Analysing High-Level Help-Seeking Behaviour in ITSs

Moffat Mathews, Tanja Mitrović, and David Thomson

Computer Science & Software Engineering,
University of Canterbury, New Zealand
{moffat,tanja,djt88}@cosc.canterbury.ac.nz

Abstract. In this paper, we look at initial results of data mining students' help-seeking behaviour in two ITSs: SQL-Tutor and EER-Tutor. We categorised help given by these tutors into high-level (HLH) and low-level help (LLH), depending on the amount of help given. Each student was grouped into one of ten groups based on the frequency with which they used HLH. Learning curves were then plotted for each group. We asked the question, *"Does a student's help-seeking behaviour (especially the frequency with which they use HLH) affect learning?"* We noticed similarities between results for both tutors. Students who were very frequent users of HLH showed the lowest learning, both in learning rates and depth of knowledge. Students who were low to medium users of HLH showed the highest learning rates. Least frequent users of HLH had lower learning rates but showed higher depth of knowledge and a lower initial error rate, suggesting higher initial expertise. These initial results could suggest favouring pedagogical strategies that provide low to medium HLH to certain students.

A primary aspect of researching and developing adaptive systems is to try and understand the behaviour of those using the system. Being able to comprehend various types of behaviour gives us the basis to form strategies to adequately, effectively, and even adaptively aid users of the system. This is particularly the case in Intelligent Tutoring Systems (ITSs), where understanding each student's behaviour is critical to creating and implementing suitable pedagogical strategies to appropriately guide each student adaptively through their learning tasks in order to maximise their learning. One such type of behaviour is the way in which a student requests and utilises help. Help-seeking behaviour has been studied in various contexts; from traditional teaching methods in the classroom to e-learning applications. It has long been noted in education literature that seeking help and way in which it is sought affects learning [1]. Certain aspects of help-seeking behaviour (such as gaming [2]) have been researched in the context of ITSs [3]. In an adaptive system, one method of studying users' behaviour is by mining data collected from users (e.g. user models and logs).

In this paper, we discuss the initial results of data mining student logs and user models for help-seeking behaviour in two ITSs, namely SQL-Tutor [4] and EER-Tutor [5]. We grouped students by the frequency with which they used help

W. Nejdl et al. (Eds.): AH 2008, LNCS 5149, pp. 312–315, 2008.

and tried to determine if there were differences in learning between the groups of students. We did this by plotting learning curves for each group and seeing if any trends existed, and if these trends were similar between the two ITSs.

SQL-Tutor is a constraint-based modelling (CBM) ITS that provides intelligent and adaptive guidance in the domain of SQL database querying. SQL-Tutor has been used since 1998 in tertiary undergraduate database courses. The student spends the majority of time solving problems in the task environment. The task environment contains the problem text, solution workspace, feedback pane, and problem context information (e.g. information about the schema). On submission of their solution, a student can receive help from six levels of problem-related feedback. These levels increase in the amount of help, and are 1. Simple Feedback, 2. Error Flag, 3. Hint, 4. Partial Solution, 5. List All Errors, and 6. Complete Solution. On each incorrect submission, the help level automatically increments to a maximum of 3 (i.e. hint). Help levels are selected via a combo box and the student has the ability to override the current selection at any time by selecting a different level. We divided the help into two categories depending on the amount of help given: low-level help (LLH) for the first three help levels and high-level help (HLH) for levels four, five, and six. Furthermore, LLH automatically increments on incorrect submissions whereas HLH has to be selected by the student.

EER-Tutor (Enhanced-Entity Relationship Tutor) is a CBM ITS that teaches conceptual database design using the Enhanced Entity Relationship Model, and provides students with problems to practise their entity relationship modelling skills in a coached environment. Developed initially as KERMIT (Knowledge-based Entity Relationship Modelling Intelligent Tutor) then ER-Tutor (Entity-Relationship Tutor) and now EER-Tutor, this ITS has also had many years of successful use with students in tertiary undergraduate database courses. The help-levels in EER-Tutor are similar to SQL-Tutor and thus make it easy for comparison. As with SQL-Tutor, we divided help into LLH and HLH.

Although these two ITSs deal with database related areas, each domain is very different, Furthermore, the method of solving problems (even to the point of *text* versus *diagrammatic)* is considerably different.

1 Method

In both datasets, data for students who made less than five attempts was omitted from the analysis. The SQL-Tutor dataset consisted of 1,803 students who made a total of 100,781 attempts, and spent just over a total of 1,959 active hours on the system. EER-Tutor dataset consisted of 936 students who made a total of 43,485 attempts, and spent just over 2,830 active hours on the system.

To enable us to compare the frequency of HLH use among students, we calculated an HLH-Ratio $(\frac{\text{Number of HLH attempts}}{\text{total number of attempts}})$ for each individual. For example, a student with an HLH ratio of one used HLH on every attempt; in contrast a student with an HLH ratio of zero never used HLH. For comparison between groups of users with similar HLH, ten groups ($A1 - A10$) were formed, each with an HLH ratio range of 0.1. Students were placed into groups depending on their

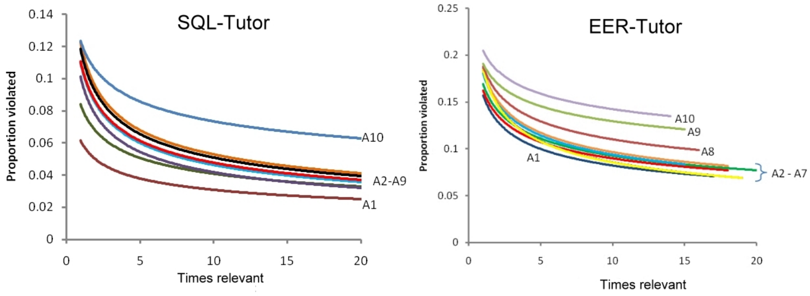

Fig. 1. Learning curves for the HLH groups (A1-A10) for SQL-Tutor and EER-Tutor

HLH ratio, such that students in group A10 (who used HLH 90–100% of the time) were the most frequent users of HLH while the least frequent HLH users were in group $A1$. Learning curves were then plotted for each group (Figure 1).

Table 1. Power curve equations and fits (R^2) for the ten HLH groups ($A1 - A10$) in SQL-Tutor and EER-Tutor

		SQL-Tutor			EER-Tutor		
Group	HLH ratio	Users	Curve equation	R^2 (Fit)	Users	Curve equation	R^2 (Fit)
A1	0.0 - 0.1	222	$y = 0.061x^{-0.30}$	0.844	210	$y = 0.156x^{-0.28}$	0.959
A2	0.1 - 0.2	214	$y = 0.084x^{-0.31}$	0.956	186	$y = 0.162x^{-0.25}$	0.963
A3	0.2 - 0.3	248	$y = 0.101x^{-0.38}$	0.955	120	$y = 0.169x^{-0.26}$	0.963
A4	0.3 - 0.4	315	$y = 0.109x^{-0.37}$	0.956	103	$y = 0.185x^{-0.33}$	0.938
A5	0.4 - 0.5	295	$y = 0.110x^{-0.36}$	0.965	89	$y = 0.181x^{-0.28}$	0.967
A6	0.5 - 0.6	211	$y = 0.123x^{-0.39}$	0.961	57	$y = 0.183x^{-0.28}$	0.947
A7	0.6 - 0.7	122	$y = 0.122x^{-0.36}$	0.953	51	$y = 0.184x^{-0.33}$	0.978
A8	0.7 - 0.8	88	$y = 0.115x^{-0.35}$	0.953	51	$y = 0.189x^{-0.23}$	0.954
A9	0.8 - 0.8	44	$y = 0.118x^{-0.36}$	0.912	34	$y = 0.190x^{-0.16}$	0.858
A10	0.9 - 1.0	44	$y = 0.123x^{-0.22}$	0.956	35	$y = 0.204x^{-0.15}$	0.878

2 Results and Discussion

The power curve equations and fits are shown in Table 1. The results discussed are similar for both tutors.

All learning curves have a very good fit (R^2), with the lowest fit being just 0.844. The degree of fit usually indicates level of transferability of the skills that were learned. For example, a low fit indicates high variability in the error rates (i.e. high deviation of points from the power curve) indicating that error rates still vary each time a particular concept is encountered (i.e. low transferability). This result indicates that whatever skills students are learning are also transferable. This does not indicate that all students are learning the same skills.

The exponent in the equation indicates the learning rate. As can be seen from Table 1, the learning rates are highest for students who are low to medium users of HLH. Students that are extremely high users of HLH (e.g. A10) have the lowest learning rates. These students also display shallow learning. This can be seen from the point at which the slope of the learning curves approximates zero. For extremely high HLH users, this point still shows a high error rate, indicating that the concept has not been learned to any great depth. This could be because students that rely heavily on HLH do not actively think for themselves or engage in deliberate practice, and therefore do not get the opportunity to learn from their mistakes.

The coefficient of x (known as χ) shows the initial error rate. Low χ usually means the presence of expertise or previous experience and vice-versa. The χ value for group $A1$ in SQL-Tutor shows this expertise or prior knowledge. Manual inspection of logs indicated the presence of students with higher expertise in this group. As a consequence, students who have a higher χ find the domain more difficult than those with lower χ values. From Figure 1, we can see that the students who used the least help had the least χ, whereas students who used the most help had the highest χ and therefore found the domain more difficult.

Although these initial results provide a good basis for understanding one aspect of help-seeking behaviour, and thus aids in creating pedagogical strategies, it cannot be construed from these results that providing low to medium help to students will automatically increase learning. Other factors such as the meta-cognitive ability (e.g. help-seeking skills) of students, their upbringing, and even their cultural influences also need to be considered. It could also be that students who are slower to learn are less confident and therefore seek HLH more often.

In the near future, we intend to analyse the effect of other variables such as time spent on attempts, number of problems solved, and difficulty of problems solved on these groups of students.

References

1. Gall, S.N.: Help-Seeking Behaviour in Learning. Review of Research in Education 12, 55–90 (1985)
2. Baker, R., Corbett, A., Koedinger, K., Roll, I.: Detecting When Students Game the System, Across Tutor Subjects and Classroom Cohorts. In: Ardissono, L., Brna, P., Mitrović, A. (eds.) UM 2005. LNCS (LNAI), vol. 3538, pp. 220–224. Springer, Heidelberg (2005)
3. Aleven, V., Koedinger, K.: Limitations of Student Control: Do Students Know when They Need Help? In: Gauthier, G., VanLehn, K., Frasson, C. (eds.) ITS 2000. LNCS, vol. 1839, pp. 292–303. Springer, Berlin (2000)
4. Mitrović, T.: An Intelligent SQL Tutor on the Web. IJAIED 13, 173–197 (2003)
5. Suraweera, P., Mitrović, T.: An Intelligent Tutoring System for Entity Relationship Modelling. JIJAIED 14, 375–417 (2004)