Long Term Retention of Programming Concepts Learned Using Software Tutors

Amruth N Kumar [0000-0002-1951-3995]

Ramapo College of New Jersey, Mahwah NJ 07430, USA
amruth@ramapo.edu

Abstract. Do students retain the programming concepts they have learned using software tutors over the long term? In order to answer this question, we analyzed the data collected by a software tutor on selection statements. We used the data of the students who used the tutor more than once to see whether they had retained for the second session what they had learned during the first session. We found that students retained over 71% of selection concepts that they had learned during the first session. The more problems students solved during the first session, the greater the percentage of retention. Even when students already knew a concept and did not benefit from using the tutor, a small percentage of concepts were forgotten from the first session to the next, corresponding to transience of learning. Transience of learning varied with concepts. We list confounding factors of the study.

Keywords: Retention of Learning, Programming, Code-tracing tutor, Transience of learning.

1 Introduction

Long term retention of learning is of great interest to educational researchers. In the context of computer programming, long term retention has been studied for various educational interventions (e.g., [1,2]). We wanted to investigate long term retention of programming concepts due to the use of software tutors. Typically, long-term retention is measured using delayed post-tests (e.g., [3]) or retention tests (e.g., [4]). Instead, we tried to document long-term retention using data collected by a software tutor used repeatedly by students in introductory programming courses. The results of our study have implications for the usefulness of tutoring software used *in natura* for learning programming concepts.

In our study, we analyzed the data collected by a code-tracing tutor that was made available to students of introductory programming courses as after-class assignments. Students were allowed to use the tutor as often as they wished. A large number of students chose to use the tutor more than once, often, several weeks apart. The data of these repeat-users, especially as it pertains to whether they had retained until the second tutoring session what they had learned during the first session provided evidence of long-term retention of learning. We treated the concepts each student knew before the first session as control group and those that the student did not know as test group, resulting in a within-subjects evaluation.

2 Retention of Learning

For this study, we analyzed the data collected by selection tutor over five years. The tutor is part of a suite of code-tracing tutors called problets (problets.org) available online for general use. Selection tutor poses code-tracing problems, wherein, the student is presented a complete program and asked to identify its output. It covers 9 concepts in Java/C# and 12 concepts in C++. Built into the tutor are 10-15 parameterized problem templates per concept. The tutor dynamically generates a problem by randomly instantiating the parameters in a template such as variable names, literal constants, and identifiers. So, a student never sees the same problem twice. Because problems are randomly generated from parameterized problem templates, memorization of the correct answer from the first session can be ruled out as explanation for performance of a student during the next session. The tutor administered pre-test-practice-post-test protocol [5]:

Pretest – The tutor presented one problem per concept. If a student solved the problem correctly, no more problems were presented on the concept during practice or post-test. If the student skipped the problem or solved it partially or incorrectly, feedback was presented to the student and additional problems on the concept were scheduled for the subsequent practice stage. The feedback included line-by-line explanation of the correct solution to the problem.

Adaptive practice – Once a student had solved *all* the pretest problems, practice problems were presented on only the concepts on which the student had skipped the pretest problem or solved it partially/incorrectly. Problems were presented on each concept until the student had mastered the concept, i.e., solved a minimum percentage of problems correctly. Line-by-line explanation of the correct solution was provided after each problem.

Focused Post-test - During this stage, which was interleaved with practice, posttest problems were presented on only the concepts that the student had mastered during adaptive practice. If the student solved the post-test problem correctly, no more problems were presented on the concept. If the student solved the post-test problem incorrectly, the tutor repeated practice and post-test stages for the concept.

Pretest, practice and post-test were administered by the tutor back-to-back without interruptions. The entire protocol, administered over the web, was limited to 30 minutes.

Each concept covered by the tutor can be classified as known, tested, practiced or learned for each student, as shown in Table 1. Note that if the student solves the pretest problem incorrectly, line-by-line explanation of the correct solution is presented, which provides an opportunity for the student to learn the concept. The student may end up not solving any practice problems on a concept even after solving the pretest problem incorrectly because of the 30-minute limit on the duration of the tutoring session. After a typical tutoring session, a student may end up with a few concepts in each category: known, tested, practiced and learned.

Table 1. Types of Learning Experience with the Tutors.

Pretest	Practice	Posttest	Type of Learning
Correct			Known
Incorrect	None		Tested
Incorrect	Some		Practiced
Incorrect	Mastered	Incorrect	Practiced
Incorrect	Mastered	Correct	Learned

A student is said to **retain** a concept from one session to the next if the student solves the pretest problem on the concept correctly during the second session. Otherwise, the student is said to have **forgotten** the concept. The following are the eight possible retention behaviors of a student on a concept based on the student's learning experience during the first tutoring session and pretest performance on the second tutoring session:

- For a concept **known** during the first session, **known-retained** if the student solves the pretest problem correctly during the second session and **known-forgotten** otherwise. Known-retained concept is part of the prior knowledge of the student. Known-forgotten category is attributable to forgetting of learning over time, called transience of learning. Neither of these is affected by tutor use.
- For a concept **tested** during the first session, **tested-retained** if the student solves the pretest problem correctly during the second session, and **tested-forgotten** otherwise. The student could have learned the concept from the explanation provided after the pretest problem was solved incorrectly during the first session.
- Similarly, **practiced-retained** and **practiced-forgotten** for concepts **practiced** during the first session, and **learned-retained** and **learned-forgotten** for concepts **learned** during the first session. In these cases, explanations provided after pretest and practice problems provided opportunities for students to learn the concept.

Known-retained and known-forgotten concepts served as control group in a **within-subjects controlled study**, since neither is affected by tutor use. In contrast, tested-retained, practiced-retained and learned-retained all provide evidence in support of long-term retention of what was learned using the tutor, whereas tested-forgotten, practiced-forgotten and learned-forgotten all provide evidence refuting retention. These served as test group in the within-subjects controlled study.

3 The Study

The Tutor: Among the 12 concepts covered by the tutor, concepts 1, 2, 6 and 7 pertain to basic understanding of the behavior of selection statements. Concepts 3, 4 and

11 are about nesting, and are harder concepts. Concepts 5 and 12 are about multiple back-to-back selection statements in a program – tracing their execution is more labor-intensive. Concepts 8, 9 and 10 are “pathological uses” and apply only to C/C++.

The Subjects: The tutor was used by introductory programming students as after-class assignment. We used the data collected by the tutor over five years: Fall 2012 – Spring 2017. Students from 44 high schools, 8 community colleges and 51 universities used the tutor during this period. Since we were interested in retention of learning, we used data only from students who had used the tutor at least twice and gave us permission to use their data for research purposes.

In all, 1,243 students used the tutor more than once. Collectively, repeat users solved problems at least twice on 10,213 concepts. These 10,213 student concepts are categorized into the eight types of retention behavior in Table 2. Also shown in the table are the mean pretest score on the first and second pretests, and the mean time between the two sessions in hours. The score on each problem was normalized to the range 0 → 1.0, 1.0 corresponding to correct solution and 0 to incorrect solution.

Table 2. Selection tutor - Descriptive Statistics for the Types of Retention Behavior.

Retention Behavior Type	N	Pretest1	Pretest2	Time (hours)
Known-Retained	7286	1.0	1.0	507
Known-Forgotten	476	1.0	0.06	579
Tested-Retained	1203	0.03	1.0	85
Tested-Forgotten	359	0.05	0.06	128
Practiced-Retained	277	0.18	1.0	262
Practiced-Forgotten	109	0.11	0.20	341
Learned-Retained	433	0.08	1.0	548
Learned-Forgotten	70	0.07	0.06	483

Known-retained accounted for 71.34% of all the concepts. These are the concepts for which students did not need to use the tutor, although they used it more than once. **Known-forgotten** are concepts students knew before the first session, got no feedback on from the tutor, and forgot by the time of the second session. They represent transience of learning, the deterioration of learning over time. They constitute **6.13%** of all known concepts. We used this as the baseline against which to compare other ‘forgotten’ categories.

Tested-retained represented 77.01% of tested concepts and **practiced-retained** represented 71.76% of practiced concepts. Students retained 86.08% of the concepts they had learned during the first session (**learned-retained)**, and the duration over which they retained the concepts was 548 hours, or 22.8 days. So, *students retained over 71% of the concepts covered by selection tutor*. Conversely, **tested-forgotten** represented 22.98% of tested concepts, **practiced-forgotten** represented 28.23% of practiced concepts and **learned-forgotten** represented 13.91% of learned concepts.

All these percentages are greater than 6.13% attributable to transience of learning. So, there is room for improvement of the tutor to promote retention of learning.

Table 3 lists the distribution of selection concepts among known-retained, tested-retained, practiced-retained and learned-retained categories. In each category, the number of student concepts retained (N) and the retention percentage (R%) are listed, e.g., 786 students knew concept 1 and retained it through subsequent use of the tutor. They were 85.8% of all the students who knew concept 1 before using the tutor for the first time.

From Table 3, we can compute the transience of learning for each concept as the complement of its known-category retention percentage (R%), e.g., transience of learning of concept 1 is 1 – 0.858 = 0.142 (14.2%). From the table, it is clear that *the transience of learning varied by concept.* We also find that for all the concepts except concept 4, percentage retention of learned concepts (Learned R%) was within a few points of percentage retention of known concepts (Known R%). Concept 4 is nesting in both if-clause and else-clause of an `if-else` statement and is a harder concept for novices to trace, which may explain why its retention percentage of learned concepts was lower than that of known concepts. Finally, note in the table that retention percentages were lower for tested and practiced categories than for learned categories on most concepts. Since students in learned category solved more problems than those in tested and practiced categories, we can state that *retention percentage improved with practice*.

Table 3. Selection tutor - Distribution of Concepts among the Types of Retention

Retained concepts	**Known**		**Tested**		**Practiced**		**Learned**	
	N	**R%**	**N**	**R%**	**N**	**R%**	**N**	**R%**
Basic concepts								
1	786	85.8	625	72.7	22	68.8	73	80.2
2	955	95.7	154	82.4	27	73.0	60	95.2
6	824	97.2	56	93.3	5	100	10	100
7	785	98.1	61	88.4	15	83.3	10	90.9
Multiple statements								
5	584	91.0	71	67.0	109	70.8	71	89.9
12	622	96.0	23	63.9	28	68.3	76	89.4
Pathological use								
8	55	73.3	33	75.0	8	47.1	44	72.1
9	125	89.3	12	80.0	2	100	10	100
10	126	96.9	6	100	5	62.5	9	90.0
Nested statements								
3	861	91.7	82	91.1	20	76.9	40	88.9
4	824	93.4	62	91.2	21	72.4	21	72.4
11	739	99.3	18	85.7	15	88.2	9	100

Students who use a tutor repeatedly without being explicitly instructed to do so are typically self-motivated. This self-selection bias is a confounding factor of this study in terms of being able to generalize the results to the general population.

The mean duration between tutoring sessions in this study was several weeks. During this time, it is inconceivable that students did not have the opportunity to practice the tutored concepts through other activities such as reading a text book, writing a program or discussing with friends. Such activities would affect retention of the learned concepts. These extraneous influences are confounding factors in the evaluation of any educational intervention carried out in a real course with actual students, and are unavoidable.

The tutor provided error-flagging feedback, which gave students the opportunity to recover from slips. It used a reified interface for entering the solution, making it harder for students to guess the correct answer. So, we did not make allowances for slips and guesses when calculating the correctness of student solutions.

For analysis purposes, we treated the concepts covered by the tutor as being independent. But, on code-tracing exercises, the line-by-line explanation provided by the tutor could have facilitated transfer of learning from one concept to another, i.e., a student could have learned one concept from the line-by-line explanation provided for another concept. This may have added to the numbers of tested-retained and practiced-retained concepts, but did not affect our conclusions about long-term retention of learning. In the future, we plan to repeat this study with code-debugging tutors, wherein, such transfer of learning is less likely to occur since explanation provided for each bug is specific to the bug.

Acknowledgments: Partial support for this work was provided by the National Science Foundation under grant DUE-1432190.

References

1. Ya-Feia Yang, Chien-Ia Lee and Chih-Kaia Chang. (2016). Learning motivation and retention effects of pair programming in data structures courses. Education for Information. 3(3). 249-267
2. Federici, S., Medas, C. and Gola, E. (2018) Who Learns Better - Achieving Long-term Knowledge Retention by Programming-based Learning. In Proceedings of the 10th International Conference on Computer Supported Education (CSEDU 2018) - Volume 2, pages 124-133
3. Butcher, Kirsten R., and Vincent Aleven. "Diagram interaction during intelligent tutoring in geometry: Support for knowledge retention and deep understanding." Proceedings of the 30th annual conference of the cognitive science society. Austin, TX: Cognitive Science Society, 2008.
4. Jackson, G. Tanner, et al. "The impact of Why/AutoTutor on learning and retention of conceptual physics." International Conference on Intelligent Tutoring Systems. Springer, Berlin, Heidelberg, 2004.
5. Kumar, A.N. A Model for Deploying Software Tutors. IEEE 6th International Conference on Technology for Education (T4E). Amritapuri, India, 12/18-21/2014, 3-9.