Wayang Outpost: A web-based multimedia intelligent tutoring system for high stakes math achievement tests

Carole R. Beal, Ivon Arroyo, James M. Royer, & Beverly P. Woolf
University of Massachusetts-Amherst

Context. The theme of the 2003 annual meeting is "accountability for educational quality". The emphasis on accountability reflects the increasing need for educators, students and politicians to demonstrate the effectiveness of educational systems. As part of the growing emphasis on accountability, high stakes achievement tests have become increasingly important and a student's performance on such tests can have a significant impact on his or her access to future educational opportunities. At the same time, concern is growing that the use of high stakes achievement tests, such as the SAT-Math exam and others (e.g., the Massachusetts MCAS exam) simply exacerbates existing group differences, and puts female students and those from traditionally underrepresented minority groups at a disadvantage (Willingham & Cole, 1997). New approaches are required to help all students perform to the best of their ability on high stakes tests.

Background: Intelligent tutoring systems. Computer-based intelligent tutoring systems (ITS) provide one promising option for helping students prepare for high stakes assessments. Research on intelligent tutoring systems has clearly shown that users of tutoring software can make rapid progress and dramatically improve their performance in specific content areas. Although much ITS research focuses on military, industry, and other non-academic training situations, evaluation studies of several ITS for school mathematics also show benefits to student users in school settings. Specifically, studies of the Carnegie Tutor (CMU) for algebra and the AnimalWatch tutor (UMass-Amherst) for arithmetic indicate that student users successfully master specific skills and that their attitudes towards math become more positive as a result of working with the software (< www.carnegielearning.com >; Arroyo, Beck, Schultz & Woolf, 1999; Beal & Arroyo, in press). However, existing ITS are limited in several ways. First, as stand-alone applications that must be installed on individual computers, they are difficult to disseminate, update and extend. Second, although they provide effective instruction, few ITS have really taken advantage of the instructional possibilities of dynamic multimedia techniques such as animation, techniques that are common in commercial software (Beal, Beck, Westbrook, Atkin, & Cohen, 2002). For example, both the Carnegie Tutor and AnimalWatch systems are basically interactive electronic worksheets. Third, although current ITS model the student's knowledge on an ongoing basis in order to provide effective help when needed, there have been only preliminary attempts to incorporate knowledge of student group characteristics (e.g., profile of cognitive skills, gender) into the tutor and to use this profile information to guide instruction (Shute, 1995).

Objectives. The goal of the current project is to design, implement and evaluate an intelligent tutoring system for high-stakes math achievement tests. The system, called Wayang Outpost, currently provides tutoring for SAT-Math problems, but has been designed so that problem content appropriate for other math achievement tests (e.g., the MCAS in Massachusetts) can be added to the existing architecture. Wayang Outpost is

designed to provide web-based access to a) tutoring on SAT-Math problems, using information about the student's cognitive profile to customize instruction and improve performance on high stakes assessments and b) modules to assess and enhance the student's profile of cognitive strengths and weaknesses. More detail about the design of these components will be provided next.

Tutoring via web-based multimedia. Wayang Outpost provides instruction via a web site, ensuring easy access to students either at home or from any school connected to the Internet. Use of the web delivery format allows easy expansion and frequent updating of the system, without requiring users (students or teachers) to download or reinstall the program. The student begins a session by logging into the site and receiving a math problem. The setting is a Flash animated classroom based in a research station in Kalimantan (formerly, Borneo), populated by researchers, visitors, and students interested in environmental issues, which provides rich real-world content for mathematical problems.

Each math problem (currently, a battery of SAT-Math problems provided by the College Board) is presented in a Flash movie, including an animated character based on the traditional Indonesian art form of shadow puppetry (Wayang; hence the name of the system). If the student answers incorrectly, or requests help, the teacher character provides step by step instruction and guidance in the form of Flash animations with audio. For example, on a geometry problem, the student might see an angle with a known value rotate and move over to the corresponding angle with an unknown value on a parallel line, thus emphasizing the principle of correspondence, while the teacher describes the relation. The explanations and hints provided in Wayang Outpost therefore resemble what a human teacher might provide when explaining a solution to a student, e.g., by drawing, pointing, and gesturing while talking, in contrast to previous mathematics ITS which relied heavily on screen-based text and static examples.

The selection of specific math problems for the student and the provision of hints is based on the tutor's representation of the student's knowledge and on the student's cognitive profile (e.g., spatial ability and math fact retrieval speeds). For each SAT-Math problem, data from paper-and-pencil test performance have been used to identify the most critical cognitive skills predicting successful solution, which may differ by gender. For example, for a specific geometry problem, high spatial ability may be strongly predictive of success for females. Based on these data, each problem includes two types of hints: one based on an algebraic, step by step approach, the second based on mental rotation and visual estimation (a strategy that past research suggests is more often used by male students). The choice of hint type can be customized for individual students on the basis of their cognitive profile, to help them develop strategies and approaches that may be more effective for particular problems. For example, women who score low on the spatial ability assessment might receive a high proportion of hints that emphasize mental rotation and estimation, approaches that women often avoid even though they are generally more effective in a timed testing situation.

As the student works through a problem, performance data (e.g., latency, answer choice) are used to update the student model on an ongoing basis and to select problems at the appropriate level of challenge. Student models and performance data

reside on a central server; communication between Flash and the Java-based tutor is via an XML protocol. A novel feature of the Wayang Outpost site is the link between the animated content observed by the student user, and the intelligent selection and delivery of the content by the tutor. Flash problem movies are initially downloaded to the student user's local computer, but the full set of hints for each problem can include dozens of movie and sound files. To minimize download times, the tutor uses its estimate of the student's knowledge (e.g., this student does not yet understand that angles on parallel lines have corresponding values) to download only the files it believes the student is going to need, i.e., the tutor delivers web content intelligently and in real time, customized for each student.

Macroadaption of instruction. One of the advantages of intelligent tutoring software is the ability to represent the knowledge of individual users, and to select problems to extend the user's current level of understanding. For example, in the area of mathematics, the student model includes a representation of the student's estimated knowledge relative to the tutor's network of domain skills (e.g., knowledge of right triangles; parallel lines; perimeter, etc.). This student model representation is continually updated as the student works and is used to micro-adapt instruction to the individual user by presenting problems at the appropriate level of challenge, just as a skilled human tutor would choose problems that are neither too easy or too difficult for the student.

More recently, research in the ITS community has begun to move towards macro-adapting instruction on the basis of more global student characteristics, such as level of cognitive development and preference for help that is more or less concrete, structured and algorithmic (Arroyo et al., 1999). Wayang Outpost extends this macro-adaptive approach by representing information about the student's skills in areas that have been shown to predict performance on math achievement tests, specifically, spatial ability and math fact retrieval speed (MFR). Gender and other group differences have been observed in these component skills; in particular, on average, women tend to score significantly lower on spatial ability than males, and to have slower math fact retrieval times. Young women also tend to under-perform on high stakes math achievement tests, relative to their performance in the classroom.

Wayang Outpost includes on-line assessments of these cognitive skills as well as activities to enhance these abilities. E.g., for spatial ability, students complete an on-line version of the Vandenberg test of mental rotation. Students who perform poorly are directed to an area of the site that offers training in spatial visualization via animated problems involving paper folding and 3-D rotation, and training in specific strategies for tackling spatial problems (e.g., rotate on one axis at a time). Our hypothesis is that students with weak spatial ability (more of whom are likely to be female) will benefit from the opportunity to practice spatial visualization skills and that this experience will transfer to improved performance on certain SAT-Math problems for which we have shown spatial ability to be important.

The Wayang Outpost design also includes a module to assess and enhance students' math fact retrieval skill (MFR). The automatic and efficient retrieval of math facts is important because it frees cognitive capacity for problem representation and solution activities. Past research suggests that rapid MFR times are a stronger predictor of high

stakes math achievement test performance than spatial ability. MFR times are assessed by speed and accuracy of responses to a battery of simple addition, subtraction and division problems (Royer, Tronsky, Chan, Jackson, & Merchant, 1999). Times can be dramatically improved through practice, and practice has in turn been shown to reduce gender differences on math achievement tests. Therefore, students who score poorly on the MFR assessment may be directed, on the basis of the student model representation, to an area of the Wayang Outpost system that provides MFR practice.

Evaluation. Large-scale evaluation studies of Wayang Outpost will be conducted in the coming academic year. The primary measures will be performance on SAT-M problems, in a pre and post test design, as a function of the type of tutoring provided by the system (more spatial versus more algebraic) and the student's cognitive profile. In addition, data on student performance as a function of hints will also be used to assess the effectiveness of the tutor's help. Specifically, a student who receives help on a problem of one type should achieve correct performance with less or no help on a subsequent problem requiring the target concepts and skills. Use of the hint effectiveness measure (reduction in errors across problems of the same type) allows for evaluation of system effectiveness even though there is wide variability in the pattern of problems and performance across student users.

Future goals: Transfer. Although it is important to help achieve their potential on high stakes tests, the ultimate goal for educators is to enhance students' conceptual understanding of mathematical concepts, and their ability to draw on their skills to solve novel problems. Tnext step in the Wayang Outpost research project is to incorporate animated adventures which present the student with a challenging mathematical problem that will require multiple steps and skills to solve. These adventures are being designed to incorporate real-world content appropriate to the setting, e.g., an animated character (based on Dr. Anne Russon of York University, an expert on the orang-utan) will invite the student to help her observe an illegal logging operation in the rainforest and estimate the number of orang-utans that might be affected. Performance on the adventure math problems will become a measure of transfer of math skills from the SAT-Math tutor to more complex mathematics.

References

Arroyo, I., Beck, J., Schultz, K., & Woolf, B. P. (1999). A useful prediction variable for student models: Cognitive development level. Poster presented at the Seventh International Conference on User Modeling, Banff Canada.

Beal, C. R., & Arroyo, I. (in press). The AnimalWatch project: Creating an intelligent computer mathematics tutor. In S. Calvert, A. Jordan, & R. Cocking (Eds.), Children and the digital age. Greenwood: Westport CT.

Beal, C. R., Beck, J., Westbrook, D., Atkin, M., & Cohen, P. (2002, April). Intelligent modeling of the user in interactive entertainment. Paper presented at the AAAI Spring Symposium, Stanford CA.

Royer, J. M., Tronsky, L. N., Chan, Y., Jackson, S. J., & Merchant, H. (1999). Math fact retrieval as the cognitive mechanism underlying gender differences in math test performance. Contemporary Educational Psychology, 24, 181-266.

Shute, V. (1995). SMART: Student Modeling Approach for Responsive Tutoring. User Modeling and User-Adapted Interaction. 5, 1-44.

Willingham, W. W., & Cole, N. S. (1997). Gender and fair assessment. Mahwah NJ: Erlbaum.