

The feasibility of automatic assessment and feedback

Mikko-Jussi Laakso and Tapio Salakoski
University of Turku
Department of Information Technology

Ari Korhonen
Helsinki University of Technology
Department of Computer Science and Engineering

{milaak,tapio.salakoski}@it.utu.fi, archie@cs.hut.fi

Abstract

In this study, we report on the results of studies in which two randomized groups of students are monitored while they solved exercises in a Data Structures and Algorithms (DSA) course. The first group did the exercises on the web and the second one in the classroom sessions. A web based system was employed that was able to give feedback and automatically assess the exercises. The research question was to find out how we should introduce the self study material and automatically assessed exercises to the students in order to maximize their learning experience and to avoid drop outs. In addition, we surveyed the students' attitude towards www-based exercises by using questionnaires. The students were asked what kind of exercises they would prefer to do in DSA courses as well as how they would assess their own learning experience in the three different setups (human guided, web based or mixed). All these studies were carried out simultaneously in two different universities.

It is not surprising that the results suggest to introduce easy and human guided exercises at the very beginning of the course. However, we conclude that currently there is an emerging need for both web-based and classroom exercises. The recommended way to introduce the web-based exercises in DSA courses is by combining these two approaches. There is a set of exercises that are best suitable to be solved and automatically assessed on the web while the rest of the exercises are best suitable for traditional classroom sessions. We believe that the results of this study can be generalized to cover also other similar learning environments than that used in this research to give automated feedback for the students, and thus improve the learning experience.

Keywords: automatic assessment, feedback, computer science education, evaluation

1 Introduction

Computers are now being used to ease the learning process in a variety of disciplines, including computer science and engineering. There are several methods to apply. One such method is animation of the behavior of complex systems and the other is simulation exercises in which the learner reproduces the animation trace.

Today, the primary use of algorithm animation has been for teaching and instruction. The instructor executes the animation and the learner has a quite passive role in following the representation. From the pedagogical point of view, however, we believe that this is not enough. A large number of systems and reports written on this [1, 2, 3, 4, 5, 10, 11, 12] indicates that we should aim at *activating and engaging the learner* in order to promote the learning process. This, however, requires *feedback on students' performance*. The problem is how to provide feasible feedback in large courses that typically have limited resources.

Fortunately, the formal nature of algorithms and data structures allows us build learning environments in which we can compare the student's solution to the correct model solution easily. This gives an opportunity to produce systems that not only portray a variety of algorithms and data structures, but also distribute

tracing exercises to the student and then evaluate the student's answer to the exercises. This is called *automatic assessment and feedback of simulation exercises*. One such a system is TRAKLA2 that we have employed with good results during the past few years [9].

Our previous research show that there seem to be no significant differences between learner groups that solve the *same* exercises on the web and in the class room [7]. However, not all the exercises are such that they can be automatically assessed. Thus, the question is how far we can count on the automatic assessment today, and what would be the most feasible way to establish such a course organization that can best utilize the current learning environments in practise.

In this paper, we report on the preliminary results of the intervention study carried out in two different universities in Finland during the spring term in 2005. In both universities, the TRAKLA2 learning environment was employed in the data structures and algorithms courses. The students ($N = 133 + 134$) were divided into two randomized exercise groups in both universities. The first group started their exercises on the web with the TRAKLA2 learning environment while the second group did their exercises in the class room sessions. The research aimed at repetition of the intervention study carried out at the Helsinki University of Technology (HUT) in 2001 [7]. However, there were some differences as well. This time, we wanted to provide the same learning experience for all the students in order to prevent the drop out consequent upon the research setup. Thus, at the midpoint of the course, the groups transposed their places. The first group continued in the class room and the second group on the web. Moreover, we also report on the repetition of the attitude survey carried out at the University of Turku (UTU) in 2004 [9]. This time, the same survey was carried out in both of the aforementioned universities.

In Section 2, we describe the TRAKLA2 learning environment as well as the course syllabus for both of the university courses. In Sections 3 and 4, we report the results of the intervention study, and the attitude survey carried out, respectively. Finally, in Section 5 we make the conclusions based on the results achieved.

2 Exercises

2.1 TRAKLA2

TRAKLA2 is a learning environment providing algorithm simulation exercises in which students simulate the working of algorithms covered in the data structures and algorithms course. In algorithm simulation, the student must show in terms of conceptual diagrams how an algorithm changes given initial data structures during its execution.

The exercise could be, for example, *“Insert the following keys in this order into an initially empty binary search tree: O, H, T, F, R, K, M, A, U, B, Q, J, Y, C, S, W, F, and S. Draw the tree after each insertion.* The exercises are solved with a Java applet, i.e., students change the contents (keys and references) of data structures visualized on the screen by performing drag-and-drop operations supported by the TRAKLA2 environment [8]. An example exercise is seen in Figure 1. The student should drag and drop the keys from the array into the appropriate positions in the binary search tree. In the example, the first 7 keys are already inserted into the tree. The model answer in the front window shows the next state where also the key A is inserted.

Each student is given a unique initial data structure to work with like the stream of keys in the previous example. The sequence of performed operations are recorded and the submitted answer is assessed automatically by comparing it to the model sequence generated by a real implemented algorithm. The student receives immediate feedback on his/her solution that indicates the number of correct steps out of the maximum number of steps. In addition, the model solution that is represented as an algorithm animation can be reviewed at any time.

If the student did not get the exercise correct, the exercise can be solved and resubmitted again. However, each time the exercise is reseted, a new instance of it with new initial data set is created. In general, the number of resubmissions can be unlimited as one cannot continue with the same data set. Moreover, the solution space of the exercises is far too large to allow brute force solution with simple trial and error method.

Currently, about 30 difference exercises are included in the TRAKLA2 environment. These cover basic data structures, priority queues (heap), sorting algorithms, hashing, dictionaries and graph algorithms. We emphasize here that students do not need to code anything while solving the simulation exercises. They

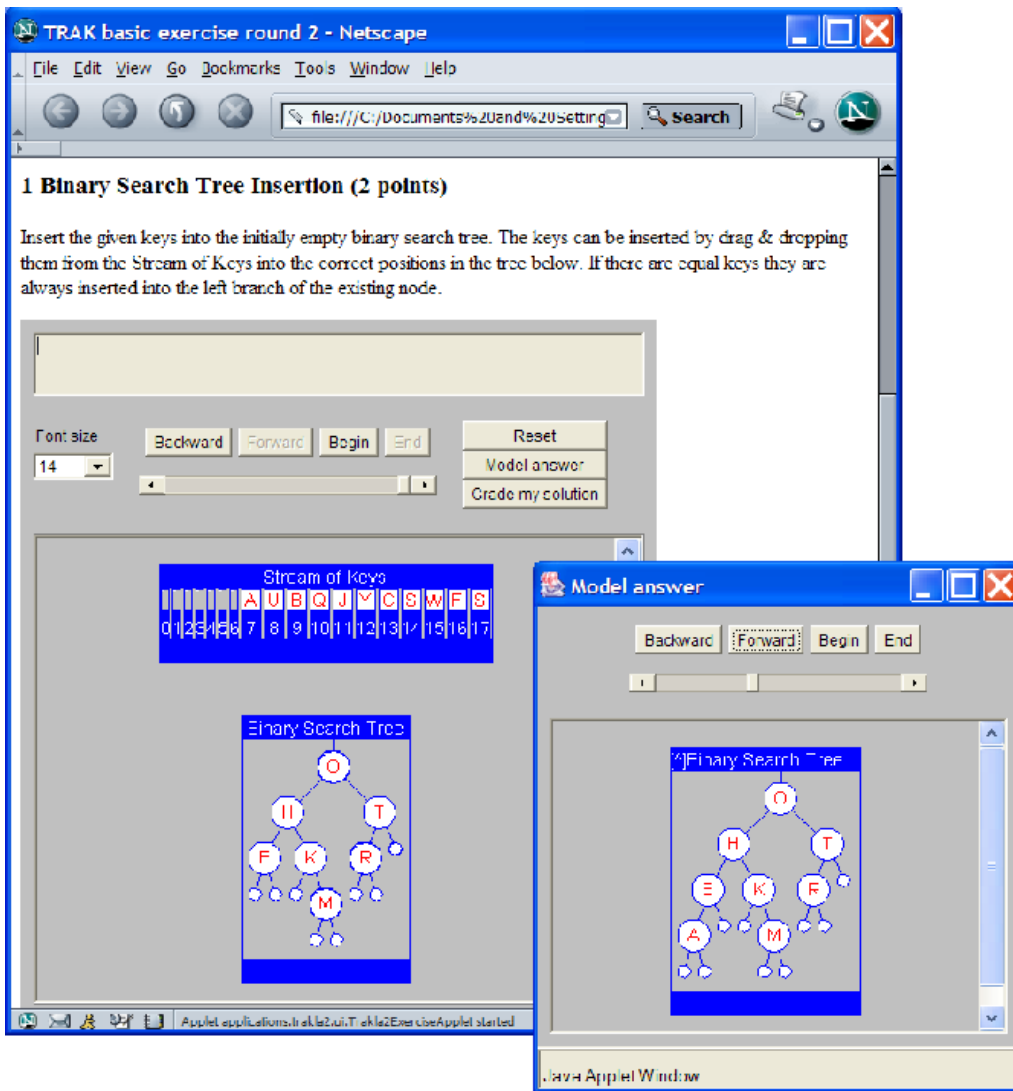


Figure 1: TRAKLA2 applet. The exercise window comprises the data structures and push buttons. The model solution window is open in the front.

operate purely at a conceptual level. However, to be able to solve these exercises, they have to understand the key principles of the corresponding algorithms.

In the learning environment, the exercises are divided into exercise rounds as it is typically the case in the classroom exercises as well. Each exercise round covers one topic in the course. Moreover, the exercise rounds have deadlines as in traditional classroom exercises that are returned at a certain time in a certain place.

2.2 Course at UTU

In spring 2004, the DSA-UTU course (UTU04) included 56 lecture hours, 10 classroom sessions (each 2 hours), and 22 TRAKLA2 exercises. The classroom sessions consist of four to six single exercises, such as illustrating exercises, proof exercises, etc.

The previous study [9] reports very positive results on utilizing TRAKLA2 system in DSA-UTU course. Due to this, it was decided to use TRAKLA2 more extensively in DSA-UTU course in spring 2005 (UTU05). Altogether, UTU05 course included 56 lecture hours, 10 classroom sessions (each 2 hours), and 14 or 15 TRAKLA2 exercises depending on the group. The TRAKLA2 exercises replaced 4 class-

room sessions (out of the 10), thus the number of classroom exercises decreased from 50 to 30 and were replaced by TRAKLA2 exercises. However, TRAKLA2 exercises did not cover all the areas of the replaced classroom exercises such as algorithm analysis, design exercises, proofs, and so on.

There were 2 classroom sessions for both groups at the beginning of the UTU05 course. After that the student were randomly divided into two groups A and B. Group A did 14 TRAKLA2 exercises in three rounds on the web, and at the same time, Group B did 4 classroom sessions (20 single exercises) covering the topics in the first half. The roles of the groups were changed in the middle of the course after which Group A did 4 classroom sessions, and Group B did 15 TRAKLA2 exercises in two rounds covering the rest of the topics. However, also in this case, the classroom exercises included simulation exercises similar to the TRAKLA2 exercises as well as more challenging exercises such as analysis, design, and proof exercises.

The nature of the classroom sessions is such that every student has to do the given exercises beforehand and at the beginning of the classroom session every student informs the demonstrator, which single exercises he has done. Then every single exercise is presented by a selected student in the blackboard. After that the presented solution is analysed by the demonstrator which gives feedback on students' performance. Furthermore, the demonstrator presents the correct solutions as well as clarifies the most important aspects of the exercises.

2.2.1 Grading and requirements at UTU

It was possible to get 40 course points from the examination(s). In addition, any student can get additional 6 course points based on their activity in TRAKLA2 exercises and classroom exercises. These points were summed up to get the total number of course points.

The 6 extra course points were determined by the TRAKLA2 exercise points (maximum 35) and classroom exercise points (maximum 35) that were summed up. This value is the student's total exercise points (TEP), and there were the maximum of 70 TEP in share. The conversion of TEP to the course points was linear between the minimum requirements 40% (pass with zero course points) and 100% points (6 course points, i.e. 15% of the maximum). The minimum points were required in order to get the examination.

There were two ways of passing the course: taking the final examination (0–40 course points) or taking two midterm-examinations (both 0–20 course points). In either case, the students still had to fulfill the minimum requirements: i) to get at least 40% of the classroom exercise points (maximum 35), ii) to get at least 40% of the TRAKLA2 points (maximum 35 points), and iii) to get at least 20 course points out of the total of 46 in share. The first midterm-examination was arranged in the middle of the UTU05 course and the second one at the end of the course.

The final grading was on a scale from one to three with 0.25 steps. By getting 20 course points the student got the lowest grade (one).

2.3 Course at HUT

The course syllabus at HUT was similar to that at UTU, but there are many differences as well. At HUT, the students have passed only one programming course before attending data structures and algorithm course. At UTU, however, they have a little bit more preliminary courses that aid taking the course.

HUT05 course included 40 lecture hours and 7 rounds of exercises. Some 140 students were enrolled in the course including a small number of foreign students that are excluded from the research setup as they could not follow the lectures that were given in Finnish only. Each student selected an exercise group from the 5 possible alternatives. The first exercise session was common for all of the students, but after that each group was randomly split into two subgroups A and B (based on the last digit of the study book number). Group A did the exercise rounds 2–4 on the web with TRAKLA2 and 5–7 in the classroom, and group B vice versa. Each exercise round comprises some 4–6 different simulation exercises. The exercises were the same in both groups, i.e., also the classroom exercise group did simulation exercises and not, for example, analysis or design exercises as they did in UTU.

2.3.1 Grading and requirements at HUT

The midterm examinations and the compulsory exercises were assessed separately. Both had the 50% influence to the final grade. The conversion of points received from the exercises was linear between the minimum requirements 50% (pass with the grade 1) and 90%-100% points (pass with the maximum

grade 5). However, the students did not have to fulfill the minimum requirements before attending the examinations as it was the case in UTU. They had the chance to complete the exercises also after the examination in HUT.

2.4 Drop out and performance evaluation

In this study, we have monitored two randomized groups of students (A and B) while they solved exercises in course of data structures and algorithms. In group A, the solving procedure started with web based exercises while in B they practised in a classroom. In the middle of the course, the procedure was changes: group A continued in the classroom while the group B went on the web.

The same evaluation was carried out both in UTU and HUT during the spring term 2005. Both courses started with a pretest in which questions on the basic data structures and algorithms were asked. There is no differences between the groups A and B in the learning results of this test. In addition, some of the questions were repeated in the first midterm examination in HUT. For example, a question concerning the visiting order of nodes in tree traversal algorithms shows that only 26% of students was able to connect the algorithms (pre-, post-, inorder as well as level order) to the corresponding list of nodes with the given binary tree representation. However, in the first midterm examination this number was over 90%.

We conclude that actual learning has taken place in this setup. Thus, the research question concerns more with the other quality aspects of the learning besides the amount of learning: the overall throughput of the course (i.e., passed students) as well as the fine tuning of the course difficulty. We try to gather evidence to adjust the course difficulty at the level that is the most feasible according to the use of web based exercises and overall learning results. Moreover, we study the question how should we introduce the self study material and exercises to the students in order to avoid large drop out in courses.

2.5 Attitude study

Students' opinions about the TRAKLA2 system were gathered through a web-based survey at the end of the HUT course in 2003. 364 students answered. 80% of them gave an overall grade of 4 or 5 to the system on the scale 0.5, where 5 was the best grade. The system was almost unanimously considered to aid learning and to be easy to use. In addition, a survey was arranged at UTU in spring 2004. This survey included three sets of questionnaires about students' attitudes and opinions about www-based systems, one at the beginning, one at the middle, and one at the end of UTU04 course.

The purpose of the survey was to gather information about students' attitudes towards and experiences and changes in those during the UTU04 course. The first and the last questionnaire also included a question about what kind of exercises the students would prefer to do in DSA-UTU courses as well as a question on how they assessed their own learning. There where also possibility to give free comments about the TRAKLA2 system. Indeed, the free feedback from this survey was well in line with the observations from the HUT survey.

The results from study in 2004 [9] shows that the TRAKLA2 system was well approved by students and it can be said that the students' attitudes became more positive towards www-based systems. Moreover, the results indicated that web-based exercises constitute a good amendment to courses on DSA and suggest that the mixed-alternative is far the most appropriate way to learn topics of courses on DSA.

As a continuation on this survey we want to confirm these observed results and it was decided to arrange a repetition study about the attitudes and opinions of the UTU and HUT students using questionnaires in spring 2005. In UTU05 and HUT05 course, a survey was arranged with two sets of questionnaires, one at the beginning and one at the end of each course. Both questionnaires included the very same questions than in the UTU04 survey. In both questionnaires, the students were asked to state how well they thought web-based exercises could suite to the course, and how much students believe that www-based learning could help them to understand DSA topics better. In addition, in the second questionnaire, there was a question in which students was asked to give their preferration about the role of the TRAKLA2 exercises in DSA courses.

3 Results of performance evaluation

In the following, we focus on the overall throughput of the passed students in the two courses involved (UTU and HUT). The null hypothesis is that there is no significant differences between the groups A and

Table 1: Course statistics of students taking the course exercises (CE) and midterm examinations (MTE). In addition, there is the number of students that passed/took the MTE. Finally, the throughput indicates the percentage of students that passed both CE and MTE.

	UTU A	UTU B	UTU Total	HUT A	HUT B	HUT Total
Nr of students (#)	67	66	133	72	62	134
# passed CE	48 (72%)	39 (59%)	87 (65%)	52 (72%)	48 (77%)	100 (75%)
# took MTE	42 (63%)	38 (58%)	80 (60%)	52 (72%)	49 (79%)	101 (75%)
# passed MTE	29 (69%)	26 (68%)	55 (69%)	42 (81%)	42 (86%)	84 (83%)
Throughput	43%	39%	41%	58%	65%	61%

B. In addition, we have assumed that the learning results must be the same in both of the groups. This is due to the fact that in HUT, the students did the very same exercises regardless of the solving procedure. Even though in UTU the class room exercises were more challenging, the students had the chance to do both kind of exercises, thus giving them almost an equal learning experience. Performed tests supported this assumption. In UTU, for example, the average points received from the first midterm examination were 9.81 (group A) and 10.36 (group B). The same figures from the second midterm examination were 13.22 and 12.04, respectively. The midterm examinations revealed no statistical differences (t-tests, $p_1 = 0,58$, $p_2 = 0,33$), respectively.

Table 1 summarizes the basic data from both of the two courses. The total number of enrolled students in both courses (UTU 133, and HUT 134) is almost equal. In addition, there is no significant differences among the passed students of TRAKLA2 exercises except in UTU-B (t-test, $p = 0.02$). Thus, there is some evidence to reject the hypothesis that the level of difficulty in exercises makes no difference. Actually, it is quite natural that students (in group B) starting with more challenging exercises drop the course early more easily. However, it is interesting to note that there is no similar effect in HUT (t-test, $p = 0,80$). Contrary to this, the relative number of passed students is higher in group B. More students drop the course in group A, i.e., this confirms our previous results that in self study, there is a slight tendency to drop the course more easily than if the students were allowed to attend class room sessions (if the exercises are the same) [7].

In UTU, the students are not allowed to take the 2nd midterm examination until they have passed all the exercises. This is not the case in HUT, however. Actually, there is 9 students more taking the midterm examinations than passing the exercises in HUT (the number is not visible in Table 1). It should be noted, however, that in both courses there is an option to complete the performance in exercises by solving more exercises. Thus, the surplus students in examination are possibly planning to do the exercises after the examination. This seems to be a bad strategy seeing that only 2 out of the 9 students passed the midterm examinations.

The midterm examinations were more difficult in UTU than in HUT. Only 69% passed in UTU while 83% passed in HUT. However, there is no differences between the groups A and B in either of the courses. In addition, it should be noted that midterm examinations was the first possibility to pass the course. There are a couple of more upcoming examinations in both of the courses later this year. Thus, for example in HUT, the overall throughput will be around 80%.

4 Results of attitude study

In the following, we present the detailed analysis of the attitude survey results in HUT and UTU in spring 2005. In addition, the opinion on learning results based on students' self evaluation are also presented.

At UTU, there were 108 answers to the first questionnaire ('Start') and 68 to the second questionnaire ('End'). At HUT, the same numbers were 129 and 89, respectively.

At the Start and End, the students were asked about their opinion on the suitability of www-based exercises for learning DSA. The scale was from 1 to 7 (7 being the best). The Start averages were quite high, 5.44 (UTU) and 5.54 (HUT), and the End averages were even higher, 5.59 (UTU) and 6.15 (HUT). The same phenomenon was observed in UTU04 course [9]. Furthermore, the students were asked about how much students believe that www-based learning could help them to understand DSA topics easier (same scale as above). The Start averages were also quite high, 4.80 (UTU) and 4.89 (HUT), and the End averages

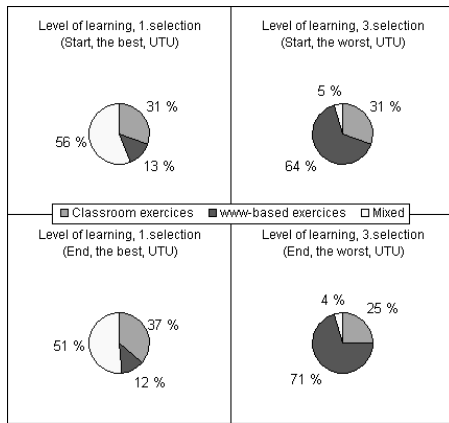


Figure 2: The first and last selection in the level of learning (UTU05). The upper pie charts illustrate the attitude at the beginning of the course and the lower pie charts at the end.

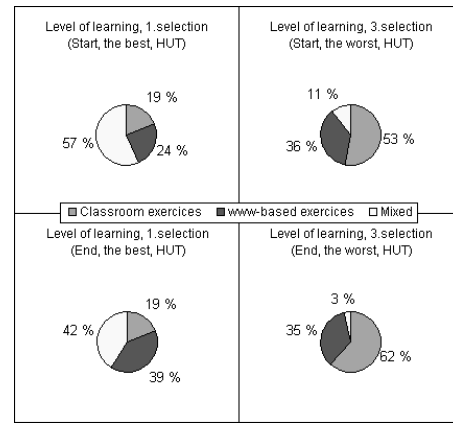


Figure 3: The first and last selection in the level of learning (HUT05). The upper pie charts illustrate the attitude at the beginning of the course and the lower pie charts at the end.

also higher, 4.97 (UTU) and 5.49 (HUT).

These observations clearly indicate that web-based exercises are perceived to be suitable for learning DSA course topics and those exercises are well accepted and approved by students both at UTU and HUT. Moreover, it can be also said that www-based exercises aids students' learning process of DSA course's topics.

In addition, the students were asked to assess the level of their learning on three alternative ways of doing exercises: traditional classroom exercises, web-based exercises or a mix of these two approaches (see Figure 2 and Figure 3).

In Figures 2 and 3 we can see that at the Start, there is approximately the same amount of students in UTU and HUT courses that choose the mixed alternative on the 1. selection when looking the most appropriate way of learning. Mixed alternative was considered the most appropriate way to learn DSA course's topics and the least appropriate was doing only www-based exercises. However, there was a quite big difference between UTU and HUT on the 3. selection, which indicates that HUT students are more open minded towards www-based exercises or learning. The natural reason for this is the fact that there have been much more usage of www-based exercises at HUT (since 1993) than at UTU (since 2004) in the previous DSA courses. Moreover, there has been only classroom exercises in DSA-UTU courses before UTU04 course that first time introduced TRAKLA2 system.

At the End, changes were quite different at UTU than at HUT. At UTU, mixed alternative loosed some ground to classroom exercises while www-exercises stood at the same level. And the same statistic from UTU04 course (see Figure 4) shows that mixed alternative actually gained a lot of ground (from 48% to 73%). We suggest that this different behaviour can be explained by the fact that in UTU04, the www-based exercises replaced only such exercises that are much more elegantly done in TRAKLA2 system. For example, all the tracing or illustrative type of exercises (i.e. how a specific algorithm works on given input). In UTU05 course, however, TRAKLA2 exercises did not cover all the areas of replaced classroom exercises. Therefore, www-based exercises gave a little bit narrower conception of the topics compared with the classroom exercises that included also analyzing and proof exercises as well.

At HUT, the share of the www-based exercises grew from 24% to 39% (1. selection) almost catching up the mixed alternative. The growth is explained by the nature of the classroom exercises at HUT. Students did exactly the same exercises both in TRAKLA2 system and in the classroom with pen and paper. However, the most feasible way to do illustrative and tracing type of exercises is to do them in TRAKLA2 system. This same effect is observed also in [9] and explains why www-based exercises growth its share.

In the same manner as the level of the learning was studied, the students were asked to give their preference on the three alternative ways of doing exercises. The preferred alternative at Start was towards www-based exercises (1. selection: 44% (UTU) and 66% (HUT)). It turns out to be even stronger biased towards www-based exercises at the End (1. selection: 51% (UTU) and 75% (HUT)). In the UTU04 course,

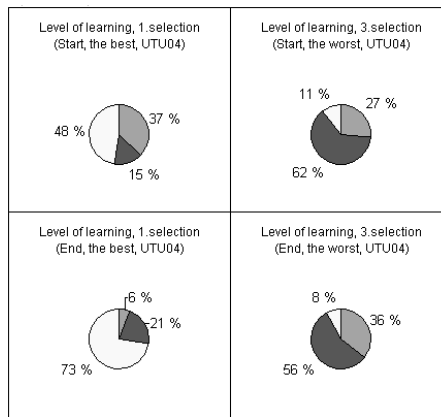


Figure 4: The first and last selection in the level of learning (UTU04). The upper pie charts illustrate the attitude at the beginning of the course and the lower pie charts at the end.

the change was about the same order from 38% to 44%. We can argue that the www-based exercises are the most preferred alternative to do the exercises according to the students.

Finally, at the End, the students were asked about the role of TRAKLA2 exercises. At UTU, 85% of students responded that TRAKLA2 exercises should be compulsory part of DSA-UTU course. The same percentage at HUT was 96%. In addition, at UTU, 73% of students responded that TRAKLA2 exercises should be compulsory extra exercises or compulsory exercises to replace some of the classroom exercises. The same percentage at HUT was 55%. Still, 40% at HUT responded that they should be compulsory exercises to replace all of the class room exercises. This observation is in line with other observation in HUT05. Only 15% (UTU) and 4% (HUT) of students responded that TRAKLA2 exercises should be voluntary extra exercises (besides class room exercises). This concludes that there is a need for both TRAKLA2 exercises and classroom exercises in DSA course. Moreover, while the simulation exercises are well suitable to be performed in the TRAKLA2 system, there are still many exercise types (i.e. analysis and proof exercises) that are better suited to the classroom environment, where human interaction can occur.

5 Conclusions

In this study, we have monitored two randomized groups of students (A and B) in two different universities while they solved exercises in course on data structures and algorithms. The first group did the exercises on the web and the second one in the class room sessions in both universities. In addition, we have surveyed their attitude towards www-based exercises. The students were asked whether they prefer to do the exercises in the class room or on the web. The third alternative was mixed of these two.

There is no difference between groups A and B when looking at the overall throughput, i.e., passed students in the courses. Thus, the final examination (two midterm examinations in this case) seems to be equally difficult for both groups. We conclude that the procedure of doing exercises, i.e., web or class room does not have any influence on the learning results. However, if we look at the number of students passed the exercises, there is a difference at UTU. This is due to the fact, that the students starting with more challenging class room exercises drop the course more often in the beginning of the course compared with those starting with the web based exercises. This goes against the previous results indicating that self learning causes the students to drop more easily. Thus, the difficulty level and nature of the exercises have more influence in this respect.

Based on the results from the prior attitude study in 2004 [9] together with this years study, it can be said that the mixed alternative is the recommended way to introduce the web based exercises. The most preferred way to do simulation exercises is on the web, which indicates that the TRAKLA2 system is very easy to use and suited for its purpose. In addition, we argue that the role of www-based exercises should be compulsory and the exercises should replace only those classroom exercises that are suitable for TRAKLA2.

We believe that the results of these studies can be generalized to cover also other similar learning environments that can give automated feedback for the students. It seems that automatic feedback can be

adequate enough to compensate its drawbacks compared with human guidance due to the fact that it is available all the time during the exercise session, thus allowing the students to study at their own pace. Moreover, it is not surprising that the results suggest us to introduce easy and human guided exercises first. After this, the students are more engaged into the course, thus they pass the whole course more probably, and can be directed to self learning environments such as TRAKLA2.

The results also imply that the learning in such a learning environment is as good as in traditional sessions if the exercises are the same. Thus, keeping in mind the limits of automatic assesment, we can recommend the use of the web based learning environments. In our case, this means that we are going to deliver all the simulation exercises through the web based learning environment, but parallel to this, we have traditional classroom sessions as well to cover the design and analysis exercises that are beyond the scope of automatic assessment.

We have concluded that currently there is a need for both TRAKLA2 and traditional classroom exercises. There is a set of exercises that are best suitable to be automatically assessed by TRAKLA2. These include not only *tracing exercises*, but also *exploration exercises* [6] such as coloring a binary tree to a red black tree. Thus, an important challenge of the future is to develop novel types of TRAKLA2 exercises to cover more of the tradional exercises. Our future plans are to do this in collaboration between Helsinki University of Technology and University of Turku.

References

- [1] S. Bridgeman, M. T. Goodrich, S. G. Kobourov, and R. Tamassia. PILOT: An interactive tool for learning and grading. In *Proceedings of the 31st SIGCSE Technical Symposium on Computer Science Education*, pages 139–143. ACM Press, New York, 2000.
- [2] M. H. Brown and R. Raisamo. JCAT: Collaborative active textbooks using Java. *Computer Networks and ISDN Systems*, 29(14):1577–1586, 1997.
- [3] J. Carter, J. English, K. Ala-Mutka, M. Dick, W. Fone, U. Fuller, and J. Sheard. ITICSE working group report: How shall we assess this? *SIGCSE Bulletin*, 35(4):107–123, 2003.
- [4] S. R. Hansen, N. H. Narayanan, and D. Schrimpsheer. Helping learners visualize and comprehend algorithms. *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning*, 2(1), May 2000.
- [5] D. J. Jarc, M. B. Feldman, and R. S. Heller. Assessing the benefits of interactive prediction using web-based algorithm animation courseware. In *The proceedings of the thirty-first SIGCSE technical symposium on Computer science education*, pages 377–381, Austin, Texas, 2000. ACM Press, New York.
- [6] A. Korhonen and L. Malmi. Taxonomy of visual algorithm simulation exercises. In A. Korhonen, editor, *Proceedings of the Third Program Visualization Workshop*, pages 118–125, Warwick, UK, July 2004.
- [7] A. Korhonen, L. Malmi, P. Myllyselkä, and P. Scheinin. Does it make a difference if students exercise on the web or in the classroom? In *Proceedings of The 7th Annual SIGCSE/SIGCUE Conference on Innovation and Technology in Computer Science Education, ITiCSE'02*, pages 121–124, Aarhus, Denmark, 2002. ACM Press, New York.
- [8] A. Korhonen, L. Malmi, P. Silvasti, J. Nikander, P. Tenhunen, P. Mård, H. Salonen, and V. Karavirta. TRAKLA2. Computer program, 2003.
- [9] M.-J. Laakso, T. Salakoski, L. Grandell, X. Qiu, A. Korhonen, and L. Malmi. Multi-perspective study of novice learners adopting the visual algorithm simulation exercise system TRAKLA2. *Informatics in Education*, 4(1):49–68, 2005.
- [10] T. L. Naps, J. R. Eagan, and L. L. Norton. JHAVÉ: An environment to actively engage students in web-based algorithm visualizations. In *Proceedings of the SIGCSE Session*, pages 109–113, Austin, Texas, Mar. 2000. ACM Press, New York.

- [11] T. L. Naps, G. Rößling, V. Almstrum, W. Dann, R. Fleischer, C. Hundhausen, A. Korhonen, L. Malmi, M. McNally, S. Rodgers, and J. Ángel Velázquez-Iturbide. Exploring the role of visualization and engagement in computer science education. *SIGCSE Bulletin*, 35(2):131–152, June 2003.
- [12] R. J. Ross and M. T. Grinder. Hypertextbooks: Animated, active learning, comprehensive teaching and learning resource for the web. In S. Diehl, editor, *Software Visualization: International Seminar*, pages 269–283, Dagstuhl, Germany, 2002. Springer.