A Developmental Model for Algebra Symbolization: The Results of a Difficulty Factors Assessment

Neil T. Heffernan (Neil@cs.cmu.edu)
Kenneth R. Koedinger (Koedinger@cmu.edu)
School of Computer Science; 5000 Forbes Avenue
Carnegie Mellon University; Pittsburgh, PA 15213 USA

Abstract

Given that the single most important mathematical skill for students who wish to study beyond arithmetic is the ability to take a problem situation (usually stated in words) and formulate a mathematical model (usually an equation), we are working on a cognitive developmental model of this skill to be used in an intelligent tutoring system. We call this skill *symbolization*. High school students do it poorly and improve slowly. We are using a Difficulty Factors Assessment as an efficient methodology for identifying the critical cognitive factors that distinguish competent from less competent symbolizers. We present a developmental model identifying three major transitions through which a student must pass. Underlying the developmental model are empirical results which suggest, contrary to prior research and common belief, the difficulty in algebra word problem solving is less about the difficulties of comprehending the word problems, and more about the difficulty of speaking in the foreign language of algebra. Many of students' errors are analogous to the errors people make when learning to speak in a new language. While it may be that mathematically algebra symbolization is a *generalization* of arithmetic, cognitively it is more accurate to say algebra symbolization is the *articulation* of arithmetic.

Introduction

When a student is presented with an algebra word problem such as problem *CS* in Table 1 and asked to provide a symbolic expression (rather than a numerical answer), she is doing what we refer to as *symbolizing*. For instance, the symbolic expression for *CS* is "(72-m)/4." In studying symbolization skills we have focused on algebra story problems but our results may also be relevant more generally to symbolization skills needed in using a calculator or programming a spreadsheet or computer. As these computational devices take over more of the symbol manipulation of algebra, symbolization deserves, and is receiving, increasing instructional emphasis. We are trying to understand how students learn to symbolize. To test that understanding we are building a cognitive developmental model which will be used in an intelligent tutoring system.

Much of the prior work (Cummins et. al., 1988; LeBlanc & Weber-Russell; 1996, Lewis & Mayer, 1987; Paige & Simon, 1979) on word problem solving has focused on students' comprehension abilities. Cummins et. al. "suggest that much of the difficulty children experience with word problems can be attributed to difficulty in comprehending abstract or ambiguous language." The general conclusion from the above research is that comprehension rules are key knowledge components students must acquire to become competent problem solvers.

Generalization/Comprehension Hypothesis

This brings us to the first naturally intuitive hypothesis about what makes algebra symbolization difficult. What we call the *generalization hypothesis* flows naturally from the previous literature that has focused on comprehension difficulties caused by abstract language. An important conceptual leap for students is to move from the concrete grounded world of arithmetic problems to the more abstract world of algebra problems with variables. According to this hypothesis, what is hard about algebra is the generalization ability needed to think about a problem with an abstract variable, as opposed to a concrete known.

But more recent research by Koedinger & Anderson (in press) found evidence that acquiring such comprehension skills is not sufficient for symbolization competence. They reported that in 36% of problems that students comprehended well enough to find a numerical answer, students nevertheless failed to correctly symbolize. We call the difference between the students' performance on symbolization problems and on similar arithmetic problems the *symbolization effect*.

Previous Results on this Problem

To determine if the symbolization effect is well-explained by the generalization hypothesis we (1997) performed an assessment in which we compared students' performance on symbolization problems (like problem *CS*) with their performance on *arithmetic symbolization* problems. We could turn problem *CA* into an arithmetic symbolization problem simply by changing the final question to "Write an single expression that shows all the computations you would need to do in order to find how much she can spend on each sister." The correct answer for this problem would be "(72-32)/4." Arithmetic symbolization problems are a special type of symbolization problem that have no variables. We found no statistically significant difference between students performance on these two types of problems. This suggests that the presence of a variable did not increase difficulty, thus