Work in Progress: A Markov Chain Method for Modeling Student Behaviors

Dr. Corey T. Schimpf, The Concord Consortium

Corey Schimpf is a Learning Analytics Scientist with interest in design research, learning analytics, research methods and under-representation in engineering, A major strand of his work focuses on developing and analyzing learning analytics that model students' cognitive states or strategies through fine-grained computer-logged data from open-ended technology-centered science and engineering projects. His dissertation research explored the use of Minecraft to teach early engineering college students about the design process.

Ms. Molly H. Goldstein, Purdue University, West Lafayette (College of Engineering)

Molly Goldstein is a Ph.D. student in the School of Engineering Education at Purdue University, West Lafayette with a research focus on characterizing behaviors in student designers. She previously worked as an environmental engineer specializing in air quality influencing her focus in engineering design with environmental concerns. She earned her B.S. in General Engineering (Systems Engineering & Design) and M.S. in Systems and Entrepreneurial Engineering from the University of Illinois in Urbana-Champaign.

Dr. Robin Adams, Purdue University, West Lafayette (College of Engineering)

Robin S. Adams is an Associate Professor in the School of Engineering Education at Purdue University and holds a PhD in Education, an MS in Materials Science and Engineering, and a BS in Mechanical Engineering. She researches cross-disciplinarity ways of thinking, acting and being; design learning; and engineering education transformation.

Dr. Jie Chao, The Concord Consortium

Jie Chao is a learning scientist with extensive research experience in technology-enhanced learning environments and STEM education. She completed her doctoral and postdoctoral training in Instructional Technology and STEM Education at the University of Virginia. Her past research experiences ranged from fine-grained qualitative mental process analysis to large-scale quantitative and longitudinal investigations. She is currently focusing on learning analytics research in open-ended domains such as engineering design and authentic scientific inquiry. With insights in learning sciences and a strong, computationally oriented mindset, she hopes to utilize learning analytics to investigate important questions with unprecedented granularity and generate actionable knowledge for the design of technology and curriculum.

Dr. Senay Purzer, Purdue University, West Lafayette (College of Engineering)

enay Purzer is an Associate Professor in the School of Engineering Education. She is the recipient of a 2012 NSF CAREER award, which examines how engineering students approach innovation. She serves on the editorial boards of Science Education and the Journal of Pre-College Engineering Education (JPEER). She received a B.S.E with distinction in Engineering in 2009 and a B.S. degree in Physics Education in 1999. Her M.A. and Ph.D. degrees are in Science Education from Arizona State University earned in 2002 and 2008, respectively.

Dr. Charles Xie, The Concord Consortium

WIP: A Markov Chain Method for Modeling Student Design Behaviors

Abstract
Students from a middle school (N=152) and from a high school (N=33) completed the same energy-efficient home design challenges in a simulated environment for engineering design (SEED) supported by rich design tool with construction and analysis capabilities, Energy3D. As students design in Energy3D, a log of all of their design actions are collected. In this work-in-progress a subsample of the five most engaged students from both the middle and high school samples are analyzed to identify similarities and differences in their design sequences through Markov chain models. Sequence learning is important to many fields of study, particularly fields that have a large practice component such as engineering and design. Design sequences represent micro-strategies for developing a design. By aggregating these sequences into a model we aim to characterize and compare their design process. Markov chains aid in modeling these sequences by developing a matrix of transition probabilities between actions. Preliminary results suggest we can identify similarities and differences between the groups and that their design sequences reflect important considerations of the design problem. We conclude that Markov chains hold promise for modeling student practices.

Keywords: engineering design, Markov chains, large learner data, research methods

Introduction

Learning skills or practices in many domains often involves learning the strengths and weaknesses of taking actions in different temporal order, which is sometimes called sequence learning[1]. This type of learning is particularly applicable in fields that are practice-oriented such as design. Studies of experts and students suggest that distinct stages of design at a high level might be thought of as: problem scoping, information gathering, idea/alternative generation, modeling, analysis, iteration, implementation, and others[2-3]. Other studies have investigated time spent in each of these stages, comparing freshmen and senior college students[4] and comparing students to professional designers[2]. Breaking design stages into smaller units, or at a finer resolution results in design operations[5] that typically happen within stages. This work-in-progress aims to aggregate beginning designers' design operations into a comprehensive matrix of sequences in order to develop models of their overall design process. These sequences are like micro-strategies representing more or less useful chains of actions and therefore can aid us in understanding how designers navigated a design challenge. We use Markov chain analysis, a method for analyzing series of states or actions, to develop the model and then apply to two distinct groups of beginning designers to test its ability to elucidate similarities and differences in students' design processes.

Markov chains model the probability of moving to some new system state determined by what state the system currently is in. The order of a Markov chain refers to how many previous states are used to estimate a new state. For instance, a first-order Markov chain takes into account the present state for estimating the next state that will occur. In applying this technique to design it is important to recall that design is an ill-structured[6-7], highly iterative[8-9] and evolutionary activity[10] which therefore makes it challenging to compare across students or teams' approaches even as they work on the same design challenge. By chaining designers' actions into sequences we are able to model their collective actions as a matrix of transitions between action-states or what is called a *transition matrix* in Markov chains[11]. A transition matrix for a first-order Markov chain, is simply an n x n matrix where n is number of unique states and each cell contains the probability of going from one state to another. Thus, Markov chains and their sequences can model designers' actions in way that can easily be compared across individuals or groups for similar design activities.

The two groups of beginning designers compared in this study completed a design challenge in which they designed an energy efficient home using a computer simulation tool that logged their design operations. Beginning designers are people with minimal experience in design[12]. More specifically, the two groups were 8th grade middle school students in a lower financial and technology resourced district and 9th grade high school students in a higher financial and technology resourced district. While the age difference between them is small, the context of their schools and levels of support and resources make them distinct groups. In analyzing the similarities and differences between the students we seek to inspect if there is a spectrum or distribution in design processes across beginners.

Research Questions

This research seeks to understand:
RQ1: What are the common design patterns (seen through operation sequences) of middle school students engaged in the engineering design project?
RQ2: What are the common design patterns (seen through operation sequences) of high school students engaged in the engineering design project?
RQ3: What are the similarities and differences in operation sequences between middle and high school students?

These research questions are addressed by using Markov chain constructs to represent the sequential pattern of student design behaviors.

Literature Review

Sequential learning is a topic that has been studied extensively by psychologists and educational psychologists[1,13-14]. Researchers in these fields have argued that sequences are essential to human cognition across a variety of abilities and skills[1,14,15]. Given that much of this work examines general cognitive functions such as memory[15], implicit and explicit processing[16-18], and learning channels[1, 14], many of the sequences studied represent general tasks instead of learning context specific sequences. Nonetheless, becoming proficient in some field often requires learning key skills and techniques, which are often sequential in nature[19], particularly when fields have a strong practice component such as design or engineering more broadly.

Markov models, including Markov chains have seen a limited but growing application in the study of design. Researchers have used Markov models for empirical studies of design behavior[5,20--21], to analyze the actions of an intelligent agent that models cognitive and memory functions in a design context[22] and for other empirical studies in domains related to design including innovative thinking[23]. For example, McComb, Cagan & Kotovsky[20] used first-order Markov chains to identify beneficial operation sequences in a truss and home cooling system design challenge. The authors then simulated the behavior of non-sequenced actions and first-order Markov chain design teams with Cognitively Inspired Simulated Annealing Teams (CISAT) framework (see McComb, Cagan & Kotovsky[24]), an agent-based modeling approach. Comparison of the final performance of the simulated teams' artifacts for both design scenarios revealed that the first-order Markov chain outperformed the non-sequenced teams, indicating that sequenced behavior lead to better design solutions than strictly independent behavior.

Following past researchers, we employ Markov chains to study a new population, beginning designers and connect this work to the learning and use of sequences in design.

Research Methods

Research Participants & Classroom Context
This study was conducted in two separate locations, a middle school and high school. The middle school study involved four classrooms of 8th grade students (ages 13-14) in a Midwest urban setting for a total sample of N=152 students. The high school study involved 2 classrooms of 9th grade students (ages 14-15) in a New England urban setting for a total sample of N=33 students. It is important to note that despite there only being a one-year difference between the middle school and high school students their school contexts are markedly different which likely accentuates students' preparedness for complex tasks like design projects. The high school is a heavily resourced school in wealthy district with below national average free or reduced lunch rates and many advanced classes and curriculum whereas the middle school is an under-resourced school in a less wealthy district and higher than national average free or reduced lunch rates that does not have as many advanced classes or curriculum. Therefore, we believe the institutional differences between the school sharpen the contrast in academic preparedness between the two groups.

Students at both schools participated in an in-class design project using Energy3D (http://energy.concord.org/energy3d/), a CAD simulation environment[25]. Energy3D is developed by the Concord Consortium as "a computer-aided engineering tool for designing, analyzing, and constructing green buildings and power stations that utilize renewable energy". The user-friendly software works in a way that allows students to see the effects of each design and specifications they choose to their overall design specifics. It offers a simple 3D graphical user interface for drawing buildings, and evaluating their performance using cost and energy (solar and heat) simulations (see Figure 1, below).

Over the course of approximately two weeks, students at both schools participated in the same design challenge. At both schools, the challenge was formally introduced through a presentation that discussed engineering design in general and the challenge in specific. Students were then given time to learn the software through free-play or a small design challenge. When it came time to begin the design challenge, all students were given two-page design specifications sheet that summarized the details from the presentation. For the challenge students were asked to use Energy3D to create single-family homes that (1) minimized energy consumption, (2) minimized construction cost, while (3) designing a house large enough for a family of four, and (4) maintaining an attractive appearance.

Data Sources & Feature-Based Action Schema

While designing in Energy3D, each student operation was recorded in the background of the program in JSON files. These operations, detailed in Table 1, along with a timestamp of the operation allow a full reconstruction of each student's design activity. Energy3D captures atomic actions or the smallest actions that may affect a design or designer. For example, actions such as add a wall, add a solar panel, or conduct a daily energy efficiency analysis are all recorded with timestamps and metadata like position or kilowatt hour consumed. A full schema of the actions Energy3D records can be found at Energy3D's website[26].

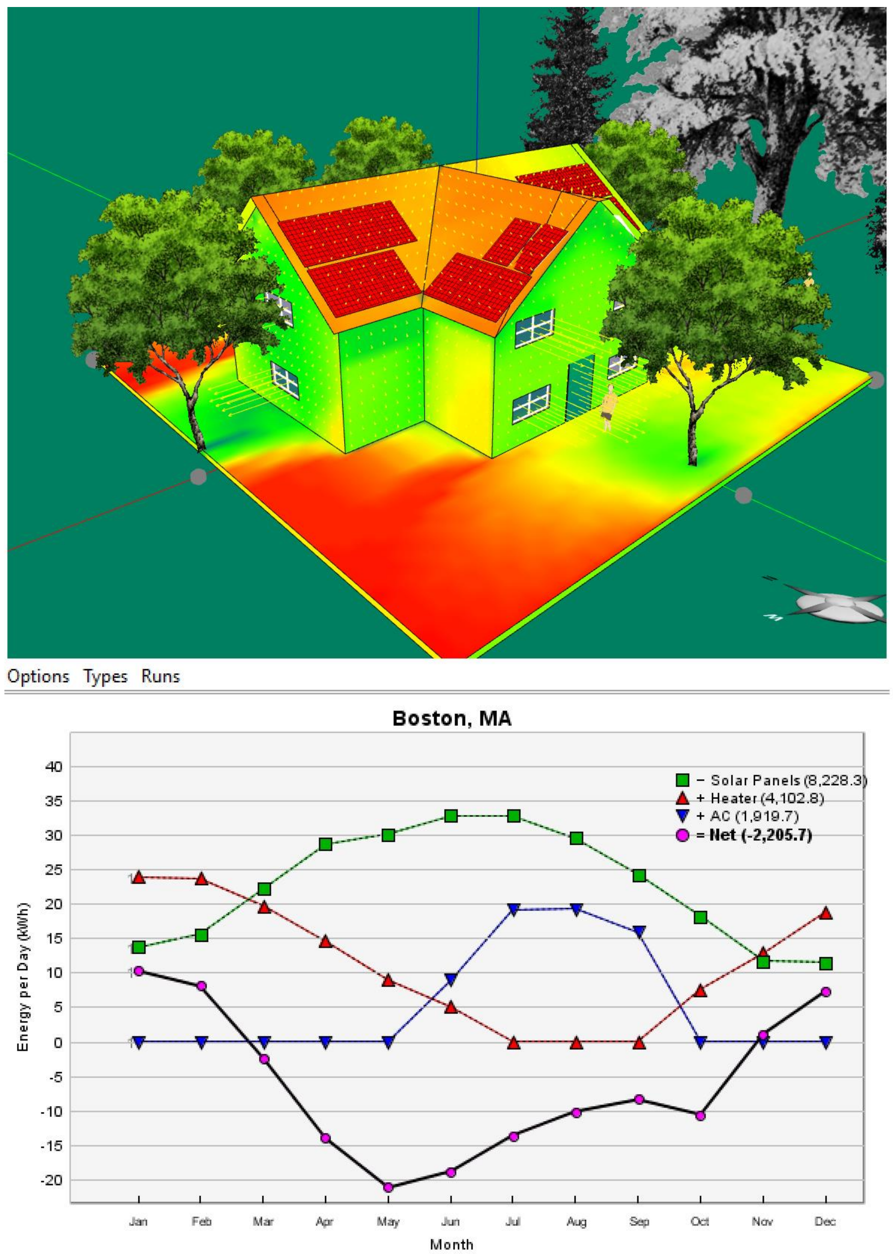

Figure 1. Energy3D example design and analysis

Table 1. Energy3D Collapsed and Removed Action Schemas

Schema 1		
Categories	Summary of Included Actions	Removed from Schema
Building	Move/Rotate/Remove Building, Add Components	None
Door	Add/Edit/Remove Door	
Floor	Add/Edit/Remove Floor	
Foundation	Add/Edit/Paste/Remove Foundation	
Roof	Add/Edit/Remove Different types of Roof, Convert to Gable, Resize overhang	
Solar Panel	Add/Edit/Paste/Remove/Rotate Solar Panel, Change Cell Efficiency, Change Micro-Inverter	
Wall	Add/Edit/Remove Wall, Change Type of Wall	
Window	Add/Edit/Paste/Remove Window	
Schema 2		
Façade	Color Change for House Components, Texture Change	Camera
Human	Add/Edit/Paste/Remove Human	
Numeric Analysis	Daily/Annual Analysis, Remove or Show past data curves	
Sensor	Add/Edit/Remove Sensor	
Tree	Add/Edit/Paste/Remove Tree	
Visual Analysis	Heliodon, Animate Sun, Shadow, Daily Cumulative Radiation	
Schema 3		
Insulation	Change Solar Heat Gain Coefficient, Change U-Value	Spin View, Top View, Zoom

An initial review of student data logs showed 95 and 121 distinct student actions captured in the process data at the high school and middle school, respectively. Any resulting n x n matrix resulting would have over 9000 or over 14000 cells many of which would be sparsely populated. In order to make the matrices amenable to analysis, researchers reviewed each of the lists of actions to determine how to classify individual actions into categories of actions. For example, the actions of "Move Building," "Resize Building," and "Rotate Building" were categorized into one family of "Building". Energy3D allows users to add five different types of roofs and also allows edits to those roofs so the family of "Roof" is comprised of 17 unique actions. The process of developing the current action scheme involved multiple iterations to determine which actions served similar functionality, e.g. a series of actions related to altering roofs, both in Energy3D and with student design intent. Table 1 captures the cumulative changes to the schema over 3 iterations. The first column displays the names of the categories sets of actions were grouped into while the column beside it displays a summary of the actions that fall into the category. Note the schemas are cumulative so iteration 3 incorporates the categories from iteration 1 and 2. The third column lists actions that were removed from the sequence analysis. These actions primarily involved view controls within the software and tended to be high in number while having only an ambiguous connection to designer's intentions. None of these actions transformed the design in a lasting way. The researchers therefore determined they could be safely removed, allowing for sequences to bypass or ignore them. Lastly, some meaningful actions within the schema that were not collapsed into categories include note-taking and viewing artifact information on the graph tab, which are not in Table 1.

Table 2. Summary of Key Characteristics from Two Datasets

	Middle School	High School
Number of operation types	121	95
Student academic level	8^{th}	9^{th}
Total students in analysis	5	5
Average operations per student	1723	1104

In this work-in-progress, we selected a subset of the entire dataset for analysis. Our goal for this analysis was to first look at students who were very engaged in the design project. We estimated a student's degree of engagement by the activity of their log files and selected the five most active students from each school. The average of design actions was 1104 and 1723 from the middle school and high school, respectively. Given that the time for the design challenge and learning context were kept as similar as possible it is not clear why the average operations are notably different between schools. This may reflect different levels of engagement between the high school and middle schools' students or that the high school students felt more comfortable with the software and therefore made more design actions. Note that camera and note actions were not included in these tallies as both tend to be exceptionally high as an artifact of the logging manager.

Data Analysis

Students' Energy3D log files were run through a python script that identified and tallied sequences between any two sequential actions students took. These two-action sequences represent a first-order Markov chain, with one current action followed by the next action taken[9]. As discussed in the data source section, Camera related actions such as rotating or panning the camera are ignored, for example, a sequence of Add Solar Panel → Rotate Camera → Add Solar Panel would be counted as a Add Solar Panel → Add Solar Panel sequence.

We follow Purzer and Fila[23], as we also looked individuals, in calculating the transition probabilities between 'states' as $P(s_i \rightarrow s_j) = \sum s_{ij} / \sum s_i$ or the number of times students went from state i to state j divided by the total number of state i sequences. This represents students' overall probability of transitioning from some state or action to another based on their final full set of time sequenced actions.

Results & Discussion

The results have three components: 1) the four most frequent actions for each student in this sample, 2) the general frequency and transition probabilities of repeat actions or within-action matrix components for core construction, analysis, information seeking and note taking behaviors and 3) the across-action sequences that included the same set of behaviors as 2). Probabilities represent a student's project-wide likelihood of going from state i to state j. Within-action probabilities are of interest because they reflect sequences when students were focused on particular subsystems of the house or tools (e.g., roof to roof). In contrast, across-actions represent transitions between various subsystems and/or tools. For the final component of the results, we used two criteria for inclusion of sequences: the base action of the sequence needed 20 or more instances and the particular sequence of interest needed to be at least 20% or more of the total sequences starting with the same action. Only cases that meet his criteria are reported to avoid cases that are too small or may represent primarily noise.

In what follows we first present the results for the sub-sample of Middle School, followed by the sub-sample of High School and close the section with a comparison of the two.

Middle School Students

RQ1: What are the common design patterns (seen through operation sequences) of middle school students engaged in the engineering design project?

Turning to left part of Table 3 first, which shows the most common actions for each of the highly active middle school students, a few points become apparent. Four out of five students had a high emphasis on solar panels, which suggests they were trying to increase energy production to get their house to meet net-zero energy. Four out of five students also had a high emphasis on windows or trees, which suggests they were trying to use passive solar improve the energy efficiency of their design. Only two students had a high level of wall manipulations suggesting there was less emphasis in the group on manipulating the size and shape of the house. Most students (4/5) had a high amount of notes.

The right section of Table 3 displays any across-action transitions that included numeric or visual analysis, note-taking or information seeking (primarily the graph tab). Three of the five students

Table 3 Middle School Students' Top Four Design Actions and Across-Action Sequences

Student ID	1st	2nd	3rd	4th	Across-Action Sequence(s)	Prob.
A001	Note 347	S. Panel 245	Tree 239	Wind.73	numeric analysis → tree	.33
A003	Wall 168	S. Panel 153	Tree 138	Build. 120	numeric analysis → graph tab	.49
					visual analysis → numeric analysis	.23
B013	Note 367	S. Panel 270	C. Time 236	Tree 161	numeric analysis → graph tab	.26
B031	Note 347	Build. 220	Wall 185	Wind. 118	visual analysis → numeric analysis	.30
					save → numeric analysis	.21
D15	Note 480	Build. 232	C. Date 168	S. Panel 107	visual analysis → save	.23
					graph tab → numeric analysis	.43

had some connection between either graph tab and numerical analysis, suggesting that information seeking and analysis were a common sequence. Two students had some connection between visual analysis and numeric analysis, suggesting they sought out different levels of analytical feedback in tandem.

Next, Table 4 displays the within-action transition probabilities for the subset of middle school students. Note that shaded cells indicate that less than 20 actions happened for that category. A common pattern is that many of the construction-related actions pertain to sub-systems of the house (e.g. walls or windows). This seems to make sense as we might expect students to focus their work on a specific sub-system at a time. Numerical analysis and Graph tab generally showed lower levels of within-action transitions. This seems to be a positive pattern as analysis should happen between construction, note-taking and information gathering behaviors. Visual analysis seems to show higher levels of within-action transitions, however. There are not any sequences going from insulation to insulation and generally students in the middle school did not have many sequences with insulation. This limits their designs as insulation can be a critical factor on reducing energy consumption.

High School Students
RQ2: What are the common design patterns (seen through operation sequences) of high school students engaged in the engineering design project?

Turning first to the left side of Table 5 which displays the most common actions for the high school students subsample we see all of them had solar panels as one of their top actions, suggesting they all put an emphasis on energy production. Three of the five students manipulated windows or trees heavily, suggesting these students had a balance of emphasis on energy production and passive solar strategies, such as using tree shading to reduce solar radiation through windows. The two students who do not have as many passive solar actions both had a larger number of wall actions, suggesting they put more of an emphasis on construction and manipulating the shape/size of their building. One student put a heavy emphasis on change date and visual analysis which may reflect spending time performing different kinds of visual analysis such as analyzing shadows and animating the sun to view shadow patterns over a day.

On the right side of Table 5 high school students across-action sequences are displayed. Only three out of five high school students had transition probabilities between their analysis, information seeking and note taking behavior high enough to be reported here. Of those 3, two showed connections between changes to insulation in their house and numerical analysis. While in general a connection between these two is positive, seeing this sequence occur at a high rate may represent trial and error behavior between insulation and energy savings. Graph tab and numerical analysis also was sequence above the threshold probability for all 3 students, showing a connection between information seeking and analysis. One student had a common sequence between visual analysis and change date, suggesting they explored patterns of the sun or shadows across different days.

Next, Table 6 displays the transition probabilities for repeated actions. A common pattern is that many of the construction related actions occur within sub-systems of the house (e.g. windows or walls). Numerical and Graph tab generally had less within-action transitions, suggesting these

Table 4 Middle School Students Repeat Sequences

Student	Building	Door	Foundation	Graph Tab	Insulation	Notes	Numerical	Roof	Solar Panel	Tree	Visual	Wall	Window
A001	0.00	0.65	0.60	0.30	0.00	0.97	0.04	0.50	0.89	0.82	0.62	077	0.85
A003	0.78	0.67	0.55	0.52	0.00	0.00	0.49	0.57	0.83	0.83	0.58	0.74	0.66
B013	0.83	0.33	0.18	0.33	0.00	0.96	0.20	0.35	0.86	0.83	0.64	0.69	0.62
B031	0.95	0.32	0.50	0.00	0.00	0.95	0.17	0.50	0.69	0.61	0.43	0.85	0.66
D005	0.94	0.56	0.63	0.29	0.00	0.97	0.19	0.00	0.87	0.87	0.38	0.74	0.72

actions were linked with other types of actions. Visual analysis was not consistent in the group with three of the students exhibiting very high repeat action sequences and two with low transition probabilities. The high school students made limited changes to their foundations; only one of the five students did a few times. This means for the most part they self-imposed an area constraint on themselves. If they had changed the foundation more they could have increased or decreased the amount of space their house could occupy, thus affecting energy consumption.

Table 5 Summary of High School Students' Design Actions

Student ID	1st	2nd	3rd	4th	Across-Action Sequence(s)	Prob.
B06	S. Panel 627	Notes 317	Wind. 218	Build. 203	numeric analysis → note	.35
					numeric analysis → save	.2
					graph tab → numeric analysis	.25
B07	Wall 335	Notes 315	S. Panel 243	Build. 220	numeric → insulation	.28
					insulation → numeric analysis	.34
					graph tab → numeric analysis	.24
B08	Wind. 307	C. Date 227	S. Panel 215	V. Anal. 183	None	
D13	Tree 428	Notes 329	Build. 264	S. Panel 254	numeric analysis → insulation	.2
					change date → visual analysis	.22
					visual analysis → change date	.27
					graph tab → numeric analysis	.44
D15	Build. 840	Notes 712	Wall 369	S. Panel 215	None	

RQ3: What are the similarities and differences in operation sequences between middle and high school students?

Both the high school and middle school students had high levels of within-action sequences for construction actions, in general. Middle school students did not interact with insulation much but did manipulate the foundation of their houses, unlike the high school students, who manipulated insulation but not foundation. Thus, even among these two groups of beginning designers, there were differences in how the most active members approached improving the performance of their designs. Both groups showed a mix of students who put more emphasis on energy production, or a mix of energy production and passive solar. The high school students had more positive sequences connecting numerical or visual analysis to other actions beyond information seeking, including note-taking, insulation changes or changing the date of analysis.

Conclusions & Future Work

Key take-away points from this pilot study suggest there are differences in how even beginning designers approach their design process. We were able to see patterns in design actions, providing a view into the similarities and differences among the most engaged students. These patterns (i.e. top actions, within-action and across-action analyses) make sense in light of larger design actions or principles (e.g. reduce energy, link analysis and graphs, etc.). For example, from the across-action results in Table 5, only one student showed a marked connection between running analysis and taking notes. It would be unlikely and not preferable that students always transition from analysis to notes, but we would expect some connection between analysis and

Table 6 – High School Students Repeat Sequences

Student	Building	Door	Foundation	Graph Tab	Insulation	Notes	Numerical	Roof	Solar Panel	Tree	Visual	Wall	Window
B06	0.00	0.11	0.00	0.60	0.60	0.60	0.10	0.34	0.95	0.50	0.79	0.83	0.89
B07	0.02	.51	0.00	0.19	0.29	0.58	0.11	0.55	0.86	0.64	0.69	0.84	0.75
B08	0.92	0.58	0.50	0.40	0.83	0.00	0.00	0.83	0.92	0.00	0.82	0.85	0.92
D13	0.01	0.33	0.00	0.09	0.57	0.66	0.18	0.78	0.82	0.96	0.29	0.88	0.67
D15	0.16	0.20	0.00	0.13	0.38	0.61	0.00	0.61	0.93	0.06	0.17	0.90	0.86

recording information or thoughts as a matter of good design practice. A similar case can be made for other across-actions such as relationships between construction and analysis that demonstrate regular testing of a system.

Given we can see patterns make sense conceptually, we see promise in the technique of Markov chain analysis for both characterizing design patterns and for being able to see similarities and differences in how students navigate the design process. We see potential utility in this Markov chain modeling technique for researchers in comparing design practices across individuals as well as groups. Future work will expand this pilot study through more advanced Markov chain analysis (e.g. statistical estimation of Markov chains, higher order chains). We will use these models to examine the relationships between students' design sequences and design performance.

Acknowledgements

We are grateful for the students who participated in this study and for their teachers who supported data collection efforts. This work presented in this manuscript is based upon work supported by the National Science Foundation (Grant DUE #**1348530** and DUE #**1348547**: Large-Scale Research on Engineering Design Based on Big Learner Data Logged by a CAD Tool). Any opinions, findings, and conclusions or recommendations expressed in this paper, however, are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. Atman, C. J., Adams, R. S., Cardella, M. E., Turns, J., Mosborg, S., & Saleem, J. (2007). Engineering design processes: A comparison of students and expert practitioners. *Journal of engineering education*, *96*(4), 359-379.
2. Cross, N., & Cross, A. C. (1998). Expertise in engineering design. *Research in Engineering Design*, *10*(3), 141-149.
3. Atman, C. J., Cardella, M. E., Turns, J., & Adams, R. (2005). Comparing freshman and senior engineering design processes: an in-depth follow-up study. *Design studies*, *26*(4), 325-357.
4. Crismond, D. P., & Adams, R. S. (2012). The informed design teaching and learning matrix. *Journal of Engineering Education*, *101*(4), 738–797.
5. Zimring, C. & Craig, D.L. (2001). Defining Design between Domains: An Argument for Design Research á la Carte. In C. Eastman, M. McCracken & W. Newstetter (Eds.), Design *Knowing and Learning: Cognition in Design Education* (79-103). Elsevier: Oxford, UK.
6. Simon, H. (1973). The Structure of Ill Structured Problems. *Artificial Intelligence, 4, 181-201*.
7. Adams, R. & Atman, C. (2000). Characterizing Engineering Student Design Processes: An Illustration of Iteration. *107th ASEE Annual Conference & Exposition*. Presented at the American Society for Engineering Education Annual Conference, St. Louis, MI.
8. Schimpf, C. & Xie, C. (2017). Characterizing Students' Micro-Iterations Strategies through Data-Logged Design Actions. *124th ASEE Annual Conference & Exposition.* Presented at the American Society for Engineering Education Annual Conference, Columbus, OH.
9. Dorst, K. & Cross, N. (2001). Creativity in the design process: the co-evolution of problem-solution. *Design Studies, 22(5),* 425-437.
10. Cinlar, E. (2013). *Introduction to Stochastic Processes.* Prentice Hall: Englewood Cliffs, NJ.
11. Daltrozzo, J., & Conway, C. M. (2014). Neurocognitive mechanisms of statistical-sequential learning: what do event-related potentials tell us? *Frontiers in Human Neuroscience*, *8*.
12. Keele, S. W., Ivry, R., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychological Review*, *110*(2), 316–339.
13. Clegg, B. A., DiGirolamo, G. J., & Keele, S. W. (1998). Sequence learning. *Trends in Cognitive Sciences*, *2*(8), 275–281.

14. Boyer, M., Destrebecqz, A., & Cleeremans, A. (2005). Processing abstract sequence structure: learning without knowing, or knowing without learning? *Psychological Research*, *69*(5–6), 383–398.
15. Remillard, G. (2008). Implicit Learning of Second-, Third-, and Fourth-Order Adjacent and Nonadjacent Sequential Dependencies. *Quarterly Journal of Experimental Psychology*, *61*(3), 400–424.
16. Destrebecqz, A. (2005). The neural correlates of implicit and explicit sequence learning: Interacting networks revealed by the process dissociation procedure. *Learning & Memory*, *12*(5), 480–490.
17. Destrebecqz, Arnaud, & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review*, *8*(2), 343–350.
18. Cornwell, B. (20150. *Social Sequence Analysis: Methods and Applications.* Cambridge University Press, New York, N.Y.
19. McComb, C., Cagan, J., & Kotovsky, K. (2017). Utilizing Markov Chains to Understand Operation Sequencing in Design Tasks. In *Design Computing and Cognition'16* (pp. 401–418). Cham: Springer.
20. McComb, C., Cagan, J., & Kotovsky, K. (2017). Capturing human sequence-learning abilities in configuration design tasks through markov chains. *Journal of Mechanical Design*, *139*(9).
21. McComb, C., Cagan, J., & Kotovsky, K. (2017). Mining Process Heuristics from Designer Action Data via Hidden Markov Models. *Journal of Mechanical Design*, *139*(11).
22. Gero, J. S., & Peng, W. (2009). Understanding behaviors of a constructive memory agent: A Markov chain analysis. *Knowledge-Based Systems*, *22*(8), 610–621.
23. Purzer, S. & Fila, N.D. (2013). Innovation Process Mapping Protocol: An Approach to Assessing Students' Understanding of Innovation as a Process. 120th ASEE Annual Conference & Exposition. Presented at the American Society for Engineering Education Annual Conference, Atlanta, GA.
24. McComb, C., Cagan, J., and Kotovsky, K. (2015). "Lifting the Veil: Drawing Insights About Design Teams From a Cognitively-Inspired Computational Model," *Design Studies* 40, pp. 119–142.
25. Xie, C. Schimpf, C. Chao, J. Nourian, S. & Massicotte, J. (2018). Learning and teaching engineering design through modeling and simulation on a CAD platform. *Computer Applications in Engineering Education*, pp. 1-17.
26. The Concord Consortium. (2016). The JSON Data Schema That Encodes Energy3D Design Processes. http://energy.concord.org/energy3d/schema/schema.pdf