Student Assessment in Small Groups: A Spectral Clustering Model

Wanli Xing, University of Missouri, Columbia
Sean Goggins, University of Missouri, Columbia

Abstract
Enabling the formative assessment of students while limiting demands on teachers' time is a significant concern for technology mediated learning in small groups. Previous approaches have either required extensive time commitments on the part of teachers or relied on the development of special computational models of behavior. Oftentimes, these models overlook the way in which traces of student interaction in a learning system also constitute traces of human behavior, and, instead of providing an account of human behavior, act only as "blunt instruments" relying only on the automated accounting of student activities. In this article, we employ activity theory to categorize traces of student behavior captured from a Virtual Math Teams (VMT) geometry class in an online, synchronous environment. From this, six semantically-grounded measures are generated for each student. Using these, a recently-developed clustering algorithm – spectral clustering – is coded to identify students who have similar behavior patterns. Structured in such a fashion, the theoretical and computational approach taken allows for an automated and meaningfully-grounded assessment of student performance, enabling teachers to offer concrete and personalized help in a timely format. A preliminary discussion for assessment in CSCL is then discussed.

1   Introduction

Automated assessment of student performance in online environment is a long hoped for, often strived for and probably distant objective for educational research. A more pressing and achievable objective for learning analytics is to provide teachers with insights into student interactions. This work provides an activity theory-informed computational method for producing a holistic view of individual student participation in CSCL activities.
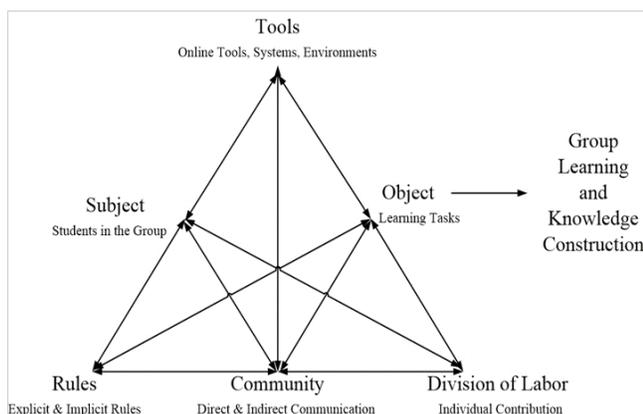
2   Theoretical Framework

Figure 1. Activity theory for students in CSCL

The Activity System model developed by Engestrom (1987) offers a way to comprehensively frame the collaborative knowledge development process while linking together social behavior and its interdependencies (de Laat, 2006). Engestrom's (1987) activity model includes six interacting components: subjects, tools, objects, rules, community and division of labor (see Fig. 1). Contradictory

tensions between the elements of the Activity System serve to produce the outcome of the activity. The activity of learning (Basharina, 2007) is "the joint activity of a student, physical/symbolic tool(s), and another person(s) performing together as a working social system to achieve some outcome under constraints such as rules." In our CSCL student assessment context, the outcome and process of this transformation may both be seen as learning and knowledge. It is the sum of the system components and the tensions among them that make up the learning and influence the learning outcomes. Activity Theory helps us to address the complex interactions and see into student performance in a socio-technical CSCL environment (see Table 1).

Measure-metric	Definition
Object	Completed learning tasks together such as the solution of a problem or production of an artifact.
Subject	Students' individual effort that contributes to problem solving.
Tools	Computers, online tools, and environments that mediate the learning and collaboration activities.
Community	Direct and indirect communication enables the group of students to maintain a sense of community and belonging.
Rules	Implicit and explicit rules and guidelines that constrain the activity.
Division of Labor	Concrete contribution from each group member.

Table 1. Description of Activity Theory operationalization
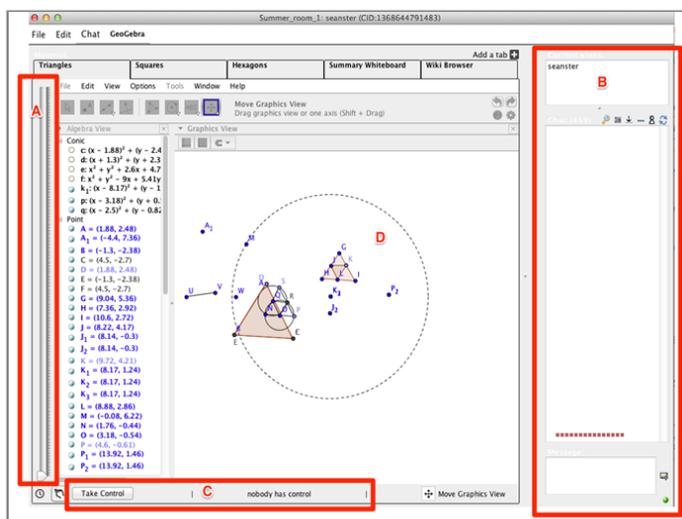
3   Research Context

Figure 2. VMTwG, an analytical tool for collaborative math discourse

In this study, we operationalize Activity Theory in order to make sense of the electronic trace data left behind in group learning interactions in order to assess students using a synchronous math discourse tool. One topic of a 2013 course – "Constructing Triangles" – was chosen in which to assess students' performance with 13 groups and 41 students in total. Figure 2 provides us with a guide for understanding cognitive learning discourse in VMT. Section A of Figure 2, the VMT replayer bar, reveals the time dimension. Each action within VMTwG is logged with a timestamp. Section B is the chat window. Sections C and D are related to Geogebra actions. C is the "Take Control" button, which gives an individual user control of the tools. Section D is the GeoGebra window itself. Here, students are working to create an equilateral triangle within an equilateral triangle using multiple approaches.

We collected log data in .txt format from five modules in a single course (Figure 3). The data centers on specific event types from the CSCL environment (VMT): Chat, Awareness, Geogebra, System, and WhiteBoard (Wb). The Chat event type logs all the messages that students communicate with each other. Awareness records the actions of erasing the chat messages when the student realizes they are full on the chat bar. Geogebra logs information on how students visually construct a geometry artifact

(e.g. add a point, update a segment, etc.). The System event type records information on how the VMT environment is accessed.

Community	Subject	Topic	Room	Source	Target	Time	Finish Time	Event Type	Event
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	madison_	emma_r	0:00:48	26:52.8	chat	2013-03-08 13:26:52.799 - madison_m -> what do we do now im confused what tab are we in??
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	madison_	0:0:-44	26:08.6	Geogebra:I	2013-03-08 13:26:08.643 - emma_r -> added line:Line i
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:00	26:13.0	awareness	2013-03-08 13:26:12.987 - emma_r -> [fully erased the chat message]
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:17	26:31.6	chat	2013-03-08 13:26:31.614 - emma_r -> i just made the line segment between my new points
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	amina_p	emma_r	0:00:00	26:40.7	chat	2013-03-08 13:26:40.697 - amina_p -> k
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	amina_p	0:00:23	27:04.0	Geogebra:I	2013-03-08 13:27:03.997 - emma_r -> added line:Line j
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:03	27:15.0	chat	2013-03-08 13:27:15.045 - emma_r -> bisector
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:21	28:03.7	chat	2013-03-08 13:28:03.744 - emma_r -> no i cant construct the other line segments. maybe you guys can
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:00	28:04.1	Geogebra:I	2013-03-08 13:28:04.139 - emma_r -> tool changed to Move Graphics View
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:06	28:21.2	chat	2013-03-08 13:28:21.155 - emma_r -> who wants control now?
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	madison_	emma_r	0:00:12	28:33.9	system	2013-03-08 13:28:33.946 - madison_m -> Now viewing tab TEAM 4 TAB
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	madison_	madison_	0:00:03	28:37.3	system	2013-03-08 13:28:37.313 - madison_m -> Now viewing tab Bisector
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	madison_	madison_	0:00:19	29:01.9	chat	2013-03-08 13:29:01.877 - madison_m -> ive added a tab its called....TEAM 4 TAB
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	madison_	0:00:08	28:52.3	chat	2013-03-08 13:28:52.349 - emma_r -> why did you add a new tab madison>
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:01	28:54.8	chat	2013-03-08 13:28:54.817 - emma_r -> ?
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:00	28:57.3	chat	2013-03-08 13:28:57.335 - emma_r -> ?
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:02	28:59.6	system	2013-03-08 13:28:59.602 - emma_r -> Now viewing tab TEAM 4 TAB
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:16	29:16.2	wb	2013-03-08 13:29:16.190 - emma_r created a scribble
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:00	29:16.6	system	2013-03-08 13:29:16.629 - emma_r -> Now viewing tab Bisector
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:02	29:18.7	Geogebra:I	2013-03-08 13:29:18.654 - emma_r -> tool changed to Move
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	madison_	emma_r	0:00:34	30:04.0	chat	2013-03-08 13:30:03.979 - madison_m -> i feel abandond whats wrong guys?????? ;(
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	amina_p	madison_	0:00:22	29:55.0	chat	2013-03-08 13:29:55.023 - amina_p -> are finish with the instructions emma
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	amina_p	0:00:00	29:48.5	chat	2013-03-08 13:29:48.485 - emma_r -> i get it
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:01	29:52.6	chat	2013-03-08 13:29:52.573 - emma_r -> geeeezzz
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:15	30:08.1	Geogebra:I	2013-03-08 13:30:08.068 - emma_r -> tool changed to Intersect Two Objects

Figure 3. Sample logs from VMT

4   Methodology

4.1   Measure Construction

Subject: In the CSCL activity of learning, Subject represents the individual student. When mapped to our log data, it represents all actions one student makes during the whole training under the individual modules, and is the sum of all actions of individual endeavors.

Rules: In this VMT context, students have to perform actions that the VMT environment offers. Therefore, the Rules are reflected by the sum of actions the student uses in this module.

Tools: The tools are the System and Wb event types where the student's action for tool usage is registered. Thus, Tools are presented as the sum of participation frequency in System and Wb events.

Community: All the communications that help maintain the community structure. Therefore, Community can be presented as the sum of chat messages and awareness.

Division of labor: Contributions of the student made to the collaborative learning modules. The only concrete contribution to the geometry object construction is from the Geogebra dimension. Therefore, we use the student's participation in Geogebra events to represent the student's Division of Labor aspect.

Object: The CSCL activity is to achieve the object of a student's active involvement in the whole class. In order to quantify whether the student is active, we incorporate the frequency of participation and the number of event types.

4.2   Student Assessment

An activity system is characterized by the internal tensions among its components. Therefore, it is hard to compute one value as functions of the six dimensions to indicate the learning or performance result of the learner, especially considering the complexity of the nature of group learning in a technology mediated environment. In this exploratory study, spectral clustering (Luxburg, 2007) is coded to chunk students with similar learning behaviors. Cluster analysis brings into consideration all of the six dimensions in the activity system rather than accounting for only one dimension. Using cluster analysis is also especially useful when learning takes place in large online context.

5   Results

Cluster analysis is not only able to facilitate scaling up in online context but also offer actionable intelligence for the teachers. Table 2 presents the results of the spectral clustering analysis when k is set to 3. The table further shows the standardized means of each cluster over the six dimensions and the number of students in each cluster, as well as the standardized range of each dimension. To illustrate the actionable power, we will look at Cluster 1 as an example. Looking at this table, a teacher can see that

Cluster 1 has 15 members in it, that the student scores best in Subject (0.48), Tools (0.22), Division of Labor (0.26) and Community (0.21), but has lowest ranking for Object (-0.16) and only a medium value on Rules (0.01). The low ranking on Object reflects that students in this cluster might be very active in one of the subtasks but not others. The average value on Rules is a proof implying that students are not investigating all of the available functions deeply enough. Based on this, the teacher could remind these students to try use different kinds of geometric functions. He or she can also communicate with these students and offer a remediation or training program if the low engagement with the VMT functions results from the unfamiliarity with the environment. In addition, the teacher can encourage students to keep consistent involvement through the subtasks or reflect on the curriculum design for this task and see whether students are having difficulty understanding the requirements or instruction for the task. In the context of a large number of online groups, the teacher can send one reminder or report to all the students of a cluster at the same time to help them improve their performance. While we focus here on the teacher's perspective, the table can also be shared with the students themselves in order to encourage self-awareness and reflection on their performance status.

Cluster Size		Cluster1	Cluster2	Cluster3
		15/41	20/41	6/41
Measure	Range (Standardized)	Cluster Means		
Subject	[-0.73, 3.48]	0.48	-0.15	-0.21
Rules	[-2.55, 0.39]	0.01	0.39	-1.08
Tools	[-1.12, 2.19]	0.22	0.20	-0.47
D. of Labor	[-1.63, 1.58]	0.26	0.03	0.23
Community	[-0.68, 5.11]	0.21	-0.10	-0.14
Object	[-0.52, 4.79]	-0.16	0.04	0.49

Table 2. Sample spectral cluster k = 3 results

6   Discussion & Conclusion

There are multiple perspectives on assessment in the CSCL literature related to group learning. In this paper we explore a different way of developing automated assessment that is coherent with theory and is scalable. Founding our assessment on the model of group dynamics found in Activity Theory, this work approaches assessment in a more holistic manner than other approaches which seek to only provide a data-centric result. Of course, this theoretically-enabled mode of development is unique to the VMT system on which the automated assessment is based. Given the affordances of other learning systems, other theoretical models may have to be used to achieve a comparable result. One thing that is certain is that by producing a theoretically-informed system for automated assessment, we are able to provide teachers with the kind of real-time and in-depth insight into student performance that has not been previously available. Modulating a data-driven approach with established theoretical insight provides a useful level of abstraction for teachers as they work to guide student learning.

References

Basharina, O. K. (2007). An activity theory perspective on student-reported contradictions in international telecollaboration. *Language Learning & Technology*, *11*(2), 82-103.

Daradoumis, T., Martínez-Monés, A., & Xhafa, F. (2006). A layered framework for evaluating on-line collaborative learning interactions.*International Journal of Human-Computer Studies*, *64*(7), 622-635.

De Laat, M.F. (2006). Networked Learning. Apeldoorn: Politie Academie. *PhD dissertation*, Utrecht Univ., the Netherlands.

Engeström, Y. (1987). Learning by expanding: an activity-theoretical approach to developmental research. Helsinki: Orienta-Konsultit Oy.

Gress, C. L., Fior, M., Hadwin, A. F., & Winne, P. H. (2010). Measurement and assessment in computer-supported collaborative learning. *Computers in Human Behavior*, *26*(5), 806-814.

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research*, *60*(1), 1-64.

Lonchamp, J. (2010). Customizable Computer-based Interaction Analysis for Coaching and Self-Regulation in Synchronous CSCL Systems. *Educational Technology & Society*, *13*(2).

Reimann, P. (2009). Time is precious: Variable-and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning, 4*(3), 239-257.

Stahl, G. (2013). Interaction analysis of a biology chat. In *Productive multivocality in the analysis of group interactions* (pp. 511-539). Springer US.

Strijbos, J. W. (2011). Assessment of (computer-supported) collaborative learning. *Learning Technologies, IEEE Transactions on*, *4*(1), 59-73.

Von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and computing*, *17*(4), 395-416.

Table of Figures

Table of Tables