

Identifying Learning Behaviors by Contextualizing Differential Sequence Mining with Action Features and Performance Evolution

John S. Kinnebrew
Department of EECS and ISIS
Vanderbilt University
1025 16th Ave S, Ste 102
Nashville, TN 37212
john.s.kinnebrew@
vanderbilt.edu

Gautam Biswas
Department of EECS and ISIS
Vanderbilt University
1025 16th Ave S, Ste 102
Nashville, TN 37212
gautam.biswas@
vanderbilt.edu

ABSTRACT

Our learning-by-teaching environment, Betty's Brain, captures a wealth of data on students' learning interactions as they teach a virtual agent. This paper extends an exploratory data mining methodology for assessing and comparing students' learning behaviors from these interaction traces. The core algorithm employs sequence mining techniques to identify differentially frequent patterns between two predefined groups. We extend this technique by contextualizing the sequence mining with information on the student's task performance and learning activities. Specifically, we study transformation of action sequences using action features, such as activity categorizations, relevance and timing between actions, and repetition of analogous actions. We employ a piecewise linear segmentation algorithm in concert with the action transformation and differential sequence mining techniques to identify and compare segments of students' productive and unproductive learning behaviors. We present the results of this methodology applied to a recent middle school class study, in which students learned about climate change. Our primary focus in this analysis is the effectiveness and variation in the reading behaviors of high-versus low-performing students. These results illustrate the potential of this iterative methodology in identifying and interpreting learning behavior patterns at multiple levels of detail.

1. INTRODUCTION

Cognitive scientists have established that metacognition and self-regulation are important components for developing effective learning in the classroom and beyond [5; 18]. In developing a computer-based learning environment (CBLE) called Betty's Brain, we have adopted a self-regulated learning (SRL) framework to help students develop learning strategies. As they explore hypermedia resources on a science topic, they construct a causal map to teach Betty, their virtual Teachable Agent (TA) [4]. Betty only knows what she has been taught by the student, but, once taught, she can use this information to answer questions like "if *deforestation* increases, what effect does it have on *polar sea ice*?"

and explain her answers as a chain of causal relations [9]. The student can also ask their TA to take quizzes, which are a set of questions created and graded by a Mentor Agent named Mr. Davis. The TA's quiz performance helps the students to assess and reflect on their TA's, and, therefore, their own learning performance. This assessment and subsequent reflection can help guide them as they continue their learning and teaching tasks. Previous studies have shown that observing Betty's quiz performance (which is actually a reflection of their own understanding) motivates students to learn more in order to help Betty improve her quiz score [4]. Overall, the combined learning and teaching task is complex, open-ended, and choice-rich, so learners must employ a number of cognitive and metacognitive skills to achieve success. At the cognitive level, they need to identify and understand relevant information from the resources in the system, represent that information in the causal map format to teach their agent, and use questions and quizzes to explore Betty's understanding and assess her overall progress. At the metacognitive level, they need to set goals and choose strategies related to their knowledge construction and monitoring tasks. In other words, they must decide when and how to acquire information, build and modify the causal map, check Betty's progress, and reflect on their own understanding of both the science knowledge and the evolving causal map structure. Their cognitive and metacognitive activities are scaffolded through dialogue and feedback provided by Mr. Davis. This feedback aims to help students progress in their learning, teaching, and monitoring tasks. Betty's Brain is designed to track many details of students' learning interactions along with their teaching performance. This wealth of data provides opportunities to assess, model, and understand student learning behaviors and strategies more accurately. Realizing these opportunities requires effective methods for identifying interesting learning behavior patterns in the activity trace data. For example, sequential pattern mining [2] can be employed to identify frequent patterns in students' activity trace data. However, this can also result in a very large number of patterns¹ To overcome this

¹Sequential pattern mining with activity traces of 16 8th-grade students working in Betty's Brain identified over 1,000 patterns that occurred in at least 80% of the traces, when allowing gaps of one action to account for noise introduced by random or inconsequential actions.

problem, we have developed an algorithm that employs a novel combination of sequence mining techniques to identify differentially frequent patterns between groups of students (*e.g.*, experimental versus control conditions or high- versus low-performers) [8]. Further, this technique can be contextualized with information about the student’s performance (*e.g.*, productive and counter-productive phases) over the course of their learning interactions [8].

In this paper, we extend these techniques by incorporating them in an iterative, exploratory methodology and further contextualizing the differential sequence mining with action features, such as activity categorizations, relevance and timing between actions, and repetition of analogous actions. We apply this exploratory data mining methodology to learning trace data gathered during a recent Betty’s Brain study run in a middle school classroom. Previous analyses have shown that reading the resources occupies a significant portion of the students’ learning activities. Therefore, we delve deeper than previous analyses by exploring reading action features (*e.g.*, short versus long reads and first reads versus rereads of a page) and analyze student behaviors and performance using this more detailed characterization of reading actions.

2. RELATED WORK

In this section, we briefly review relevant past work on using sequence mining techniques to analyze students’ learning behaviors. For example, Perera *et al.* [14] investigate trace data from mirroring and feedback tools that support effective teamwork among students collaborating on software development using an open source professional development environment called TRAC. In their approach, they help all groups improve their work by observing and emulating the behaviors of the strong groups. They use k-means clustering to find groups of similar teams and similar individuals, and then employ a modified version of the Generalized Sequential Pattern (GSP) mining algorithm [16] to show that leadership and group interaction are important to success. Martinez *et al.* [12] discovered frequent sequences of actions that differentiate high-achieving groups from low-achieving groups of learners, who collaborate around a shared tabletop to answer an open question posed as a mystery problem. They apply a clustering algorithm to group similar patterns to aid in analyzing the pattern distribution across the groups. Employing sequential pattern mining allows them to identify differences between the higher- and lower-achieving groups in their manner of information gathering to solve the problem. Like Perera *et al.* [14] and Martinez *et al.* [12], we compare sequential patterns derived from groups of student activity sequences. However, our differential sequence mining algorithm directly incorporates comparisons between groups with additional metrics to identify interesting patterns, rather than manually performing researcher-directed comparisons after data mining.

Other researchers have employed sequential pattern mining (with a single set of student activity sequences or subsequences) to understand student learning behaviors. For example, Su *et al.* [17] propose a method for creating personalized activity trees to be used in a Sharable Content Object Reference Model (SCORM) e-learning system. They use sequential pattern mining to extract frequent learning patterns as part of a larger process that creates a decision tree to predict the group/category for a new student. Nesbit *et*

al. [13] employ sequential pattern mining to investigate self-regulation in gStudy, which is a software application with similarities to Betty’s Brain. In this system, students learn from multimedia documents and organize their knowledge with notes, concept maps, and other objects. Using sequential pattern mining, the authors hope to step beyond the question of whether a tool helps learners construct knowledge and instead investigate when and how learners use the tool as they self-regulate their knowledge construction activities. Similarly, our work investigates learning behaviors and self-regulation by identifying sequential patterns of student activity. However, unlike all of the preceding applications of sequential pattern mining, our methodology also analyzes students’ evolving performance to identify, and group, action subsequences corresponding to productive and counter-productive phases. Further, our methodology iteratively employs action abstraction/transformation using features, such as activity categorizations, relevance and timing between actions, and repetition of analogous actions.

3. DIFFERENTIAL SEQUENCE MINING METHODOLOGY

To effectively perform sequential data mining on learning interaction traces, raw logs must first be transformed into an appropriate sequence of actions. Since these logs can contain a significant quantity of information about each student interaction with the system, as well as other system book-keeping information, raising the level of abstraction from raw log events to a canonical set of distinct actions is a vital first step in effective analysis. Our methodology incorporates iterative refinement of this action abstraction step to focus the analysis on various learning activities and actions.

3.1 Action Abstraction with Context Summarization

Action abstraction is the first step of our data mining methodology, in which researcher-identified categories of actions define an initial alphabet (set of action symbols) for the sequences. This step filters out irrelevant information (*e.g.*, cursor position) and combines qualitatively similar actions (*e.g.*, querying an agent through different interfaces or about different concepts in a given topic).

To apply the abstraction process, log events captured by the CBLE are mapped to a sequence of canonical actions taken by each student. As in previous work, we abstract student activities in five primary categories [8]:

- *READ*: students access a page in the resources;
- *LINK* or *CONCept Edit*: students edit the causal map, with actions further divided by: (i) whether they operate on a causal link (“LINK”) or concept (“CONC”) and whether the action was an addition (“ADD”), removal (“REM”), or modification (“CHG”), *e.g.*, *LINKREM* or *CONCADD*;
- *QUER*: students use a template to ask Betty a question, and she uses a causal reasoning method to answer the question [9];
- *EXPL*: students probe Betty’s reasoning by asking her to explain her answer to a question, and she uses dialogue and animation on the causal map to demonstrate her use of causal reasoning to answer the question;

- *QUIZ*: students assess how well they have taught Betty by having her take a quiz, which is a set of questions chosen and graded by the Mentor agent.

However, abstracting the raw log traces through action categorization, also strips potentially important context associated with the actions in the traces. For example, with the LINK-ADD action, the particular link added can provide important context information, such as whether this link relates to resource material the student read in a previous action. However, if the details of the exact link added are used to differentiate each edit action, we would end up with an unwieldy number of distinct actions, making it hard to discover and interpret behavior pattern sequences. To maintain a balance between the number of distinct actions and retaining relevant context information, we employ metrics that summarize context in order to distinguish actions. For example, we employ a relevance summarization metric, which establishes whether the content/object of an action is related to a small number of recent activities, where recent is defined by a configurable window of previous actions [3]. This relevance metric splits each categorized action into two distinct actions: (1) relevant to at least one of the recent actions (with the “-REL” suffix) and (2) irrelevant to any of the recent actions (with the “-IRR” suffix).

In this methodology, the choice of specific context-summary metrics and their application to different categories of action is iteratively refined over repeated analyses of the interaction traces. This allows the researcher to focus the analysis, providing more detail and context associated with specific learning activities or strategies. In previous work, we presented an initial analysis of student action sequences applying only the relevance metric, which illustrated some interesting map editing and monitoring behaviors distinguishing high-performing and low-performing students [8]. However, that analysis did not differentiate between reading actions (*e.g.*, long versus short, or reading pages in sequence versus using keyword search), which are frequent and vital to student learning in Betty’s Brain [8; 15].

In this paper, we present the results of a subsequent iteration in this extended methodology, in which we apply additional metrics to distinguish different types of READ actions. As a continuation of the exploratory methodology, a future iteration might instead focus on actions related to editing the causal map by applying additional editing metrics (*e.g.*, whether the edit increased or decreased the correspondence between the student’s map and the expert map, or whether the edit introduced a cycle, continued a chain of causal relationships, or added a branch to a chain of causal relationships). However, to maintain a reasonable number of distinct actions in that hypothetical iteration (such that sufficiently frequent patterns could still be identified), the number of reading-related metrics would be correspondingly reduced.

In the analysis and iteration of the methodology presented in this paper, we apply three reading-related metrics to the student action sequences:

- *Source (TOC/HLNK/HIST)*: how the student reached the page he/she is reading - by selecting a page in the table of contents (TOC) always displayed on the left of the resources, from a hyperlink (HLNK) on another page, or using the backward or forward button to move through their history of pages (HIST) like a

web browser;

- *Time (SHRT/FULL)*: a determination of whether the student spent enough time on the page to have read a significant amount of the material² (FULL) or only spent a brief period of time on the page (SHRT), possibly skimming the material or checking whether the page was one for which they were searching;
- *Repetition (FRST/REPT)*: a determination of whether the student had never done a FULL read of the page (FRST) or this was a reread of the page (REPT) because the student had previously done a FULL read;

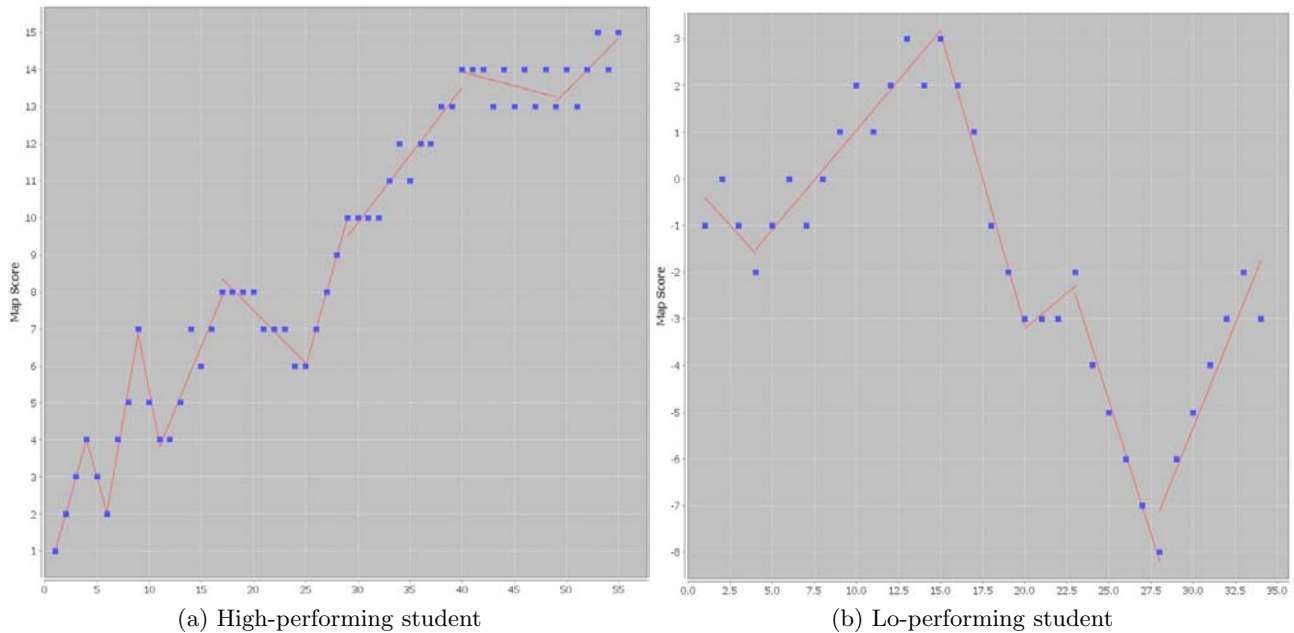
In addition to metrics related to individual actions, we also apply another, general transformation to the action sequence. In an environment like Betty’s Brain, there are cases in which students often perform a particular type of action (*e.g.*, adding concepts) repeatedly in sequence, which can result in a variety of frequent patterns that differ only by the number of repetitions of that action. To improve this exploratory analysis, our action abstraction step distinguishes a single action from repeated actions, which are condensed to a single action with the “-MULT” suffix. Using the re-transformed sequences, our differential sequence mining technique can more efficiently identify trends that could otherwise be hidden by the multitude of frequent patterns differing only in the length of a repeated action sequence [8].

3.2 Differential Sequence Mining

To identify important activity patterns in a comparison between two sets of action sequences, our methodology employs a novel combination of sequence mining techniques. Sequential pattern mining [2] methods find the most frequent action patterns across a set of action sequences, while episode mining [11] discovers the most frequently used action patterns within a given sequence. However, finding the patterns most important for interpreting learning behaviors or differentiating between groups of students is challenging, because of the need to limit the large set of frequent patterns to ones that are interesting and important (*i.e.*, our focus is on the *effectiveness* of mining techniques in identifying these important patterns, rather than the *efficiency*, or speed, in calculating the frequent patterns [1]).

In comparing across groups of action sequences, such as high- versus low-performing students, the differences between the groups provide a natural criterion for identifying important patterns that may elucidate differences in learning behavior. To use this criterion for mining important frequent patterns, we define two measures of frequency and the corresponding differences calculated across the groups. The sequential pattern mining frequency measure (*i.e.*, the number of sequences in which the pattern occurs, regardless of how many times) is important for identifying patterns common to a group of action sequences. We refer to this as the “sequence support” (*s-support*) of the pattern, following the convention of [10], and we call patterns meeting a given *s-support* threshold *s-frequent*. The second metric is the

²Based on the length of typical resource pages and the reading abilities of the students in the study, we set the threshold between short and full reads to be 30 seconds. Further, the large majority of reads in the short category were actually under 5 seconds and most of the reads in the full category were over a minute.



1: Example student performance evolution with identified phases

episode frequency, defined as the number of times the pattern is repeated *within* an action sequence. We refer to this frequency measure as the “instance support” (*i-support*), following [10]. To calculate the *i-support* of a pattern in a *group* of traces, we use the mean of the pattern’s *i-support* values across all sequences in the group.

The details of our differential sequence mining algorithm are presented in [8], but we briefly outline the main steps of the algorithm. First, a sequential pattern mining algorithm (SPAMc [6]) identifies the patterns that meet a minimum *s-support* constraint within each group, employing a maximum gap constraint to account for noise, which is interpreted as a small number of irrelevant actions that may be interspersed in a pattern. In this paper, we employ a gap constraint of 1, *i.e.*, we allow at most one irrelevant action between each consecutive action in a pattern. To compare the identified *s-frequent* patterns across groups, we calculate the mean *i-support* of every pattern for each group. In order to identify patterns whose usage more clearly differ between the two groups, we also filter the patterns based on the *p* value of a *t-test* comparing pattern *i-support* between the groups.

This comparison produces four distinct categories of frequent patterns: two categories where the patterns are *s-frequent* in only one group, illustrating patterns primarily employed by that group, and two categories where the patterns are common to both groups but used more often in one group than the other. The patterns in each of these qualitatively distinct categories are (separately) sorted by the difference in mean group *i-support* to focus the analysis on the most differentially frequent patterns.

3.3 Performance Evolution Phases

In the Betty’s Brain environment, a student’s work can be assessed in terms of their performance on the learning

task, which we define as the student’s current map score³. By tracking the evolution of students’ map scores, we can quantify how their learning and map-building performance develop as they work on the system. To more effectively identify and contextualize learning behavior patterns, we consider phases of productive (increasing map score) and counter-productive (decreasing map score) activity over the course of learning by tracking their map scores, as illustrated in Figure 1.

These phases are identified by generating a piecewise, linear representation (PLR) for a sequence of two-dimensional points. In this representation, the *x-value* is a cumulative measure of student editing activity (*i.e.*, the number of edit actions the student has performed thus far) and the *y-value* is the student’s total map score after the corresponding edit action [8]. Figures 1(a) and 1(b) illustrate these performance phases with plots of map score versus number of edits for a high-performing and a low-performing student, respectively. To generate this representation, we employ a standard bottom-up, time-series linear segmentation algorithm [7] with the sum-squared-error (SSE) of the segments as the criterion metric [8].

3.4 Summary of Methodology

Our iterative methodology consists of four major steps to identify learning behaviors contextualized by performance evolution between groups of students:

1. *Action abstraction*: Logfiles are processed to produce a sequence of actions for each student by mapping sets of interaction events to canonical actions. Each canonical action is contextualized and split into distinct actions by applying metrics, such as the relevance

³The map score is defined as the number of correct links (based on the expert map) in the student’s map minus the number of incorrect links.

metric and the reading metrics. At each iteration additional metrics can be applied, as well as previous metrics removed, based on the results of previous iterations. Finally, any subsequences of a repeated action are condensed into a single “action” identified with the “-MULT” suffix.

2. *Performance phase identification*: Student action sequences are split into subsequences using the time-series segmentation algorithm. These subsequences are filtered to produce two sequential datasets: a) productive action sequences corresponding to segments with a positive progress slope above a given cutoff, and b) counter-productive action sequences corresponding to segments with a negative progress slope below a given (negative) cutoff.
3. *Differential sequence mining*: The student groups, as well as productive and counter-productive action subsequences within those groups, are compared to identify differentially frequent patterns of action.
4. *Interpretation*: The differentially frequent sequential patterns of action are interpreted in terms of effective and ineffective learning behaviors exhibited by students during the learning task. Investigation of pattern details (i.e., raw event details for instances of these patterns) may yield further insights into student cognition and metacognition, as well as potential flags and triggers for adaptive feedback/scaffolding in the system.

4. RESULTS

We illustrate our methodology using interaction trace data from a recent study with 40 8th-grade students taught by the same teacher in a middle Tennessee school. At the beginning of the study, students were introduced to the science topic (global climate change) during regular classroom instruction, provided an overview of causal relations and concept maps, and given hands-on training with the system. For the next five days, students taught their agent about climate change and received feedback on metacognitive strategies from the Mentor agent. In this version of the system, the majority of the metacognitive feedback was related to knowledge construction strategies [15]. However, the Mentor agent also provided advice on monitoring strategies to help students recognize and correct errors in their causal maps.

The results of this study presented an interesting dichotomy in student performance at constructing their causal concept maps. 16 of the students taught their agent a correct, complete map or one very close to it (these students achieved map scores between 11 and 15, inclusive, where 15 was the maximum possible score). Another 18 students taught their agents relatively poor maps with a map score of 5 or below. Only 6 students had a map score in between these groups (i.e., a map score of 6 to 10, inclusive). Therefore, we focus on an analysis and comparison of the learning activities of the high-performing (“Hi”) student group and the low-performing (“Lo”) student group. An initial analysis of the activity traces from this study was presented in [8]. Here we focus on the effectiveness and variation in students’ reading behaviors by refining the action abstraction step in our exploratory methodology with additional (reading-related)

metrics, discussed in Section 3.1. We should note that students in the “Hi” group had higher pre-test scores in all of the categories as compared to the “Lo” group. However, a detailed analysis shows that 40% of the links added by the “Hi” group were initially incorrect (this number was 58% for the “Lo” group). This shows that the “Hi” group had to put in significant effort into discovering errors in their maps and correcting them, and the final results show that they were quite successful in their monitoring and correction tasks. This was not the case for the “Lo” group. Therefore, a comparison of the learning behaviors of the two groups should demonstrate an important dichotomy in the strategies employed by the two groups that mirrors the dichotomy in their performance. To further differentiate behaviors associated with high and low performance, we compared productive and counter-productive phases of student activities. We discuss the results of our analyses in greater detail below. To assess students’ overall learning gains, calculated as normalized gains⁴ in pre- to post-test scores, we categorize the pre- and post-test questions into three groups: (i) definition questions about the science topic in multiple choice (MC) format, (ii) questions requiring reasoning about the science topic that students had to answer by writing sentences (“short answer”), and (iii) questions about causal reasoning using a causal map that was not related to the science topic. Table 1 presents the students average scores (and standard deviations). The results of an ANOVA comparing the Hi and Lo student groups on each of the pre-post gains show significant differences between the Hi and Lo groups only for the definitional MC questions. Table 1 also presents ANOVA analyses of the difference in performance for the map-building metrics: (i) link accuracy - the percentage of links added to the map that were correct; (ii) link creation effort - the total number of student actions divided by the number of correct link edits, a measure of the effort by the student in order to produce a correct link edit; and (iii) action relevance - the percentage of student actions that were relevant (as described in Section 3.1) to at least one of the three previous actions show significant differences in favor of the Hi group with moderate effect sizes. These results indicate that students in the Hi group were more accurate in their map edits and generally more efficient in their learning and teaching activities. Further, they tended to employ a more systematic approach to the task, as indicated by their higher action relevance score. Overall, students who achieved success in teaching Betty accurate causal maps also learned significantly more factual information, but their gains in causal reasoning and short answer questions were not significantly different from the low-performing group [8; 15].

As a first analysis to elucidate broad differences in reading behavior between the Hi and Lo groups, Table 2 presents the relative proportion of reading activities categorized by each metric presented in Section 3.1. Both groups performed roughly equal numbers of read actions on pages they had previously read in-depth (“Repeat” (REPT)) compared to ones they had not read in-depth (“First” (FRST)). The Lo group relied slightly more on short (SHRT) reads (74%) than the Hi group (69%), and the ratio of short to full page (FULL) reads was approximately 3:1 for the Lo group and 2:1 for the Hi group. Similarly, the Lo group’s read actions

⁴The normalized learning gain was calculated as: $\frac{(post-pre)}{(max-pre)}$

1: High vs. Low Performers - Learning Gain and Map Score

Metric	Hi Group	Lo Group	F	Sig.	Effect Size (Cohen's f)
Definition MC Norm. Gain	0.535 (0.344)	-0.202 (0.769)	12.448	0.001	0.624
Causal Reasoning Norm. Gain	0.130 (0.614)	-0.029 (0.414)	0.799	0.378	0.157
Short Answer Norm. Gain	0.027 (0.241)	-0.028 (0.134)	0.700	0.409	0.146
Map Score	14.500 (1.155)	2.780 (1.592)	590.171	0.001	4.314
Link Accuracy (LA)	60.3% (7.1%)	42.8% (10.6%)	31.528	0.001	0.992
Link Creation Effort (LCE)	11.630 (8.196)	20.665 (5.942)	12.745	0.001	0.652
Action Relevance	53.3% (8.4%)	40.6% (12.2%)	12.401	0.001	0.622

were deemed more irrelevant (IRR) to recent actions (again the ratio of IRR to REL (relevant) reads was 3:1). The same ratio for the Hi group was 2:1.

2: Relative Proportion of Actions by Reading Metrics

Group	First : Repeat (FRST : REPT)	Short : Full (SHRT : FULL)	Irrel. : Rel. (IRR : REL)
Hi	51% : 49% (SD 12%)	69% : 31% (SD 15%)	67% : 33% (SD 9%)
Lo	55% : 45% (SD 14%)	74% : 26% (SD 12%)	74% : 26% (SD 7%)

To analyze specific reading behaviors illustrated by these students' interaction traces, we applied the differential sequence mining technique described in Section 3.2. This allowed us to identify a variety of interesting learning behaviors related to reading that were not apparent from the higher level analyses of behavior patterns we had conducted in the past [8; 15]. Table 3 presents the top five patterns in each of the differential categories detailed in Section 3.2. For the analysis, we employed an s-support threshold of 50% to analyze patterns that were evident in the majority of either group of students and employed a standard statistical significance cutoff of $p < 0.05$. In all of the differential sequence mining results presented here, we employed a maximum gap threshold of 1, to allow for "noise" from irrelevant or interchangeable actions in the learning activity sequences, as described in Section 3.2.

All reads in the differentially frequent patterns distinguishing reading behaviors between high and low students were pages selected from the table of contents (TOC) rather than from hyperlinks within pages or the (backward/forward) history mechanism. This is unsurprising since raw frequencies of these different types of reading activities indicated that in both the high and low group, the large majority of reading activities involved selecting pages from the table of contents. Table 3 shows that the high group was much more likely to add a link (both relevant (REL) or irrelevant (IRR) links with respect to recent actions) following a full-length (FULL) re-read (REPT) of a page that was relevant (REL) to recent actions. This greater reliance on extended re-reads before adding links suggests the high group employed a more careful approach to identifying causal links in the resources, which may have helped increase their accuracy in teaching correct links, and also their ability to correct previously taught incorrect links.

Further, the high group more frequently employed reading activities in a monitoring context (*i.e.*, in conjunction with quiz actions). Besides following extended re-reads by adding links, the high group was also more likely to follow them with

quizzes, possibly in an attempt to connect what they were reading with their TA's right and wrong answers based on the current map. Following quizzes, they were more likely to do a quick re-read of a relevant page, which suggests another monitoring strategy, such as confirming links used by the TA in quiz answers. The differentially frequent patterns employed more by the low group were various combinations of reading, especially short reads and ones not relevant to recent actions. This may be indicative of a less consistent approach to reading and of strategies that do not systematically combine reading with other knowledge construction and monitoring activities.

To further investigate which reading behaviors may have contributed to the high performers' success, we identified differentially frequent patterns when students were productive as opposed to being counter-productive during their map building activities. The method for extracting the productive versus counter-productive phases was described in Section 3.3, and we included all segments with a slope greater than or equal to 0.4 in the productive set and all segments with a slope less than or equal to -0.4 in the counter-productive set⁵ [8]. For the differential sequence mining with performance evolution subsequences analysis, we employed a lower s-support threshold because the sequences were significantly shorter than the complete student activity sequences. Specifically, we employed an s-support threshold of 20% to analyze patterns that occurred with some regularity (*i.e.*, in at least one out of every five subsequences). Similarly, given the limited length and number of sequences, we employed a relaxed cutoff on the t-test comparison of $p < 0.10$.

In comparing the Hi group's productive to counter-productive periods, the only differentially frequent pattern observed was that extended, relevant rereads (READ-TOC-REPT-FULL-REL) occurred approximately twice as frequently ($p = 0.034$) in productive segments ($i - support = 0.65$) than counter-productive segments ($i - support = 0.38$). This reliance on extended, relevant re-reads, especially during productive periods, provides further evidence that a more careful, systematic approach to reading may have been particularly beneficial for the high-performing students. In comparing the Lo group's productive to counterproductive periods, the only differentially frequent pattern observed was that extended, relevant reading of a page for the first time occurred approximately five times as frequently ($p = 0.039$) in productive segments ($i - support = 0.28$) than counter-productive segments ($i - support = 0.06$). This suggests that when the low-performing students read a page in-depth for the first time

⁵The slope cutoff of 0.4/-0.4 was determined by qualitative analysis of a sample of student map score plots to distinguish generally productive/counter-productive segments [8]

3: High vs. Low Performers - Differentially Frequent Patterns

Pattern	I-Support (Hi)	I-Support (Lo)	t-test (p value)	S-Freq Group
READ-TOC-REPT-FULL-REL → LINKADD-REL	2.19	0.61	0.005	Hi
READ-TOC-REPT-FULL-REL → QUIZ	1.50	0.39	0.010	Hi
QUIZ → READ-TOC-REPT-SHRT-REL	1.44	0.33	0.018	Hi
READ-TOC-REPT-SHRT-REL → LINKADD-IRR	1.13	0.17	0.008	Hi
READ-TOC-REPT-FULL-IRR → LINKREM-IRR	1.13	0.28	0.016	Hi
READ-TOC-REPT-FULL-REL → LINKADD-IRR	2.00	0.94	0.022	BOTH
READ-TOC-FRST-SHRT-IRR → READ-TOC-REPT-SHRT-IRR-MULT	0.31	0.89	0.043	Lo
READ-TOC-REPT-SHRT-IRR → READ-TOC-REPT-SHRT-IRR-MULT	0.31	1.00	0.040	Lo
READ-TOC-REPT-FULL-IRR → READ-TOC-FRST-SHRT-IRR	0.44	1.17	0.033	Lo
READ-TOC-REPT-FULL-IRR → READ-TOC-REPT-SHRT-IRR	0.63	1.67	0.013	Lo
READ-TOC-REPT-SHRT-REL → READ-TOC-FRST-SHRT-REL	0.50	1.78	0.047	Lo

they tended to extract useful information, especially if the page was relevant to their recent activities. However, they did not seem to be as effective as the high-performing students in extracting additional information when re-reading pages.

To gain further insight into differences in productive reading behaviors exhibited by students in Betty’s Brain, we applied the differential sequence mining technique to compare the productive subsequences between the Hi and Lo groups. This analysis, presented in Table 4, illustrates four reading activities more frequently employed by the low-performing students during their productive activity periods. The greater frequency of irrelevant, extended rereads (both from the table of contents and from hyperlinks) in the Lo group, suggests that although they tended to be less systematic in their approach, they were still able to gain some useful information from rereading the resources. However, their greater reliance on initial, in-depth reads (READ-TOC-FRST-FULL-REL) in productive periods suggests they may have had a harder time identifying the less obvious causal relations often found after multiple reads. Finally, the repeated use of short, irrelevant reads during productive periods by the Lo group may indicate an inefficient, but ultimately effective, searching behavior.

5. CONCLUSIONS

In this paper, we extended an exploratory data mining methodology for identifying important learning behaviors with an iterative approach to action abstraction using a variety of action features and presented results analyzing reading behaviors of students in a learning-by-teaching environment. The exploratory methodology combines iterative action abstraction, a sequence mining technique to identify differentially frequent activity patterns, and piecewise linear segmentation of activity phases with respect to the evolution of a performance or progress measure. Results from a recent classroom study with Betty’s Brain illustrate the effectiveness of this methodology and iterative action abstraction for identification of important learning behaviors at multiple levels of analysis, including behaviors distinguished by their relationship to productive or counter-productive phases of activity.

Although the majority of frequent action patterns that could be identified in these interaction traces are common to both high-performing students and low-performing students (and occur throughout the course of students’ interaction with

the system), the analysis employing our differential sequence mining methodology elucidated some important learning behaviors. In this paper, we extended previous results with a focus on reading behaviors related to successful performance in the learning environment and differentiating them from ones employed by less successful students. Overall, high-performing students differentially employed reading behaviors that indicated a more careful and systematic strategy of reading. Their activity patterns more frequently involved re-reading pages from the resources, such as employing full-length re-reads of the resources before adding a link. Further, the reading activity patterns distinguishing high-performers from low-performers usually involved reading pages that were relevant to recent actions, suggesting a more systematic reading behavior overall. Productive periods were particularly distinguished in high performers by a larger number of full-length, relevant re-read actions.

Performance was also linked to monitoring behaviors that incorporated re-reading of the resource material. In particular, high-performers were more likely to employ various types of reading actions both before and after assessments of progress/correctness using the quiz. Low performers, on the other hand, had a differential tendency to use irrelevant, extended re-reads of pages in the resources during productive periods. They may have also had more difficulties identifying the less obvious causal relations, as suggested by their greater reliance on initial reads of a resource page during productive periods.

In future work, we also intend to expand upon the presented data mining techniques through a variety of enhancements and additional applications. We will enhance the existing summarization of action-relevance to include determination of the relationship between the specific actions in all subsequences matching the identified patterns. For example, this enhanced action-relevance summarization will allow us to determine how frequently a pattern like READ-TOC-REPT-FULL-REL → LINKADD-REL involves adding a correct link and whether the specific link added was discussed in the resource page from the reading action. Relating identified patterns of action back to specific details and context in the interaction traces could provide significant benefits for more efficient and effective interpretation of learning behaviors. Based on this analysis, we will expand and revise the feedback triggering conditions and student modeling to improve learning behavior feedback from the Betty’s Brain agents.

4: High vs. Low Performers (Productive Segments) - Differentially Frequent Patterns

Pattern	I-Support (Hi)	I-Support (Lo)	t-test (p value)	S-Frequent Group
READ-TOC-REPT-FULL-IRR	0.51	1.00	0.070	BOTH
READ-TOC-FRST-FULL-REL	0.10	0.28	0.091	Lo
READ-HLNK-REPT-FULL-IRR	0.08	0.28	0.057	Lo
READ-TOC-REPT-SHRT-IRR-MULT	0.15	0.39	0.079	Lo

6. ACKNOWLEDGEMENTS

This work has been supported by NSF-IIS Award #0904387.

7. REFERENCES

- [1] R. Agrawal, T. Imieliński, and A. Swami. Mining association rules between sets of items in large databases. In *Proceedings of The ACM SIGMOD International Conference on Management of Data*, volume 22, pages 207–216. ACM, 1993.
- [2] R. Agrawal and R. Srikant. Mining sequential patterns. In *Proceedings of the Eleventh IEEE International Conference on Data Engineering (ICDE)*, pages 3–14, 1995.
- [3] G. Biswas, H. Jeong, J. Kinnebrew, B. Sulcer, and R. Roscoe. Measuring Self-regulated Learning Skills through Social Interactions in a Teachable Agent Environment. *Research and Practice in Technology-Enhanced Learning (RPTEL)*, 5(2):123–152, 2010.
- [4] G. Biswas, K. Leelawong, D. Schwartz, N. Vye, and T. Vanderbilt. Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence*, 19(3):363–392, 2005.
- [5] J. Bransford, A. Brown, and R. Cocking, editors. *How people learn*. National Academy Press Washington, DC, Washington, D.C., 2000.
- [6] J. Ho, L. Lukov, and S. Chawla. Sequential pattern mining with constraints on large protein databases. In *Proceedings of the 12th International Conference on Management of Data (COMAD)*, pages 89–100, 2005.
- [7] E. Keogh, S. Chu, D. Hart, and M. Pazzani. Segmenting time series: A survey and novel approach. In *Data mining in time series databases*, volume 57, pages 1–22. World Scientific, 2004.
- [8] J. S. Kinnebrew, K. M. Loretz, and G. Biswas. A contextualized, differential sequence mining method to derive students’ learning behavior patterns. *Journal of Educational Data Mining*, In Press, 2012.
- [9] K. Leelawong and G. Biswas. Designing learning by teaching agents: The Betty’s Brain system. *International Journal of Artificial Intelligence in Education*, 18(3):181–208, 2008.
- [10] D. Lo, S. Khoo, and C. Liu. Efficient mining of recurrent rules from a sequence database. In *Proceedings of the 13th International Conference on Database Systems for Advanced Applications*, pages 67–83. Springer-Verlag, 2008.
- [11] H. Mannila, H. Toivonen, and A. Inkeri Verkamo. Discovery of frequent episodes in event sequences. *Data Mining and Knowledge Discovery*, 1(3):259–289, 1997.
- [12] R. Martinez, K. Yacef, J. Kay, A. Al-Qaraghuli, and A. Kharrufa. Analysing frequent sequential patterns of collaborative learning activity around an interactive tabletop. In *Proceedings of the Fourth International Conference on Educational Data Mining*, Eindhoven, Netherlands, 2011.
- [13] J. Nesbit, M. Zhou, Y. Xu, and P. Winne. Advancing log analysis of student interactions with cognitive tools. *12th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI)*, 2007.
- [14] D. Perera, J. Kay, I. Koprinska, K. Yacef, and O. Zaiane. Clustering and sequential pattern mining of online collaborative learning data. *IEEE Transactions on Knowledge and Data Engineering*, 21(6):759–772, 2009.
- [15] J. R. Segedy, J. S. Kinnebrew, and G. Biswas. Supporting cognitive and metacognitive skills in complex, open-ended learning environments. *Journal of Educational Psychology*, In Review, 2012.
- [16] R. Srikant and R. Agrawal. Mining Sequential Patterns: Generalizations and Performance Improvements. In *Proceedings of the 5th International Conference on Extending Database Technology (EDBT): Advances in Database Technology*, pages 3–17. Springer-Verlag, 1996.
- [17] J.-M. Su, S.-S. Tseng, W. Wang, J.-F. Weng, J. Yang, and W.-N. Tsai. Learning portfolio analysis and mining for scorm compliant environment. *Journal of Educational Technology and Society*, 9(1):262–275, 2006.
- [18] B. Zimmerman. Theories of self-regulated learning and academic achievement: An overview and analysis. In B. Zimmerman and D. Schunk, editors, *Self-regulated learning and academic achievement: Theoretical perspectives*, pages 1–37. Erlbaum, Mahwah, NJ, 2001.