Adding duration-based quality labels to learning events for improved description of students' online learning behavior

Matthew Guthrie and Zhongzhou Chen

May 27, 2019

Adding duration-based quality labels to learning events for improved description of students' online learning behavior

1st author
1st author's affiliation
1st line of address
2nd line of address
Telephone number, incl. country code
1st author's E-mail address

2nd author
2nd author's affiliation
1st line of address
2nd line of address
Telephone number, incl. country code
2nd E-mail address

ABSTRACT
Many existing studies analyzing log data from online learning platforms model events such as accessing a webpage or problem solving as simple binary states. In this study, we combine quality information inferred from the duration of each event with the conventional binary states, distinguishing abnormally brief events from normal or extra-long events. The new event records, obtained from students' interaction with 10 online learning modules, can be seen as a special form of language, with each "word" describing a student's state of interaction with one learning module, and each "sentence" capturing the interaction with the entire sequence. We used second order Markov chains to learn the patterns of this new "language," with each chain using the interaction states on two given modules to indicate the interaction states on the following two modules. By visualizing the Markov chains that lead to interaction states associated with either disengagement or high levels of engagement, we observed that: 1) disengagement occurs more frequently towards the end of the module sequence; 2) interaction states associated with the highest level of learning effort rarely leads to disengaged states; and 3) states containing brief learning events frequently lead to disengaged states. One advantage of the current approach is that it can be applied to log data with relatively small numbers of events, which is common for many online learning systems in college level STEM disciplines. Combining quality information with event logs is a simple attempt at incorporating students' internal condition into learning analytics.

Keywords
Markov chains, online learning modules, log data analysis

1. INTRODUCTION
Understanding and predicting students' learning behavior by mining the log files of online and computerized learning systems has been the focus of a significant body of research in educational data mining. For example, many studies have modeled student learning as a chain of ordered events such as opening a page, viewing a video, or solving a problem [8, 10]. Oftentimes, each event is represented by a binary variable, i.e. whether the student accessed a webpage or answered a problem correctly [8, 12, 17]. While describing events using binary variables can significantly reduce the complexity of the data, doing so also risks eliminating important information from the event logs. For example, a page visit lasting 10 seconds is most likely a qualitatively different event than a page visit of 10 minutes.

Can we include information about the "quality" of learning events into our data analysis approach to gain new insights into students' online learning behavior?

One readily available indicator of the quality of an event is its duration. For example, abnormally short problem-solving attempts have been associated with either random guessing due to low test-taking effort [6, 23] or answer copying [2, 15]. In two earlier studies [authors3, authors4], we demonstrated that students' learning events can be separated into "Brief" (B) and "Normal" (N) categories by applying a mixture model clustering algorithm using the time-on-task data alone, since most other measurements such as the number of practice problems answered are highly correlated with time-on-task.

In this study, we combine duration based categorical quality labels such as "Brief" or "Normal" with conventional binary event states, such as "Pass" or "Fail," to analyze the log data obtained from students' interaction with 10 Online Learning Modules (OLMs). Explained in detail in multiple previous papers [authors1-4], OLM is a new form of online instructional design in which students progress through a sequence of learning modules in a pre-determined order. Students are required to attempt the assessment problems at least once before accessing the accompanying learning resources. The restrictive structure of OLMs has two major advantages for data analysis. First, it provides more accurate estimations of duration information, since the start and end of each attempt or learning event is clearly marked by navigation events. More importantly, because these assessments and learning events are closely coupled on each module, the OLM structure improves the interpretability of log data events. For example, a single "Brief" learning event may result from either low levels of engagement with learning or high levels of incoming knowledge (and thus a student would not need to engage with instructional material). Yet if it is preceded by a failed brief attempt and followed by another failed brief attempt on the same module, it is much more likely that the student is not fully engaged with learning.

As will be explained in detail in the next section, the event logs combined with categorical quality labels can be interpreted as a simple artificial language. Each event, such as "Brief Pass" or "Normal Learning," becomes part of a "word" that captures students' interaction with either the assessments or the learning components of a module. Four such "words" form a "phrase" that describes a student's state of interaction with one section of the

OLM, and each 10 word "sentence" corresponds to a student's interaction with the entire OLM sequence. Can we gain insight into the patterns in students' learning behavior by understanding the underlying "grammar" of this artificial language? In particular, are there certain combinations of words in these phrases that are frequently followed by other words which indicate that the student is either disengaged or highly engaged with the learning from the OLM sequence?

We will answer this question by utilizing a second order Markov chain, a common technique used in natural language processing. First order Markov chains can be thought of as weighted random walks in some state space x over discrete timestep i where the current state is dependent only on the state occupied at time $i-1$. The probabilities of the state-to-state transitions are predetermined using some known information about the state space. For Markov chains of second order, the probability of being in some state S at time $i > 2$ is dependent on the previous two states,

$$P(S = x_i | x_{i-1}, x_{i-2}, \dots, x_0) = P(S = x_i | x_{i-1}, x_{i-2}).$$

Markov chains have been utilized in the analysis of large text corpuses such as the complete works of William Shakespeare or Arthur Conan Doyle [19]. In doing so, the data corpus is converted into a probabilistic representation, where words have a chance of following some given word or words depending on the original text and the depth (order) of the model. When unsupervised, this process almost always results in preposterous dialogue between characters from completely different plays or books, such as Friar Laurence (from *Romeo and Juliet*) discussing Fulvia's death with Cleopatra (both from *Antony and Cleopatra*), while the generated exchange is about a common theme (the death of crucial characters in both stories). When restricted to generating short phrases, predictive smartphone keyboards can use Markov chains of various order to suggest words based on the current input in a text field [1]. Given the input "I want," my phone suggested that I continue to type "a refund."

One advantage of our approach is that the Markov model can be trained using a relatively small number of events from a student population of approximately 250, owing to the increased information by introducing the quality labels. This advantage is critical for modeling student learning behavior for many STEM disciplines such as college level physics, where solving one problem can take 5 to 10 minutes. A typical weekly online homework assignment of 10 problems may only generate around 30 to 50 major events (excluding minor log events such as scrolling or navigating between pages).

In this paper, we will first demonstrate that it is possible to construct the Markov model using log data from two weeks' worth of homework assignments for a college physics class of approximately 250 students. Second, the resulting Markov chains can be visualized to reveal combinations of "words" that will lead to states associated with either disengagement from or high levels of engagement in subsequent modules.

It must be pointed out that "engagement" is a highly complex concept that has different definition depending on the context and the measurement method [3]. For this study we will adopt a very narrow and pragmatic definition of engagement to indicate that students spent a normal or extended amount of time (and likely also cognitive resources) in consecutive events on consecutive modules. Such a definition emphasizes the cognitive and behavioral aspect of engagement, which bears some similarity to the definition proposed by Miller [14].

2. METHODS

2.1 Structure of OLM and OLM Data

Data analyzed in this study were collected from student interaction with 10 OLMs assigned as homework to be completed over a period of two weeks in a calculus-based college physics class. Students are not required to finish the entire sequence in one session and are free to leave and return to the modules during the two weeks. However, the modules must be completed in the order given. As described in detail in several earlier papers [authors1-4], each OLM consists of an assessment component (AC) and an instructional component (IC). Students are required to attempt the AC at least once before being able to learn from the IC and can make additional attempts after interacting with the IC. Therefore, the majority of students' interactions with each OLM can be divided into three stages: ***Pre-Learning***: attempting the AC once or twice before accessing the IC, ***Learning***: interacting with the IC after one or two initial failed AC attempts, and ***Post-Learning***: making additional attempts on the AC after learning from the IC. If a student passes the AC during the Pre-Learning stage, they will not have the following two stages as the student will immediately proceed to the next OLM, as illustrated in Figure 1.

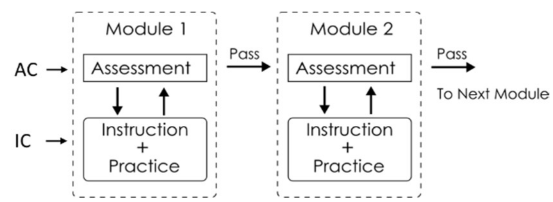

Figure 1: Schematic illustration of the structure of OLMs.

2.2 Combining Quality Labels with Event Logs

For each Pre-Learning and Post-Learning stage, students' interaction with the AC is captured by the attempt outcomes: "Passing" (P) or "Failing" (F). In addition, the quality of each attempt is estimated from the duration of the attempt, t, which is classified into three categories: "Brief" (B: $t < 40$ s) "Normal" (N: $40 \text{ s} \leq t < 180$ s), or "Extensive" (E: $t \geq 180$ s). The cutoffs are determined based on mixture-model clustering method applied to log-transformed attempt duration data. Combining the three quality categories with the two attempt outcomes resulted in six different states: BF, BP, NF, NP, EF, and EP. The EF and EP states are only assigned to the Post-Learning stages, since there were significantly fewer attempts with $t \geq 180$ s in the Pre-Learning stages, and it was less clear whether those longer attempts resulted from longer problem-solving time or students leaving the system. For a detailed explanation of those categories and cutoffs, please see [authors4].

For the Learning stage, students' interaction with the IC was modeled as a single learning event described by a binary variable. The duration of the learning event is classified as "Brief" (B) or "Normal" (N) according to cutoffs determined for each module by a mixture-model clustering analysis of learning time distribution [authors4]. Note that an isolated "Brief" category does not necessarily imply that the event is of lower quality. For example, brief learning can result from a student having a high level of incoming knowledge and only needed to quickly view the learning resources to answer the problem.

In the majority of cases, a student's interaction state in one module can be classified by a triplet of combined quality and event labels in the three stages. For example, a student could have made a very brief and failed initial attempt on the AC, spent a normal amount of time learning from the IC, then spent a normal amount of time attempting and passing the AC. This student's interaction with the OLM is classified as BF-N-NP. There are 26 possible triplet states, including BP- - and NP- - states which indicate that a student passed a module on their initial attempt.

Finally, in a small number of cases, students made 3 or more failed attempts on the AC before accessing the IC or kept attempting the AC until all attempts were used up without accessing the IC. Those cases are classified as "Other." In even fewer cases, due to either a corrupted log file or other system glitches, some students were able to proceed to the next module without finishing the current module. Those cases were classified as "NAOther," making a total of 28 possible states, listed in Table 1.

2.3 Defining States Associated with Either Disengagement or High Level of Engagement

For most of the interaction states, it is impossible to estimate the level of engagement associated with the state, and probably the same state can be observed from students with different levels of engagement. However, there are several states that are clearly more likely to be associated students with either a very low or a very high level of engagement with the learning process.

For example, since "Brief" problem solving occurs in under 40 s, it has a high probability of resulting from a guessing attempt, or answer copying event which will result in a BP- - state. Additionally, "Brief" learning events are more likely than "Normal" learning events to come from students who skimmed through the content. When a student displays consecutive "Brief" events on the same module, such as in state BF-B-BF, they are highly likely to be not fully engaged with the learning process.

Similarly, consecutive "Normal" or "Extensive" events, such as NF-N-EF, are more likely to come from students who are highly engaged with the learning process, as they devoted adequate or extensive amount of time to every stage of the learning process. While individual "Extensive" events may be caused by a student leaving the computer without exiting from the module, it is much less likely that three such events occur on the same module.

Table 1: List of all possible interaction states. D: Disengaged. E: Highly Engaged

State	Rank	Indication	State	Rank	Indication
NAOther	0	D	BF-N-EP	14	E
Other	1	D	NP- -	15	E
BP- -	2	D	NF-B-BF	16	
BF-B-BF	3	D	NF-B-BP	17	
BF-B-BP	4	D	NF-B-NF	18	
BF-B-NF	5		NF-B-NP	19	
BF-B-NP	6		NF-B-EF	20	
BF-B-EF	7		NF-B-EP	21	
BF-B-EP	8		NF-N-BF	22	
BF-N-BF	9		NF-N-BP	23	
BF-N-BP	10		NF-N-NF	24	E
BF-N-NF	11		NF-N-NP	25	E
BF-N-NP	12		NF-N-EF	26	E
BF-N-EF	13	E	NF-N-EP	27	E

As listed in Table 1, in this study we assumed that states with three consecutive "B" labels, or BP- - are more likely associated with disengagement. Similarly, states with three consecutive "N" or "E" labels or NP- - are likely associated with higher levels of engagement. Note that we did not distinguish between productive and unproductive engagement, as failed attempts are also included in high engagement states.

There are two exceptions to these rules: First, the "Other" state is classified as "Disengaged," since most engaged students should at least look at the instructional resources after 2 failed attempts. Second, the BF-N-EF and BF-N-EP states are classified as highly engaged, since it is possible that the student quickly decided that the assessment problem was too difficult and immediately engaged in the learning process.

2.4 Training of the Markov Model

Initial construction of the text corpus required asserting that each state occurs in temporal order; the ordering of states coincides with each student's trajectory through the modules. This assures that eventual training of a model using the module data will give temporally possible results. The construction was accomplished by appending the module number to each state: a state for a student in module 7 could look like BF-N-NP7.

We define four sequential states to be a "phrase." These phrases can be analyzed in the context of the specific modules in which they occurred. Modules 1-4 of the OLM sequence were, on average, less difficult than the final four modules of the sequence. As such, the phrases created by students in modules 1-4 and 7-10 are qualitatively different, on average, as student behavior adjusts to the module content. Ten sequential states are defined as a "sentence." Each student contributed a complete ten state sentence to the text corpus.

After the module information was appropriately formatted, the Markovify python library [13], described as a "simple, extensible Markov chain generator," was used to parse the text corpus. Markovify constructs Markov chains from text data and has been previously used for many purposes, such as the construction of titles and abstracts of hypothetical papers from the 18,000 most cited science articles according to the Web of Science [21], the analysis of debates in the United States 2016 presidential election [22], for creating twitter bots to simulate a social media attack [7], and in the creation of synthetic data to supplement a smaller data set [20]. These examples display the utility of Markovify for mimicking language expressed in an input text corpus and as a result it was deemed appropriate for the current study.

We utilized second order Markov chains in this study due to the consideration that students' interaction record on a single module (as would be modeled with a single order Markov chain) is unlikely to adequately account for the complexity of their subsequent behavior. The Markov model created with Markovify was used to build second order Markov chains for every possible combination of initial state in three starting positions: modules 1 and 2, modules 6 and 7, and modules 7 and 8. Many of the initial state combinations were not present in the original text corpus. For example, while most students start their module sequence with the combination of NP- -1 and NP- -2 states, there were no students who followed the pattern of NP- -1 and NF-N-BP2. The second order Markov chains were used to investigate the behavior of students as they progressed through the modules and to infer how changes in behavior can be related to student engagement levels.

3. RESULTS

3.1 Outcomes of Second Order Markov Chain

For a given pair of states on the two input modules, we use the Markov model to return the probability of observing different interaction states on each of the two following modules. An example case is visualized as a Sankey diagram [18] shown in Figure 2. The diagram shows that for the pair of input NP- -7 and NF-N-EP8 states, the Markov model showed two possible states on module 9 and 6 possible states on module 10.

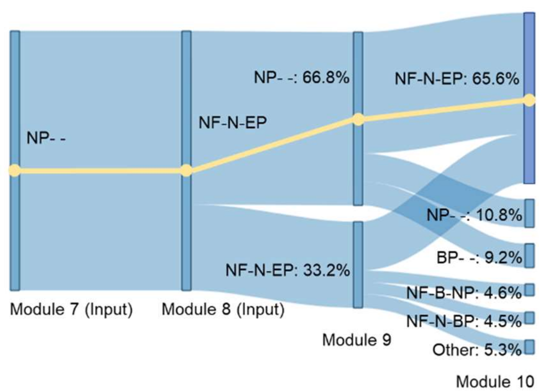

Figure 2: Sankey diagram of an example of resulting Markov chains given two input states on Modules 7 and 8. The yellow line indicates a probable chain.

For each Markov chain consisting of four states, we consider the chain as "probable" if the possibility of the last two states adds up to more than 100%. The only probable chain in Figure 2 is highlighted with a yellow curve. On average, 0.2% of all chains generated are considered probable.

The three cases for which we chose to run the Markov model are listed in Table 2. Those three cases are of particular interest because a previous analysis of the data [authors4] revealed that more students have lower levels of engagement in modules 3, 8, 9, and 10. The number of probable chains for each case is also listed in Table 2.

Table 2: Combinations of input states which were analyzed in this study.

Case	Input	Predict	All chains	Probable chains	Disengaged chains	Highly Engaged chains
I	M1+M2	M3+M4	55664	169	3	66
II	M6+M7	M8+M9	45472	65	18	15
III	M7+M8	M9+M10	72912	108	31	16

3.2 Markov Chains Leading to Consecutive Disengaged or Highly Engaged States

For this study, we are interested in chains that could indicate whether a student is either disengaged or highly engaged with the learning process. We only consider a student likely to be disengaged from the learning process if their interaction states are indicative of disengagement on both of the last two modules in the chain. Similarly, if a student's interaction states are associated with high levels of engagement on the last two modules, the student is considered as "highly engaged." The number of chains that lead to consecutive disengaged or highly engaged states on the last two modules are also listed in Table 2. The rest of the chains are not included in this analysis because the relation between students' level of engagement and the states on the last two modules weren't as clear as the chains included.

We plot all the chains indicating disengagement or high engagement for each of the three cases in Figure 3. In each case, the 28 interaction states are arranged on the y-axis according to the order listed in Table 1.

This ordering groups similar states according to their similarities in Pre-learning, Learning, and Post-learning stages, and listed from low to high following the order of "B," "N," and "E" in quality labels. States with passing events are assumed to be of higher engagement than those with failing events. States associated with disengagement are placed at the bottom and states associated with high engagement are placed at the top, with the exception of states 13-15 in Table 1. Note that for states near the middle of the pack, the ranking does not reflect the learning effort required for each state, as it is difficult to estimate whether 11: BF-N-NP requires more or less effort than 19: NF-B-NP.

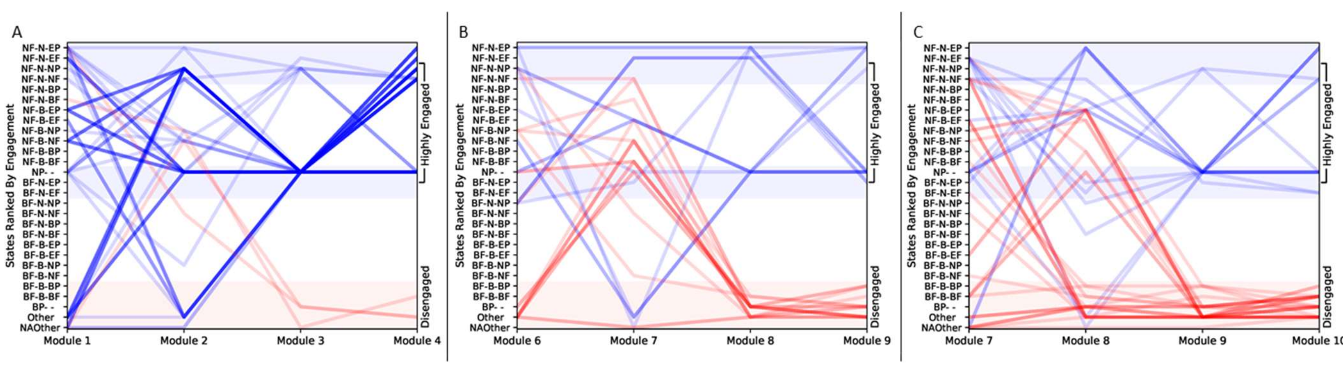

Figure 3: Probable chains that lead to consecutive disengaged states (red) or highly engaged states (blue) on the following two modules. Darker lines indicate where multiple chains overlap; blue and red zones highlight states associated with disengagement or high engagement, respectively.

For Case I (M1-M4), significantly more chains lead to highly engaged states on M3-M4 than disengaged states. On M1, those chains started from either a variety of states above NP- -, or from Other and NAOther. On M2, most of the chains concentrated on three states: 25: NF-N-NP, 15: NP- -, and 1: Other. While 25 and 15 were highly populated states in the original text corpus, state 1 was scarcely populated for M2. The results reflected that very few students were consistently disengaged on both M3 and M4, which appear early in the learning sequence and cover less difficult concepts.

For Case II (M6-M9), there were almost an equal number of chains leading to either highly engaged or disengaged states on M7 and M8. Notably, most of the chains leading to disengagement passed through one of the states between 15 and 24 on M7. More than half of those chains started in states 2 and 3 on M6. In contrast, several chains leading to high engagement started with high effort states on M6 and passed through Other or NAOther on M7. Finally, all of the chains (except one) starting from or passing through one of the top three states (25-27) lead to high engagement, and all of the chains starting with disengaged states led to disengagement on M8 and M9.

Finally, for Case III (M7-M10), there were more chains leading to disengaged states than highly engaged states on M9 and M10. The chains leading to disengagement started at a variety of states on M7, and forms two clusters on M8. The first cluster passes through disengagement between states 0 and 4, and the second cluster passes through states between 15-21. On the other hand, nearly all of the high engagement chains either started from state 26 on M7 or passed through state 27 on M8.

4. DISCUSSION

By inspecting and comparing the three graphs in Figure 3, we can identify four common patterns.

Disengagement happens late: chains leading to disengagement occur much more frequently on later modules in the sequence. This type of behavior is expected since the difficulty of the modules increase towards the end, yet each module is worth the same amount of course credit. Therefore, students have less incentive to devote effort on the harder modules.

Disengagement-free states: three states requiring the highest level of learning effort, 25-28, are seemingly "immune" to chains leading to disengagement. In all three cases, only two of those chains pass through these states, while most of the chains leading to high engagement involve those states on at least one input module. This shows that students who are observed to spend an extensive amount of effort on one module are more likely to be more persistent, especially on difficult modules towards the end.

Disengagement "hot zone": states 15-24 (white area between two blue bands in Figure 3) seem to be a "hot zone" for chains leading to disengagement in all three cases, especially when the state appeared on the second module in the sequence. Those states have a "Brief" label on either the learning stage or the post-learning stage, such as NF-B-NF. This could be evidence that a brief event in the learning and post learning stages is more indicative of subsequent disengagement behavior than a brief event in the pre-learning stage.

V-shaped high engagement chains: several chains leading to high engagement states started with states beyond 15 and passed through either 1: Other or 0: NAOther on the second module, forming a V-shape on Figure 3C. This may suggest that even highly engaged students may occasionally display disengaged states on certain modules. It also may suggest that the two "other" states are not always associated with disengagement as previously thought.

These patterns can be valuable for future developments of an intelligent and personalized learning system that recommends different learning resources appropriately to the correct student populations [12, 15]. For example, for students in the disengagement "hot zone" on key modules such as M7, the system could present encouraging messages or recommend supplementary learning resources to facilitate learning. On the other hand, for students in the "disengagement immune" zone, the system could recommend more advanced content by adding M11 and M12 to the sequence. The patterns can also help instructors in prioritizing future efforts in improving the OLMs, focusing more on modules such as M7 and M8 that are critical in students' decision to either persist or disengage from the process.

More importantly, we demonstrated that the Markov model can be trained using a small number of events collected from approximately 250 students over the period of two weeks. Adding duration-based quality labels is crucial for our approach, otherwise the model would only have 5 possible states, and would have treated distinct states such as NF-N-EP and BF-B-BP as identical. Moreover, since the majority of students attempted all the problems assigned for course credit, and accessed all the learning materials, only 3 of the 5 states would be heavily populated. Such a model, even if functional, would likely produce trivial or non-informative outcomes. The addition of duration-based quality labels to log events can be considered as a (very simple) attempt to consider the "effects of students' internal condition" in learning analytics, proposed by Gasevic et.al. [9]. In addition, the ability to train the model using data from students in the same class completing assignments on a single topic significantly reduces the effect of difference in instructional conditions on the results, which was also suggested by Gasevic et.al. in the same work.

On the other hand, as an exploratory first attempt, the current study has several notable caveats that needs to be investigated and addressed in more detailed future studies.

First, due to the limitation of computational capacity, only three pairs of modules were analyzed. Whether the patterns observed such as the "safe" and "hot" zones are general to all modules or specific to the selected cases can be answered by future studies analyzing all 7 pairs of input modules.

Second, as previously mentioned, we adopted a very narrow definition of "engagement," which simply means that students are spending an expected or extended amount of time on completing each component in a single module. While this crude definition is sufficient for the purpose of the current study, future work is needed to investigate the relation between time-on-task and engagement, and in finding new categorical labels that can better reflect students' levels of engagement.

Third, the current Markov model "predicts" student behavior based solely on their interaction states on preceding modules. In reality, students' decisions to engage or disengage from learning involves highly complex metacognitive processes influenced by a number of external factors including incoming knowledge, instructional condition, metacognitive skills, and emotional states [4, 5]. Just as more sophisticated predictive keyboards consider external conditions such as the currently open application and specific text field, future predictive models can achieve more accurate outcomes by including more factors that influence students' metacognitive processes.

Fourth, the current study utilizes a simple set of qualitative labels obtained from clustering algorithms on time-on-task data. More elaborate future studies are needed to investigate the validities of those labels, as well as find new and better-quality indicators for both the existing events and new events in other learning systems such as discussion forum posts.

Finally, the approach in the current paper relies on the restrictive structure of OLMs, which provides a regular and simple data structure, and allows for straightforward interpretation of some interaction states. However, the log file events and quality labels used to generate the artificial language can be obtained from essentially any online learning platform, and more sophisticated Markov models are capable of learning languages with many more irregularities. A valuable future research direction is to investigate to what extent the current method can be modified and applied to other more common learning systems that are more accessible to the average instructor.

5. ACKNOWLEDGMENTS

The authors would like to thank the Learning Science and Technology team at [Anonymous university] led by Dr. [Anonymous] for developing the [name redacted] platform and providing technical support for the current study.

6. REFERENCES

[1] Agarwal, S. 2007. Context based word prediction for texting language. *Large Scale Semantic Access to Content (Text, Image, Video, and Sound)*. (2007), 360–368.

[2] Alexandron, G. et al. 2017. Copying @ Scale: Using Harvesting Accounts for Collecting Correct Answers in a MOOC. *Computers & Education*. 108, (2017). DOI:https://doi.org/10.1016/j.compedu.2017.01.015.

[3] Azevedo, R. 2015. Defining and Measuring Engagement and Learning in Science: Conceptual, Theoretical, Methodological, and Analytical Issues. *Educational Psychologist*. 50, 1 (Jan. 2015), 84–94. DOI:https://doi.org/10.1080/00461520.2015.1004069.

[4] Azevedo, R. et al. 2018. Understanding and reasoning about real-time cognitive, affective, and metacognitive processes to foster self-regulation with advanced learning technologies. *Handbook of self-regulation of learning and performance (2nd ed.)*. (2018), 254–270. DOI:https://doi.org/10.4324/9781315697048.

[5] Azevedo, R. et al. 2015. Using Trace Data to Examine the Complex Roles of Cognitive , Metacognitive , and Emotional Self-Regulatory Processes During Learning with Multi-agent Systems. April 2013 (2015). DOI:https://doi.org/10.1007/978-1-4419-5546-3.

[6] Barry, C.L. et al. 2010. Do Examinees Have Similar Test-Taking Effort? A High-Stakes Question for Low-Stakes Testing. *International Journal of Testing*. 10, 4 (2010), 342–363. DOI:https://doi.org/10.1080/15305058.2010.508569.

[7] Bossetta, M. 2018. A simulated cyberattack on Twitter: Assessing partisan vulnerability to spear phishing and disinformation ahead of the 2018 U.S. midterm elections. *First Monday*. (2018). DOI:https://doi.org/10.5210/fm.v23i12.9540.

[8] Faucon, L. and Kidzi, Ł. 2016. Semi-Markov model for simulating MOOC students. *Proceedings of the 9th International Conference on Educational Data Mining (EDM)*. (2016).

[9] Gaševic, D. et al. 2015. Let ' s not forget : Learning Analytics are about Learning. *TechTrends*1, 59. (2015), 64–71. DOI:https://doi.org/10.1007/s11528-014-0822-x.

[10] Hansen, C. et al. 2017. Sequence Modeling for Analysing Student Interaction with Educational Systems. *Proceedings of the 10th International Conference on Educational Data Mining, Edm 2017, Wuhan, Hubei, China, June 25-28, 2017* (2017).

[11] Hoic-Bozic, N. et al. 2016. Recommender System and Web 2.0 Tools to Enhance a Blended Learning Model. *IEEE Transactions on Education*. 59, 1 (2016), 39–44. DOI:https://doi.org/10.1109/TE.2015.2427116.

[12] Karumbaiah, S. et al. 2018. Predicting Quitting in Students Playing a Learning Game. *EDM* (2018).

[13] Markovify: 2015. *https://github.com/jsvine/markovify*.

[14] Miller, B.W. and Miller, B.W. 2015. Using Reading Times and Eye-Movements to Measure Cognitive Engagement Using Reading Times and Eye-Movements to Measure Cognitive Engagement. 1520, (2015). DOI:https://doi.org/10.1080/00461520.2015.1004068.

[15] Palazzo, D.J. et al. 2010. Patterns, correlates, and reduction of homework copying. *Physical Review Special Topics - Physics Education Research*. 6, 1 (Mar. 2010), 010104. DOI:https://doi.org/10.1103/PhysRevSTPER.6.010104.

[16] Ricci, F. et al. 2015. *Recommender systems: Introduction and Challenges*.

[17] Sahebi, S. and Brusilovsky, P. 2018. Student Performance Prediction by Discovering Inter-Activity Relations. *Proceedings of the 11th International Conference on Educational Data Mining (EDM)* (2018).

[18] Schmidt, M. 2008. The Sankey diagram in energy and material flow management - Part II: Methodology and current applications. *JOURNAL OF INDUSTRIAL ECOLOGY*. (2008). DOI:https://doi.org/10.1111/j.1530-9290.2008.00015.x.

[19] SCHMIPSUM: *schmipsum.com*.

[20] Sibolla, B. et al. 2018. An Automated Approach to Mining and Visual Analytics of Spatiotemporal Context from Online Media Articles. (2018).

[21] topscience: 2017. *https://github.com/schollz/topscience*.

[22] Understanding Markov Chains with 2016 Presidential Debate Transcripts: 2016. *http://danicodes.github.io/2016/exploring-markov-chains/*.

[23] Wise, S.L. and Kong, X. 2005. Response Time Effort: A New Measure of Examinee Motivation in Computer-Based Tests. *Applied Measurement in Education*. 18, 2 (Apr. 2005), 163–183. DOI:https://doi.org/10.1207/s15324818ame1802_2.