E-Assessment of Mathematical Proofs

Chances and Challenges for Students and Tutors

Susanne Gruttmann
European Research Center for Information Systems
WWU Münster, Germany
gruttmann@ercis.de

Dominik Böhm
Department of Mathematics and Computer Science
WWU Münster, Germany
dominik@dbedv.de

Herbert Kuchen
European Research Center for Information Systems
WWU Münster, Germany
kuchen@ercis.de

Abstract—**Conventional e-assessment systems often are not suited for examining mathematical calculations or proofs, and the few existing ones have very limited functionality. On this account we have developed the e-assessment system EASy which focuses on mathematical proofs. As we are not aware of any documented experiences concerning the applicability, usability and acceptance of a corresponding system for students and tutors, we have evaluated EASy in a lecture on data structures and algorithms with more than 200 students.**

E-assessment; mathematical proofs; mass lectures; evaluation of e-learning

I. INTRODUCTION

Assessments help to identify and to evaluate teaching success. But due to mass lectures, decreasing resources and low personnel capacities, continuous and frequent assessments of learning progresses may cause excessive efforts and costs. For this reason, e-assessment systems have become of increasing interest in academic e-learning infrastructures [1].

Today's e-assessment systems meet a variety of popular types of items. Typically, they focus on multiple choice, inserting short text, and other simple forms of examination. But most systems are not suited for examining creative tasks and particularly they are not able to assess mathematical calculations or proofs. Thus, we have developed an e-assessment system, called EASy, which focuses on mathematical proofs. Moreover, we have evaluated this system in a lecture on data structures and algorithms with more than 200 students, to gain information on efficiency, effectiveness, usability and acceptance of e-assessment in mathematics.

The rest of this paper is structured as follows. In Section 2, we discuss requirements of an e-assessment system for proofs and mention related work. In Section 3, we briefly describe the design, functionality and handling of EASy. The results of an empirical evaluation of EASy concerning the two main user groups, students and tutors, will be presented in Section 4. In Section 5, we conclude and point out future work.

II. E-ASSESSMENT IN MATHEMATICS

There are several forms of assessments in higher education: self-assessment by the student, periodic exercises (e.g. accompanying a lecture), and formal examinations (e.g. at the end of a course). Since we want to gain experience with a prototypical e-assessment system, we will focus on the use in weekly exercises. This allows us to concentrate on the core mathematics and ignore e.g. organizational, security, and legal issues which would have to be taken into account when considering examinations [2].

A. Requirements

E-assessment systems that are supposed to support exercises consisting of the formulation of mathematical theorems, their proof and revision have to fulfill a wide range of specific demands.

There is a large number of different proof strategies in mathematics such as e.g. different induction schemes, case distinction, estimations, and transformations. Additionally, there is a considerable quantity of theorems and rules. As it is impossible and not reasonable to provide all existing rules and strategies, an e-assessment system has to provide a selection that is relevant for the considered course or exercise. Moreover, an adequate system has to support step-by-step proofs to ensure that the students' mathematical skills are measured in deep. Each step corresponds to the application of a predefined rule or theorem at a certain position of the currently considered proof.

There are different possibilities to cope with the requirement of mathematical correctness in an e-assessment system. If the students are free to apply any rule this includes allowing an erroneous application. Another option is to disable the use of a rule at an inappropriate position. So, students cannot generate wrong proofs or proof wrong theorems. The only way to terminate a proof unsuccessfully is that the student gets stuck and does not know how to continue. Especially for an introductory course, we prefer the second option, since it teaches the student to apply rules correctly.

The complexity of producing a proof should not be increased due to the application of a tool. Hence, an e-assessment system should provide a good usability and support an intuitive handling of strategies and rules. The acceptance of an e-assessment system by students and teachers enormously depends on the ease of use and the required time and costs. If a system enhances complexity, users will stick to conventional assessment.

B. Related work

Cognitive skills and application of methods cannot be assessed via multiple-choice tests and equivalent forms of basic assessment items [3]. This means the majority of existing e-assessment systems is inappropriate for use in mathematics.

There exist few systems exclusively designed for handling assignments in mathematics. For instance, the AiM system (Assessment in Mathematics) uses the computer-algebra system Maple for specification of assignments and verification of the user's solutions [4]. Thus it provides support for numeric and symbolic calculations, but it does not support mathematical proofs. The system Euclid Avenue allows executing proofs in propositional calculus [5]. The user transforms an expression by manually inserted rules and Euclid Avenue checks the submitted solution for correctness. As the tool is restricted to propositional calculus, Euclid Avenue is not applicable for other sorts of proofs.

To sum up, we are not aware of any system supporting e-assessment of general mathematical proofs in a convincing way. Accordingly, there are not any documented evaluations on the use of e-assessment for mathematical proofs.

C. The e-assessment system EASy

Precisely because there are very few systems for assessing mathematical purposes with a very limited, unsatisfactory functionality, we decided to develop the e-assessment system EASy, which focuses on the processing of mathematical proofs [6].

For the purpose of a first evaluation of convenience and acceptance of e-assessment in math-related subjects the EASy prototype does not need to support the entire range of administrative functionality. Aspects like database connection, special authoring tools or automatic feedback are neglected.

We made the key functionality available via a Java applet integrated in a static web page [7]. The front end provides the graphical user interface for presentation, processing and review of proofs. EASy's interface is divided into several areas. The central area contains the theorem which has to be proven and the current status of the proof. On the left hand side, the student can perform administrative steps such as storing a proof or undoing the last step. Moreover, the strategy which shall be used next can be selected. On the right hand side, the student can select the next rule and the position in the last formula of the proof, where this rule should be applied. A video demonstrating the use of EASy can be found on our webpage [7]. We strongly recommend watching it.

The backend provides the mathematical core which is based on conditional term rewriting [8]. Each theorem induces a set of rewrite rules to transform terms. Such a rewrite rule consists of a number of preconditions and a conclusion. The rules engine ensures that the current proof context satisfies the theorem's preconditions. If the engine does not succeed, it cancels the application of the rule. This mechanism guarantees the mathematical correctness of EASy.

III. EMPIRICAL EVALUATION OF EASY

As there is poor information about the usability, acceptance, efficiency and effectiveness of e-assessment in a math-related subject an evaluation of EASy was considered reasonable to improve the system. Based on a formative evaluation strengths and weaknesses of the system should be identified in an early stage of its application [9]. The first experience with the EASy prototype has been made in the bachelor course "Data structures and Algorithms" at the University of Muenster which was attended by about 200 students in summer term 2008 [7]. To ensure that EASy meets the demands of all main user groups, we evaluated both, tutors and students.

A. Evaluation of the students' perspective

In order to get an impression about the applicability and benefits of EASy in academic scenarios, in a previous work, we have focused on an evaluation of the students' perspective [6].

Several exercises of the course "Data Structures and Algorithms" have been formulated in EASy such as:

- $\sum_{i=0}^{n} i = \frac{n \cdot (n+1)}{2}$
- Quick sort $\mathrm{O}(n \cdot \log n)$ needs steps in best case.

Afterwards they were asked to fill in a questionnaire to communicate their experiences with EASy. The participating students are predominantly bachelor students in computer science (24%) and information systems (49%) with good computer skills and little experiences with e-assessments. Although the students were free to use EASy or to prove the theorems by hand, most of the survey participants decided to use EASy (nearly 80% in exercise 1).

EASy helps to identify opportunities to solve an exercise by providing proof strategies and applicable rules for each exercise. In addition, EASy supervises the correct use of rules. Especially students with less developed mathematical knowledge appreciated these features. Many students are interested in the EASy concept and think that it can support to learn elementary proving skills. Some of them even asked for a possibility to generate new exercises on their own to use EASy for training purposes.

On the other hand, the students pronounced they needed more time for a proof than by hand. This is partly caused by the fact that they initially were not familiar with EASy as e.g. they had problems to find certain rules in the rule base. Another reason for needing more time is that EASy requires small steps to be executed one by one, which normally could be combined in a paper-based proof. Some students mentioned that a computer-supported proof could lead to a reduction of mathematical skills.

We asked the students to estimate the use of EASy in terms of weekly exercises and self-study arrangements. Figure 1 shows that most students have a positive attitude towards EASy, while a minority was skeptical.

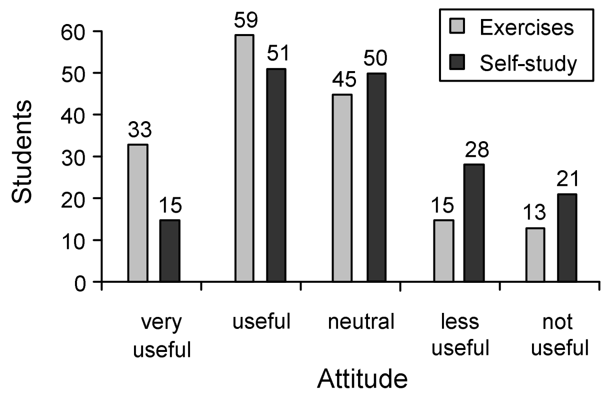

Figure 1. Attitude of students towards an application of EASy in academic assessments.

In addition, the students suggested improvement concerning organizational aspects owing to the prototypical status of EASy. For example, students demand a desktop version of the system enabling to work offline; they miss the ability to re-open a completed proof and they request a more detailed demo or help function such that it is easier to get started.

B. Evaluation of the tutors' perspective

As mentioned above it is important to gain the experiences and attitudes of all significant user groups. Thus, in a second survey we concentrated on the tutors of "Data Structures and Algorithms".

First, we asked for the organizational framework of the tutorials in summer term 2008. The students are supported by nine tutors in groups with about 23 participants each. Without any computer-support, the tutors spend up to nine hours (six hours on average) with the correction of exercises. It is obvious that manual correction in view of this large quantity of students is time consuming. Since EASy has been implemented, all tutors stated that the correction effort reduced significantly.

Although students were not forced to solve the exercises with EASy, only few students decided to submit manual proofs (see Figure 2).

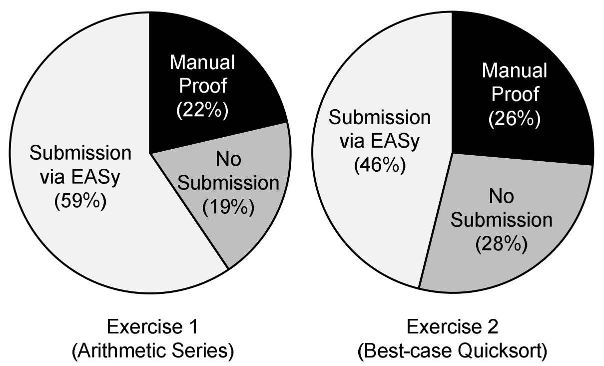

Figure 2. Proof submissions.

After obtaining the students' attitudes towards an application of EASy in several academic assessment scenarios, we also wanted to receive the tutors' opinions on this topic. Figure 3 shows that the tutors' attitudes to EASy. Although some tutors seem to be skeptical to some extent, most tutors estimate the use of EASy in exercises and self-study as beneficially.

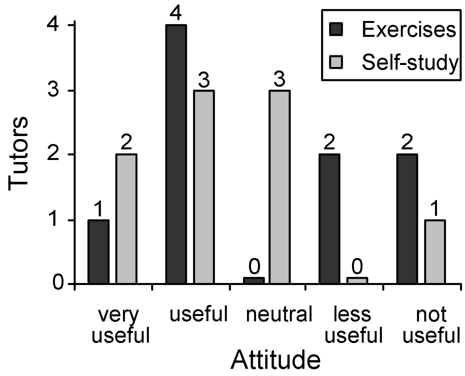

Figure 3. Attitude of tutors towards an application of EASy in academic assessments.

In general, the tutors solely mentioned problems that refer to EASy's functionality for students: Solving a proof with EASy takes more time than solving it manually, the familiarization with EASy was slightly complex and the detection of the required rules and theorems was not intuitive in some cases. The tutors also suggested a possibility for students to reopen and to share proofs.

Remarkably, the applicability and effectiveness of EASy concerning specific tutor functionality has not been criticized. All tutors evaluated the reduction of correction time as very helpful. Instead of approximately 20 minutes correction time per student, now few seconds are enough to detect whether the proof is correct. Additionally the tutors appreciate the quality of submissions, as EASy does not allow the misuse of theorems and provides erroneous proofs. They also stated that EASy is useful for the acquisition of proving skills.

To sum up, the tutors seem to be confident with their perspective of EASy, but they are aware of the challenges students have to face when they use EASy.

IV. CONCLUSIONS

To profit from computer-support in mass lectures in mathematics, we have developed the e-assessment system EASy for mathematical proofs. As we are not aware of any documented experiences with the application of e-assessment tools in math-related subjects, we have formatively evaluated the system in a "Data Structures and Algorithms" course with more than 200 students. In terms of two surveys we investigated applicability, effectiveness, usability and acceptance of EASy. The results were promising and encouraging.

In summary, most students were able to solve the given tasks after a period of familiarization, although they had to face several problems due to the prototypic status of EASy.

Especially the tutors profit from the implementation of EASy as the correction process is much less time consuming and sample solutions of the proofs are self-explanatory. Both, students and tutors, were interested in the concept and most of them would appreciate further application of EASy in academic assessments. Clearly, EASy needs to be extended for this.

In future work we intend to improve the EASy prototype to establish an enhanced application of e-assessment for mathematic purposes, which corresponds to the users' requirements. To facilitate the use of EASy for tutors we intend to implement a special authoring tool for a more comfortable design of assignments and the initiation of a direct presentation of automatic, situational feedback. Students will profit from an offline version of EASy, a more detailed demo and a possibility to store, reload und comment proofs. In addition we plan an improved and more intuitive rule base, the integration of a theorem prover and, the implementation of several administrative functions.

V. REFERENCES

[1] [1] Scottish Qualifications Authority, SQA Guidelines on Online Assessment for Further Education, SQA 2003.

[2] B. Eilers, S. Gruttmann, and H. Kuchen, „Konzeption eines integrierbaren Systems zur computergestützten Lernfortschrittskontrolle“, H. L. Grob, J. vom Brocke & C. Buddendick (Eds.) E-Learning-Management, Vahlen Verlag, München, 2008 (to appear).

[3] D. Chalmers and W. D. M. McAusland, Computer-assisted Assessment, Glasgow Caledonian University, 2002.

[4] C. J. Sangwin, "Assessing mathematics automatically using computer algebra and the internet", Teaching Mathematics and its Applications, Oxford University Press Oxford, 2004, 1-14.

[5] B. Lukoff, Automated Proof Assessment, Cornell University, Ithaca, 2004.

[6] S. Gruttmann, D. Böhm and H. Kuchen, An E-Assessment System for Mathematical Proofs, Proceedings of IASTED – CATE, ACTA Press 2008 (to appear).

[7] H. Kuchen, Lecture Data Structures and Algorithms, URL: http://www.wi.uni-muenster.de/pi/lehre/ss08/info2/, version: 2008, accessed 2008-06-30.

[8] F. Baader and T. Nipkow, Term Rewriting and All That, Cambridge Univ. Press, Cambridge 1999.

[9] B.N. Flagg, Formative Evaluation for Educational Technologies, Lawrence Erlbaum Assoc, Mahwah 1989.