Integrating Model-Driven and Data-Driven Techniques for Analyzing Learning Behaviors in Open-Ended Learning Environments

John S. Kinnebrew, James R. Segedy, and Gautam Biswas, *Fellow, IEEE*

Abstract—Research in computer-based learning environments has long recognized the vital role of adaptivity in promoting effective, individualized learning among students. Adaptive scaffolding capabilities are particularly important in open-ended learning environments, which provide students with opportunities for solving authentic and complex problems, and the choice to adopt a variety of strategies and approaches to solving these problems. To help students overcome their difficulties and become effective learners and problem solvers, we have to develop methods that can track and interpret students' open-ended learning and problem-solving behaviors. The complexity of the problems and the open-ended nature of the solution processes pose considerable challenges to accurately interpret and evaluate student behaviors and performance as they work on the system. In this paper, we develop a framework that combines model-driven strategy detection with data-driven pattern discovery for analyzing students' learning activity data in open-ended environments. We present results from an in-depth case study of multiple activity patterns identified in data from the *Betty's Brain* learning environment. The results illustrate the benefits of combining model- and data-driven techniques to precisely characterize the learning behavior of students in an open-ended environment.

Index Terms—Open-ended learning, environments, modeling learner, behaviors, hierarchical task modeling, activity sequence mining, coherence analysis, strategy modeling

1 Introduction

Adapting to learners' needs and providing timely and useful individualized feedback to help them succeed has been a hallmark of many computer-based learning environments (e.g., [1]). In addition to providing a structure and resources that facilitate learning and problem solving, these systems take explicit actions [2], such as reminding learners of relevant information or modifying the learning activity to support learning processes [3], [4]. To promote deep learning, critical thinking, and problem-solving skills in STEM disciplines, researchers have been developing open-ended learning environments (OELEs) [5]. These systems provide students with a learning goal, usually in the form of a complex problem, and a set of tools that support the problem-solving task [6].

However, open-ended problem solving can present significant challenges for novice learners [7], [8]. To succeed, they need to make choices on how to structure the solution process, explore alternative solution paths, develop awareness of their own knowledge and problem-solving skills, and develop strategies that support more effective learning and problem solving [9], [10]. In other words, to become successful problem solvers, students need to employ metacognitive processes for planning, monitoring, controlling, and reflecting on relevant cognitive processes as they search for information, interpret it, and apply it to construct and test potential solutions. However, students often lack proficiency in using the system tools, as well as the experience and understanding they need to explicitly regulate their own learning and problem solving in these environments [11].

To help students overcome their difficulties and become effective learners and problem solvers, we have to develop methods that can track and interpret students' open-ended learning and problem-solving behaviors. Traditionally, learning behaviors in OELEs have been assessed with model-driven metrics and context-driven hypotheses about the students' learning tasks [12], [13], [14]. More recently, researchers have been developing exploratory mining techniques that provide the basis for discovering behavior patterns (*c.f.*, [15], [16]). Together, these techniques may provide a more expansive framework for analyzing how students learn while tackling open-ended problems.

This paper extends and formalizes an approach we have developed for analyzing students' learning activity data collected from log files in OELEs. The framework brings together two approaches we have developed: (1) model-based analysis of students' activity sequences that assesses the coherence among students' actions [14]; and (2) data-driven pattern discovery methods [15] to provide a more complete analysis of students' behavior as they work in an OELE. In previous work, we described an initial integration of a task-modeling approach that captures a relevant set of problem-solving tasks and a data-driven pattern discovery approach [17], [18]. In this paper, we augment the task model with a strategy model that formalizes the manner in which the model-driven portion of the framework directly

• *The authors are with the Institute for Software Integrated Systems, EECS Department, Vanderbilt University, Nashville, TN 37240.*
E-mail: {john.s.kinnebrew, james.segedy, gautam.biswas}@vanderbilt.edu.

Manuscript received 4 June 2014; revised 8 Dec. 2015; accepted 12 Dec. 2015. Date of publication 30 Dec. 2015; date of current version 16 June 2017.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TLT.2015.2513387

influences the data-driven analysis. Strategies are defined as consciously-controllable processes for completing tasks [19], and they are known to comprise a significant portion of metacognitive knowledge [20]. We present a formal definition of a *strategy model* but restrict the definition to strategies represented as a sequence of tasks and sub-tasks that can be accomplished in the learning environment. These strategies are further characterized by the context in which they are applied and the specific relations among component activities in the strategy.

In addition to using task and strategy models to provide a more specific analysis and characterization of the action patterns discovered by the data-driven analysis, our framework incorporates a novel extension that operates in the reverse direction. In other words, we also use data-driven discovery of frequent patterns to iteratively extend existing strategy models. To enable an effective and systematic application of this approach, we extend a lift measure [21], initially developed for analysis of association rules, to the analysis of sequential patterns. This sequential lift analysis identifies activity patterns that occur more frequently than expected from a random behavior model and are, therefore, candidates for defining additional strategies not yet covered by the strategy model.

We apply this combined model- and data-driven analysis framework to *Betty's Brain* [22], an OELE where students learn science by constructing causal models. A detailed case study illustrates the benefits of applying our analytic framework to characterize the learning behaviors of students in an OELE. The results show that this approach provides valuable information about differences among students that may employ similar patterns of actions but in different ways and for different purposes. Further, we identify aggregate behavior differences between student groups that provide a better understanding of their approach to learning and problem solving in OELEs. Finally, we identify and analyze additional strategies that were employed by students but had not previously been included in the strategy model.

2 Background: Metacognition

Metacognition describes the ability to reason about and explicitly manage one's own cognitive processes [23]. From an information-processing perspective, Winne [24] describes cognition as dealing with knowledge of objects and operations on objects (the object level), while characterizing metacognition as the corresponding *meta level* that contains information about when to use particular cognitive processes and how to combine them to accomplish larger tasks. Metacognitive monitoring brings the two levels together, as it describes the process of observing and evaluating one's own execution of cognitive processes and, when necessary, exercising control over one's cognition in order to improve one's overall effectiveness in accomplishing tasks.

When applied to learning, metacognition is often considered a subset of the more encompassing conception of self-regulated learning (SRL). SRL is an active theory of learning that describes how learners set goals, create plans for achieving those goals, continually monitor their progress, and reflect and revise their plans when necessary to become more successful at achieving their overall goals [25]. Within this SRL framework, metacognition deals directly with the regulation of cognition without explicitly considering its interactions with emotional or motivational constructs [26]. Despite this, models of self-regulation are valuable in depicting key metacognitive processes.

For example, Winne and Hadwin [27], [28] have proposed a model of SRL called COPES. Learning according to this model occurs in four weakly sequenced and recursive stages: (1) task definition, where the student develops an understanding of the learning task, (2) goal setting and planning, which follow the task definition phase and represent the student's approach to working on the learning task, (3) enactment of tactics, during which the student carries out plans for learning, and (4) adaptations to metacognition, which are linked to both in-the-moment adjustments of the student's tactics and post-hoc evaluation of the student's approach based on successes and failures during enactment.

Like COPES, we adopt a task-oriented framework to interpret students' learning activities and behaviors in OELEs. In particular, our focus on metacognition is centered on students' understanding of and use of *strategies*, which, as discussed earlier, have been defined as consciously-controllable processes for completing tasks [19]. Strategies consist of declarative, procedural, and conditional knowledge that describe their purpose and how and when to employ them [20]. The research community has identified several types of strategies based on the tasks for which they are designed. For example, strategies may be (1) cognitive (e.g., a strategy for solving an addition problem); (2) metacognitive (e.g., a strategy for monitoring one's own cognitive operations when working on an equation-solving task), (3) focused on management (e.g., managing one's environment to promote focused attention); (4) directed toward learning (e.g., a strategy that facilitates learning of feedback processes); or (5) a combination of the four strategy types discussed above [29], [30]. For example, a metacognitive learning strategy may involve activating prior knowledge before learning about a topic by consciously bringing to mind information one already knows about the topic [31]. When faced with a complex task, students may invoke known strategies, or invent one using their current cognitive and metacognitive knowledge.

An important characteristic of a strategy is its *level of generality*. Some strategies apply to very specific situations (e.g., an approach to adding two-digit numbers) while other strategies apply to a broader set of situations (e.g., summarizing recently learned information to improve retention). An understanding of more general strategies, and their specific implementations when applied to concrete tasks, is important for developing one's ability to adapt existing strategies to new situations or even invent new strategies. Thus, an important goal in developing adaptive support for students' working in OELEs is to explicitly teach students general strategies for regulating their learning as they solve complex, open-ended problems. Doing so can prepare students for future learning [32] by developing their ability to independently investigate and solve open-ended problems.

3 Background: Strategy Understanding

Measuring students' understanding of strategies by observing their behavior is a difficult task; it requires determining whether their behaviors are consistent with a strategy. For small tasks the interpretation is straightforward. For example, to assess a student's understanding of an algebraic problem-solving strategy, a learning environment can present multiple problems and observe the steps that a student takes to solve these problems. If the student consistently carries out a sequence of steps that matches the sequence prescribed by the strategy, the learning environment can reasonably assume that the student understands the strategy. This approach is employed in many step-based intelligent tutoring systems, such as Cognitive Tutors [33], [34].

However, for complex tasks, such as those presented by OELEs, interpretation is more difficult. Succeeding in these tasks involves breaking up the overall task into sub-tasks, setting goals, and applying strategies for completing each sub-task. Thus, learning environments need methods for inferring: (1) the student's chosen task decomposition; (2) the goal and corresponding sub-task that the student is currently working on; (3) the strategy being used to complete that sub-task; and (4) whether or not they are executing the strategy correctly. The open-ended nature of OELEs further exacerbates the measurement problem; students may constantly change their chosen task decomposition, their chosen goal, and, therefore, the tasks they are currently attempting or the strategies they are utilizing. The corresponding analysis techniques must have the ability to detect and interpret these switches as they track students' problem-solving activities.

Researchers have approached this problem from multiple angles. For example, *MetaTutor* [31] adopts a very direct approach with some similarities to self-report; it provides interface features through which students explicitly state the tasks they are attempting (called subgoals) and the strategies they are using to complete those tasks. This allows the system to directly capture students' strategy use without having to make inferences based solely on their activities in the system. However, this relies on students accurately communicating their intent to the system.

Another approach involves first defining a model of behavior as prescribed by the strategy and then checking students' behaviors to see if they are consistent with that model. For example, Zhang and colleagues [30] instructed students in the use of a *target node strategy* while constructing system dynamics models and then measured their use of that strategy. Results showed that in one experiment, 34 and 39 percent of students' steps (in two different groups) were consistent with the strategy. In a second experiment, two additional groups of students achieved consistency scores of 66 and 70 percent, indicating a wide variation in how often students were successfully using the strategy between experiments.

Previous work in educational data mining has been used to understand the strategies students use while learning in OELEs. For example, Perera et al. [35] used sequence mining to derive frequently-used learning behaviors of more- and less-successful student groups. They then provided mirroring and feedback tools to support effective teamwork among students, using the derived feedback based on more-successful groups' behaviors. Results showed that mirroring and feedback helped all groups improve their work by emulating the behaviors of the strong groups. In previous work, we have compared sequential patterns derived from student activity sequences to identify ones that differ in use between two or more groups of students [15] and over time [36]. Nesbit et al. [37] use sequential pattern mining to find the longest common subsequences across a set of action files from the gStudy learning environment and study how students self-regulate as they learn.

A variety of other researchers have also employed sequential pattern mining to generate student models for customizing learning to individual students [38], [39]. In this body of work, however, researchers are not explicitly tracking students' understanding and/or use of strategies. Rather, they are characterizing students' behaviors in order to understand how learning happens in complex learning environments.

4 Framework Integrating Model- and Data-Driven Analysis

Our framework for analyzing OELE learning activity data integrates task and strategy modeling with data-driven sequential pattern discovery, as illustrated in Fig. 1. The model-driven component, shown in the top of the figure, is based on a hierarchical task model that describes the relationships between top-level domain-general OELE tasks to domain-specific tasks and subtasks. At the lowest level, these subtasks link to tool-specific actions in the particular OELE that facilitate learning and problem solving in a particular learning domain. The strategy model complements this task model by describing how related actions combined across tasks and subtasks define learning and problem-solving strategies. By specifying a temporal order and conceptual relationships among elements of the task model that define a strategy, the strategy model codifies the semantics that provide the basis for analyzing students' actions beyond the categorical information available in the task model.

The complementary, data-driven portion of the framework, outlined in the bottom of the figure, describes how model-driven analytics and relations can be combined with sequence mining techniques to improve the detection and interpretation of behavior patterns that can be linked to students' learning and problem solving behaviors. As shown, data-driven analysis of these behaviors first employs sequential pattern mining techniques to identify frequent action patterns. However, it goes beyond a traditional application of pattern mining by then mapping the identified patterns back into the activity sequences to analyze specific instances with respect to the task and strategy models and the context in which these actions and strategies are applied. This integrated analysis allows us to map the discovered behavior patterns to cognitive and metacognitive processes that can be associated with students' activities in the learning environment. In addition, behaviors that do not match known strategies can be identified to fill gaps in the coverage of the current models. They may also imply suboptimal strategies and processes that students employ when working on their learning and problem solving tasks in these environments [7].

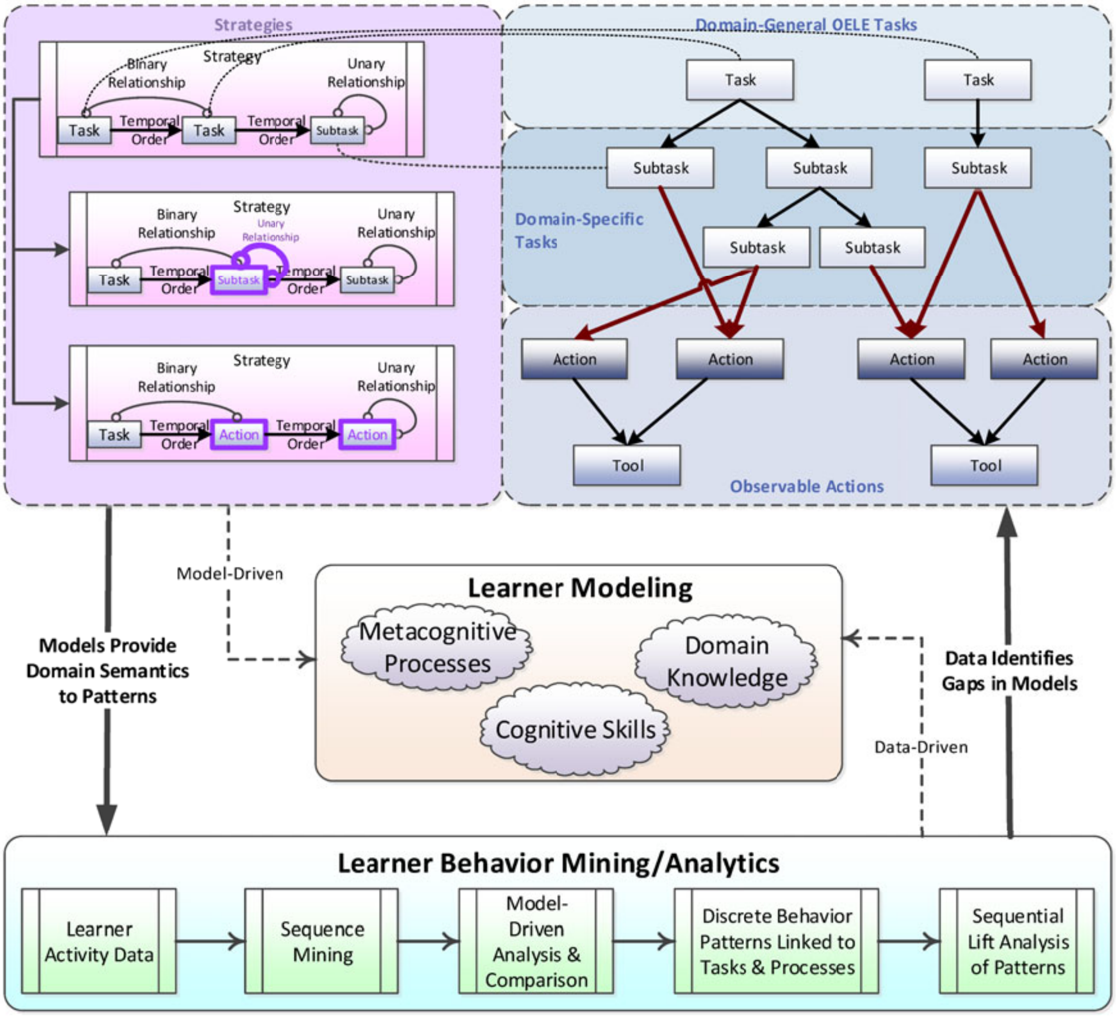

Fig. 1. Integrated analytic framework.

In the rest of this section, we describe the task and strategy modeling representations, context relations that we instantiate in this framework for defining cognitive and metacognitive strategies, and the pattern discovery techniques and measures employed for analyzing learning and problem-solving behavior. The combined model- and data-driven framework provides the basis for dynamic learner modeling (the central part of Fig. 1), but an extended discussion of learner modeling is outside the scope of this paper.

4.1 Model-Driven Strategy Detection

In our integrated framework, learning and problem-solving activities are modeled through linked task and strategy models, as illustrated at the top of Fig. 1. The task model is represented as a directed (acyclic) graph, which provides a successive, hierarchical breakdown of the tasks into their component subtasks[1] in the OELE. At the lowest levels of the hierarchy the tasks are linked to the observable actions that can be performed with the tools provided in the OELE. The links in the task model are categorical in nature, indicating, for example, that two subtasks like "Identifying Relevant Information" and "Interpreting Information" are part of a (more general) parent task like "Information Acquisition." However, the task model does not indicate whether (or in what circumstances) both subtasks need to be completed for effective information acquisition, nor whether there are any necessary relations (such as ordering) among them. Similarly, links from a task/subtask to actions indicate have to be executed to complete the task or a subset of the actions might suffice.

Instead, it is the strategy model that captures this information in a form that can be directly leveraged for analysis. This model describes how actions, or higher-level tasks and subtasks, can be combined to provide different approaches or strategies for accomplishing learning and problem-solving goals. In our approach, strategies manifest as partially-ordered sets of elements from the task model with additional relationships among those elements determining whether a particular, observed behavior can be interpreted as matching the specified strategy. Fig. 1 illustrates unary relationships for additional features or characterizations of a single strategy element, binary relationships among pairs of elements, and the temporal ordering among elements of the strategy. Further, if a relationship is not specified between any two elements in a strategy, then the strategy is

1. While we use the terms "task" and "subtask" for describing different granularities of tasks in our description, there is no explicit distinction in the task model representation between "tasks" and "subtasks."

agnostic to the existence or non-existence of that relationship.[2] Because the elements of the task model used in the definition of strategies are hierarchically related, strategies may also naturally be related from more general strategy definitions to more specific variants. In this representation, specifying additional relationships, additional elements, or more specific elements (e.g., a specific action replacing a more general task/subtask) derive a more specific strategy from a general one.

The structure of our task and strategy models incorporates two key insights that relate to understanding students' cognitive and metacognitive processes in an OELE. First, higher-level tasks may comprise multiple lower-level tasks. This implies that students may fail to complete higher-level tasks for multiple reasons based on their ability (or lack of ability) to complete corresponding lower-level tasks. For example, students may struggle to successfully evaluate their solution in an OELE because: (1) they do not evaluate their solution with the provided system tools; or (2) because they are unable to interpret the results of such an evaluation. Second, the combination of multiple tasks/subtasks in the strategy representation illustrates the coordination of multiple learning and problem-solving activities, with corresponding skills and cognitive processes, which is a vital component of metacognitive monitoring and regulation.

Many relations can be defined among tasks/actions for strategy representation in this framework, such as those defined for *coherence analysis* (CA) [14], which analyzes learners' behaviors to produce measures describing the coherence of OELE activities. Multiple CA relations and measures (described in [14]) could be employed in the strategy representation, but here we will focus on the CA relation called *support*:

> *Two ordered actions ($x \rightarrow y$), i.e., x before y, taken by a student in an OELE exhibit the* support *relationship if the second action, y, is based on information generated as a result of the first action, x. In this case, x provides* support *for y, and y is* supported *by x. Note that while x must occur before y, they need not be consecutive actions.*

The assumption behind describing the coherence of actions with the support measure is that students are more likely demonstrating effective metacognitive regulation when an action they perform is *supported* by information that was generated from one of their previous actions [14]. In addition, to assess an aspect of student's cognitive skills, we also define a unary relation of *effectiveness* that judges whether a change to a problem solution moves it closer to a correct solution. Overall, students with higher proportions of effective actions are considered to have a higher mastery of the related cognitive skills.

4.2 Integrating Data-Driven Discovery

The task and strategy models defined in the previous section describe the learning and problem-solving activities for an OELE in a form that can be directly related to observed patterns in interaction traces. The task model links any task to a set of observable actions, and strategies are defined with respect to the task model, which ultimately allows the automated interpretation of any given pattern of actions observed in the OELE as an instance of any matching strategies.[3]

To identify patterns of actions for model-driven analysis and interpretation, we employ a data-driven approach based on sequential pattern mining [40], [41]. The mining algorithm discovers common behaviors as action patterns from the sequences of student actions in the OELE, which are collected in log files. To generate meaningful results from the sequential data mining, raw activity logs must first be transformed into an appropriate *abstracted* sequence of actions [15]. The set of possible actions are defined in the task model and the set of action relationships used in the strategy model identifies additional context information that may be needed for strategy matching. Pre-processing log files into sequences of these actions, augmented with the necessary information for identifying the action relationships, yields a representation amenable to mining and further analysis while filtering out unnecessary detail (e.g., cursor position) and combining qualitatively-similar event types (e.g., when an action can be performed through either of two available interface features). A more detailed discussion on pre-processing and defining the set of actions that make up the action sequences is presented in [15], [42].

After pre-processing, the resulting action sequences are mined for common patterns. A sequential pattern mining algorithm (e.g., Pex-SPAM [43] in the analysis presented here) is used to identify all patterns that meet a given sequence frequency threshold (i.e., the identified patterns are observed for at least a given percentage of students). One difficulty in mining frequent patterns in learning interaction traces is that the action sequences are "noisy." Students may coordinate their activities in a particular behavior pattern, but they may also perform additional actions interspersed with the actions that constitute the pattern making it more difficult to identify. To increase the number of behavior patterns that can be identified in such data, we allow up to one irrelevant or variable action between consecutive actions in identifying or matching a pattern (i.e., a maximum gap of 1 in sequence mining).[4] Further, to identify sufficiently common patterns, i.e., patterns observed for the majority of the students, we apply a sequence frequency threshold of 50 percent with the sequential pattern mining algorithm.

In general, common behavior patterns identified by sequence mining algorithms then have to be interpreted and analyzed by researchers to identify a relevant subset of important patterns that provide a basis for generating actionable insights (e.g., how to support users and encourage specific, more productive learning and problem-solving behaviors). Our framework systematizes and automates a

2. Explicitly requiring the *lack* of a relationship in a strategy definition is accomplished by defining an additional relationship that simply indicates the lack of the original relationship.

3. The hierarchical structure of the task and strategy models implies that a given pattern of actions may match multiple strategy descriptions, defined at different levels of detail/specificity.

4. The choice of gap size depends both on the data being analyzed and the goals of the analysis performed. In some situations, other sizes of gaps or no gap at all may be the most appropriate choice.

significant portion of this process by integrating the model-driven and data-driven perspectives. The first step in relating mined patterns to modeled strategies is to map the patterns back into student sequences to identify the individual *instances* of each pattern (i.e., each occurrence of the pattern, including the details specific to the actions taken in that occurrence). The patterns are then analyzed in the context of these individual instances to calculate all relationships (defined by the strategy model) that hold among the specific actions performed in each instance. This enables more effective interpretation and differentiation among behaviors that result in the same action pattern but have different relationships among the instances of those actions. For example, the sequential pattern mining algorithm might identify the pattern "*A brief reading action followed by adding a component to the solution*." In order to identify common patterns like this example, the definition of distinct actions necessarily omits a variety of details, such as the specific page read and the particular component that was added to the solution in a given instance. As a result, this pattern may occur a number of times for a single student and across students, and each time it may involve reading a different page and adding a different component. The specific details of each instance of this pattern are then used to determine whether a particular relationship holds among the actions. For example, the support relation is satisfied if the information available on the page read is linked to the component that was added to the evolving solution structure.

By taking into account relations, such as support, our framework can distinguish between different variants of a strategy (e.g., effective versus ineffective) that are defined by the same action pattern but differ in their instantiation (e.g., whether one of the component actions supports another). While a binary relation like support can often only be calculated after identifying a specific pattern instance, unary relations (e.g., whether an action is effective) apply to individual actions and could be used to refine the definition of canonical actions in the task model. However, though this information can be very useful in contextualizing the meaning and use of derived patterns that contain these actions, that approach of defining finer-grained actions also has the effect of reducing the frequency of observed patterns. In particular, the qualification of actions by an additional feature may reduce the occurrence of some patterns containing these actions to below the mining frequency threshold, preventing those patterns from being discovered at all. Our approach circumvents this problem by calculating both binary relationships (of necessity) and unary relationships (by choice) with respect to instances of patterns defined and discovered with actions that were initially undifferentiated by those relationships.

In addition to using task and strategy models to enhance the data-driven analysis, our framework enables the reverse, in which data-driven pattern discovery can be used to iteratively extend existing models. Specifically, action patterns that have no matches in the strategy model can suggest candidates for defining new strategies to increase model coverage. However, many patterns may be common simply due to the frequency of their component actions, rather than representing an intentional process combining those actions. Therefore, our approach goes one step further by extending a heuristic from the analysis of association rules, called the lift measure [21], to sequential patterns. The original lift measure compares the frequency with which an association rule is matched in the data to the expected frequency using a baseline random model that assumes independence of the left- and right-hand sides of the rule. We extend this approach to a *sequential lift* measure, which is straightforward for length-2 sequential patterns: we calculate the ratio of the observed frequency of each pattern to the expected frequency from a random model that assumes that each action is chosen independently. In the random model, the probability of a particular action occurring at any point in a sequence is simply its a priori probability (which is approximated by the observed action frequency across all students) without respect to previous actions performed by the student. A high lift (i.e., the ratio of observed pattern frequency to the expected frequency from the random model) implies that the pattern is more likely to correspond to an explicit strategy or procedure, rather than a random combination of the component actions. In this case, we refer to the independent random model as a "level-1" model because it employs only the frequency of length-1 patterns (i.e., single actions) in the data.

However, using the same level-1 random model for calculating lift in length-3 or longer patterns may be less effective. The high lift patterns with this approach would include a variety of extensions to the high lift length-2 patterns, where an additional frequent action is added at the beginning or end. In other words, the lift in many of the length-3 patterns can largely be attributed to a length-2 sub-pattern with high lift. To account for this, we use an incremental modeling approach in which we employ a level-2 random model for calculating lift with length-3 patterns. The level-2 random model uses knowledge of length-2 pattern frequency from the data to approximate the conditional probability of a second action following a first action. Then the expected frequency of the length-3 pattern for this level-2 random model is based on the probability of the first action, the conditional probability of the second action given the first, and the conditional probability of the third action given the second. This incremental approach can be extended to length-4 patterns by using a level-3 random model (i.e., one in which probabilities of subsequent actions are conditioned on the preceding two actions), and so on.

Thus, our framework integrates the model- and data-driven perspectives to iteratively improve both the models and the analysis. Illustrated from the top down in Fig. 1, the integration of a model-driven perspective addresses two important challenges in effective use of pattern discovery with student activity traces: (1) in the context of a specific sequence of actions performed by a student, comparison to known strategies can improve the ability to distinguish chance occurrences of a particular pattern from instances that imply effective learning and problem-solving behaviors; and (2) matches of a portion of a pattern to strategies in the model can help determine whether and where to split discovered patterns into distinct behaviors for more accurate analysis and interpretation. Illustrated from the bottom up in Fig. 1, the

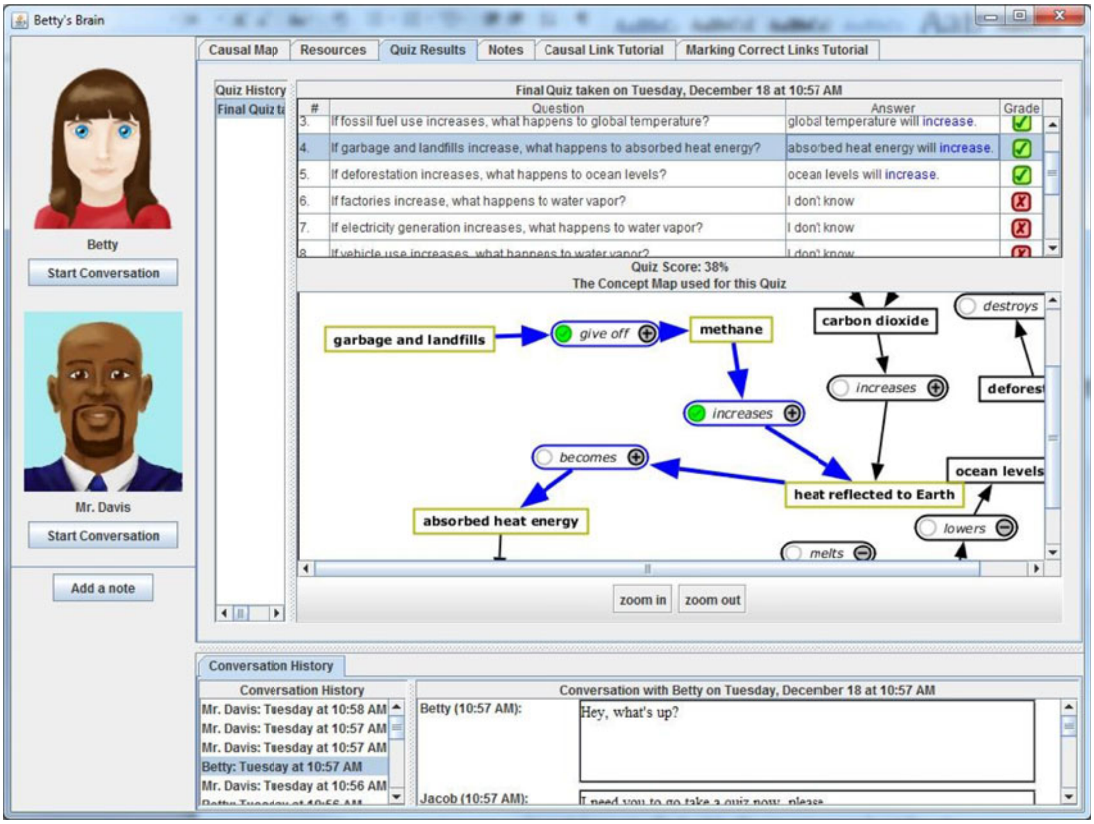

Fig. 2. Betty's Brain system showing the quiz interface.

integration of a data-driven perspective addresses an important challenge in developing task and strategy models: with appropriate heuristics like the sequential lift analysis, patterns that are more likely to correspond to explicit strategies rather than coincidental combinations of actions can be identified in order to extend the coverage of existing models in a data-driven manner.

5 Applying the Framework to Betty's Brain

The Betty's Brain learning environment [22], [44] presents students with the task of teaching a virtual agent named Betty, a science topic by constructing a visual causal map that represents the relevant science phenomena as a set of entities connected by directed links that represent causal relations. Once taught, Betty can use the map to answer causal questions and explain those answers. The goal for students using Betty's Brain is to teach Betty a causal map that matches a hidden, expert model of the domain. The students' learning and teaching tasks are organized around three activities: (1) reading hypertext resources, (2) building the map, and (3) assessing the correctness of the map. The hypertext resources describe the science topic under study (e.g., climate change) by breaking it down into a set of sub-topics. Each sub-topic describes a system or a process (e.g., the greenhouse effect) in terms of entities (e.g., absorbed heat energy) and causal relations among those entities (*absorbed heat energy* increases the *average global temperature*). As students read, they need to identify causal relations and then explicitly teach those relations to Betty by adding them to the current causal map. Fig. 2 illustrates the Betty's Brain system interface.

Learners can assess the quality of their current map in two ways. First, they can ask Betty to answer a cause-and-effect question using a template. After Betty answers the question, learners can ask Mr. Davis, another pedagogical agent who serves as a mentor, to evaluate her answer. If the portion of the map that Betty uses to answer the question matches the expert model, then Betty's answer is correct. Learners can also have Betty take a quiz on one or all of the sub-topics in the resources. Quiz questions are selected dynamically by comparing Betty's current causal map to the expert map. Since the quiz is designed to reflect the current state of the student's map, a set of questions is chosen (in proportion to the completeness of the map) for which Betty will generate *correct* answers. The rest of the quiz questions produce either *incorrect* or *incomplete* answers. These answers help students determine the correctness of causal links. Students may also realize that Betty is unable to answer questions because they have not taught her certain links. Should learners be unsure of how to proceed in their learning task, they can ask Mr. Davis for help via a menu-based conversation that allows the user to choose from a set of pre-specified options. Mr. Davis responds by asking learners about what they are trying to do and responds with suggestions appropriate to the user's indicated goals [44].

5.1 Task and Strategy Modeling for Betty's Brain

In Betty's Brain, we instantiate the task portion of the model as shown in Fig. 3. The top level of the model identifies the three broad classes of OELE tasks related to: (i) *information seeking and acquisition*, (ii) *solution construction and refinement*,

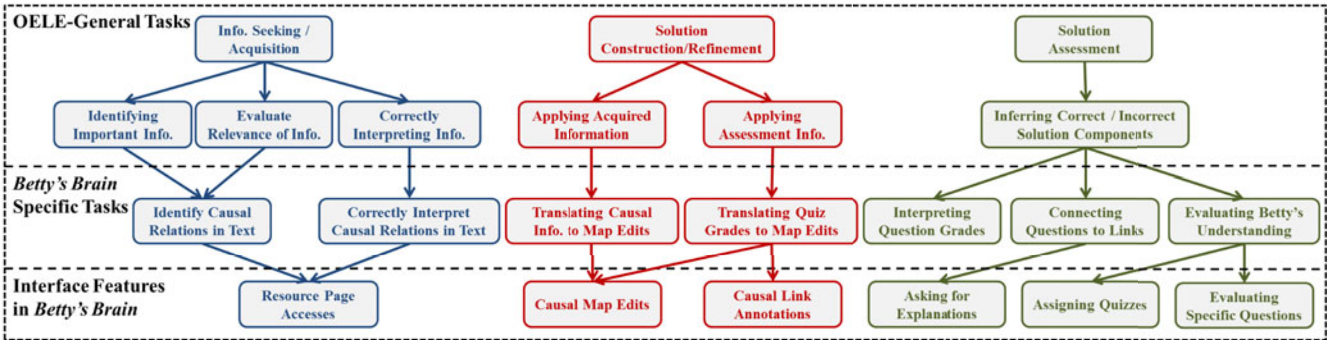

Fig. 3. Task model for *Betty's Brain*.

and (iii) *solution assessment*. Each of these task categories is further broken down into three levels that represent: (i) general task descriptions that are common across many OELEs; (ii) Betty's Brain-specific instantiations of these tasks; and (iii) actions using tools and interface features in Betty's Brain through which students can accomplish their tasks.

Information seeking and acquisition involves identifying, evaluating the relevance of, and interpreting information in the context of the overall task. Solution construction and refinement tasks involve applying information gained both by conducting information seeking tasks and by analyzing the solution assessment results to construct and refine an existing solution (e.g., an evolving model of a science process). Finally, solution assessment tasks involve interpreting the results of solution assessments as actionable information that can be used to refine the solution in progress. In order to accomplish these general tasks in Betty's Brain, students must understand how to perform the related Betty's Brain specific tasks by utilizing the system's interface features and tools. For example, successful information seeking in Betty's Brain involves identifying and correctly interpreting the causal relations described in the resource pages. Students must then translate this information into an equivalent "causal relation" form that can be used to construct and refine the causal map. Similarly, solution evaluation tasks involve interpreting Betty's quiz results as actionable information that can be used to evaluate and refine the causal map.

To specify and match strategies to instances of action patterns (with respect to this task model), we instantiate the support relation for two cases of causal link edits: (i) those that are *supported* by one or more information seeking actions, e.g., reading resources pages where the concepts associated with the link edit appear on those pages; and (ii) those that are *supported* by solution evaluation actions, e.g., having Betty take a quiz or asking Betty to explain her answer to a quiz question, and these evaluation actions provide evidence that a particular causal link is correct or maybe incorrect.[5]

6 OELE Study and Data

The analyses presented in this paper used data from a recent classroom study with Betty's Brain. The study was designed to assess the effect of two support modules that used the pedagogical agents to scaffold students' understanding of skills and strategies important for success. One module (*Mod-1*) provided guidance and practice on identifying causal relations in the resources and using the information to structure the causal map. A second module (*Mod-2*) provided guidance and practice on monitoring Betty's progress using the quiz results to identify correct and incorrect causal links in the map. Participants were divided into four treatment groups with one for each of the support modules, one with both support modules, and a control with neither support module and only the baseline level of support from the agents that was available in all groups.

6.1 Participants

Ninety-eight sixth-grade students from four middle Tennessee science classrooms, taught by the same teacher, participated in the study. We did not collect demographic information for this group of students, but the school website reports that the student population is 51.4 percent female. We believe that this demographic holds for the students who participated in our study. The current version of Betty's Brain requires that students' possess the ability to independently read and understand the resources, therefore, this study was not suitable for students with limited English proficiency. Also, the system does not currently support students who are visually impaired or have cognitive disabilities. As a result, we did not analyze data from English Language Learners or those with special education needs. We also excluded data of students who missed more than two class periods of work on the system. The experimental analysis reported in this paper used data from 68 students.

6.2 Topic Unit and Text Resources

Students used the Betty's Brain system to learn about climate change. The expert map contained 22 concepts and 25 links representing the greenhouse effect (solar energy, absorbed light energy, absorbed heat energy, global temperature, and heat reflected to earth), human activities affecting the global climate (deforestation, vegetation, vehicle use, factories, electricity generation, fossil fuel use, carbon dioxide, garbage and landfills, and methane), and impacts on climate (sea ice, ocean level, coastal flooding, carrying capacity, condensation, water vapor, precipitation, and drought). The resources were organized into

5. In the analysis presented in this paper, an action only provides support for a causal map edit if both the action and the edit take place within 10 minutes of each other.

TABLE 1
Performance [Mean (s.d.)] by Treatment
(Max Pre- and Post-Test Score = 16, Correct Map Score = 25)

Group	Pre-Test	Post-Test	Gains	Best Map
Control	3.07 (1.83)	6.53 (1.55)	3.47 (2.64)	8.87 (8.20)
Mod-1	3.10 (1.94)	6.25 (2.05)	3.15 (2.13)	9.55 (6.64)
Mod-2	2.65 (1.27)	6.59 (2.00)	3.94 (2.41)	9.53 (7.55)
Mod-1&2	3.06 (1.24)	5.88 (2.34)	2.81 (2.29)	7.25 (6.36)

one introductory page, three pages covering the greenhouse effect, four pages covering human activities, and two pages covering impacts on climate. Additionally, a glossary section provided a description of some of the concepts, one per page. The complete resources was made up of 31 hypertext pages (4,188 words and accompanying figures) with a Flesch-Kincaid reading grade level of 8.4.[6]

6.3 Learning and Performance Metrics

Learning was assessed using a pre-post test design. Each written test was made up of five questions that asked students to consider a given scenario (e.g., a significant increase in the use of road vehicles) and explain its causal impact on climate change. These questions were challenging because they required students to remember information from the science resources, follow a chain of causal relations from the expert model on climate change, and then express those ideas in writing. The score for a question was computed by counting the number of correct causal relations that students used in their answers. Correctness was evaluated by comparing the set of causal relations used in the students' answers to the causal relations that would be generated from the expert map. Students received 1 point for each link in their answers that corresponded to an expert causal link. The maximum combined score for the five questions was 16.

Two coders independently scored the same test for a small subset of students and then discussed any differences in their scoring with each other and a researcher. They repeated this process with additional subsets of tests until they achieved at least 85 percent agreement when independently scoring a set of tests. At this point, the two coders split the remaining tests, and scored them independently. Performance on the system was assessed by calculating the number of correct links (the links in the student's map that appeared in the expert map) minus the number of incorrect links in the student's final map.

6.4 Study Procedure

The study was conducted for nine school days, with students participating for a 60-minute class period each day. During the first class period, students completed the pre-test. During the second and third class periods, researchers taught students how to build and reason with causal models and how to identify causal relations while reading text passages. Students also worked on simple paper-and-pencil exercises that required them to identify causal relations, build simple maps, and answer questions using their causal map. During the fourth class period, students were provided with hands-on system training by the researchers. Students then spent four class periods (days 5-8) working with their respective versions of the Betty's Brain system with minimal intervention by the teachers and the researchers. On the ninth day, students completed the post-test that was identical to the pre-test.

6.5 Log Analysis

To extract the activity sequences for mining, log events captured by the learning environment were mapped to sequences of canonical actions as described in Section 4.2. As in previous work analyzing Betty's Brain log data, we abstracted student activities into a few primary categories with some additional subcategories [15]. Further, each of these actions was linked to relevant tasks from the task model (at both the OELE-general and Betty's Brain-specific level) so strategies specified in the strategy model could be matched at the corresponding level of detail to student actions. The primary actions extracted from the logs to generate the action sequences were:

- *Read*: students access a page in the resources;
- *Note*: students use a note-taking tool to create or edit a note;
- *Edit*: students edit the causal map, with actions further divided by: (i) whether they operate on a causal *link* or *concept* and whether the action was an addition (*Add*), removal (*Remove*), or modification (*Change*), e.g., *LinkAdd* or *ConceptRemove*;
- *Query*: students use a template to ask Betty a question, and she answers the question using a causal reasoning algorithm [22];
- *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; and
- *Explain*: students probe Betty's reasoning by asking her to explain her answer to a question (either from the quiz or from a query).

7 Results

To illustrates the benefits of our combined model- and data-driven analysis framework, we applied it to the Betty's Brain data from the study described in Section 6. As background, we first present results describing the overall outcomes of the intervention. To determine if our intervention helped students learn the science content and causal reasoning skills, we computed: (i) student pre-to-post learning gains, and (ii) students' causal map scores.[7] Table 1 presents these results for each treatment in the intervention. A repeated measures ANOVA performed on the pre- and post-test data revealed a significant effect of time on pre-to-post-test scores (F=28.656, $p<.001$, $\eta_p^2 = 0.481$), but it failed to

6. The Betty's Brain system can be downloaded from http://www.teachableagents.org/downloadsoftware.php.

7. We use the highest causal map score a student achieved at any time during the intervention rather than their final map score to avoid penalizing students for incorrect guesses or deletions made in haste near the end of the intervention when there was insufficient time to correct them.

reveal a significant effect of treatment ($F = 1.402$, $p =$ n.s., $\eta_p^2 = 0.044$). Similarly, an ANOVA revealed no significant effect of the treatment on the map scores ($F = 0.044$, $p =$ n.s., $\eta_p^2 = 0.011$).

Clearly the students learned as the result of the intervention and several students produced a significant portion of the correct causal map (for the class as whole, the mean map score was 8.85, s.d. 7.08). However, the small sample sizes and the large variations in performance within groups (much more so than across groups) make detailed analysis of the experimental treatments difficult. Therefore, in this paper, we focus on analyzing the different action patterns corresponding to strategies of interest across the entire sample of students. Further, we compare behaviors of students with high map scores with those of students who had low map scores, without regard to treatment. The median map score was 7.5 (i.e., there were an even number of students with the median falling between a student with a map score of 7 and one with a map score of 8, since map scores are whole numbers), so we consider the students below the median (a map score of 7 or lower) as the "LowMap" group and the ones above the median (a map score of 8 or higher) as the "HiMap" group.

As a case study of our model- and data-driven framework, we analyze some basic strategies that use information seeking (IS) and solution assessment (SA) to drive solution construction (SC) activities. Specifically, we consider several variants of two high-level strategies defined by $IS \Longrightarrow SC$ and $SA \Longrightarrow SC$, where $\Longrightarrow$ captures both temporal and support relationships. Therefore, $IS \Longrightarrow SC$ indicates that the SC activity follows the IS activity and uses information that was generated by the IS activity. In short, IS supports SC, following the definition in Section 4.1.

Following the methodology described in Section 4, each pattern identified by the data-driven sequence mining is analyzed with respect to these strategies and the more specific variants described below. Because we present an exploratory analysis of behavior and strategy use, rather than an attempt to verify experimental hypotheses, we do not present statistical significance tests that could be misinterpreted in this context. Instead, we only describe differences in strategy use with descriptive statistics, effect sizes (as Cohen's d), and data visualizations. Further, in analyzing features beyond counts of strategy and activity pattern occurrences (e.g., effectiveness of an SC activity or length of an IS activity in the occurrences), we only include students with at least three occurrences of the given strategy/pattern to avoid giving undue weight to single occurrences.

Analyzing the basic strategy of applying information from resources to constructing the causal model (i.e., $IS \Longrightarrow SC$), we find that HiMap students exhibit this general strategy much more often than LowMap students (d = 1.32), although both groups make a relatively large percentage of errors when doing so: the percentage of $IS \Longrightarrow SC$ instances that are effective (i.e., result in an improvement in map score) for the HiMap group is only 62 percent (s.d. = 9%) and for the LowMap group is only 53 percent (s.d. = 16%).

To better understand how students employ this general strategy, we consider two specific variants possible in Betty's Brain: $IS \Longrightarrow Add[ition]$ of a causal link to the map and $IS \Longrightarrow Correct[ion]$ by changing or removing an incorrect causal link in the map. Results indicate that $IS \Longrightarrow Add$ is used more often than $IS \Longrightarrow Correct$ in both the HiMap and LowMap groups, as shown in Table 2. In addition, these results indicate that employing information from the resources to correct the solution ($IS \Longrightarrow Correct$), while less frequent, is more often effective than using information from the resources to add to the solution ($IS \Longrightarrow Add$) in both the HiMap and LowMap groups.

TABLE 2
$IS \Longrightarrow SC$ Behaviors (Effectiveness Calculated for Students with at Least Three Occurrences)

Group	$IS \Longrightarrow Add$		$IS \Longrightarrow Correct$	
	Occurrence	Effective	Occurrence	Effective
LowMap	8.9 (7.5)	48.4% (22.2%)	2.9 (3.6)	81.1% (23.5%)
HiMap	23.9 (12.0)	59.3% (13.1%)	5.3 (4.0)	75.1% (17.2%)
All	*17.3 (12.7)*	*55.2% (17.7%)*	*4.2 (4.0)*	*76.8% (18.9%)*

Further, we analyzed the more common of these two strategies ($IS \Longrightarrow Add$) to illustrate how a purely data-driven perspective (even with a task model relating individual actions to IS and $Add[ition]$ tasks) might differ from our approach that integrates an explicit strategy model. To do this we compared instances of the corresponding logfile activity pattern ($Read \rightarrow LinkAdd$) identified by sequence mining, depending on whether the individual occurrence matched the strategy. For this particular strategy the distinction between matching the basic activity pattern and matching the strategy was whether the $LinkAdd$ action was supported by the preceding $Read$ action for a given occurrence within the student's logged sequence of actions. In general, a variety of other relationships or specific combinations of relationships in longer patterns, as well as the hierarchical relationship of tasks to multiple corresponding actions, might distinguish a match to the strategy model versus a simple match to an activity pattern.

Fig. 4 illustrates that there were important differences between the occurrences of this activity pattern that matched the strategy model and those that did not. The pattern occurrences matching the $IS \Longrightarrow Add$ strategy tended to involve longer reading times (mean 46s [s.d. 17s] versus mean 35s [s.d. 22 s] for non-strategy occurrences) and were more often effective (mean 60 percent [s.d. 14 percent] versus mean 33 percent [s.d. 20 percent] for non-strategy occurrences), regardless of which group the student belonged to. Further, of the $Read \rightarrow LinkAdd$ activity pattern occurrences, HiMap students tended to have a higher percentage that matched the $IS \Longrightarrow Add$ strategy than the LowMap students. While some $IS \Longrightarrow Add$ strategy matches could still be coincidence rather than intended applications of the corresponding strategy, it allows us to disregard many $Read \rightarrow LinkAdd$ activity pattern occurrences that *cannot* be coherent applications of the strategy (because the $LinkAdd$ action is not supported by the $Read$ action). This allows a more precise characterization of how students employ this strategy (e.g., in terms of associated reading time and effectiveness of the causal map additions) and can highlight

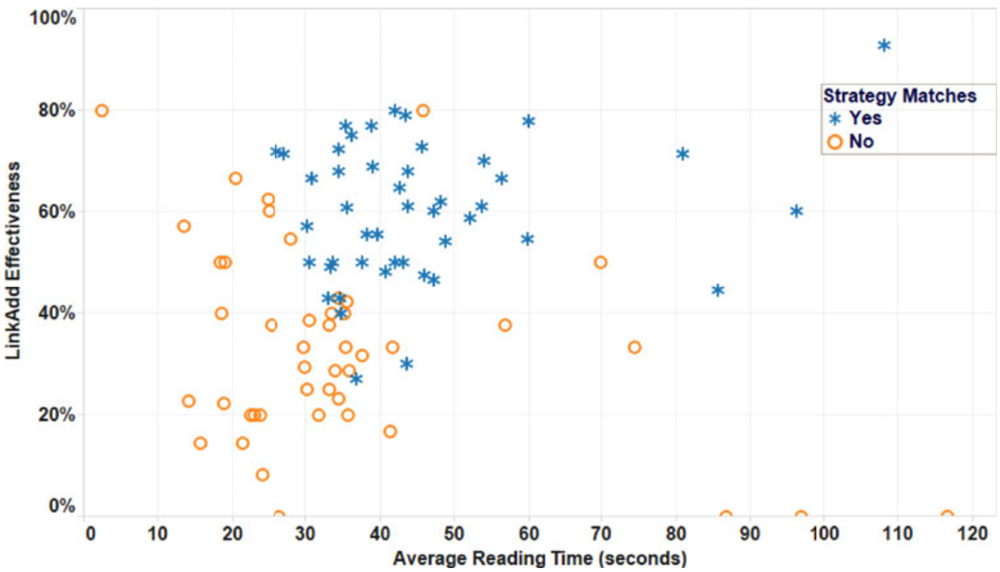

Fig. 4. Average reading time and percentage effectiveness of $Read \rightarrow LinkAdd$ occurrences per student.

important differences across students (e.g., the relatively more frequent application of the strategy, as opposed to coincidental or other instances of $Read \rightarrow LinkAdd$, in the HiMap group).

Finally, we consider how uses of causal map correction activities differ depending on whether they are driven by information seeking or solution assessment activities. Specifically, we analyze the frequency of the strategies $IS \Longrightarrow Correct$ and $SA \Longrightarrow Correct$. Overall, both correction strategies are much less frequent in the LowMap group than the HiMap group (IS-based d = 0.66, SA-based d = 0.90). Further, the HiMap group employs SA-based correction (relative to total use of both correction strategies) more than the LowMap group (between groups d = 0.41). This analysis suggests that the lower-performing students may not be as aware of, or as skilled in, correcting their causal map using the quiz assessment functionality. Given that both groups add a variety of incorrect links to their map, the strategies for identifying and correcting mistakes are vital for eventually getting to the correct map. Therefore, these strategies are important to tracking and provide support for when students show a lack of proficiency with them.

In addition to leveraging the task and strategy models for analysis of activity data, our framework completes the modeling-analysis cycle by using the data-driven results to identify gaps in the coverage of the strategy model. To identify patterns that may represent these previously unconsidered strategies (i.e., ones not represented in the strategy model), we employ the sequential lift analysis method described in Section 4.2. When considering the 2-length patterns with high lift first, many of the top patterns are repetitions of the same activity. This seems reasonable as many activities lend themselves to repetitive use, making their observed repetition frequency higher than what an independent random model would predict. For example, patterns of marking links in the model (as correct, maybe wrong, or unknown) one after another provide four of the top five lift patterns, and all have a lift of at least 5 (i.e., five times more frequent than what the level-1 random model predicts). Pairs of searching actions (in the resources) and of note-taking actions also have high lift values (22.0 and 6.3, respectively). The especially high lift value of $Search \rightarrow Search$ suggests that students rarely find the information they are searching for on the first try, indicating that the search functionality provided in the system (or other forms of support that help students identify good keywords for searches) could be improved. Most of the other high lift 2-length patterns involve related activities that illustrate simple or partial strategies like getting an explanation for a quiz question after taking a quiz or taking a note after reading a page in the resources.

Of more interest, are the longer patterns with high lift. The 3-length pattern with the highest lift value, as compared to the (level-2) random model, is $TakeNote \rightarrow Read \rightarrow TakeNote$ (lift = 7.6). This indicates that even accounting for the frequency of reading following note-taking and of note-taking following reading, this pattern of repeated alternation between note-taking and reading is more frequent than what we would expect. Therefore, it should be included in our strategy model. Even more interesting are the high lift patterns that take the form $SA \rightarrow LinkAdd \rightarrow SA$ with two specific variants: $QuizExplanation \rightarrow LinkAdd \rightarrow Quiz$ (lift = 2.2) and $Quiz \rightarrow LinkAdd \rightarrow Quiz$ (lift = 2.0). These patterns suggest an informed *guess-and-check* strategy in which the quiz results (either the overall results or the information gleaned from a $QuizExplanation$ for a specific quiz question) is used to suggest a potentially-missing link, which is then added to the map. This is followed by checking the correctness of the "guess" by taking another quiz. As expected, for a *guess-and-check* strategy, the link added was usually incorrect (average percentage of correct additions per student = 19%, s.d. = 14%).

Given the more detailed information available from the quiz question explanation, we initially expected better performance in adding a correct link for that variant of the strategy. However, this variant actually had a marginally lower percentage of correct additions compared to the other

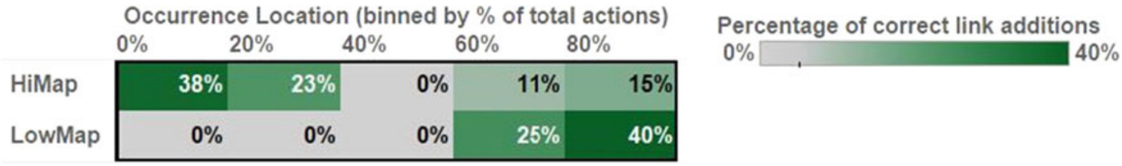

Fig. 5. Heatmap of percentage correct links added in $SA \rightarrow LinkAdd \rightarrow SA$ over time.

variant. Further, analysis of effectiveness over the course of students' work in the environment illustrated that performance (correctness percentage) with $SA \rightarrow LinkAdd \rightarrow SA$ was better early and late while being especially poor in the middle. The performance heatmap shown in Fig. 5, indicates that while the HiMap students performed best with this strategy early and (to a lesser extent) late in the intervention, the LowMap students did not make correct link additions with this strategy until relatively late (after at least 60 percent of their total actions on the system). This may imply that it took the LowMap students until late in the intervention to understand how to interpret and use the quiz results. On the other hand, the HiMap students used this strategy with more success in the early phases of map building. The HiMap students' effectiveness with this strategy may have dropped off once they started dealing with the more difficult material (for which they had little prior knowledge) toward the middle of their activities, finally rebounding some as they gained proficiency. In addition to illustrating the importance of incorporating the overall informed *guess-and-check* strategy in the strategy model, analysis of this high lift pattern suggests that there may be additional interactions with prior knowledge and skills worth investigating through further experiments.

8 Conclusions

In this paper, we presented a framework for analyzing students' activity data in open-ended learning environments that integrates model-driven behavior characterization and data-driven pattern discovery. In this framework, the model-driven approach uses linked task and strategy models to provide more precise interpretation of students' activity sequences as learning and problem-solving strategies, while the pattern mining approach enables the identification of new variations of strategies and of gaps in the coverage of the current strategy model. This analytic framework extends an initial combination of task models and pattern discovery [17], [18] by providing a systematic integration of the model- and data-driven approaches to address challenges from each direction. This integration is achieved through the design and incorporation of (1) an explicit strategy model used to analyze individual instances of discovered action patterns (in the context of a student's complete action sequence) and (2) a sequential lift analysis that uses the data to augment the set of modeled strategies.

We instantiated and applied this framework for the Betty's Brain OELE where the task and strategy models include (1) the primary tasks of knowledge acquisition, application of this knowledge to problem solving (in this case, model building), and verification of solutions using information generated through assessment tools in the environment; (2) strategies combining these tasks with respect to the support relation from coherence analysis [14]; and (3) the impact of these activities on performance in the learning task. A case study of activity patterns identified through this approach illustrated that the framework enabled a more precise interpretation of behaviors and found additional strategies employed by the students.

The results showed potentially important differences between high- and low-performing students in terms of their strategy use, which were not apparent from analysis with either the model-driven measures or the action pattern mining in isolation. The sequential lift analysis identified strategies for addition to the strategy model that included an informed guess-and-check approach and a systematic reading and note-taking approach. Overall, these results illustrated the benefits and power of incorporating model- and data-driven techniques to precisely characterize the learning behavior of students in OELEs and to iteratively refine and extend strategy models.

An effective analysis framework applied to the rich behavioral data produced by OELEs has the potential to enable deeper analyses of students' cognitive and metacognitive behavior in complex learning tasks. Ultimately, we believe that this analysis framework can form the basis for designing richer learner modeling schemes that characterize students' activities by analyzing their learning behaviors and performance with respect to both cognition and metacognition. This can in turn help researchers identify opportunities for providing relevant scaffolds that are triggered based on students' recent behavior, as well as performance. In future work, we will incorporate pattern detectors derived from the application of our integrated analysis framework into the Betty's Brain system to directly test its efficacy in improving dynamic learner scaffolding.

Acknowledgments

This work was supported by IES CASL grant # R305A120186.

References

[1] O. Park and J. Lee, "Adaptive instructional system," in *Handbook of Research for Education Communications and Technology*, 2nd ed. D. Jonassen, Ed. Mahwah, NJ, USA: Erlbaum, 2004, pp. 651–684.
[2] S. Puntambekar and R. Hübscher, "Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed?" *Educational Psychologist*, vol. 40, no. 1, pp. 1–12, 2005.
[3] S. Lajoie and R. Azevedo, "Teaching and learning in technology-rich environments," in *Handbook of Educational Psychology*, 2nd ed., P. Alexander and P. Winne, Eds. Mahwah, NJ, USA: Erlbaum, 2006, pp. 803–821.

[4] J. Segedy, K. Loretz, and G. Biswas, "Model-driven assessment of learners in an open-ended learning environment," in *Proc. 3rd Int. Conf. Learn. Analytics Knowl.*, 2013, pp. 200–204.

[5] S. Land, "Cognitive requirements for learning with open-ended learning environments," *Educational Technol. Res. Develop.*, vol. 48, no. 3, pp. 61–78, 2000.

[6] D. H. Jonassen and J. Hernandez-Serrano, "Case-based reasoning and instructional design: Using stories to support problem solving," *Educational Technol. Res. Develop.*, vol. 50, no. 2, pp. 65–77, 2002.

[7] J. S. Kinnebrew, J. R. Segedy, and G. Biswas, "Analyzing the temporal evolution of students' behaviors in open-ended learning environments," *Metacognition Learn.*, vol. 9, no. 2, pp. 1–29, 2014.

[8] J. Metcalfe and B. Finn, "Metacognition and control of study choice in children," *Metacognition Learn.*, vol. 8, no. 1, pp. 19–46, 2013.

[9] J. E. Brophy, *Motivating Students to Learn*. Evanston, IL, USA: Routledge, 2013.

[10] P. H. Winne, "Improving measurements of self-regulated learning," *Educational Psychologist*, vol. 45, no. 4, pp. 267–276, 2010.

[11] B. J. Zimmerman, "Becoming a self-regulated learner: An overview," *Theory Practice*, vol. 41, no. 2, pp. 64–70, 2002.

[12] B. P. Woolf, *Building Intelligent Interactive Tutors: Student-Centered Strategies for Revolutionizing e-Learning*. San Mateo, CA, USA: Morgan Kaufmann, 2010.

[13] R. Baker and G. Siemens, "Educational data mining and learning analytics," in *Cambridge Handbook of the Learning Sciences*. Cambridge, U.K.: Cambridge Univ. Press, 2014.

[14] J. R. Segedy, J. S. Kinnebrew, and G. Biswas, "Using coherence analysis to characterize self-regulated learning behaviours in open-ended learning environments," *J. Learn. Analytics*, vol. 2, no. 1, pp. 13–48, 2015.

[15] J. S. Kinnebrew, K. M. Loretz, and G. Biswas, "A contextualized, differential sequence mining method to derive students' learning behavior patterns," *J. Educational Data Mining*, vol. 5, no. 1, pp. 190–219, 2013.

[16] C. Romero and S. Ventura, "Educational data mining: A review of the state of the art," *IEEE Trans. Syst., Man, Cybern.*, vol. 40, no. 6, pp. 601–618, Nov. 2010.

[17] G. Biswas, J. S. Kinnebrew, and J. R. Segedy, "Using a cognitive/metacognitive task model to analyze students' learning behaviors," presented at the 16th Int. Conf. Human-Comput. Interaction, Heraklion, Crete, Greece, Jun. 2014.

[18] G. Biswas, J. R. Segedy, and J. S. Kinnebrew, "A combined theory- and data-driven approach for interpreting learners' metacognitive behaviors in open-ended tutoring environments," in *Design Recommendations for Intelligent Tutoring Systems: Volume 2 - Instructional Management*, R. Sottilare, A. Graesser, X. Hu, and B. Goldberg, Eds. Orlando, FL, USA: U.S. Army Res. Laboratory, 2014, p. 135.

[19] M. Pressley, F. Goodchild, J. Fleet, R. Zajchowski, and E. Evansi, "The challenges of classroom strategy instruction," *Elementary School J.*, vol. 89, pp. 301–342, 1989.

[20] G. Schraw, K. Crippen, and K. Hartley, "Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning," *Res. Sci. Educ.*, vol. 36, no. 1, pp. 111–139, 2006.

[21] S. Brin, R. Motwani, J. D. Ullman, and S. Tsur, "Dynamic itemset counting and implication rules for market basket data," in *ACM SIGMOD Rec.*, vol. 26, no. 2, pp. 255–264, 1997.

[22] K. Leelawong and G. Biswas, "Designing learning by teaching agents: The Betty's Brain system," *Int. J. Artif. Intell. Educ.*, vol. 18, no. 3, pp. 181–208, 2008.

[23] J. Flavell, "Metacognition and cognitive monitoring: A new area of cognitive—Developmental inquiry," *Am. Psychologist*, vol. 34, no. 10, p. 906, 1979.

[24] P. Winne, "A metacognitive view of individual differences in self-regulated learning," *Learn. Individual Differences*, vol. 8, no. 4, pp. 327–353, 1996.

[25] B. Zimmerman and D. Schunk, Eds., *Handbook of Self-Regulation of Learning and Performance*. New York, NY, USA: Routledge, 2011.

[26] D. Whitebread and V. Cárdenas, "Self-regulated learning and conceptual development in young children: The development of biological understanding," in *Metacognition Science Education*. The Netherlands: Springer, 2012, pp. 101–132.

[27] P. H. Winne and A. F. Hadwin, "Studying as self-regulated learning," in *Metacognition in Educational Theory and Practice*. Mahwah, NJ, USA: Lawrence, 1998, pp. 277–304.

[28] P. Winne and A. Hadwin, "The weave of motivation and self-regulated learning," in *Motivation and Self-Regulated Learning: Theory, Research, and Applications*, D. Schunk and B. Zimmerman, Eds. New York, NY, USA: Taylor & Francis, 2008, pp. 297–314.

[29] A. Donker, H. de Boer, D. Kostons, C. Dignath van Ewijk, and M. van der Werf, "Effectiveness of learning strategy instruction on academic performance: A meta-analysis," *Educational Res. Rev.*, vol. 11, pp. 1–26, 2014.

[30] L. Zhang, K. VanLehn, S. Girard, W. Burleson, M. E. Chavez-Echeagaray, J. Gonzalez-Sanchez, and Y. Hidalgo-Pontet, "Evaluation of a meta-tutor for constructing models of dynamic systems," *Comput. Educ.*, vol. 75, pp. 196–217, 2014.

[31] F. Bouchet, J. Harley, G. Trevors, and R. Azevedo, "Clustering and profiling students according to their interactions with an intelligent tutoring system fostering self-regulated learning," *J. Educational Data Mining*, vol. 5, no. 1, pp. 104–146, 2013.

[32] J. Bransford and D. Schwartz, "Rethinking transfer: A simple proposal with multiple implications," *Rev. Res. Edu.*, vol. 24, no. 1, p. 61, 1999.

[33] K. VanLehn, "The behavior of tutoring systems," *Int. J. Artif. Intell. Educ.*, vol. 16, no. 3, pp. 227–265, 2006.

[34] V. Aleven, B. McLaren, I. Roll, and K. Koedinger, "Toward meta-cognitive tutoring: A model of help seeking with a Cognitive Tutor," *Int. J. Artif. Intell. Educ.*, vol. 16, no. 2, pp. 101–128, 2006.

[35] D. Perera, J. Kay, I. Koprinska, K. Yacef, and O. Zaïane, "Clustering and sequential pattern mining of online collaborative learning data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 6, pp. 759–772, Jun. 2009.

[36] J. S. Kinnebrew, D. L. Mack, and G. Biswas, "Mining temporally-interesting learning behavior patterns," in *Proc. 6th Int. Conf. Educational Data Mining*, 2013, pp. 252–255.

[37] J. Nesbit, M. Zhou, Y. Xu, and P. Winne, "Advancing log analysis of student interactions with cognitive tools," in *Proc. 12th Biennial Conf. Eur. Assoc. Res. Learn. Insruction (EARLI)*, Doi: 10.1.1.150.6671&rep=rep1&type=pdf

[38] S. Amershi and C. Conati, "Combining unsupervised and supervised classification to build user models for exploratory learning environments," *J. Educational Data Mining*, vol. 1, no. 1, pp. 18–71, 2009.

[39] T. Tang and G. McCalla, "Student modeling for a web-based learning environment: A data mining approach," in *Proc. 18th Nat. Conf. Artif. Intell.*, 2002, pp. 967–968.

[40] R. Agrawal and R. Srikant, "Mining sequential patterns," in *Proc. 11th IEEE Int. Conf. Data Eng.*, 1995, pp. 3–14.

[41] R. Srikant and R. Agrawal, "Mining Sequential Patterns: Generalizations and Performance Improvements," in *Proc. 5th Int. Conf. Extending Database Technol.: Adv. Database Technol.*, 1996, pp. 3–17.

[42] J. S. Kinnebrew and G. Biswas, "Identifying learning behaviors by contextualizing differential sequence mining with action features and performance evolution," presented at the 5th Int. Conf. Educational Data Mining, Chania, Greece, Jun. 2012.

[43] J. Ho, L. Lukov, and S. Chawla, "Sequential pattern mining with constraints on large protein databases," in *Proc. 12th Int. Conf. Manage. Data*, 2005, pp. 89–100.

[44] J. R. Segedy, J. S. Kinnebrew, and G. Biswas, "The effect of contextualized conversational feedback in a complex open-ended learning environment," *Educational Technol. Res. Develop.*, vol. 61, no. 1, pp. 71–89, 2013.

John S. Kinnebrew received the BA degree in computer science from Harvard University and the PhD degree in computer science from Vanderbilt University. He is a research scientist at the Institute for Software Integrated Systems, Vanderbilt University. His research interests include data mining, user modeling, intelligent agent design, and coordination in multiagent systems. His research in open-ended learning environments focuses on data mining and machine learning for modeling human learning behaviors, including metacognition and self-regulated learning strategies.

James R. Segedy received the BA degree in computer science from Goucher College and the MS degree in computer science from Vanderbilt University. He is currently working toward the PhD degree studying educational technology and learning sciences at the Institute for Software Integrated Systems, Vanderbilt University. His research focuses on developing analytic frameworks for open-ended computer-based learning environments in order to influence the scaffolding decisions made by teachers, tutors, and computer-based agents as they help students develop their capabilities for open-ended problem solving. He regularly contributes to international research communities; he was recently instrumental in organizing an international research conference (AIED 2011), workshop (at AIED 2013), and special track (ITS special track at FLAIRS-27).

Gautam Biswas is a professor of computer science, computer engineering, and engineering management in the EECS Department and a senior research scientist at the Institute for Software Integrated Systems, Vanderbilt University. He conducts research in intelligent systems with primary interests in modeling and simulation, analysis of complex embedded systems, data mining, and computer-based learning environments (CBLEs) for STEM disciplines. Two primary CBLEs developed by his group are Betty's Brain, a learning by teaching system, and CTSiM, that exploits synergies between computational thinking and science to support learning by model building and simulation. His data mining projects combine model-based and data-driven approaches for diagnosis, prognosis, and discovering student learning behaviors in open-ended learning environments. He is a fellow of the IEEE.