Modeling problem-solving strategy invention (PSSI) behavior in an online math environment

Nidhi Nasiar[1], Ryan Baker[1], Yishan Zou[1], Jiayi Zhang[1], Stephen Hutt[1]

[1] Graduate School of Education, University of Pennsylvania, PA, USA
nasiar@upenn.edu

Abstract. This study uses Knowledge Engineering (KE) to develop an automated model of problem-solving strategy invention (PSSI) behavior (defined as inventing a new strategy for solving a math problem, outside of system-offered default strategies). The PSSI model identified the students inventing new strategies, and examined the relationship between PSSI behavior and existing fine-grained detectors of self-regulation. The findings suggest that students inventing new strategies to use for problems, are more likely to transform the information provided in the question, and to reason around the problem's contextual information.

Keywords: Strategy Invention, Self-Regulated learning, SMART model

1 Introduction

It is essential to learn to choose problem-solving appropriate strategies, and online learning platforms often scaffold students by recommending strategies. However, these supports are just a starting point for students to learn the use of strategies before creating their own, a valuable 21st century skill involving creativity and critical thinking. The invention of new strategies relies heavily on self-regulation, a key component of effective problem-solving [7]. Previous research in modeling self-regulated learning (SRL) behavior has typically considered high-level strategies ([4] is an exception), but fine-grained behaviors can provide insights into the underlying processes [11] and help us understand the role cognitive operations play in the emergence of SRL strategies.

The increasing availability of fine-grained interaction data made it feasible to model student behavior and strategies. Behavior modeling can be done using Knowledge Engineering (KE), also known as rational modeling; Machine Learning (ML) based on prediction modeling; or as an output from a bottom-up approach such as cluster analysis or sequential pattern mining. In this study, we use Knowledge Engineering (KE) to develop an automated model of problem-solving strategy invention (PSSI) behavior, defined as the learner inventing new strategies to solve the math problem, in addition to default strategies provided by the learning system. We adopt this term from [8], which defines invention activities as problem-solving tasks where learners invent novel procedures as solutions to unfamiliar problems, focusing on the development of novel problem-solving strategies. This study is conducted with CueThink, a learning platform designed to support middle school students in developing mathematical problem-solving skills. A KE model was used to identify the students exhibiting PSSI behavior in CueThink, and study the relationship of PSSI with other indicators of SRL behavior.

1.1 Approaches to behavior modeling in AIED systems

Supervised Machine Learning has emerged as the predominant approach to behavior modeling. Despite their efficacy and ease of validation, the resultant detectors often lack interpretability, although explainable AI has attempted to remedy that. Another approach is Knowledge Engineering (KE), or rational modeling, which has a rich history of modeling complex behavior in AIED systems. For example, [1] developed a KE model of help-seeking within Cognitive Tutor with production rules. The transparency of KE models has enabled insights into behaviors such as gaming the system [6]. One popular method to build KE models is Evidence-centered design (ECD), which gathers evidence of competencies within assessment of complex behavior [5]. ECD involves expert-designed tasks and statistical models to evaluate learner competencies. ECD can handle complex input data while maintaining transparency and interpretability [9].

A KE model relies on domain experts' knowledge to establish a set of rules based on existing literature, and/or formalizations of experts' decision-making processes to capture the behavior [1]. KE models have higher levels of transparency and interpretability in decision-making, and capture deeper underlying features of behavior [6], making them sometimes able to generalize better to new contexts. KE models are usually validated by verifying that the rules fully capture the behavior, by comparison to theory, and iteratively by researchers, domain experts, designers, and teachers, which is different from calculating performance metrics for machine-learned models– but construct validity is important to both paradigms. PSSI is apt for a KE model as it is a straightforward construct with an explicit decision-making process. The system logs the choice of students to create a new strategy. The decisions taken to identify PSSI behavior are direct, independent, and each one is conceptually explainable. No arbitrary cut-offs or calculations, which are difficult to identify rationally, are needed to identify PSSI.

1.2 SRL behavior, and existing SMART models on CueThink

Self-regulated learning (SRL) is a series of learner-generated thoughts and behaviors for goal attainment through information seeking, strategy planning, and effort alignment with objectives [10]. Log data from AIED systems can be used to model SRL behaviors such as help-seeking [1]. This paper focuses on PSSI, an SRL behavior involving an individual-level decision-making process to create and utilize new strategies beyond those provided by the system. PSSI is situated in the SMART operations of Winne & Hadwin's COPES model [10], as the process of inventing new strategies, which requires problem identification, monitoring and evaluation of available default strategies against the requirements to solve the problem, and creating new strategies if necessary. Learners engage in translating their existing knowledge and manipulating known information into a new representation to find a solution. A recent paper by Zhang et al. [11] built ML detectors of four SRL constructs also based on the SMART model [10] for the CueThink system: numerical representation (NR), contextual representation (CR), outcome orientation (OO), and data transformation (DT). NR and CR are Assembling behaviors within the SMART framework, assembling information into new representations, and occurring early in problem-solving. OO involves estimating the final answer before starting to solve, and DT is manipulating information into different representations to support solving problem. The automated detectors were successful at capturing the 4 constructs; achieving AUC ROC from 0.76 – 0.89 under 10-fold student-level cross-validation [11]. We analyzed the relationships between the PSSI behavior and the predictions of these 4 detectors on the same students, to explore the fine-grained cognitive operations that students employ when inventing strategies.

2 Methods

2.1 Math Learning Environment: CueThink

The study uses a dataset from CueThink, an online learning platform that presents problems as Thinklets, to be solved as a four-phase process—Understand, Plan, Solve, and Review. The platform design is based on Winne & Hadwin's SMART model of cognitive operations [10], facilitating analyses using that model. The Understand phase involves reading the problem, extracting information, and creating a representation. The Plan phase involves selecting predefined strategies (for instance, model with an equation; work backwards), or create new strategies, and outlines their use. In the Solve phase, the student gives a solution and explanation, while Review phase is for reflection on the answer's clarity and logic. This study focuses on the Plan phase where PSSI behavior occurs. The dataset, also used by [11], consists of 79 grade 6 and 7 students at a diverse suburban school in the southwestern U.S. during 2020-2021. The log data captured action-level student usage of the application and their text entries. Students spent an average of 5.2 hours using CueThink, spending 1.8 hours per Thinklet. After removing duplicates, 181 attempted Thinklets on 24 unique problems remained.

Building automated models of PSSI using KE included a formal operationalization of PSSI, establishing conditions and rules for detecting, and identifying edge affectin detection. The resultant operational definition was straightforward and did not involve subjective interpretation. Consequently, as with other studies employing similar operationalizations in KE models [1], inter-rater coding was unnecessary.

2.2 Developing the KE model

Problem-solving strategy invention (PSSI) is defined as the behavior where a learner invents a new strategy to solve a given problem. In CueThink, where the learner is provided an initial set of 8 strategies to select from during the Plan phase, PSSI is operationally defined as the student inventing a new strategy not in the provided list and not suggested by the teacher. If a new strategy is created, the log data records this choice along with the exact text of that new strategy.

KE model development was an iterative process involving SRL and KE experts and CueThink developers. The rules and conditions that were developed to describe PSSI behavior went through multiple rounds of revision to ensure robustness and validity. The rules emerged through discussion and were conceptually coherent both in theory and practice, with the potential to generalize to future problems (via an automated model), and then the final set of rules was formalized and programmed.

Data extraction & preparation. Data extraction and preparation began by extracting log data from the Plan phase. Initially, 3380 logged actions were recorded; 2326 strategy-specific actions were extracted for further analysis.

Making the rules & Conditions. The following rules identified PSSI and edge cases:

Compare with existing strategies: In the 1st step, every student strategy was compared against the 8 default strategies: "Draw a picture," "Make a table," "Solve with an easier problem," "Work backwards," "Guess, check and revise," "Model it with manipulatives," "Look for a pattern," "Model with an equation." The log data had instances of differently capitalized versions of default system strategies; these were treated as system strategies. Non-matching strategies were tagged as new.

Spelling Errors: Some added strategies were the same default strategies with spelling errors. Misspellings were checked automatically by computing the Levenshtein distance, which measures character differences between words, and removing strategies with a distance of 4 or less from the default list.

Filter out class-wide strategies: In many cases, the same new strategies were added by multiple students from one class, likely due to the teacher discussing and sharing strategies in class. As our focus is on student invention, strategies added by more than 5 students in one class were tagged as 'class-wide', and tagged as non-PSSI behavior. However, one class's class-wide strategy could still be considered new in another class.

3 Results

3.1 PSSI behavior and using new strategies

The KE model classified total of 37 students (out of 79 students) as exhibiting PSSI behavior; 70.27% of these students added more than 1 unique strategy. Students collectively added 85 new strategies on 55 math problems. Examples of invented strategies included "Use direction arrows when adding and subtracting", "Use multiplication rules & relationships", "and "Chunk the problem into smaller problems." Creating a new strategy alone is insufficient for improving problem-solving; students must also use the new strategy. The 1st and 3rd authors qualitatively coded the recorded videos from the Solve phase, where students make their process visible using the virtual whiteboard to write and draw their solutions, to see if students used their new strategies. Inter-rater reliability (IRR) was acceptable (kappa=0.79). After removing two instances with a poor-resolution video and a missing file, students used their newly added strategies in 97.7% of cases (85 out of 87 instances). Thus, we can say that students showing PSSI behavior are quite likely to use their invented strategy in solving problems.

3.2 Association with SMART models

The predictions of four SRL detector from Zhang et al [11] were initially confidence probabilities, which were converted into a binary variable for comparison to the binary assessments of PSSI. The outputs were analyzed at the level of an entire math problem. Risk ratio (also called relative risk) was calculated to examine the associations between variables. The risk ratio value indicates how many times more likely construct 2 is if construct 1 is present, than if construct 1 is absent (e.g. a risk ratio of 1.0 indicates that construct 2 is neither more nor less likely if construct 1 is present). The results of the co-occurrence of PSSI with SMART models from [11] are shown in Table 1. PSSI behavior showed very strong association with DT; with learners engaging in DT 3.62 times as likely to show PSSI behavior. Learners engaging in CR had just over twice the probability of exhibiting PSSI behavior. On the other hand, the association between PSSI and NR, and PSSI and OO were not significantly different from chance.

Table 1. Risk Ratio values of association between PSSI and SMART models

Construct 1	Construct 2	Risk Ratio	Confidence Interval (95%)
PSSI	Numerical Representation (NR)	0.96	0.61 - 1.51
PSSI	Contextual Representation (CR)	2.12	1.12 - 4.01
PSSI	Outcome Orientation (OO)	1.05	0.61 - 1.80
PSSI	Data Transformation (DT)	3.62	1.39 - 9.40

4 General Discussion, Applications, and Limitations

We present a KE model of inventing new strategies (PSSI), a key skill for effective problem-solving in math. Since less than half of students show this behavior, it is crucial to offer support, such as scaffolds to explain the role of strategies and invention practices. Multiple reasons may drive PSSI behavior, including seeking efficiency, recognizing limitations in default strategies, or using familiar techniques. Correlating PSSI with other SRL constructs, we found a high co-occurrence between PSSI and DT, where learners manipulate information that is presented to them to find solutions. The link for PSSI - DT is plausible, as both involve the cognitive operation of translation (from the SMART model). As the learners invent a new strategy, we can expect learners to manipulate how information is presented to them to use a new strategy to reach solution. Additionally, a strong association was found between PSSI and CR, when students create their internal problem representation using contextual details (e.g., settings, characters, situations, etc.), a process that typically occurs early in problem-solving [10]. Though the exact causality of this relationship remains unclear, representing a problem contextually might imply a deeper understanding of the task, thereby equipping learners to assess if the strategies available are insufficient and invent new ones.

The KE model for PSSI identifies students who are creating strategies. It informs formative feedback for teachers to promote asset-based approaches like sharing students' creative strategies with the class. It can also be used to scaffold students who struggle with strategy invention (as in [2]), and provide metacognitive prompts for reflection and planning, like pop-up messages encouraging students to deliberately choose the strategies before solving, reminders that they can create their own ones, or reminding students of the new strategies they decided to use. Further scaffolds based on SRL processes of orientation and reflection, can improve the quality of students' invented strategies [3]. Investigating cases where students don't use their invented strategies could be informative. Identifying problems with less frequent use of default strategies indicates a lack of suitable default strategies, which could be added in the future.

Limitations of this study include operationalizing PSSI only in the Plan phase, and not considering other spontaneous strategy inventions that might have occurred in the Solve phase. Future work should explore broader PSSI definitions and assess PSSI's impacts on learning. Additionally, utilizing NLP for semantic analysis could provide insights into new strategy types and their effectiveness in problem-solving.

In conclusion, inventing new strategies is essential for effective math problem-solving, as PSSI develops transferrable skills for tackling new contexts. It is crucial to foster PSSI skills since less than half of learners engage in it. This study used KE to develop an automated PSSI model to identify students engaging in this behavior, and findings indicate that the self-reported strategy invention led to the actual use of those strategies. The model enabled us to analyze associations of PSSI with other SRL behaviors, to inform teacher reports and scaffolds for strategy invention.

References

1. Aleven, V., Mclaren, B., Roll, I., & Koedinger, K. (2006). Toward meta-cognitive tutoring: A model of help seeking with a Cognitive Tutor. Int'l J of AIED, 16(2), 101-128.
2. Gidalevich, S., & Kramarski, B. (2019). The value of fixed versus faded self-regulatory scaffolds on fourth graders' mathematical problem solving. Instructional Science, 39-68.
3. Holmes, N. G., Day, J., Park, A. H., Bonn, D. A., & Roll, I. (2014). Making the failure more productive: scaffolding the invention process to improve inquiry behaviors and outcomes in invention activities. Instructional Science, 42(4), 523-538.
4. Hutt, S., Ocumpaugh, J., Andres, J., Bosch, N., Paquette, L., Biswas, G., & Baker, R.(2021). Investigating SMART Models of Self-Regulation and Their Impact on Learning. EDM.
5. Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. Educational measurement: issues and practice, 25(4), 6-20.
6. Paquette, L., de Carvalho, A.., & Baker, R. S. (2014). Towards Understanding Expert Coding of Student Disengagement in Online Learning. Learning and instruction, 15(2), 123-139.
7. Perels, F., Gürtler, T., & Schmitz, B. (2005). Training of self-regulatory and problem-solving competence. Learning and instruction, 15(2), 123-139.
8. Roll, I., Aleven, V., & Koedinger, K. R. (2010, June). Analysis of students' actions during online invention activities-seeing the thinking through the numbers. In Proceedings of the 9th International Conference of the Learning Sciences-Volume 2 (pp. 49-52).
9. Shute, V., Kim, Y. J., & Razzouk, R. (2010). ECD for dummies. Working Examples.
10. Winne, P. H., & Hadwin, A. F. (1998). Metacognition in educational theory and practice, chapter in Studying as self-regulated learning, 277–304.
11. Zhang, J., Andres, J., Hutt, S., Baker, R. S., Ocumpaugh, J., Nasiar, N., Mills, C., Brooks, J., Sethuraman, S., & Young, T. (2022). Using Machine Learning to Detect SMART Model Cognitive Operations in Mathematical Problem-Solving Process. Journal of Educational Data Mining, 14(3), 76-108