WEBsistments: Enabling an Intelligent Tutoring System to Excel at Explaining Teaching Other Than Coaching

Yue Gong, Joseph E. Beck, Neil T. Heffernan

Computer Science Department, Worcester Polytechnic Institute
100 Institute Road, Worcester, MA, 01609, USA
{ygong, josephbeck, nth}@wpi.edu

Abstract. Most step-based Intelligent Tutoring Systems (ITS) are well suited for providing problem solving practice, and are well-tailored to help students solve specific items. Consequently, many ITS typically fail to perform as strong media for conveying conceptual and procedural *instruction*, rather than coaching. In order to overcome this deficiency, we leverage existing web-based resources, as many existing resources are well-designed for providing instruction. By combining external web pages with the ASSISTments tutoring system, we have created a stronger intervention that we have dubbed *WEBsistments*. A preliminary study found that students who were wrong on a problem and received a web page as assistance, improved more, relative to students who did not see a web page. In addition, our results suggest that weaker students seem to benefit more from using web pages as extra help.

Keywords: WEBsistments, Web-based Resources, Conceptual Instruction

1 Introduction

After their first emergence over two decades ago, Intelligent Tutoring Systems (ITS) have attracted researchers from a variety of disciplines. Many research studies have been showing that ITSs resulted in substantial successes in improving student learning in different domains, such as mathematics [1], physics [2], and reading [3]. These systems' two major advantage over traditional classroom practicing is that students can get immediate feedback on correctness, and the ability to request help [4].

A common type of ITS is step-based [5]. Once a student enters a step, the tutor can provide feedback or help. Following this architecture, many ITS were designed to help students solve problems step by step. They generally provide several different forms of help, such as worked examples, hint messages and scaffolding questions (e.g. [1, 2]). Such assistance, independent of the many forms it takes on, is tactical with the goal of directing students to the solution for this problem. Thus the systems are well suited for coaching students. On the other hand, we notice that due to such design, the systems lack the ability to perform as a strong medium for conveying instruction for students who lack the background knowledge to benefit from coaching. We seek to address this problem of low-knowledge students not by authoring new content, but by

adfa, p. 1, 2011.

instead utilizing web-based resources which are already on the Internet. There are three reasons we see a benefit from this integration.

First, ITS are effective in assisting students with problem-solving practice. Web-based resources, however, are often not designed for a specific problem, but rather illustrate concepts, introduce vocabulary, and explain procedural solutions of a skill at a more general level. Including such material extends the repertoire of instruction that an ITS can deliver to students.

Second, web-based resources cheaply extend the range of media available for tutoring. Traditionally, for reasons of cost and expertise, much assistance in an ITS is text-based. Web-based resources are able to convey knowledge in a variety of modalities such as videos of a human teaching the skill that serve as a human tutor within a computer tutor, or animations that allow students to manipulate some components in order to teach students interactively. Intuitively, these new features could possibly help students learn by broadening the types of interactions.

Third, there are lots of good educational resources on Internet already. Rather than spending effort to create such resources, it is more cost effective to search for existing content, select content that appears to be effective, and integrate it into the ITS.

It appears that computer tutors and web-based recourses each address one aspect of education: coaching on problem-solving and general instruction, respectively. However, neither of them alone offers a complete solution. Towards the goal of finding an efficient means of constructing an intervention that covers both aspects of education, this paper presents our early-stage effort of combining web pages with ASSISTments: WEBsistments.

2 Methodology

The ASSISTments (www.assistments.org) system is a web-based tutor, primarily used for middle school mathematics by tens of thousands of students. Its standard method of instruction is to provide hints to help the student solve the problem, or scaffolding, which breaks the problem down into smaller steps. We enhanced its functionality to enable it to provide a button “Show me a web page,” which allows students to request a web page while solving a problem.

Students are allowed to request a web page in any stage of problem-solving, even before their first attempts. When a student clicks the request button, WEBsistments displays a web page associated with the skill tested by the problem. When there are multiple skills required in a problem, the web pages associated with the most advanced skill will be used to select a web page. A student cannot ask for multiple web pages while solving a problem, but he can use original assistance (hints and scaffolding) for the problem. WEBsistments collects information, such as how long the student spent on a web page, his next immediate action after seeing a web page was, whether he got the question correct right after seeing a web page, etc.

In the 2011-2012 version of WEBsistments, web resources were selected by two Worcester Polytechnic Institute undergraduates and a few volunteer middle school Math teachers. They ensured that each of the 147 Math skills that ASSISTments tracks had 2-5 web pages that provide instruction on the skill. Then they tagged those web pages with the skill, indicating that the web page is relevant to that skill. Most problems in ASSISTments have already been tagged with one (or more) of 147 skills by domain experts. Therefore, through the skill mapping, there is a connection established between a problem and a set of relevant webpages. Since this is our first implementation of WEBsistments, we do not have a basis to prefer any page that has passed our screening process. Therefore, when deciding which web page to show, WEBsistments uses random selection.

WEBsistments has been used by 1121 8th grade (approximately 13 years old) students since July 2011. Since not all students chose to see a web page, we had to decide upon a comparison group, and selected students who were classmates of those who did request web pages. We also restricted our comparison set to those problems on which a student requested a web page, and only considered cases where students made an incorrect response. As a result of these restrictions, our data set consists of 9,983 problems solved by the students. The Web group includes cases where the student requested a web page (1104 problems); the No-web group includes the cases where students did not request a web page (8879 problems). Note that a student can be a member of both groups, if, for example, he requested a web page in one instance but decided not to in another.

3 Results

Each instance in our data set represents a student's wrong response to an initial problem, which we denote as P_1. We then measure the student's performance on the next item using the same skill; we denote this problem as P_2. To measure the learning gain, rather than just taking the difference of $P_2 - P_1$, we instead normalize the result by the population's average performance on each item. If P_2 is extremely easy, we should give not treat that as strong evidence of learning relative to a student getting a difficult question correct. In addition, we also considered the easiness of P_1. This is because it tells whether one group has lower incoming knowledge than the other, as they may fail to respond to P_1 correctly even if P_1 was an easier problem. Therefore, we used the percent correctness of a problem across the entire population of ASSISTments students within the 2011-2012 school year to represent its easiness.

We used a performance score, shown in Equation 1, to represent how well a student performed in a problem. $\text{Correctness}_{i,j}$ is a binary value, 1 representing a correct response of student i to problem j and 0 representing incorrect. Problem easiness also ranges from 0 to 1 and a higher value means an easier problem. A performance score credits a student more when he successfully solves a harder problem, while punishes a student more when he fails an easier problem. Using performance scores, we calculated a gain score of a student between P_2 and P_1 by subtracting performance score_{p1} from performance score_{p2}. We then calculated a gain score for each of the instances in the web and no-web groups.

$$\text{performance score } (i \in \text{students}, j \in \text{problems}) = \text{correctness}_{i,j} - \text{easiness}_j. \quad (1)$$

3.1 Overall Trend from Web Pages

In this section, we present our preliminary analyses of the data, aiming to examine whether there are any trends suggesting the effectiveness of web pages.

Table 1. Comparisons of the mean gain scores between the web and no-web groups

	Web group			No-web group		
	Mean	95% C.I.	N	Mean	95% CI	N
Overall	0.50	0.49 - 0.51	1104	0.40	0.39 - 0.40	8879
No Bottom-out	0.60	0.58 - 0.62	518	0.49	0.48 - 0.49	5336
Bottom-out	0.41	0.39 - 0.43	586	0.26	0.25 - 0.27	3543

Table 1 compares the statistics of the gain scores of the two groups. First, we observed that there were fewer cases, 1104, where students requested web page resources. In most cases, students still only sought for the traditional assistances of the tutor when they were stuck in the problems as there are three times as many (3543) cases where students solely used bottom-out hints. This result possibly suggests that the students preferred receiving the answer to learning, and raises issues of whether the group requesting web pages differs in desire to learn.

Second, we found that the mean gain score of the web group is 0.1 higher the no-web group, and the 95% confidence intervals have no overlap in values, indicating that the means are different at a significance level of 0.05. This result suggests that overall students who saw a web page learned more.

In addition, we extended our study to examine how web pages work for students with different proficiencies in Math. We included a new factor, "bottom-out hint" and used that to indicate a student's proficiency. A bottom-out hint is presented as the last message in a sequence of hints for a problem, in which the answer to the problem is explicitly given. Due to its functionally, in the ITS research field, requesting a bottom-out hint presumably suggests that the student is weaker so as to need more help. We present the statistics of the two-way factorial in the last two rows of Table 1. The four means are corresponding to the factorial combinations of the use of web pages and the use of bottom-out hints. Consistent to the overall effects, at the factor level of "bottom-out", in its two levels, each mean of the web group is higher than that of the no-web group. It suggests that web page support is generally helpful for both stronger and weaker students.

We found that the impact of using web page resources may be more effective for those who request bottom-out-hints. The difference between the means of the two sub groups is 0.15 (i.e. 0.41-0.26 = 0.15), somewhat larger than the overall effect. Perhaps weaker students benefited more from getting extra web-based instruction? Moreover, the average gain for bottom-out-hints without web page support is just 0.26, suggesting that hint messages are a relatively slow means of instruction.

3.2 Modeling Effects of Web Pages

There are two issues which potentially impact the results of the previous statistical analyses. First, we did not consider whether the student saw a web page in P_2.

Consider an example where a student does not request a web page in P_1, requests a page in P_2 before responding, and as a result of the page gets P_2 correct. This student would show learning from P_1 to P_2, and the no-web group would benefit since the student saw no web page on P_1. Second, students certainly vary in their mathematics proficiency, which our first comparison did not account for.

To address these issues, we trained a model that considered multiple relevant factors simultaneously. For each instance, we used a binary value to indicate whether the student has seen a web page in P_1 and in P_2. We used how many correct responses and incorrect responses have been produced by the student for the required skill to represent student proficiency. These two variables are used in the PFA model [6] and have been shown to effectively represent student proficiency [7].

Table 2. The logistic regression model of impacts on correctness of P_2

Independent variables	β
Saw a web page in P_1	**0.393**
Saw a web page in P_2	-1.693
Problem easiness of P_1	-0.983
Problem easiness of P_2	4.808
Number of prior corrects on the skill	0.010
Number of prior incorrects on the skill	-0.023
Reached the bottom-out hint in P_1	-0.635
Intercept	-1.992

Table 2 shows the result of the multinomial logistic regression run in SPSS to create a model to predict the correctness of P_2. The regression model generated $r^2=0.17$, and all of the independent variables are reliable at $p<0.05$. Observing the coefficient value of "Saw a web page in P_1", 0.393, we found that the model suggests the same trend as our prior statistical analyses. Considering the effects of all the relevant factors together, the model still acknowledges the positive effect of seeing a web page on helping students respond correctly to the next problem.

4 Contributions, Future Work, and Conclusions

This paper discusses a common issue across many ITSs (e.g.[1, 2, 3]): most step-based Intelligent Tutoring Systems focus predominantly on problem-solving. However, in order for students to benefit from problem-solving practice, sufficient declarative knowledge is essential [8], but ITS generally leave this task to teachers. We proposed and have pilot tested a solution to the problem: using web page resources on Internet as a complementary medium. We built WEBsistments to enhance an ITS to have the best of both worlds of coaching and instruction. We have found a promising trend of the effectiveness of this solution. This solution could be easily applied for most computer tutors, and is a low-cost option for ITS designers.

There are steps that could make WEBsistments better. First, students appear reluctant to request instruction; perhaps a tutorial policy that is proactive for students the tutor observes struggling to master the material would make sense? Our current on-demand policy could also cause a selection bias of students, and is certainly a potential confound in our result, as instances in the web group are likely to be those done by motivated students who may be more eager to learn. However, the statistical model accounts for some of these individual differences. Second, a more intelligent method of selecting web pages is desired as it is likely that some web pages are more effective than others. In addition, individualizing web page recommendations is an interesting possibility. To prompt more learning and provide a web page to ensure that the most, a student's individual context could be considered as well. Possibly, student modeling and WEBsistments can make a strong join for this purpose.

In this paper, we presented our work, WEBsistments, to enhance a computer tutor to not only provide problem-solving practice, but also convey conceptual instruction to students. We conducted a pilot study to examine our hypothesis that students could learn more due to having this new form of assistance. Our results suggested that when web-based resources were used to help students in their problem-solving, it results in more gains in their performances in next problems. In a model where more factors were considered simultaneously, we also confirmed the positive effect of web pages. Moreover, bottom-out-hinting students, or weaker students in typical beliefs, seem to benefit more from receiving web-based resources as extra help.

References

1. Razzaq, L., Feng, M., Nuzzo-Jones, G., Heffernan, N.T., Koedinger, K. R., Junker, B., et al. The Assistment project: Blending assessment and assisting. In C.K. Looi, G. McCalla, B. Bredeweg, & J. Breuker (Eds.) Proceedings of the 12th Artificial Intelligence in Education, Amsterdam: ISO Press. 555-562. (2005).
2. Vanlehn, K., Lynch, C., Schultz, K., Shapiro, J. A., Shelby, R. H., Taylor, L., et al. The Andes physics tutoring system: Lessons learned. International Journal of Artificial Intelligence in Education, 15(3), 147–204, (2005)
3. Mostow, J. and G. Aist Evaluating tutors that listen: An overview of Project LISTEN. Smart Machines in Education. K. F. a. P. Feltovich, MIT/AAAI Pres: 169-234. (2001).
4. Mendicino, M., Razzaq, L., Heffernan, N. T. A Comparison of Traditional Homework with Computer Supported Homework: Improving Learning from Homework Using Intelligent Tutoring Systems. Journal of Research on Technology in Education (JRTE), 41(3), 331-358.(2009)
5. K. VanLehn. The Behavior of Tutoring Systems. International Journal AI in Education, 16(3):227–265, 2006.
6. Pavlik, P. I., Cen, H. , Koedinger, K. Performance Factors Analysis - A New Alternative to Knowledge. Proceedings of the 14th International Conference on Artificial Intelligence in Education, Brighton, UK, pp. 531-538. (2009)
7. Gong, Y., Beck, J. Items, Skills, and Transfer Models: Which Really Matters for Student Modeling? Proceedings of the 4th International Conference on Educational Data Mining, pp.81-90. (2011)
8. Anderson, J. R. Rules of the Mind. Hillsdale, NJ, Lawrence Erlbaum Associates. (1993).