Modeling Visitor Behavior in a Game-Based Engineering Museum Exhibit with Hidden Markov Models

Mike Tissenbaum
UW–Madison
225 North Mills St
Madison, WI
miketissenbaum@gmail.com

Vishesh Kumar
UW–Madison
225 North Mills St
Madison, WI
vishesh.kumar@wisc.edu

Matthew Berland
UW–Madison
225 North Mills St
Madison, WI
mberland@wisc.edu

ABSTRACT
Research has shown that supporting tinkering and exploration promotes a wide range of STEM related literacies. However, the open-endedness of tinkering environments makes it difficult to know whether learners' exploration is productive or not. This is especially true in museum spaces, where dwell times are short and facilitators lack a history of engagement with individual visitors. In response, this study uses telemetry data from Oztoc – an open-ended exploratory tabletop exhibit in which visitors embody the roles of engineers who are tasked with attracting and cataloging newly discovered aquatic creatures by building working electronic circuits. This data is used to build Hidden Markov Models (HMMs) to devise an automated scheme of identifying when a visitor is behaving productively or unproductively. Evaluation of our HMM was shown to effectively discern when visitors were productively and unproductively engaging with the exhibit. Using a Markov model, we identify common patterns of visitor movement from unproductive to productive states to shed light on how visitors struggle and the moves they made to overcome these struggles. These findings offer considerable promise for understanding how learners productively and unproductively persevere in open-ended exploratory environments and the potential for developing real time supports to help facilitators know how and when to best engage with visitors.

Keywords
Learning analytics, museums, interactive tabletops modeling.

1. INTRODUCTION
While there is evidence that digitally-augmented museum spaces can enhance science learning [36, 11], there is increased interest in how less-structured, open-ended designs can support new forms of STEM-based (science, technology, engineering, and math) reasoning and collaboration [18, 19]. Tinkering, in particular, often characterized by playful, experimental, iterative styles of engagement, and iterative, investigative processes of learning and discovery, has shown considerable promise in helping novices develop engineering and computer science literacies [5, 26].

Tinkering is an ideal complement to the kinds of learner-centered constructivist pedagogy found in many hands-on science museums [1]; however, in the open-ended and exploratory tasks that typify tinkering, assessment and feedback is particularly difficult [8]. This is especially true in museum environments, as visitors often do not have the expertise or confidence to conduct the coherent, in-depth investigations required to answer their questions on their own [2]. As such, within open-ended environments there is a growing need to develop methods for understanding learners' tinkering and exploration.

Digitally mediated museum spaces, when properly instrumented, can capture data on visitors' tinkering and experimentation in real-time (known as telemetry data), allowing researchers to identify and analyze temporal patterns in visitor interactions. We can then begin to investigate which patterns might be classified as productive (e.g., moving towards the broader learning goals of the exhibit) or unproductive (e.g., [23]). However, by their very nature, productive and unproductive states within open-ended tinkering activities are inherently difficult to classify.

One approach to understanding the state of a learner is through Markov Modeling [4]. Markov modeling is used to characterize patterns of sequential activity, but first-order Markov models only consist of sequences of known states, and we are often more interested in more complex relationships than just sequences of concrete data. One approach to finding hidden states in learners' activities is the use of Hidden Markov Models (HMM – [25]). Applying HMM to learning processes allows us to consider a learner as being in one of a fixed set of ("hidden") states at any moment in time. These models, are particularly well suited for museums as individual visitors' states are particularly hard to capture and pre- and post-tests are problematic if we want to ensure a naturalistic setting [9]. In response, the paper advances a research trajectory in which we attempt to highlight productive and unproductive patterns of visitor interactions by mining their telemetry data from an interactive tabletop exhibit at a large urban interactive science museum. In particular, this research addresses the following questions: *1) Can a Hidden Markov Model accurately predict if visitors are productively or unproductively engaged in an open-ended museum activity? 2) Can we identify the patterns of exploration and tinkering visitors engage in when they move from unproductive to productive states?*

2. BACKGROUND & PRIOR WORK
Within the context of this study, it is important to understand what we consider to be "productive" or "unproductive" patterns of practice. Within the learning sciences, there is interest in practices that can be considered productive for novices who are learning computer sciences and engineering [5]. With its focus on the *processes* of creative and improvisational exploration and making, tinkering is recognized as a means for developing a wide range of STEM literacies [22, 13]. Tinkering is predicated on engaging learners in activities centered on the use of scientific tools, processes, and phenomena to explore a problem space through experimentation, trial and error, and refinement [6, 10, 5].

With tinkering's focus on open exploration and learner-defined goals, understanding how and when a learner is engaged in productive tinkering is a challenge. For instance, making mistakes in "traditional" learning environments is often viewed as failure, but in tinkering environments, failure is not only tolerated but celebrated [26]. At their core, tinkering-focused environments enculturate the notion that learners should be allowed to persevere through initial struggles. However, it is not simply that learners persist, but *why they are persisting* and *how they are persisting* [27]. With persistence, it is critical that learners actively move towards new solutions or problem conceptualizations, or they risk getting stuck in cycles of unproductive perseverance [23].

In museum settings, understanding when visitors are engaging in productive versus unproductive practices *and* having museum facilitators monitor these states is a challenge. This is especially true in open-ended exploratory exhibits in which multiple visitors can engage and leave at different times (rather than having well-defined beginning and end points) and can interact with the exhibit at multiple granularities (e.g., alone, in groups, or simultaneously with strangers). However, if we can develop ways for capturing visitors' hidden productive and unproductive states, we open up the possibility for understanding underlying patterns in their tinkering and learning and providing critical information to researchers, designers, and museum facilitators.

2.1 Tabletop Interfaces and Engineering

There is significant research into the role the "programming" environment plays in supporting novices in exploring and tinkering when learning computer science and engineering [20, 5]. Tangible engineering platforms, such as "snap together circuits" (e.g., *snapcircuits.net*), allow novices to physically manipulate objects as they tinker and explore engineering concepts, providing clear feedback on their process (with pieces clearly fitting together, or lighting up when properly connected). Such interfaces can reduce learner overhead, freeing them to focus on exploration.

With their ability to support multiple visitors simultaneously and in promoting social interactions, interactive tabletops are increasingly used in science and engineering museum research [9, 1]. In general, interactive tabletops are well suited for supporting engineering practices as they promote greater co-awareness of peers' work [35], and can provide increased opportunities for others to monitor and provide feedback [20, 33]. The addition of tangible blocks (blocks that are recognized by the tabletop when placed on its surface) can further support visitors' engagement with engineering practices by allowing them to quickly try out ideas [16] and more generally explore and tinker.

While tabletops are great for supporting collaborative engineering learning, they can make it more difficult for museum explainers to know the state of tinkering of any one visitor. Similar to the problems teachers face with laptop lids [29], the flat surface of the multitouch tabletop can obscure visitors' interactions, forcing explainers to "hover" in order to know what visitors are doing. Even if explainers do hover, keeping track of multiple visitors' states manually (to know when and where they are needed) would be nearly impossible. In response, we need to develop models that can give us insight into visitor sates, particularly in real-time.

2.2 Markov and Hidden Markov Models

A Markov decision process (MDP) is defined by its state set S, and transition probabilities P [41] – assuming identical actions between states, and identical rewards for each transition. This is represented as a graph, called a Markov Model, which depicts that given a state *s*, the probability of transitioning to any of the other states *s'* is T(s, s'). In a Markov model, transition probabilities are calculated given a sequence of user states. Calculating (and then visualizing) the likelihood of a transition between states has many potential uses: identifying optimal action sequences in Intelligent Tutoring Systems towards success and using these to provide hints to users [3]; or classifying and identifying common student errors and technical problems to reduce their occurrence [15].

Hidden Markov Models (HMMs), as their name suggests, are Markov Models of *hidden* states. These are not directly observed in the input sequences, but, rather, they exist as aggregated "descriptions" of a user's visible states or "action events" [17]. These have been used to classify users through their navigation or content access patterns [12] and characterize student behaviors in computer-based inquiry learning environments [17]. HMMs require: an input sequence of visible states; an initial transition table providing a starting estimate for the transition probabilities between the hidden states; and an emission table with the probabilities of each of the visible states given each hidden state. Initialization and verification for an HMM-based learning model is an important step, as inappropriate initialization might result in the model getting stuck in local minima [7]. After appropriate initialization via the transition and emission tables, the HMM labels each input state with the corresponding hidden states, and gives the transition probabilities between the hidden states.

3. DESIGNING AN OPEN-ENDED TABLETOP ENGINEERING EXHIBIT

3.1 The *Oztoc* Exhibit

In order to address our research goals, we are building upon an existing multitouch tabletop exhibit at a large urban science museum. The exhibit, named *Oztoc* [19], situates visitors as electrical engineers called in to help fictional scientists who have discovered an uncharted aquatic cave teeming with never-before documented species of aquatic life (Figure 1). The creatures who live in this cave are bioluminescent, and visitors are asked to help design and build glowing "fishing lures" to attract the "fish" so that scientists can better study them. Visitors place wooden blocks, which act as electrical components (i.e., batteries, resistors, Light Emitting Diodes or LEDs, and timers), on the interactive table to create simple circuits (which the table recognizes the blocks via fiducial symbols – see Figure 1).

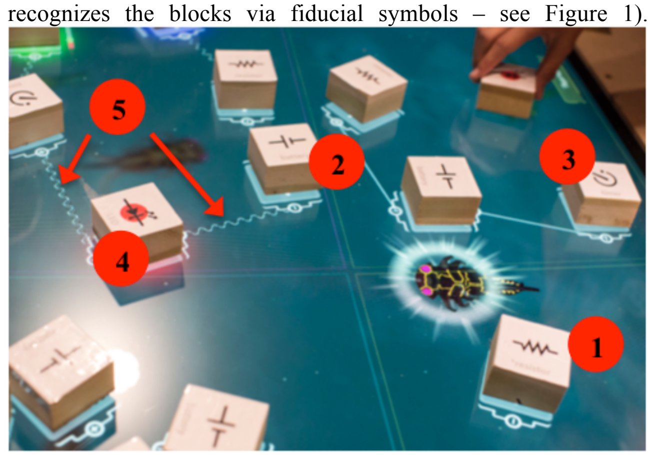

Figure 1. Visitors assemble virtual circuits using wooden blocks that represent resistors (1), batteries (2), timers (3), and different colored LEDs (4). Visitors make circuit connections (depicted as lines on the tabletop - 5) by bringing the positive and negative terminals of the blocks (augmentations displayed by the table) in contact with one another. Creating a successful circuit (one that has the correct ratio of resistors, batteries, and LEDs) causes LEDs to glow and lures creatures attracted to it for cataloging.

Oztoc's narrative aims to give learners a situated context in which to engage in engineering practices. To avoid many of the problems of other engineering and making exhibits [19], we wanted *Oztoc* to give visitors some freedom in choosing their own goals (e.g., which types of fish to target) while still giving them a common set of materials and processes.

4. METHODOLOGY AND VISITORS

Oztoc is installed in an enclosed exhibit space just off the main floor of a large urban science center. A lollipop sign just outside the exhibit space indicates when videotaping will take place in the exhibit, allowing visitors to decide to enter or to return when data collection is not active. Researchers were present for technical support to museum staff only. Video data was collected via cameras placed in the exhibit space, audio from a boundary microphone, and telemetry data using the ADAGE system [31].

Visitors in this study come from a wide range of backgrounds and SES. Visitors were also multi-generational and came to the exhibit alone, as families, and in large groups.

4.1 Establishing Visitor Start and Stop Times

Unlike many other exhibits, *Oztoc* does not have pre-determined start and stop events (such as the beginning or end of a simulation or game) – it is a continual process in which visitors enter and leave, often at different times. Therefore, in order to accurately separate visitors' sequences of activities, we developed a method for determining when visitors entered or exited the exhibit. Given all actions performed at each of the table's four "zones" over a single day, we found that if a zone was inactive and empty over a set period of time – the "inactivity interval" (InI), the next event in that zone indicated a new visitor. We evaluated an InI ranging from 10-120 seconds, and the InI did not change significantly between 45-120 seconds. As such, we validated the 45-second InI with hand-labeled data. Our 45 second InI achieved full accuracy for the 2-hour sample of video data that we hand-labeled.

4.2 Coding Visitor Events

We needed to establish a granularity of the telemetry data that would allow us to understand the state of visitors' tinkering at any moment. Based on previous research on visitors' interactions with the exhibit [19], we chose to look at the events when visitors successfully created a circuit (denoted in the logs as *MakeCircuitCreate*). This state was particularly useful as a circuit was logged in ADAGE *even if the circuit "didn't work"* (i.e., the LEDs were not supplied correct voltage), giving us insight into visitors' process exploring different circuit configurations, solution states, and goals. By leveraging visitors' histories at the table, we could mine for more complex relationships between their current circuit, previously made circuits, and those made by others at the table since their arrival. We then automatically coded each visitors' MakeCircuitCreate event using four binary codes (see Table 1).

Table 1. Binary codes for MakeCircuitCreate events

Marker	Code	Description
Is the circuit **c**omplex?	S/C	The completed circuit has 3+ components
Does the circuit **w**ork?	N/W	The circuit successfully lights up
Is the circuit **u**nique for self?	R/U	This is the first time the visitor has made this circuit
Is the circuit unique at the table?	E/O	No one else at the table has made a circuit with the same set of components

4.2.1 Is the circuit complex? (coded S or C)

Earlier analysis of visitors' interactions with the exhibit showed that most visitors (if they made *any* circuits) only made the basic three-component circuit (one LED, one resistor, and one battery) [34]. As such, the building of a complex (more than three component) circuit was a key indicator that visitors were trying out more complex configurations. If a circuit had three or less components we scored it an **S** (*indicating it was "simple"*), any circuit that had more than three components was scored a **C** (*indicating it was a complex circuit*). It is important to note that this code is not concerned with *whether or not the circuit works*, only the number of components used.

4.2.2 Does the circuit work? (coded N or W)

Understanding the relationship between the individual components and making a working circuit is a critical factor in determining the success of an exploration. As such, each completed non-working circuit was coded with an **N** and each completed working circuit with a **W**.

4.2.3 Is the circuit unique for self? (coded R or U)

Because problem solving through tinkering is characterized by exploration and iteration [26], we coded if a circuit created by a visitor was "unique" for them (i.e., had they constructed the exact same circuit earlier). A visitor who received a **W** on the *does the circuit work* code might seem to be engaging in productive tinkering; however, if they are simply repeating their first circuit over and over, this might indicate a failure to try out new ideas or expand their problem definition. To mark if a visitor's circuit was unique we coded it with a **U**, if it was a repeat of a past circuit we assigned it an **R.**

4.2.4 Is the circuit unique at the table?

Finally, *Oztoc* is designed to support visitors in collaborating with and building off others' to advance their own exploration. This use of others' constructed artifacts as a basis for one's own work has been termed "echoing" and has been shown to be an important part in open-ended and exploratory tinkering [34]. We considered a circuit to be an echo if it had the same number of each component type (battery, resisters, and LEDs). If a visitor's circuit echoed of one of their peers', we assigned it an **E** (for echo); if the circuit was unique to the table, we assigned it an **O** (for original).

The process described above resulted in every *MakeCircuitCreate* event for each visitor receiving an easily interpretable four-digit code. For instance, a *MakeCircuitCreate* that was assigned a code of **SWRO** means that it was a simple (**S**), working circuit (**W**) that was a repeat of a past circuit made by the visitor (**R**), but had not been created by anyone else at the table since this visitor started playing (**O**). These codes provided a rich and detailed source of data for passing into a Hidden Markov Model to see if we could identify if visitors were productive or unproductive at any point during their engagement with the exhibit. Since the *MakeCircuitCreate* events were chronologically ordered and separated per visitor, we could further examine which created circuits led to important state shifts.

4.3 Coding for productive behaviors

Using the coded descriptions of the circuits created by the visitors, we wanted to make an HMM that identifies when a visitor was behaving "productively", or not. For this purpose, building off of previous research [19], two members of the research team discussed and identified patterns of *MakeCircuitCreate* that were indicative of productive and unproductive tinkering.

One of the key patterns identified focuses on visitors trying out new circuit configurations to fix errors in their existing circuits or to develop new circuits (denoted by a **U** in codes). For instance, if a visitor attempted a few different non-working circuits – seen as a sequence of **SNUO**, **SNRO**, **SNUO, SNUO** (with the second circuit being a duplicate of a past circuit) – the sequence seems to indicate that while the visitor's circuits do not work (indicated by the **N**s), they are trying out new approaches and expanding their exploration. This sequence of activities was coded as productive behavior. If the visitor's continued exploration results in cycle of repeated circuits coded with **R**s (repeats) or did not eventually make a working circuit (coded with a **W**), we coded these actions as falling into unproductivity, as the visitor seems to have failed to figure out how to make a working circuit.

Similarly, a visitor might make a working circuit (indicated by a **W** in their circuit code) and repeat it over and over again (e.g., a series of circuits such as **SWRO**, **SWRO**, **SWRO**). This would seem to indicate that the visitor is repeating past success and is failing to consider new problem spaces or avenues for exploration.

A change of **SNRO** to **SNOE** – trying a new (**U** = self-unique) circuit that someone else on the table has made (**E** = table-echo), might be an attempt at gaining understanding by looking at what other visitors are doing – and was coded as productive depending on how many failed attempts the visitor had already made.

With this understanding, the first two authors first coded 200 circuit creates, and established reliability with 91% agreement. They then coded 644 of the (total of 3952) circuits made in player one's zone (one of the four game quadrants) on the table.

4.4 Training the Hidden Markov Model

We used our manually coded states to calculate appropriate values for the emission table for our HMM. The emission table was calculated by seeing how often a certain circuit code was marked as productive (or unproductive) as a proportion of all the circuits coded with the same hidden state. For instance, of all the circuits coded as productive, 5.6% of those were coded as CWUO and 6.49% were coded as CWUE (from the list of 16 circuit-codes), these values were then used to populate the HMM emission table.

We needed to identify when new visitors started playing at the table to ensure that the new visitors circuits were not considered as a continuation of earlier visitors. To do this we added events (a **0000** code) in the sequence of circuit-codes to signify new visitors. This brought up the question of whether the HMM should code new visitors as unproductive, productive, or another state altogether. To be able to show what state people tended to leave and begin at in the final transition table, we chose to make the visitor change a distinct state in our HMM even though it was not a hidden state, and is equivalent to a direct observation.

We used Python's hmmlearn package to create our HMM, which has the limitation of only looking for local optima in calculating the probabilities of transitioning from one hidden state to another. To account for this, different initial transition table values were tried. Results showed that the HMM stably converged to the final transition table (Figure 3).

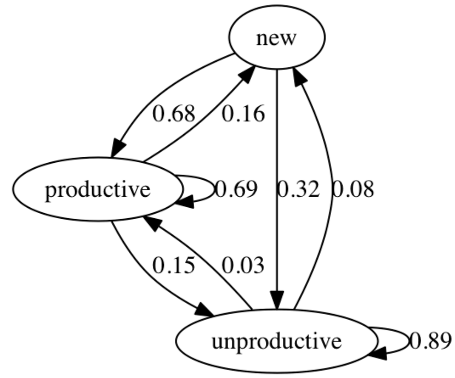

Figure 3. HMM for productive/unproductive states in Oztoc

5. FINDINGS

This study has two important findings, with the first finding acting as the scaffold for the second: First, the recognition of when visitors are engaged in productive or unproductive exploration; and second, the understanding of which sequences of events typically lead visitors from prolonged (at least three) consecutive unproductive states to a productive state.

5.1 Running HMM on Visitors' Circuits

The result of the HMM's final transition table revealed several interesting results (Figure 3). The HMM model shows that the probability of a new visitor beginning productively is 68%, versus 32% for beginning unproductively. Being unproductive appears to be a more stable state than being productive (89% versus 69%, respectively), and moving from unproductivity to productivity is also rarer than the reverse (3% versus 15%). The model also shows that the chances of leaving the table while being productive is higher than of leaving while unproductive (16% versus 8%).

To validate the predictive accuracy of the HMM's classification we used a general agreement score, the calculated the area under the curve (AUC) of the model's receiver operating characteristic (ROC) and Cohen's Kappa as compared to our 644 hand-coded labels. Our HMM had 94% agreement, scored an ROC/AUC score of 0.79, and a Cohen's Kappa of 0.59, which were

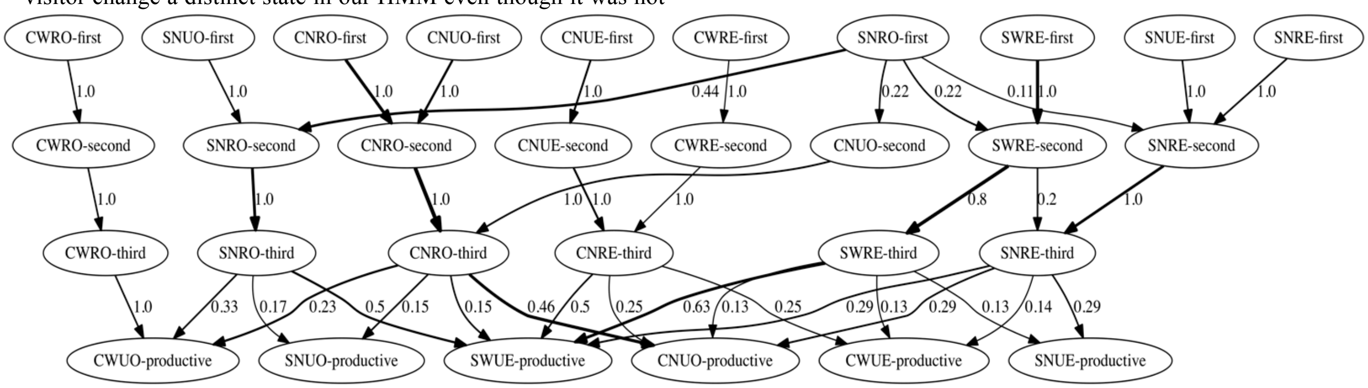

Figure 4. Markov model for visitors who transition from three consecutive unproductive states to a productive state

satisfactory measures to consider the HMM's coding reliable.

5.2 Developing Markov Models of Moving from Unproductive to Productive States

After the HMM tagged the circuits as productive or unproductive, we wanted to understand what patterns of activity preceded visitors becoming productive. We were particularly interested in sequences in which visitors struggled (had several unproductive moves) and then moved to a productive state. For this, we built a list of when a visitor had three consecutive unproductive circuits immediately followed by a productive circuit. We pruned the sequences that only happened once (as they were uninformative).

Once we had a list of the 4 step chains, we made a Markov model depicting the sequences of actions visitors followed when moving from unproductive to productive (Figure 4). This model also showed the likelihood that a visitor making a certain coded circuit would make another specific circuit next. The thickness of the lines between nodes indicates how many times a path occurred.

6. DISCUSSION

This paper outlined how the combination of Hidden Markov Models (HMMs) and Markov chains could be used to effectively predict when visitors were engaging productively or unproductively in an open-ended, exploratory museum exhibit. A closer examination of the HMM revealed several unexpected visitor behaviors. Visitors more often than not (68%) begin productively, but are less likely to stay productive (69%) than unproductive (89%) once in that state (Figure 3). The first finding is not entirely surprising, as our model considers open, thoughtful exploration as productive and it is hard to consider a visitor's "first move" as anything more than a first "exploratory step". This view is partially validated by the lower likelihood of staying productive – indicating many visitors fail to make thoughtful adjustments to their tinkering or explore new definitions of the problem space. This is compounded by instances where visitors make a successful circuit then "settle into" making the same circuit over and over. These findings are supported by the high percentage of visitors who either stay unproductive (89%) or leave the exhibit (8%). It should be noted that 69% is still a very high number of visitors staying productive and is probably further understated by the "first circuit" effect described above.

Another interesting finding is the high likelihood of leaving the table while being productive (16% compared to leaving the table while unproductive – 8%). On the surface this is surprising, as one would expect visitors to give up due to frustration more often than while 'succeeding'. The results may indicate that visitors who "figure out" multiple facets of the exhibit continue to engage productively until they leave – some of these effects have been covered in other research on this project [19]. Another possible explanation is that visitors started to engage in productive behaviors (such as trying something new that they had not done before or echoing the work of another visitor) that didn't immediately result in positive feedback from the system (e.g., capturing a fish) and they gave up.

When looking at the Markov model of unproductive to productive states we uncovered several interesting sequences (see Figure 4). For instance, unproductive circuits coded as **CNUO** (complex, not-working, unique, original) always went to **CNRO** (complex, not-working, repeated, original), followed by another **CNRO**, which finally led 15% of the time to a productive **SWUE** – a simple, working circuit that they had never made earlier, but had been made on the table in front of them by someone else! This is an interesting phenomenon – that a visitor, after some initial failures at making working circuits with a high level of complexity, likely saw a simple working circuit made by someone else, and then switched to echoing that circuit. The ability to see the work of others helped them overcome their own unproductive exploration. We see similar patterns in the Markov chain sequences **SNUO** -> **SNRO** -> **SNRO** -> **SWUE**; and **SNUE** -> **SNRE** -> **SNRE** -> **SWUE**, highlighting the role that making the work of others engaged in parallel tasks visible can serve in helping visitors move from unproductive to productive states.

7. CONCLUSIONS AND NEXT STEPS

Tinkering and exploration are powerful ways for learners to engage in science and engineering practices [24]; however, supporting leaners to productively engage in open-ended learning is inherently difficult, especially in museums [13]. Much of this has to do with the inherent chaos of the museum environment – hundreds (even thousands) of visitors interact with an exhibit in a day, coming and going at different times, and with different expectations and goals. For facilitators in exploratory exhibits, keeping track of the flow of participants and the state of their individual and collective tinkering efforts is nearly impossible.

This paper illustrates how data mining and analytics can help disambiguate the actions of visitors in such exhibits and uncover the hidden states of their tinkering. In addition to shedding light into how visitors productively and unproductively tinker, this work holds considerable potential for developing new ways to support facilitators. Knowing when and how visitors are engaging in unproductive exploration can help us develop complementary applications to help facilitators know when and how they are most needed. Knowing how visitors tend to move from unproductive to productive states can further guide us in developing strategies and scaffolds to help facilitators better engage with visitors.

While tablet applications have been used to provide added contextual information and alert museum facilitators about the visitors' interactions with exhibits in real-time [30], they have done so only using surface features, without understanding visitors' exploration 'states'. By uncovering the particular ways that a visitor is struggling, and understanding the subtle ways they can be "nudged" towards more productive exploration, there is the potential for dramatically influencing visitors' exploration and learning. By interceding at moments where visitors are struggling or are likely to give up, we may increase visitors dwell time, which has been shown to increase their collaboration with others, and domain learning [9]. In response, we are developing a tablet application that uses our models to support facilitators in real-time to understand how such applications compare to approaches that rely only on surface measures and unmodeled log data.

8. REFERENCES

[1] Allen, S. (2002). Designs for learning: Studying science museum exhibits that do more than entertain. *Sci. Ed. 88*, S1 (2004), S17-S33.

[2] Allen, S., & Gutwill, J. P. (2009). Creating a program to deepen family inquiry at interactive science exhibits. *Curator: The Museum Journal*, 52(3), 289-306.

[3] Barnes, T., & Stamper, J. (2008, June). Toward automatic hint generation for logic proof tutoring using historical student data. In *Intelligent tutoring systems* (pp. 373-382). Springer Berlin Heidelberg.

[4] Baum, L. E., & Petrie, T. (1966). Statistical inference for probabilistic functions of finite state Markov chains. *The annals of mathematical statistics*,*37*(6), 1554-1563.

[5] Berland, M., Martin, T., Benton, T., Smith, C. P., & Davis, D. (2013) Using Learning Analytics to Understand the Learning Pathways of Novice Programmers. *Journal of the Learning Sciences 22*(4), 564-599.

[6] Bevan, B., Gutwill, J. P., Petrich, M., & Wilkinson, K. (2015). Learning through STEM-rich tinkering: Findings from a jointly negotiated research project taken up in practice. *Science Education, 99*(1), 98–120.

[7] Bicego, M., & Murino, V. (2004). Investigating hidden Markov models' capabilities in 2D shape classification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *26*(2), 281-286.

[8] Blikstein, P. (2013, April). Multimodal learning analytics. In Proceedings of the third international conference on learning analytics and knowledge (pp. 102-106). ACM.

[9] Block, F., Hammerman, J., Horn, M., Spiegel, A., Christiansen, J., Phillips, B., ... & Shen, C. (2015, April). Fluid Grouping: Quantifying Group Engagement around Interactive Tabletop Exhibits in the Wild. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 867-876). ACM.

[10] Dorn, B., & Guzdial, M. (2010). Learning on the job: Charactering the programming knowledge and learning strategies of Web designers. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 703–712). New York, NY: ACM.

[11] Feder, M. A., Shouse, A. W., Lewenstein, B., & Bell, P. (Eds.). (2009). *Learning Science in Informal Environments: People, Places, and Pursuits.* National Academies Press.

[12] Fok, A. W., Wong, H. S., & Chen, Y. S. (2005, July). Hidden markov model based characterization of content access patterns in an E-Learning environment. In *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on* (pp. 201-204). IEEE.

[13] Gutwill, J. P., & Allen, S. (2010). Facilitating family group inquiry at science museum exhibits. *Science Education*, *94*(4), 710-742.

[14] Gutwill, J. P., Hido, N., & Sindorf, L. (2015). Research to practice: Observing learning in tinkering activities. *Curator: The Museum Journal*, *58*(2), 151-168.

[15] Heathcote, E. A., & Prakash, S. (2007). What your learning management system is telling you about supporting your teachers: monitoring system information to improve support for teachers using educational technologies at Queensland University of Technology.

[16] Hornecker, E., & Buur, J. Getting a grip on tangible interaction: a framework on physical space and social interaction. In *Proc. CHI 2006,* ACM Press (2006), 437-446.

[17] Jeong, H., Gupta, A., Roscoe, R., Wagster, J., Biswas, G., & Schwartz, D. (2008, June). Using hidden Markov models to characterize student behaviors in learning-by-teaching environments. In *Intelligent Tutoring Systems* (pp. 614-625). Springer Berlin Heidelberg.

[18] Land, S. M. (2000). Cognitive requirements for learning with open-ended learning environments. *Educational Technology Research and Development*,*48*(3), 61-78.

[19] Lyons, L., Tissenbaum, M., Berland, M., Eydt, R., Wielgus, L., & Mechtley, A. (2015, June). Designing visible engineering: supporting tinkering performances in museums. In *Proceedings of the 14th International Conference on Interaction Design and Children* (pp. 49-58). ACM.

[20] Maloney, J., Resnick, M., Rusk, N., Silverman, B., & Eastmond, E. (2010). The Scratch Programming Language and Environment. *ACM Transactions on Computing Education (TOCE)*, vol. 10, no. 4 (November 2010).

[21] Martin, L. (2015). The promise of the Maker Movement for education. *Journal of Pre-College Engineering Education Research*, *5*(1), 30-39.

[22] Martinez, S. L., & Stager, G. (2013). Invent to learn: Making, tinkering, and engineering in the classroom.

[23] McFarlin, D. B., Baumeister, R. F., & Blascovich, J. (1984). On knowing when to quit: Task failure, self-esteem, advice, and nonproductive persistence. *Journal of Personality*, *52*(2), 138-155.

[24] Petrich, M., Wilkinson, K., & Bevan, B. (2013). It looks like fun, but are they learning. *Design, make, play: Growing the next generation of STEM innovators*, 50-70.

[25] Rabiner, L. R., & Juang, B. H. (1986). An introduction to hidden Markov models. *ASSP Magazine, IEEE*, *3*(1), 4-16.

[26] Resnick, M., & Rosenbaum, E. (2013). Designing for tinkerability. *Design, make, play: Growing the next generation of STEM innovators*, 163-181.

[27] Ryoo, J. J., Bulalacao, N., Kekelis, L., McLeod, E., & Henriquez, B. (2015). Tinkering with “Failure”: Equity, Learning, and the Iterative Design Process. In *annual FabLearn conference. Palo Alto, CA: Stanford University.*

[28] Schweingruber, H., Keller, T., & Quinn, H. (Ed.). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*: National Academies Press, 2012.

[29] Sharples, M. (2013). Shared orchestration within and beyond the classroom. *Computers and Education*, *69*, 504-506.

[30] Slattery, B., Lyons, L., Pazmino, P. J., Silva, B. L., & Moher, T. (2014). How interpreters make use of technological supports in an interactive zoo exhibit. *In 11th International Conference of the Learning Sciences (ICLS 2014)*.

[31] Stenerson, M. E., Salmon, A., Berland, M., & Squire, K. Adage: an open API for data collection in educational games. In *Proc, SIGCHI Play 2014*, ACM Press (2014), 437-438.

[32] Sutton, R. & A. Barto. Reinforcement Learning: An Introduction, 1998, The MIT Press, Cambridge, MA.

[33] Tissenbaum, M., Berland, M., & Lyons, L. (*in review*). CCLM Framework: Understanding Collaboration in Constructionist Tabletop Learning. *International Journal of Computer Supported Collaborative Learning*.

[34] Wielgus, Tissenbaum & Berland (*in review*) Echoes paper

[35] Xambó, A., Hornecker, E., Marshall, P., Jorda, S., Dobbyn, C., & Laney, R. (2013). Let's jam the reactable: Peer learning during musical improvisation with a tabletop tangible interface. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(6), 36.

[36] Yoon, S. A., Elinich, K., Wang, J., Schoonveld, J. B., & Anderson, E. (2013). Scaffolding informal learning in science museums: How much is too much?. *Science Education*, 97(6), 848