

A framework for developing mathematical tasks for automatic formative assessment in higher education

Clara Horvath, Andreas Körner, Lana Medo

Institute of Analysis and Scientific Computing, TU Wien, Austria, Strategic Education Development Center, TU Wien, Austria.

Abstract

Mathematics education in STEM studies poses a challenge that students must practice the methods they learn to use them accurately. The mapping of teaching exercises is an important step towards creating a more engaging, interactive learning experience for students. An automatic assessment system can support such a process and help overcome obstacles towards a better learning environment for students. Developing effective support in form of high-level mathematical tasks for this type of assessment is challenging and requires careful consideration of the goals, objectives, and content of the task. With this paper, we aim to provide practical insights and references for teachers looking to develop various mathematical tasks that can provide meaningful feedback and support learning, while taking into account the limitations of automatic assessment in higher education. Additionally, this paper addresses the possibilities and challenges that arise in this educational process and provides examples of different applications.

Keywords: *Automatic assessment; e-learning methodology; innovative teaching; algorithmic randomization; STEM education.*

1. Introduction

The integration of technology into education has brought numerous benefits to the teaching and learning process. Computer-Assisted Instruction (CAI) has been shown to have a positive impact on student achievement and attitude in general as well as in higher education, but its effect size has been limited, see Kulik and Kulik (1991), Schmid et al. (2014). The effects' average size for achievement as reported by these meta-analyses ranges between 0.26 and 0.27 and therefore remains medium-low throughout the decades. Kulik and Kulik (1991) identify the developments of CAI at higher education levels to be less successful than those in elementary and secondary schools, whereas Schmid et al. (2014) identified subject matter as the most influential predictor of effectiveness and noted that the effects of technology-use were higher in non-STEM subjects.

In order to address these findings and reinforce the potential of technology-enhanced learning environments, it is important to understand the benefits and limitations of various instructional tools and assessment methods.

One such method that has gained popularity in recent years is Automatic Formative Assessment (AFA), which has the potential to support higher student achievement. Automatic assessment can be used to provide individual real-time feedback to students, allowing for an immediate response and helping to create a more engaging and interactive learning experience, addressing the conclusion of Schmid et al. (2014) that “learning is best supported when the student is engaged in active, meaningful exercises via technological tools that provide cognitive support” (p. 285).

Schmid et al. (2014) define cognitive support as “the category which encompasses various technologies that enable, facilitate, and support learning by providing cognitive tools” (p. 274). The development of effective cognitive support in form of high-level mathematical tasks for AFA is challenging and requires careful consideration of the goals, objectives, and content of the task. For example, an automatic assessment system which does not provide automatic grading beyond algebraic expressions constrains the type of tasks that can be developed. In order to maintain the benefits of AFA while still allowing for a wide selection of tasks, the teacher must, as Fisher (2006) warned, assume the role of an agent of change.

To support this role, this paper explores the development of mathematical tasks for AFA in higher education and presents a framework for its adaptation and optimization. Through this framework, we aim to provide practical insights and references for lecturers looking to develop various high-level applied mathematical tasks. By emphasizing the goal of achieving comparable difficulty for exam tasks and maximizing the randomization potential for exercise tasks, this framework helps to create meaningful feedback for students and supports their learning process, while taking into account the limitations of automatic assessment.

2. Exploring the Potential of Automatic Assessment Systems to Enhance Math Learning

According to the literature, an effective way to study mathematics at the university level is through active learning, i.e., independent solving of mathematical problems and hands-on contact with mathematical methods, see Rosenthal (1995). The challenge for teachers is to create a learning environment that enables students to solve these problems on their own and is suitable to guide and support them throughout the semester to achieve the goals of the course.

Automatic Assessment Systems (AAS) describe computer-aided applications that provide information on mathematical tasks and automatically evaluate student responses. AAS offer a conveniently accessible platform through which students can receive learning materials as well as grading and feedback services (Ihantola, Ahoniemi, Karavirta, & Seppälä, 2010).

As discussed by Barana, Marchisio, and Sacchet (2021), high-quality interactive feedback has a particularly positive effect on students, helping to improve their performance and support them in class preparation. Teachers are supported in the creation of learning materials by AAS, as these programs provide tools for the randomization of tasks, thus continuously offering new learning materials to students, as noted in the findings of Ihantola et al. (2010).

In the following subsections, we will present approaches that support the process of task development for AAS, specifically, Möbius (formerly known as Maple T.A.).

2.1. Enhancing Mathematical Tasks through Randomization

Randomization is a powerful tool that can be used to create a wide variety of mathematical objects, such as numbers, vectors, and matrices.

In Möbius, the generation of randomized variables is accomplished through the use of Maple commands. These commands offer the possibility to assign certain properties to the variables, such as restricting a parameter to be in a certain range or specifying the shape of the matrix. For example, `rand(1..10)` restricts a random number to be within the range of 1 to 10. Similarly, `LinearAlgebra[RandomMatrix](3,3, generator=-3..3, shape=diagonal)` generates a 3x3 matrix with random integer entries on its diagonal, ranging from -3 to 3. Using such Maple commands, available in libraries for Möbius, improves the precision of developed examples, allowing for more specific yet versatile tasks.

Furthermore, the possibility of drawing from a pool of predefined variables enables the randomization of text by assigning single words up to entire sentences to these variables. Consequently, the same assignment provides different tasks for the students at different attempts; for instance, specifying on continuity of rational functions or on the differentiability of logarithmic functions.

It is important to note that while the randomization approach introduces variability, it does not necessarily guarantee the comparability or solvability of the tasks. This topic will be explored in more detail in Section 3.

2.2. Enhancing Grading of Mathematical Expressions


Another essential component of any AAS is grading student responses. In Möbius, the computer algebra system Maple is utilized to efficiently and reliably grade equivalent algebraic expressions. Nonetheless, grading any other types of mathematical expressions, such as trigonometric equations, intervals, and general solution of linear systems, is an area of difficulty. To overcome this and enhance the grading functionalities, custom grading codes are developed.

A sustainable solution to manage and grade these more complex mathematical expressions conveniently would be to compile the custom grading codes in form of a library. Handily, Möbius supports the integration of Maple libraries through the repository in their so-called Maple-graded question types, allowing for more effective integration of the self-developed Maple procedures. When writing procedures, incorporating functionalities such as grading with partial credit can provide more finely-tuned feedback on the task level. This kind of grading system offers a more precise way to score a question than default scoring methods, allowing for a more nuanced understanding of a student's knowledge and skill. As an example, Figure 1 shows a problem of an under-determined system of linear equations, where the student is tasked with determining the general solution.

Let $A \cdot \mathbf{x} = \mathbf{b}$ be a linear system of equations with $\mathbf{x} \in \mathbb{R}^3$ given by

$$A \cdot \mathbf{x} = \begin{pmatrix} 2 & -2 & 1 & 3 \\ 4 & -4 & 8 & 6 \\ 2 & -2 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 8 \\ 1 \end{pmatrix}.$$

Determine the general solution \mathbf{x} of the

$\mathbf{x} =$ 

Partially Correct

Your Answer: $s*[1,1,0,0] + t*[-3,0,0,2]$

Correct Answer: $\begin{pmatrix} s_2 \\ s_2 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -\frac{3}{2} s_4 \\ 0 \\ 0 \\ s_4 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$

Figure 1. Example of Möbius question with an illustrated partial-credit evaluation in the response area of the answer field.

In this case, the provided answer is incomplete since a particular solution is missing. Typically, this would result in the response being marked incorrect; however, instead of a full deduction, partial credit is awarded.

The library allows for a more efficient and consistent approach to grading, as the same grading code can easily be reused. It can also serve as a basis for a custom grading system incorporating further functionalities and design choices, which can go as far as optimizing the library for custom syntax. In turn, this allows teachers to focus on other quality aspects of their tasks.

Furthermore, the library concept can be extended to randomization as well. Zimmermann et al. (2010) provide a concrete example of library implementation to improve the grading and randomization aspect of AAS. This publication was the starting point of the development of customized grading and randomization libraries for the AAS Möbius at TU Wien. These libraries enabled greater flexibility in designing mathematical tasks and opened new possibilities for their creation, further emphasizing the importance of custom grading and randomization as enhancement methods in designing mathematical tasks.

3. Task Differentiation Based on Use Case Scenario

AAS can be used to provide teaching material for students, as well as for examinations. These different use case scenarios require specific adaptations and considerations in the implementation of the tasks. We propose a differentiation of tasks based on their respective use cases and discuss them in the following subsections.

3.1. Exercise Tasks or Encouraging Learning and Proficiency through Practice

By using randomization with a wide range of possible parameters, lecturers can provide a large variety of examples for students to study and practice the required skills and abilities. The difficulty of a task might vary due to the parameters, which is desirable in practice scenarios as students learn to master the same skill on different levels of complexity.

According to Barana et al. (2021), “interactive feedback can be effective for the development of Mathematical knowledge ...” (p. 17) and is more elaborate than simple right or wrong feedback. It should consist of an instructive feedback guide to engage students to solve mathematical problems after failed attempts, see Marchisio et al. (2020). With this in mind and in line with Leikin's (2014) definition of the mathematical challenge as a mathematical difficulty that a person is motivated to overcome, we further endorse the use of so-called adaptive questions.

Möbius offers a predefined adaptive question type that incorporates randomization capabilities and an automated grading system, similar to standard questions. However, it also

guides students to the correct path for solving the given problem after a predefined maximum number of unsuccessful attempts. After an initial failure, the difficulty of the problem is reduced and dissected into smaller and simpler sub-problems and the students' answers in this adaptive section are graded. Depending on the correctness of the answer, more hints or instructions are given. The next problem definition to achieve the correct solution is displayed accordingly. Therefore, an adapted path to the solution of the mathematical problem is provided.

Although the proposed steps are one of usually many possible solution paths and the level of individualization and allowed creativity is limited, adaptive assignments offer students valuable first-level support when tackling tasks. Another valuable feature of adaptive assignments is that individual intermediate steps change in accordance with the initial problem, therefore every student receives a personalized guide at every attempt.

3.2. Exam Tasks or Ensuring Fairness and Consistency in Assessment

Using AAS during examinations provides the possibility to conduct exams in locations other than lecture halls. Before the global COVID-19 pandemic, this feature may not have seemed necessary; however, it is now indispensable. Stowell and Bennett (2010) report that students' anxiety can be reduced when having the opportunity of taking online exams compared to exams in classrooms or lecture halls. Additionally, AAS tests can be used to provide feedback to students away from regular exams, for example in preparation for a midterm or final test that counts towards the final grade.

Providing mathematical assignments via AAS during exams prevents students from copying solutions from neighboring students during in-person exams as well as exchanging solutions via messaging services such as WhatsApp during online exams due to randomized and thus, individualized exercises. Therefore, e-learning examinations enhance the credibility of exam results, since they address security concerns such as identifying students with their work and threats of cheating, see Karim and Shukur (2016).

However, the high variability of possible values and texts in questions poses difficulties for teachers since unwanted problematic realizations of the randomization cannot be easily prevented when using standardized randomizing Maple functionalities. Especially in exam situations, undefined values which cause the overall task to become unsolvable need to be prevented. An exemplary problematic realization would be the difference of two parameters in the denominator of a term leading to division by zero if equality is not manually excluded.

Above all, the comparability of the tasks also must remain given for reasons of fairness.

Teachers can use the same task repeatedly in exam situations since the randomization guarantees different values for each attempt. By drawing from a pool of predefined variables, cheating can still be minimized, and tasks can be reused in future exams without greatly

increasing the effort required to create them. For instance, consider a task in which a student is asked to compute eigenvalues and eigenvectors of a given matrix. In this case, the *switch* command can be used to randomly select from a predefined set of matrices that impose an equivalent level of difficulty to the task's given context.

Furthermore, Sangwin (2013) suggests using reverse engineering as an advanced approach to address the challenge of comparable task difficulty in exam situations, while also allowing for greater task versatility.

4. Discussion

With the proposed framework, we are addressing the comparatively low effectiveness of technology use in STEM subjects. In doing so, we focus on advancing the process of development of various tasks relevant to applied mathematics in higher education. We also emphasize the need to differentiate and adapt tasks according to their specific use cases. Automatic assessment of tasks which are more relevant to pure mathematics, e.g. proof examples, require different approaches and are out of the scope of this framework.

Nonetheless, AAS pose a suitable tool for lecturers to provide learning material for students. Conveniently, AAS, such as Möbius, can be integrated, e.g., into Moodle courses easily because of their Learning Tools Inoperability (LTI). Feedback from students in the various courses showed that the pool of examples as learning material was particularly praised. Before incorporating course material for AAS, teachers need to take into careful consideration the responsibility, time and effort involved. Contrary to static questions on paper, computer-algorithmic questions with incorporated randomization are more susceptible to errors as malfunctions can occur on several parts of the questions. Apart from undefined solutions originating from randomized variables, errors concerning grading algorithms or the display of variables in the text of the questions need to be addressed. One solution, we presented in this contribution, is to create custom libraries with advanced grading and randomization algorithms, as they can be reused conveniently. As a result, the possibilities of task design have advanced leading to higher engagement rates of students.

The benefits and challenges of AAS in conjunction with examinations were also discussed. In this context, the issue of available resources needs to be addressed, as students must access the exam using a computer. To prevent discrimination against certain social groups, the infrastructure and technical equipment of the university or educational building must be reviewed beforehand by lecturers to ensure access to all students. Furthermore, universities need to review their facilities providing such infrastructure.

Over the past few years, the development of randomization and grading libraries has been advanced and the related approach of designing examples has been improved and optimized

to address specialized tasks in the areas of analysis and linear algebra. A future goal is to analyze assignments and exams in one course simultaneously, providing detailed statistical correlation between them. With this statistical analysis, we plan to identify questions with a high potential for incorrect answers and adapt the framework to better meet the needs of students.

References

- Barana, A., Marchisio, M. & Sacchet, M. (2021) Interactive feedback for learning mathematics in a digital learning environment. *Education Sciences*, 11 (6). doi: 10.3390/educsci11060279.
- Fisher, T. (2006). Educational transformation: Is it like "beauty" in the eye of the beholder, or will we know it when we see it? *Education and Information Technologies*, 11,293-303. doi: 10.1007/s10639-006-9009-1
- Ihantola, P., Ahoniemi T., Karavirta V. & Seppälä O., (2010). Review of recent systems for automatic assessment of programming assignments. *Koli Calling '10 Proceedings of the 10th Koli Calling International Conference on Computing Education Research*. doi: 0.1145/1930464.1930480.
- Karim, N.A. & Shukur Z. (2016). Proposed features of an online examination interface design and its optimal values. *Computers in Human Behavior*, 64. doi: 10.1016/j.chb.2016.07.013.
- Kulik, C.C., & Kulik, J.A. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior*, 7, 75-94. doi: 10.1016/0747-5632(91)900305
- Leikin, R. (2014). Challenging Mathematics with Multiple Solution Tasks and Mathematical Investigations in Geometry. In: Li, Y., Silver, E., Li, S. (eds) *Transforming Mathematics Instruction. Advances in Mathematics Education*. Springer, Cham. doi: 10.1007/978-3-319-04993-9_5
- Marchisio M., Remogna S., Roman F. & Sacchet M. (2020). Teaching mathematics in scientific bachelor degrees using a blended approach. *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, 190-195, doi: 10.1109/COMPSAC48688.2020.00034.
- Mario, P. (2018). Online assessment, adaptive feedback and the importance of visual learning for students. The advantages, with a few caveats, of using MapleTA. *International Review of Economics Education*, 28, 11-28, doi: 10.1016/j.iree.2018.03.002.
- Sangwin, C. J. (2013). Computer aided assessment of mathematics. Oxford: *Oxford University Press*.
- Schmid, R. F., Bernard, R. M., Borokhovski, E., Tamim, R. M., Abrami, P. C., Surkes, M. A., Wade, C. A., & Woods, J. (2014). The effects of technology use in postsecondary education: A meta-analysis of classroom applications. *Computers & Education*, 72, 271–291. doi: 10.1016/j.compedu.2013.11.002
- Stowell, J., Bennett, D. (2010). Effects of online testing on student exam performance and text anxiety. *Journal of Educational Computing Research*, 42 (6). doi: 10.2190/EC.42.2.b.

- Rosenthal, J. S., (1995). Active learning strategies in advanced mathematics classes. *Studies in Higher Education*, 20 (2), 223-228. doi: 10.1080/03075079512331381723.
- Zimmermann, A., Urbonaite, V., Körner, A., Winkler, S., Krause, S., & Kleinert, M. (2010). Advanced randomization and grading in the e-learning system Maple T.A. In M. Snorek, M. Cepek, Z. Buk, & J. Drchal (Eds.), *Proceedings of the 7th Congress on Modelling and Simulation* (p. 1209–1214).