. Thus, G is classified by selecting:

$$\underset{c}{\operatorname{argmax}}\left(\hat{p}(\mathrm{c})\prod_{i=1}^{n}\hat{p}(\mathrm{x}_i \mid \mathrm{c}, p(\mathrm{x}_i))\right) \tag{7}$$

TAN uses conditional mutual information to select the parent function $p(.)$, which is developed at the time of training. This model is often a good compromise between a Naïve Bayes model and a general Bayesian network model. The general procedure for TAN includes:

- Taking the training data set and $X \setminus (c)$ as input.
- Computing the conditional mutual information $I(x_i, x_j \mid \{c\})$ between each pair of attributes, $i \neq j$.
- Developing a complete undirected maximum weighted spanning graph in which nodes are attributes $x_1, x_1, x_1, \ldots, x_n$ as shown in Figure 6.
- Transforming the developed undirected graph into a directed graph by selecting a root attribute and setting the all edges to be directed outward from it (Figure 7).

- Learning the parameters and outputting the TAN model (Figure 8) by adding a node labeled by C and adding an edge from C to each X_i.

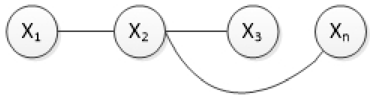

Figure 6: An Example of the undirected tree

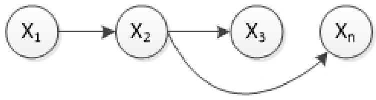

Figure 7: The directed Tree by Selecting X_2 as the root node

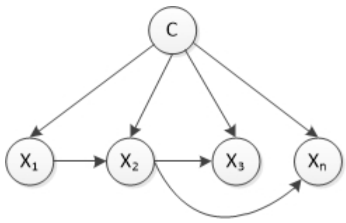

Figure 8: An Example of TAN Model

To strengthen our argument for the TAN model, we executed various models e.g. C4.5, Logistic Regression, Neural Network (Perceptron), Naïve Bayes to benchmark the proposed model for the educational dataset. All algorithms are evaluated using tenfold cross-validation, which is helpful in limiting problems such as over-fitting and increases the likelihood that the generated model can be generalized to an independent dataset.

6 Results

6.1 Complex Systems based Measures

Small groups as complex systems involve many factors that affect learning and performance. It is impossible to take all factors into account when building an assessment model, especially considering the limited data usually available. Highly dimensional datasets become even more difficult for a machine algorithm to learn. The grounding of this study in complex systems theory, specifically applied emergent complexity and emergent simplicity concepts addresses this issue. As a result, 30 features were chosen. To further reduce the data dimensionality, correlation-based feature subset selection (Hall 98) was conducted and 9 features were selected and constructed to model the group dynamics and interaction as in Table 1.

Rule	Type	Feature	Group1	Group2	Group3	Group4	…
Move Forward	'Chat'	Triangle	4.67	4.00	11.33	2.33	…
		Angle	0.50	0.80	0.33	4.60	…
	'Tool'	Circle	5.24	22.25	2.62	13.09	…
		Intersect	1.75	1.96	0.87	1.75	…
		Segment	10.47	5.45	11.31	3.49	…
Maintain Status Quo	'Chat'	Question Mark	11.75	9.50	12.33	9.33	…
		Drag	0.75	1.40	0.20	0.67	…
	'System'	Viewing tab	8.80	10.75	11.25	11.67	…
Move Away	'Tool'	Perpendi-cular	-0.14	1.00	1.00	0.04	…

Table 1: Sample Group Modeling Features Based on Complex Systems Theory

In Table 1, the first three columns show the Rule, Type and Feature and the remaining columns present values computed for each group. Therefore, each row provides the specific feature name, its type and whether it moves the group towards or away from its goal, or initiates no movement. In addition, it tells you the specific values of a feature across groups computed using the equations from the Measure Construction section. Each group is represented by a 9-dimensional vector. To some degree, each feature is an aggregation of frequency from all group members and multiplied by different weights, based on different time of usage and importance. These constructed features include a coarse time-dimension derived from weighting the features based on their significance during the different stages of problem solving progression. Compared with other ad-hoc guesswork or solely feature selection and construction methods, it is expected that this theory-based methodology will enhance the accuracy and rhetorical power of the modeling of group dynamics in a technology-mediated space. This accuracy is reflected in the next section from the TAN model application to the automatic assessment of small group learning. Furthermore, each feature selected is semantically grounded as either advancing the group's problem solving or maintaining status quo. This semantic information would enable teachers to have a rationale for understanding the assessment model and providing feedback, scaffolding and intervention accordingly, which is discussed further in the web-based tool section. In other words, how we apply our analytic approach to discern findings is designed explicitly to facilitate teacher engagement in the interpretation of analytics.

The selected features are reasonable and intuitively important for groups' problem solving. In order to build an isosceles triangle in the VMT context, Angles, Circles, and Segments are absolutely necessary. Groups working with these tools more often have a better chance of accomplishing the goal (learning). For the actions and words under the 'Maintain Status Quo' category, even though they do not directly contribute to move the group towards building the isosceles triangle or exploring the dynamic triangle relationships, these actions are also important indicators of group dynamics and collaboration. When students use the Perpendicular Line function, it shows students are off track because this particular topic does not require the Perpendicular Line function. Therefore, it is expected that these features informed by complex-systems are good indicators of whether small group learning is taking place in VMT.

6.2 Model Performance

Algorithm / Result	TAN	Naïve Bayes	C4.5	Logistic Regression	Neural Network
F-Measure	93.4%	89.5%	78.0%	78.7%	52.7%
Accuracy	95.8%	90.0%	78.8%	80.6%	55.5%

Table 2 Assessment of Model Performance: The F-measure acts as an overall evaluation of model performance. Compared with other more complex models such as tree-based (C4.5) and Neural Networks (Perceptron), the TAN model has the best performance in both F-Measure (93.4%) and Accuracy (95.8%) for this small data set,

The objective of this paper is to provide a framework for automating the assessment process for small group learning and performance, using indications from the collaboration process rather than the final product or solution of the group. The 9 features representing each group with manually assigned performance labels for that group become the dataset and input for the TAN and other algorithms to build the assessment model. In our context, there are 28 lines of data, each representing a single group as [Triangle, Angle, Circle, Intersect, Segment, Question Mark, Drag, Viewing Tab, Perpendicular, Performance Label], where Triangle, Angle, Circle, Intersect, Segment, Question Mark, Drag, Viewing Tab, Perpendicular are independent variables/features and are used to predict the dependent variable/class, the performance label of the group. The weighted results of the various models after performing 10-fold cross validation are listed in Table 2.

Table 2 presents both the accuracy and the F-measure, The F-measure acts as an overall evaluation of model performance. As shown in Table 2, compared with other more complex models such as tree-based (C4.5) and Neural Networks (Perceptron), the TAN model has the best performance in both F-Measure (93.4%) and Accuracy (95.8%) for this small data set, which is immediately followed by Naïve Bayes with F-Measure 89.5% and accuracy 90.0% respectively. The TAN model also has better performance in comparison with the Logistic Regression and Naïve Bayes model, which are more sensitive to collinearity. Therefore, based on the complex systems informed features, the relatively simpler TAN modeling technique is able to construct the most accurate and reliable assessment model for small group learning.

6.3 Web-based Tool for Time-Series Visualization

Enabling formative assessment provides actionable information for teachers. Enabling formative assessment through analytics is an important aim, and is sometimes referred to as "teaching analytics" [Ebner and Schön, 13][Taraghi et al., 14]. Simply assigning a High, Medium or Low labels automatically, based on the proposed methodology for a group performance, though helpful to teachers, does not convey actionable intelligence. We need to present information to teachers in a manner where it is easier to consume. In this way, our aim is to scaffold group progression during collaboration or provide in-time intervention for groups, especially for groups rated as low performance. The research community in learning analytics recognizes the significance of visualization design. In fact, visualization is a required component in the learning analytics cycle [Ferguson and Shum, 11]. In order to offer information for teachers to act upon in real time, we describe a web-based visualization tool to be used as shown in Figure 9.

This visualization tool is developed using a time-series perspective in order for instructors to understand how group collaboration evolves and varies over time. We focus on only one topic for the current study, but we expect this tool to work for any and all combinations of topics. Teachers may also choose to show only one group, or any combination of groups, using the check box in Figure 9 (a). Also, since the features selected matter the most for group performance, we choose to visualize 9 features (Figure 9(a)). We put words and actions in two different graphs and the appearance of these features are visualized over time. Moreover, since different topics may have different critical features, or teachers may want to choose key words or symbols which they think are important to group problem solving, this tool allows for teachers to input chosen keywords or tool actions into the system. Furthermore, when the teacher clicks on one of the dots symbolizing an occurrence, the visualization system automatically presents additional details, such as what a group member is doing at that particular moment. A screen capture of the VMT system is also obtained as shown in Figure 9(b). There are two colors representing keywords or tool usage. Blue indicates that the feature appears once either in works or tool actions, while the red represents a co-occurrence of features. For instance, when an instructor clicks on the red dot on Drag in Word part Figure 9 (a), he or she will see that the sentence contains both the Drag and Triangle features in Figure 9 (b). This is more likely to indicate that learning is occurring.

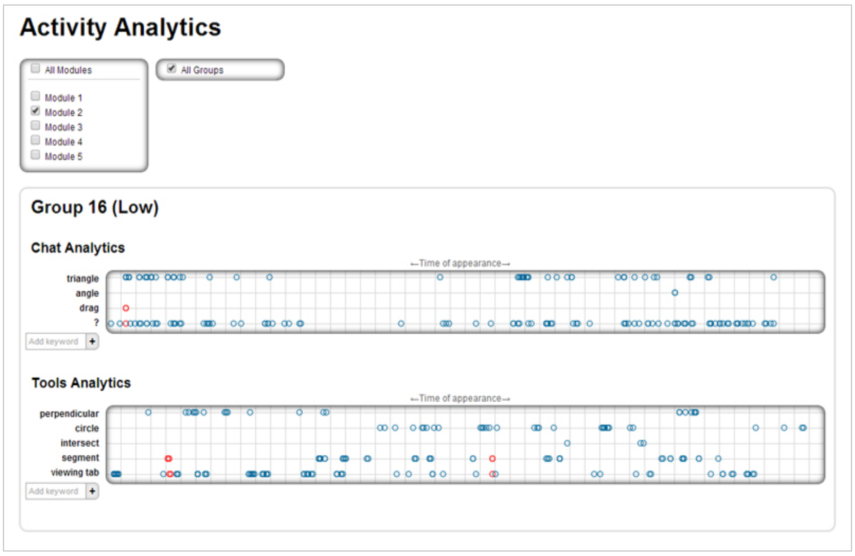

Figure 9(a): Activity Analytics

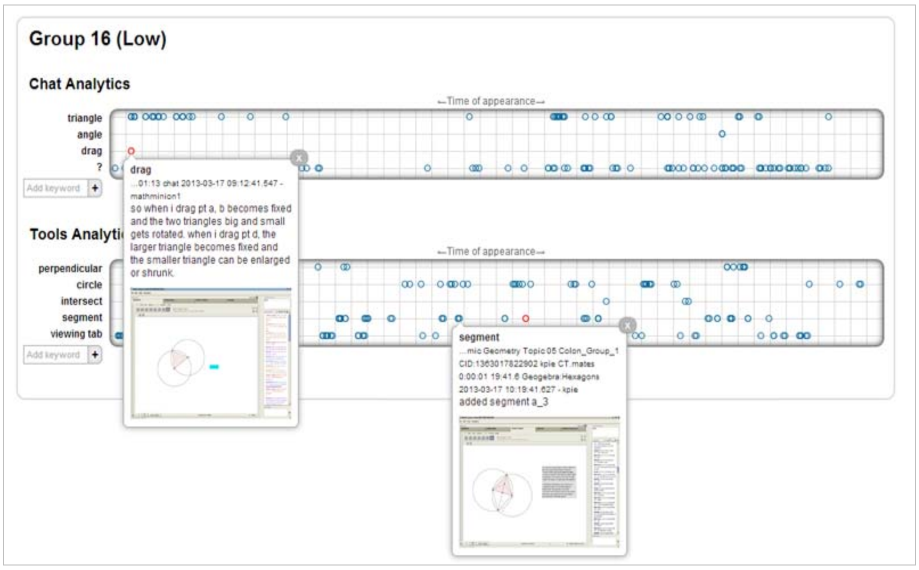

Figure 9(b): Activity Analytics, detailed view

To illustrate this tool usage, suppose students are working to build an isosceles triangle in Topic 2. Our system indicates that Group 16 is in the low performance area. When the teacher sees that this group spends a lot of time using the perpendicular line tool and that there are many questions (?) back and forth (Figure 9 (a)), they may deduce that the group is moving away from its goal (semantic background). Key actions that students employ to build a triangle, such as Circle, Intersect, and Segment (in the early part of the graph), occur few times in this group. The teacher sees the red Drag in chat and clicks on it (Figure 9(b)) and looks at the screenshot of the current environment, and realizes that this group is indeed off track from constructing the triangle. When the teacher slides over the question mark section and obtains the information that this group has a lot of questions regarding segment tool usage and circle tool usage they may then converse with the group and provide additional guidance concerning the functions of segment and circle tool (personalized help). If the personalized help offered by the teacher has been effective, the teacher may check on the group again and hover over the Segment tool to see that it is being used properly (Figure 9 (b)).

7 Discussion

Many studies of small group learning use summative assessment methods (e.g. final solution, grade) to measure performance. These assessment approaches usually overlook the collaborative process and the affordances of technology in contribution to group learning [Reimann, 09][Barab, 01]. Qualitative studies e.g. content analysis [Arnold, 01], and conversation analysis [Safin et al., 10] are typically time intensive, making them impractical to implement. Many quantitative explorations are based on ad-hoc guesswork to build their measures and do not systematically address complex small group dynamics and interactions [Mirriahi et al., 13]. We attempt to address these problems by designing an automated and process-based small group learning assessment model and then presenting a web-based tool that is informed by complex systems theory and learning analytics in order to provide actionable intelligence for teachers.

7.1 Theoretical Implications

Imagining small groups as complex systems provides a new frame for making sense of learning, and for conceptualizing future developments of learning theory. In particular, the emergent behavior rising from local/low-level interactions between multiple agents (members), and mediated by various tools and artifacts in the environment we see molecules of repeated behavioral patterns emerge in the data; this is in part the move toward simplicity that is usually overshadowed by complexity of these types of analysis. Our work provides a lens for understanding and breaking down the process of small group, technology mediated learning in new and interesting ways. While small group learning arises from and constrains the interactions and dynamics among individuals, it is a group-level property and cannot be reduced to any particular individual in the group [Kapur et al., 11]. We describe a complexity-based approach that models complex small group learning using simple interaction rules. Our identified notions of emergent simplicity and emergent complexity suggest a potential for theoretically sound rules that can be operationalized to model the interactions and dynamics among group members.

Modeling complex interaction and dynamics in small groups through a simple rule-based mechanism that is further operationalized into measures, while intriguing, might also be unsettling and counter-intuitive [Kapur et al., 11]. For example, one might reasonably be concerned about oversimplification. The underlying ontological assumption could be that complex behavior can only be explained by complex mechanisms (e.g. linguistic mechanisms); and simple mechanisms may be insufficient [Kapur et al., 11]. While this is one possibility, [Bar-Yam, 03] concepts of emergent simplicity and emergent complexity suggest that some point of learning analytics success and robustness is possible if our choice of algorithm and computational approach is closely connected to the pedagogy used (small group learning) and the phenomena being examing.

Theoretical approaches not presently applied to learning analytics elsewhere may help to directly connect our computational approach (and other computational approaches) to learning phenomena, with the aim of more efficiently surfacing trends. Complexity is critical to the theory of dynamical minimalism, which is useful for understanding complex psychological and social phenomena [Nowark, 04]. The main argument of this theory is to reconcile the scientific principle of parsimony, which is that simple explanations are more ideal than complex ones in investigating a phenomenon with an arguable loss in depth of comprehension of the phenomenon [Kapur et al., 11]. The principle of parsimony is to seek the simplest mechanism and the fewest variables to explain a complex situation. This does not necessarily sacrifice the depth in understanding in that repetitive and dynamic interactions guided by simple rules and mechanisms can generate complex behavior – the very definition of emergent complexity [Kapur et al., 11]. Hence, parsimony and complexity are not irreconcilable [Kapur and Kinzer, 09]. However, the implication here is not that complex small group learning ought to be studied via simple mechanisms; it is that the exploring and modeling complex group interaction via simple mechanisms is a promising and meaningful effort [Kapur et al., 11].

On the other hand, “One way to advance science is to progressively flesh out theories, adding experimental details and elaborating mechanistic accounts” [Goldstone, 08]. In this present study, we demonstrate that modeling the dynamics and interaction of complex small group learning by operationalizing simple-rule based mechanism, which leads to the final accurate and reliable automated assessment model for small groups. We use theory to inform learning analytics practice. As an exploratory study, it is hoped that this motivates further theory-based connections between learning analytics and computation. The learning analytics research community can explore theories in an even bigger scope (e.g. human behavior theories, learning theories etc.) to position and contribute developments for long-term viability and positive impact of learning analytics on teaching and learning.

7.2 Methodological Implications

Comprehending and modeling the evolution of interaction over time and how variation in this evolution explains learning outcomes are among the significant challenges for CSCL research [Hmelo-Silver et al., 11] [Reimann, 09]. “Temporality does not only come into play in quantitative terms (e.g., duration, rates of change), but order matters: Because human learning is inherently cumulative, the sequence in which experiences are encountered affects how one learns and what one learns” [Reimann, 09]. However, previous measures derived from qualitative interactional coding or from quantitative content analyses are limited in embodying the time and order dimensions of these measures [Kapur et al., 11]. What information can we infer if group collaboration has a high value of a particular measure or category of interaction? It is possible that these codes or interactions are equally distributed throughout the collaborative process, or they are clustered over a particular time or phase? Traditional analyses usually give the same weight to these measures for both situations or assume temporal homogeneity [Kapur et al., 08]. Such analyses, though informative, do not take the temporality of interactions into account and therefore overlook the order of interactions in the problem solving process.

In operationalizing the simple rules to construct measures, we underscored the roles of temporality and sequence. Log data generated by the VMTwG environment are processed based on a coarse timeline (four sequential parts). Depending on the tool actions and chat significance on problem solving in that particular stage, different weights are assigned to these features based on the golden ratio function. For each feature in a specific time, the value of the feature is a function of weight and frequency. The final values for each feature are an accumulation of the values over time. This measure construction process lays the groundwork for process-oriented assessment.

That is, assessment of a group performance is transformed to evaluate the contributions of these features bringing to the group problem solving. However, operationalization of these rules is relatively subjective and considered to be a design process, and therefore it may invite many questions. What essentially guides the construction of our indicators is the maximization of assessment accuracy and F-measure. Therefore, the operationalization of complex systems theory involves trial and error and incorporates mathematical and algorithmic concerns instead of purely theoretical perspectives. It is the reflexive, persistent use of algorithms to filter observations; observations to validate and inform algorithms, which defines our approach.

This integrated use of "Group Informatics" [Goggins et al., 13], which is a methodological approach and ontology for making sense of trace data using a reflexive qualitative-computational process underlies our approach. In this particular case and in the field of education in general, data sets for training machine learning algorithms are usually very small, and atheoretical approaches to computation therefore limited. For this reason we ruled out the more complex models such as tree-based algorithms and neural networks. Because regression techniques are sensitive to outliers and dependency among features or variables, both of which are common phenomena within educational datasets, we did not consider regression models.

Our results illustrate how one particular model can be optimized and show results that exceed those of other models. What we argue here is not that the TAN model is a one-fits-all algorithm; we instead propose that the TAN model or simpler Bayesian model is a good start when exploring modeling techniques for small educational datasets. While learning analytics is accustomed to incorporating analytics practices, models and algorithms from data mining, business intelligence and the "big data" fields, our work proposes another perspective – data at small scale – to expand the application of learning analytics. This approach requires a systematic research process like the one embodied by Group Informatics [Goggins et al., 13]. As the feature construction and TAN model development can be totally automated; the proposed methodology is able to accomplish the goal of building an automatic and reliable process-oriented assessment model for small group learning. In fact, the developed web-based tool has partially implemented this proposed methodology and moves a step further to visualize the identified key features from a time-series standpoint. Since these constructed features have a semantic background behind them (e.g. ‘move the group forward’), teachers could easily understand the rationale behind the model. Therefore, they can not only obtain the process-oriented assessment results (High, Medium, Low), but also provide in-time scaffolding and intervention for small group learning in a social-technical environment.

8 Conclusion

This study proposes an automatic assessment model for small group learning using a small dataset generated by a technology-mediated environment in conjunction with methods grounded in complexity theory. We first conceptualized small group learning as a complex system and explained how the group dynamics and interaction can be modeled via simple and theoretically sound rules. Then, these rules were operationalized to build temporally-embodied measures, and different weights were assigned to these measures according to their significance in that particular stage. This theory-based measure construction method in combination with a feature selection algorithm accomplishes the goal of reducing data dimensionality and contextualizes these measures on a semantic background to facilitate easy interpretation by teachers. Due to fact that the data samples generated in educational contexts are often small, TAN was coded to develop the group assessment model, which achieved the best accuracy compared to baseline models. Finally, a web-based tool was developed so that teachers could not only assess small group learning automatically, but could also provide in-time scaffolding and intervention based on the visualization of complexity-based features.

References

[Arnold et al. 09] Arnold, N., Ducate, L., Lomicka, D. and Lord, G.: “Assessing Online Collaboration among Language Teachers: A Cross-Institutional Case Study.” Journal of Interactive Online Learning, 8, 2 (2009), 121–139.

[Arrow et al. 00] Arrow, H., McGrath, J. E. and Berdahl. J. L.: “Small groups as complex systems: Formation, coordination, development, and adaptation”; Sage, (2000).

[Barab 01] Barab, S. A., Hay K. and Yamagata-Lynch, L.: “Constructing Networks of Action-Relevant Episodes: An in Situ Research Methodology.” The Journal of the Learning Sciences, 10, 1–2 (2001), 63–112.

[Bar-Yam 97] Bar-Yam, Y.: “Dynamics of complex systems”; Reading, MA: Addison-Wesley, 3(1997).

[Capra 96] Capra, F.: “The web of life: A new scientific understanding of living systems”; Random House LLC (1996).

[Chen and Monica 14] Chen, B. D., and Resendes, M.: "Uncovering what matters: analyzing transitional relations among contribution types in knowledge-building discourse." In Proceedins of the Fourth International Conference on Learning Analytics And Knowledge, ACM (2014), 226-230.

[Coffrin 14] Coffrin, C., Corrin, L., Barba, P. and Kennedy, G.: "Visualizing Patterns of Student Engagement and Performance in MOOCs." ACMPress (2014), 83–92.

[Cress 08] Cress, U.: "The Need for Considering Multilevel Analysis in CSCL research-An Appeal for the Use of More Advanced Statistical Methods." International Journal of Computer-Supported Collaborative Learning 3, 1(2008), 69–84.

[Dennen 08] Dennen, V.: "Looking for Evidence of Learning: Assessment and Analysis Methods for Online Discourse." Computers in Human Behavior, 24, 2 (2008), 205–219.

[Ebner and Schön 13] Ebner, M. and Schön, M.: "Why Learning Analytics for Primary Education Matters!" Bulletin of the IEEE Technical Committee on Learning Technology, 15, 2 (2013), 14.

[Fancsali 11] Fancsali, S. E.: "Variable construction for predictive and causal modeling of online education data." In Proceedings of the 1st International Conference on Learning Analytics and Knowledge, ACM (2011), 54-63.

[Ferguson and Shum 11] Ferguson, R., and Shum, S. M.: "Learning Analytics to Identify Exploratory Dialogue within Synchronous Text Chat." In Proceedings of the 1st International Conference on Learning Analytics and Knowledge, ACM (2011), 99–103.

[Friedman et al. 97] Friedman, N., Geiger, D. and Goldszmidt. M.: "Bayesian network classifiers." Machine learning, 29, 2-3 (1997), 131-163.

[Goggins et al. 11] Goggins, S. P., Laffey, J. and Gallagher, M.: "Completely Online Group Formation and Development: Small Groups as Socio-Technical Systems." Information Technology & People, 24, 2, (2011), 104–33.

[Goggins et al. 13] Goggins, S. P., Mascaro, C., & Valetto, G. (2013). Group informatics: A methodological approach and ontology for sociotechnical group research. *Journal of the American Society for Information Science and Technology*, *64*(3), 516-539.

[Goldstone and Wilensky 08] Goldstone, R., and Wilensky, U.: "Promoting transfer by grounding complex systems principles." The Journal of the Learning Sciences, 17, 4 (2008), 465-516.

[Gress et al. 10] Gress, C. L. Z., Fior, M., Hadwin, A. F. and Winne, P. H. "Measurement and Assessment in Computer-Supported Collaborative Learning." Computers in Human Behavior, 26, 5 (2010), 806–14.

[Hall 99] Hall, M. A.: "Correlation-based feature selection for machine learning." PhD diss., The University of Waikato, (1999).

[Hämäläinen and Vinni 06] Hämäläinen, W. and Vinni, M.: “Comparison of Machine Learning Methods for Intelligent Tutoring Systems.” In Intelligent Tutoring Systems, Springer (2006), 525–34.

[Hmelo-Silver et al. 00] Hmelo-Silver, C. E., Holton, D. L., and Kolodner, J. L.: "Designing to learn about complex systems." The Journal of the Learning Sciences, 9, 3 (2000), 247-298.

[Hmelo-Silver et al. 11] Hmelo-Silver, C. E., Jordan, R., Liu, L. and Chernobilsky, E.: "Representational tools for understanding complex computer-supported collaborative learning environments." In Analyzing Interactions in CSCL, Springer (2011), 83-106.

[Holland 98] Holland, J. H.: “Emergence-from order to chaos“; Perseus Books Cambridge, MA (1998).

[Hoyle 08] Hoyle, D. C.: “Automatic PCA Dimension Selection for High Dimensional Data and Small Sample Sizes.” Journal of Machine Learning Research, 9, 12 (2008).

[Jacobson and Wilensky 06] Jacobson, M. J. and Wilensky, U.: “Complex Systems in Education: Scientific and Educational Importance and Implications for the Learning Sciences.” The Journal of the Learning Sciences, 15, 1 (2006), 11–34.

[Kantardzic 11] Kantardzic, M.: “Data mining: concepts, models, methods, and algorithms”; John Wiley & Sons (2011).

[Kauffman 95] Kauffman, S.: “At home in the universe: The search for the laws of self-organization and complexity”; Oxford University Press (1995).

[Kapur and Kinzer 09] Kapur, M. and Kinzer, C. K.: “Productive Failure in CSCL Groups.” International Journal of Computer-Supported Collaborative Learning, 4, 1 (2009).

[Kapur et al. 08] Kapur, M., Voiklis, J. and Kinzer, C. K.: "Sensitivities to early exchange in synchronous computer-supported collaborative learning (CSCL) groups." Computers & Education, 51, 1 (2008), 54-66.

[Kapur et al. 11] Kapur, M., Voiklis, J. and Kinzer, C. K.: "A complexity-grounded model for the emergence of convergence in CSCL groups." In Analyzing Interactions in CSCL, Springer US (2011), 3-23.

[Kumar et al. 10] Kumar, V. S., Gress, C. L. Z., Hadwin, A. F. and Winne, P. H.: “Assessing Process in CSCL: An Ontological Approach.” Computers in Human Behavior, 26, 5 (2010), 825–34.

[Mennin 07] Mennin, S.: “Small-Group Problem-Based Learning as a Complex Adaptive System.” Teaching and Teacher Education, 23, 3 (2007), 303–13.

[Mirriahi et al. 13] Mirriahi, N. and Dawson, S.: “The Pairing of Lecture Recording Data with Assessment Scores: A Method of Discovering Pedagogical Impact.” In Proceedings of the Third International Conference on Learning Analytics and Knowledge, ACM (2013), 180–84.

[Nowark 04] Nowak, A.: "Dynamical minimalism: Why less is more in psychology."Personality and Social Psychology Review, 8, 2 (2004), 183-192.

[Reimann 09] Reimann, P.: “Time Is Precious: Variable- and Event-Centred Approaches to Process Analysis in CSCL Research.” International Journal of Computer-Supported Collaborative Learning, 4, 3 (2009): 239–57.

[Safin et al. 10] Safin, S., Verschuere, A., Burkhardt, J., Détienne, F., and Hébert, A. M.: “Quality of Collaboration in a Distant Collaborative Architectural Educational Setting.” International Reports on Socio-Informatics 7, 1 (2010), 40–48

[Schuldt et al. 04] Schuldt, C., Laptev, I. and Caputo, B.: "Recognizing human actions: a local SVM approach." In Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on, 3, IEEE (2004), 32-36.

[Siemens and Baker 12] Siemens, G., and d Baker, R. S . J.: "Learning analytics and educational data mining: towards communication and collaboration." In Proceedings of the 2nd international conference on learning analytics and knowledge, ACM (2012), 252-254.

[Smith 58] Smith, D. E.: “History of Mathematics”; Special Topics of Elementary Mathematics, 2 (1958).

[Stahl 12] Stahl, G.: "Theories of cognition in collaborative learning." International handbook of collaborative learning (2012), 74-90.

[Strijbos 11] Strijbos, J. W.: “Assessment of (Computer-Supported) Collaborative Learning.” IEEE Transactions on Learning Technologies, 4, 1 (2011), 59–73.

[Suthers et al. 03] Suthers, D. D., and Hundhausen, C. D.: "An experimental study of the effects of representational guidance on collaborative learning processes."The Journal of the Learning Sciences, 12, 2 (2003), 183-218.

[Tair and El-Halees 12] Tair, M. M. T., and El-Halees, A. M.: "Mining educational data to Improve student’s performance: A case study." International Journal of Information and Communication Technology Research, 2, 2 (2012), 1-19.

[Taraghi et al. 14] Taraghi, B., Ebner, M., Saranti, A., and Schön, M.: “On Using Markov Chain to Evidence the Learning Structures and Difficulty Levels of One Digit Multiplication.” ACM Press (2014), 68–72.

[Vanneschi and Poli 12] Vanneschi, L., and Poli, R.: "Genetic Programming—Introduction, Applications, Theory and Open Issues." In Handbook of Natural Computing. Springer Berlin Heidelberg (2012), 709-739.

[Webb 01] Webb, G. I.: "Candidate elimination criteria for lazy Bayesian rules." AI 2001:Advances in Artificial Intelligence. Springer Berlin Heidelberg (2001), 545-556.

[Xing et al. 14] Xing, W. L., Wadholm, B. and Goggins. S.: “Learning Analytics in CSCL with a Focus on Assessment: An Exploratory Study of Activity Theory-Informed Cluster Analysis.” ACM Press (2014), 59–67.

[Xing et al. 14] Xing, W. L., Wadholm, B. and Goggins. S.:“Assessment Analytics in CSCL: Activity Theory based Method.” In International Conference of Learning Sciences, ICLS, 2014, (pp. 1535-1536).

[Zhu 12] Zhu, C.: “Student Satisfaction, Performance, and Knowledge Construction in Online Collaborative Learning.” Journal of Educational Technology & Society 15, 1 (2012), 127–136.