Modeling Strategy Use in an Intelligent Tutoring System: Implications for Strategic Flexibility

Caitlin Tenison[1] and Christopher J. MacLellan[2]

Psychology Department[1] and Human-Computer Interaction Institute[2]
Carnegie Mellon University Pittsburgh, PA 15213, USA
ctenison@andrew.cmu.edu and cmaclell@cs.cmu.edu

Abstract. Education research has identified strategic flexibility as an important aspect of math proficiency and learning. This aspect of student learning has been largely ignored by Intelligent Tutoring Systems (ITSs). In the current study, we demonstrate how Hidden Markov Modeling can be used to identify groups of students who use similar strategies during tutoring and relate these findings to a measure of strategic flexibility. We use these results to explore how strategy use is expressed in an ITS and consider how tutoring systems could integrate a measure of strategy use to improve learning.

1 Introduction

Strategic flexibility in arithmetic problem solving is both an important reflection of knowledge [1, 2] and a recognized predictor for future learning [3, 4]. The National Mathematics Advisory Panel [5] lists flexibility along with accuracy and speed of problem solving as the core defining features of a student's math proficiency. Despite the importance placed on flexible problem-solving, work on Intelligent Tutoring Systems (ITSs) has traditionally focused on the accurate completion of problem steps rather than on the strategies used by students to complete them. The current study identifies differences in strategy use within an ITS, relates these differences to a pencil and paper measure of strategic flexibility, and explores how ITSs may be designed to support strategic flexibility.

Strategic flexibility refers to a student's knowledge of multiple strategies and their ability to choose the best of those strategies for a given problem [6, 4]. Measures of strategic flexibility correlate with both the student's procedural and conceptual knowledge [1, 7, 4]. Schneider et al. [2] found a bidirectional relationship between procedural and conceptual knowledge and hypothesized that these two types of knowledge improve strategic flexibility in an iterative fashion. Students who have high strategic flexibility are more likely to adapt their strategies, transferring their knowledge to solve new problems [8, 4]. Conversely, students who lack strategic flexibility struggle to solve more difficult or unfamiliar problems that require the use of different strategies [6]. An active area of research concerns developing pedagogical methods to foster strategic flexibility and understanding the impact of strategic flexibility on problem solving.

Studies of ITSs have shown their effectiveness for developing students' procedural and conceptual knowledge in math problem solving [9–11]. Little research, however, has explored the impact of ITSs on strategic flexibility. Unlike the traditional studies of strategic flexibility, which use pencil and paper, ITSs confine students to working within the structure of the interface. This raises the question of if and how different strategies present themselves in rigid tutoring interfaces. One study focuses on the effects of allowing strategic flexibility within a tutor [12]. Measuring the number of times students made variations from the main strategy path, Waalkens et al. found that students did not take advantage of the flexibility permitted by the interface. Without instructing students to solve problems using multiple strategies, the likelihood of a student using a divergent strategy is low. Previous research on strategic flexibility has found that although students may be aware of many strategies, they often limit their choice to the most efficient one when problem solving [6]. Acknowledging these earlier findings, we identify two ways to improve research on strategic flexibility within ITSs. First, researchers must actively encourage the use of different strategies if they wish to observe student's strategic flexibility. Second, as we will discuss, researchers should use more sophisticated methods for modeling strategies and detecting how they are used. Introducing these changes to ITS research of strategic flexibility will make it possible for researchers to explore the effect of strategically focused interventions on how well students solve future problems.

Outside of the math domain there has been some work on developing more complex methods for assessing strategy from the choices students make in ITSs. Piech et al.[13] looked at differences in the paths that introductory programmers take when completing programming assignments and found that strategic differences in two homework assignments at the beginning of the semester predicted students' midterm grades . These sequences or paths can be seen as a reflection of the strategies that a student employs during problem solving. Additionally, the area of research on meta-cognitive hint seeking identifies the strategies students use when unable to solve problems [14, 15]. Using a model of hint seeking, Roll et al. [16] found that recognizing and intervening when students are using bad hint seeking strategies improves learning. Work from these different areas demonstrates that strategy use can be identified within a tutoring system, and also suggests that these results can identify opportunities for tutoring.

1.1 The Current Study

The current study bridges research on strategic flexibility in mathematics and work identifying strategy use in ITSs. With evidence supporting the value of strategic flexibility in math, it is important that tutoring systems develop approaches for understanding and supporting strategic flexibility. In the current study, we use Hidden Markov Modeling (HMM) to cluster participants into strategically distinct groups. We present evidence supporting the hypothesis that these groups differ on a measure of strategic flexibility collected using the pencil and paper test developed by Rittle-Johnson and Star [7]. Furthermore, we use this HMM method to explore how flexibility presents itself in a math ITS as differences in tutor behavior. We conclude with recommendations concerning how ITSs can be built to measure and encourage flexible strategy use.

2 Materials and Methods

We used an observational design in which all participants completed the same tutor curriculum. Students completed a 20-minute pencil and paper math test to assess proficiency with algebraic problem solving and strategic flexibility. A week after taking this test, students spent an hour and a half in the school computer lab working with an algebra ITS.

2.1 Participants

There were 112 eighth and ninth grade Algebra I students (72 eigth grade; 57 females; mean age 13.6, SD 1.2) who took a math test and participated in the tutoring. There were seven eighth grade classes (3 advanced and 4 regular) and four ninth grade classes (2 regular and 2 remedial). All students attended the same large, urban public school. The school consisted of 60.6% Caucasian, 33.5% African American, and 1.3% Asian. Approximately 52% of students qualified for free or reduced lunch.

There were four teachers whose classes participated in this study. Each teacher taught the same grade and advancement level to students in their class. All four teachers used the same algebra curricula, which was supplemented with work on the ALEKS online math program. All classes had previously covered the distributive property and solving multi-step equations. Human subjects' approval and consent from the school was obtained prior to conducting the study.

2.2 Materials

Assessment We used a modified version of the Rittle-Johnson and Star [7] assessment, which assesses mathematical knowledge (both conceptual and procedural) and strategic flexiblity for one- and two-step algebra equation problem solving. We implemented only the procedural skill and strategic flexibility portion of the assessment. The test of strategic flexibility assessed three features;

Table 1. An example of the two strategies for solving the two problem types. Both strategies are correct, but the green strategies are those biased by the tutor.

Divide Problem Distribute Strategy	Divide Problem Both Sides Strategy	Multiply Problem Distribute Strategy	Multiply Problem Both Sides Strategy
$2(3x+5)=6$	$2(3x+5)=6$	$\frac{(8y-4)}{2}=6$	$\frac{(8y-4)}{2}=6$
$2*3x+2*5=6$	$\frac{2(3x+5)}{2}=\frac{6}{2}$	$\frac{(8y)}{2}-\frac{4}{2}=6$	$2*\frac{(8y-4)}{2}=6*2$

the ability to generate, recognize, and evaluate multiple strategies. Although we followed the organization of this assessment, we modified many of the math problems tested in this assessment in order to better accommodate an older student population and focus more directly on the problems being studied within the tutoring system. The student's procedural accuracy was calculated by percentage of problems correctly solved.

Problem Types Work on strategic flexibility has found that while students may be aware of many strategies they will often learn to use the most efficient strategy for a specific problem [4, 6]. To encourage the use of different strategies we used 6 variations of the linear equation. Students saw 6 examples of each type of problem. For this study we collapsed the 6 problem types into two categories; "divide problems," in which the problem can be solved by dividing both sides by a coefficient, and "multiply problems," in which the problem can be solved by multiplying both sides by a coefficient (Table 1). While both of these problem types can be solved using either of two correct strategies, "distribute" or "both sides", we took several actions to bias strategy use. First, the tutor hints recommended different strategies for the two problem types. Table 1 displays the recommended strategies in green. Second, distribute required fewer tutor actions than the both sides strategy for divide problems, while the both strategies required the same number of actions for multiply problems. Finally, if students chose their strategy to avoid large fractions, this would promote the use of the multiply strategy for the multiply problems. This last decision-making heuristic applies more for problems that require student calculation.

Intelligent Tutoring System We used a modified version of Cognitive Tutor [11]. The tutoring interface directs students to select the step the computer should take to solve the problem. Students can choose between actions to "transform" or "solve" the problem. The transform command directed the computer to take actions to change the structure of the equation. Table 1 shows how the transformation of "distribute," "multiply both sides" or "divide both sides" would change the equation. Solve actions instruct the computer to perform various calculations, such as combining like terms. This is an important feature of the tutor because students must indicate each step taken to solve the problem rather than combining steps in their calculations.

2.3 Data Analysis

With strategic flexibility so closely related to procedural skill [1, 2], the inclusion of off-task strategy paths would bias our clustering method to cluster based on students' use of off-task strategies. Because we are interested if the correct strategies students learn to use are related to their strategic flexibility we restrict our analysis to on-task actions only. For each student we recorded the choices made when faced with the problems shown in Table 1.

We based our analysis on the HMM clustering method described by Smyth [17] and Piech et al. [13]. Detailed justification of this method can be found in these two papers, so we will only summarize the steps of the analysis. Using the on-task student actions, we fit an HMM model to each student's action sequence; the model had a hidden state for every possible action, where each state had 100% probability of emitting the action represented by that state (this is essentially a Markov Chain). After fitting the model to a student's actions, we modified the transition probabilities between hidden states so that none of the transition probabilities were equal to zero. We did this by replacing all zero probability transitions with a small constant (1×10^{-10}) and renormalizing the transition probabilities. Next, we calculated the log probability that this model fit each of the other student's action sequences individually. After doing this for all students, we calculated the distance between two subjects sequences as the average of the log probability of one subject's model predicting the other subject's data with the log probability of the other subject's model predicting the first subject's data. Finally, we used these pairwise distances to cluster the students. We used the k-mediods clustering algorithm [18] to ensure clusters of similar size and to deemphasize the effect of outliers. We determined the number of clusters by fitting models with 2 through 10 clusters and evaluating the log probability with leave-one-subject-out cross-validation. We found four clusters best fit the data. The parameters of these 4 clusters were next used to initialize a composite HMM model, which had additional latent states for each cluster that transitioned only to their respective smaller HMMs. After initializing this new HMM using the individually computed transition probabilities, we retrained the composite HMM using all the data. The resulting estimates are considered better than those generated by separately training the model on the four smaller HMMs. This method is particularly useful for building descriptive rather than predictive models of the data [17] and is thus useful for understanding the common strategy choices displayed during tutoring.

3 Results

3.1 Flexibility Score

Our study replicates the results of Rittle-Johnson and Star [1]. We found a significant correlation between strategic flexibility and procedural knowledge $r(110) = 0.73$, $p < 0.001$.

3.2 HMM Clustering

We clustered students based on their correct problem solving paths using the method described in section 2.3. We best fit a four cluster model. We found a significant difference in the average strategic flexibility scores of students in the different clusters, $f(3, 107) = 10.1$, $p < 0.001$. Figure 1 shows the mean and standard errors for the four clusters. Post hoc comparisons using the Tukey HSD test indicated that the only significant ($p < 0.05$) differences were between the 2nd and 3rd cluster and the 3rd and 4th cluster.

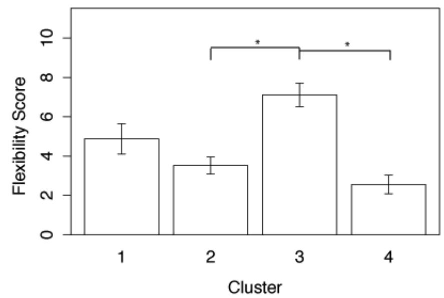

Fig. 1. The mean flexibility of the four identified clusters with standard error.

3.3 Exploring Strategic Flexibility in Tutor Behavior

We used the transition probabilities generated from the HMM training to explore the different strategies learned. The four clusters of students showed distinct patterns of strategy use. Figure 2 shows the predominate strategies used by the four groups. In cluster 2 ($n = 25$) and cluster 3 ($n = 54$), students alternate between the tutored strategies. In cluster 1 ($n = 16$) and cluster 4 ($n = 16$), on the other hand, students use the same strategy for both problem types. Students in cluster 1 distribute on both problems, whereas students in cluster 4 apply either divide or multiply to both sides of the equation in order to eliminate the coefficient.

We next wanted to test if the strategies presented by each group of students were learned over the course of tutoring. To see if students increased their use of either of the dominant strategies of each cluster (displayed in Figure 2) we fit each cluster's data to an additive factor model (AFM) [19]. For this paper we are interested in only using this model descriptively to observe if students are increasing their use of the dominant strategies. We found that as students gained practice with the problems they increased their use of the dominant strategies, these slopes are displayed in Table 2.

Table 2. The coefficient for the learning rate of each cluster's dominant strategy by problem type, as computed by the Additive Factors Model.

Cluster	Divide Problems	Multiply Problems
1	.21	.05
2	.05	.04
3	.13	.02
4	.14	.02

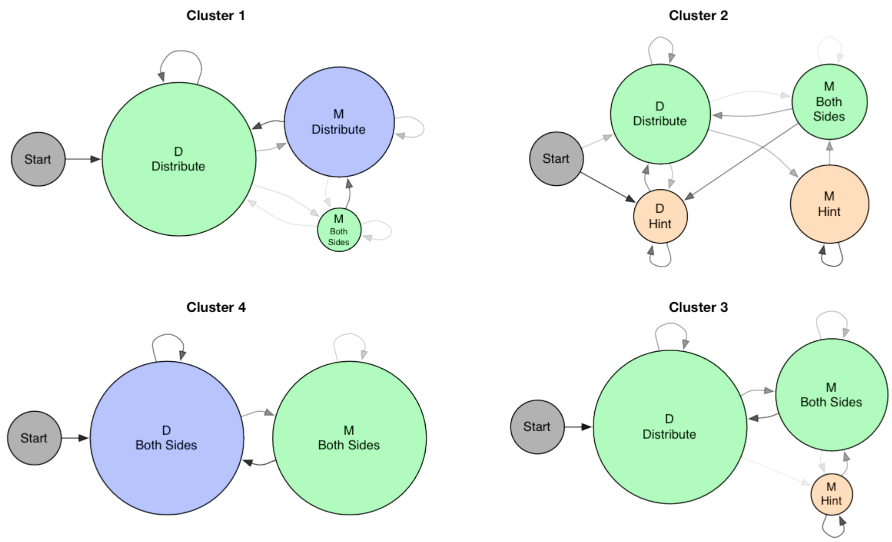

Fig. 2. The student behavior for each cluster. The Divide (D) and Multiply (M) labels represent the problem type and the Distribute and Both Sides labels denote the strategy taken. The node size denotes the number of times it was visited. Green nodes are strategies that were suggested by the tutor, blue nodes are (valid) untaught strategies, and orange nodes are hints. Arrow gradients denote transition probability.

4 Discussion

The ability to flexibly solve math problems is a valued measure of proficiency and important to future learning [5, 8]. Few studies have investigated how strategic flexibility is displayed in ITSs, and perhaps as a result no studies have directly measured math flexibility using an ITS. The current study uses a method previously employed in a different domain [13] to identify groups of students in a math ITS that differ in how they apply strategies to solving equations.

First, our findings from the behavioral pre-test replicated results from Rittle-Johnson and Star [1, 7], showing a correlation between strategic flexibility and procedural knowledge. Understanding how strategic flexibility is represented during tutoring is less clear. The Rittle-Johnson and Star measure of strategic flexibility directs students to generate multiple solution paths; however, in a tutoring setting students are not prompted to choose multiple strategies. Studies of strategic flexibility have observed that when students are not asked to generate multiple strategies they will often use only one strategy to solve problems [4, 6]. This makes it difficult to observe flexible strategy use indirectly. To combat this challenge, our study used two sets of problems, which despite being solvable by the same two methods, were set up to favor different methods. As previous research would suggest, students remain relatively consistent within problem types; however, between problem types students changed strategies. When students appropriately adjust their strategy this suggests that these students are more strategically flexible and would be better able to adapt and transfer their knowledge to similar, but novel problems– testing this hypothesis is one possible direction for future work.

In the current study we used a variation of the HMM clustering method used by Smyth [17] and Piech et al. [13] to cluster participants according to their correct action sequences. This clustering method, while conducted without information about a student's strategic flexibility measure, distinguished groups that significantly differed in strategic flexibility. This distinction suggests that the strategies students choose are reliant on their strategic flexibility. The lack of a significant difference between clusters 1 and 3 suggests that students may be less likely to use the optimal strategy on multiply problems. This can be explained by the lack of a direct benefit for using the both sides strategy on the multiply problems, as discussed in section 2.2. We ran this study to develop a method for identifying strategy use in math problem solving and while experimental manipulations must be done to learn more about the underlying causes of strategic flexibility in the tutor, the exploratory analysis from our work sets forth multiple areas of potential research.

The HMM clustering provides some insight into how the strategic flexibility scores translate into strategy use during tutoring. Students in cluster 3 show flexibility in their ability to switch between the distribute and the both sides strategy. This is echoed in their high strategic flexibility scores. Students in cluster 2 show a similar pattern, however their strategy path also shows that these students are reliant on the hints that direct them towards these strategies. The positive slopes from the AFM, indicate that while these students are learning to apply these strategies more often over the course of tutoring, this is at a slow rate. Students in cluster 1 generally use the distributive strategy to solve both problem types. Although there is some use of the both sides strategy, users of this strategy are seen returning to the distribute strategy. The positive learning gains reported from the AFM indicate that over the course of tutoring students become more rigid in their use of the distribute strategy on both problems. Students in cluster 4 use the both sides strategy to solve these problems and increase in their use of this strategy over tutoring. The behavior of cluster 4 scoring the lowest on the strategic flexibility and only using the both sides' strategy suggests that these students may not recognize distribute' as a potential strategy for solving these problems and could benefit most from an intervention.

As an observational study, we are limited in the causal claims we can make about the relationship between strategic flexibility and strategy use in the tutor. However, the methods used in this study can be applied in future experiments. First, this study demonstrates a means of observing strategy use by designing problems to specifically favor some strategies over others. Future studies could use a model of students' decision-making heuristics (i.e., a model of how they decide between possible next actions, such as avoiding fractions) to identify problems that favor different strategies. These problems could be used to triangulate the strategies that students know and do not know.

The current study used a version of the tutor in which students had to explicitly select each transformation of the problem. This tutoring format lends itself well to the method of strategy detection that we use, however, it may not apply to other tutors. Future work should extend the model to contexts in which students have more freedom in making multiple transformations in a single step, such as is seen in Waalkens et al. [12]. We expect different interfaces may foster strategic flexibility to different levels.

Two questions of great interest remain present in the field: How best can strategic flexibility be improved and to what extent does that improvement influence learning? While these questions are outside the scope of the current study, the ability to identify students based on their strategy use can establish the effects of an experimental intervention on strategy use and help identify individual differences in response to intervention. Paired with AFM, this strategy clustering method can identify the strategies students are using and if these strategies are increasing in use. Just as work by Yudelson and Koedinger [20] has demonstrated learning gains with an improved model of procedural skills, so should researchers investigate how modeling strategic flexibility impacts learning.

In conclusion, we investigated the relationship between strategic flexibility and the strategies students used in an ITS using HMM clustering. We discovered that students could be clustered into four distinct strategy groups which differed on their average flexibility scores. A closer look at the strategies used by the four groups showed that students converged in their use of these strategies over the course of tutoring. An exploration of the different strategies suggested multiple explanations for students' strategy use including learning from tutor hints, gaps in knowledge, and decision making heuristics favoring different strategies. This study is an first step in integrating models of strategic flexibility into ITSs.

Acknowledgments

This work was supported in part by a Graduate Training Grant awarded to Carnegie Mellon University by the Department of Education (#R305B090023) and by the Pittsburgh Science of Learning Center (NSF #SBE-0836012). We would like to thank Carnegie Learning, Inc., for providing the Cognitive Tutor data supporting this analysis. All opinions expressed in this article are those of the authors and do not necessarily reflect the position of the sponsoring agency.

References

1. Rittle-Johnson, B., Star, J.R.: Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. Journal of Educational Psychology **99**(3) (2007) 561–574
2. Schneider, M., Rittle-Johnson, B., Star, J.R.: Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. Developmental Psychology **47**(6) (2011) 1525–1538

3. Alibali, M., Goldin-Meadow, S.: Gesture-Speech Mismatch and Mechanisms of Learning: What the Hands Reveal about a Child's State of Mind. Cognitive Psychology **25** (1993) 468–523
4. Blöte, A.W., Van der Burg, E., Klein, A.S.: Students' flexibility in solving two-digit addition and subtraction problems: Instruction effects. Journal of Educational Psychology **93**(3) (2001) 627
5. National Mathematics Advisory Panel: The Final Report of the National Mathematics Advisory Panel. Technical report, U.S. Department of Education (2008)
6. Newton, K.J., Star, J.R., Lynch, K.: Understanding the Development of Flexibility in Struggling Algebra Students. Mathematical Thinking and Learning **12**(4) (October 2010) 282–305
7. Rittle-Johnson, B., Star, J.R.: Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. Journal of Educational Psychology **101**(3) (2009) 529–544
8. Carpenter, T.P., Franke, M.L., Jacobs, V.R., Fennema, E., Empson, S.B.: A Longitudinal Study of Invention and Understanding in Children's Multidigit Addition and Subtraction. Journal for research in mathematics education **29** (January 1998) 3–20
9. Beal, C.R., Walles, R., Arroyo, I., Woolf, B.P.: On-line Tutoring for Math Achievement Testing: A Controlled Evaluation. Journal of Interactive Online Learning **6** (2007) 1–13
10. Koedinger, K.R., Anderson, J.R.: Intelligent Tutoring Goes To School in the Big City. International journal of artificial intelligence in Education **8** (1997) 1–14
11. Ritter, S., Anderson, J.R., Koedinger, K.R., Corbett, A.T.: Cognitive Tutor: Applied research in mathematics education. Psychonomic Bulletin & Review **14**(2) (2007) 249–255
12. Waalkens, M., Aleven, V., Taatgen, N.: Computers & Education. Computers & Education **60**(1) (January 2013) 159–171
13. Piech, C., Sahami, M., Koller, D., Cooper, S., Blikstein, P.: Modeling How Students Learn to Program. In: Proceedings of the 43rd ACM technical symposium on Computer Science Education, ACM (2012) 153–160
14. Baker, R.S., Corbett, A.T., Koedinger, K.R.: Detecting student misuse of intelligent tutoring systems. In: Proceedings of the 7th International Conference on Intelligent Tutoring Systems, Springer (2004) 531–540
15. Shih, B., Koedinger, K.R., Scheines, R.: A response time model for bottom-out hints as worked examples. Handbook of Educational Data Mining (2011) 201–212
16. Roll, I., Aleven, V., McLaren, B.M., Koedinger, K.R.: Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. Learning and Instruction **21**(2) (April 2011) 267–280
17. Smyth, P.: Clustering sequences with hidden Markov models. In Mozer, M.C., Jordan, M.I., Petsche, T., eds.: Advances in neural information processing systems, Citeseer (1997) 648–654
18. Kaufman, L., Rousseeuw, P.J.: Finding groups in data: an introduction to cluster analysis. Volume 344 of Wiley Series in Probability and Statistics. John Wiley and Sons, Inc., New Jersey (1990)
19. Cen, H.: Generalized Learning Factors Analysis: Improving cognitive Models with Machine Learning. PhD thesis, Carnegie Mellon University (2009)
20. Yudelson, M.V., Koedinger, K.R.: Estimating the benefits of student model improvements on a substantive scale. In D'Mello, S.K., Calvo, R.A., Olney, A., eds.: Educational Data Mining, Memphis, TN (2013)