Use of an interactive digital assessment tool in mathematics education for engineers

Andrey Chesnokov[*1] and Siebe van Albada[†2]

[1]*Department of Mathematical Sciences, NTNU in Gjøvik, Norway*
[1]*Department of ICT and Natural Sciences, NTNU in Ålesund, Norway*

Abstract: Basic courses in mathematics for engineers often have more than several hundred participants. To achieve the learning goals it is necessary that the students work actively on specific problems. While working on the problems, students require assistance. As a supplement to guidance by lecturers and teaching assistants, NTNU has piloted use of a digital assessment system STACK, that enables constructive feedback and is to some degree interactive.

We will discuss how we integrate digital assessment into daily teaching practices, show examples of the problems and present some qualitative data collected from actual students.

1 INTRODUCTION

Historically, basic courses in mathematics for engineers had many participants and were arranged in a classical way with lectures and seminars, where a faculty member is responsible for lectures and several teaching assistants (TA's) are responsible for seminars. Many authors have indicated [1, 2] that it is difficult to hire well-qualified TA's.

According to sociocultural learning theory [3], learning happens through all kind of interactions, namely, with teachers, peers or educational environment. To facilitate learning, the majority of problems that the students work with should be in the zone of proximal development, which means that students may have difficulties with solving the problems being alone, and interactions are essential to construct solutions. So, who is capable of being a conversation partner in the classical course setup?

The faculty member responsible for lecturing clearly does not have resources to help several hundred students. Access to qualified TA's is very limited, and while the institutions try to educate their TA's through programs like LAOS ([2, 4]), their competence does not accumulate. Situation is especially difficult for smaller campuses like NTNU in Gjøvik and Ålesund, where Master's programs are few and engineering students disappear to other campuses or start their working careers right after 3 years of Bachelor studies.

As a remedy, NTNU has piloted use of a digital assessment system STACK (the System for Teaching and Assessment using a Computer algebra Kernel, [5]) that is able to provide constructive feedback to students. Use of an integrated computer algebra kernel goes beyond the classical approach with just indicating whether a solution is correct or incorrect, and opens for a detailed analysis of certain mathematical properties of the student-provided solution. By allowing to answer the same question several times and providing information on precisely the missing/present properties of a proposed solution, we make the STACK system a conversation partner. This also means that the focus in assessment shifts from being summative to being formative. Thus, rather than judging or ranking achievements, learning comes as the main purpose.

This paper presents intermediate results that come after the first year of the STACK pilot project at NTNU, and is based on qualitative interviews and own observations.

[*]andrey.chesnokov@ntnu.no
[†]siebe.b.v.albada@ntnu.no

2 PREVIOUS WORK. SELECTING AN ASSESSMENT SYSTEM

During the last years, NTNU has significantly restructured its courses in mathematics for civil engineers, and a digital assessment tool MapleTA was an important part of the change process. While supporting mathematical notation and tests for algebraic equivalence, the system did not have focus on analyzing specific properties of the answer, rather providing feedback of type correct/incorrect answer. Interviews with student panels [1] have indicated that students highly value that the feedback is instant, but miss details on what precisely is wrong in their answers.

Further, MapleTA became a part of a larger software package Möbius Assessment, which has resulted in significantly higher license costs. The sudden change in expenses was one of the factors that has motivated NTNU to start looking for alternatives to MapleTA, especially for alternatives with free licenses.

STACK (System for Teaching and Assessment using a Computer algebra Kernel, [5]) is an open source digital assessment package for mathematics. STACK is based on over a decade of research and development, and in addition to full algebraic input with validation, it allows for differentiated feedback. STACK was developed at the University of Birmingham, and the University of Edinburgh now hosts the STACK project. In Norway certain work in the context of STACK is also done at the University of Agder, as described in [6].

3 ORGANIZATIONAL ASPECTS

As a result of several University Colleges being integrated into NTNU, several courses had to be redesigned and unified to offer the same studies at different campuses. One of those unified courses – Mathematical Methods 2 for Computer Science engineers, 2nd year Bachelor – was selected as a pilot course for the STACK system at Spring semester 2020. Later, another Bachelor course – Mathematics for Informatics, 1st year Bachelor – was using STACK during the Autumn semester 2020.

We offered weekly assignments in STACK, where students were allowed to answer the same question many times without penalties, and each student's best answer counted towards the final score. The students eventually had to answer 80% of questions correctly to pass each assignment, and they had to pass 8 of 12 assignments to be admitted to the final exam.

As though assignments were completely online, students were also offered offline sessions two times a week. During the offline sessions lecturer(s) and several TA's were present at the same time.

As already motivated by sociocultural theory, a successful study setup should promote interactions. This has been achieved by the following means. First, for the most of the exercises in STACK we have programmed some form of differentiated feedback, including hints, thus enabling the system to be a conversation partner. The system never presented a complete answer, unless the student had constructed it herself. Further, the students were informally encouraged to collaborate in small colloquium groups, and tables in the exercise room were places in islands. This limited the size of groups and facilitated face-to-face conversations. While assignments being formally individual, we have several times underlined that students should help all members of the group to master questions, rather than being "first" or "best" in class.

Finally, the lecturer and TA's were guiding students through the problems in the cases where help from peers or feedback from the STACK system was not available or was not sufficient.

A flat pass/fail grading system for the exercises is known to facilitate collaboration between students, has been shown to reduce experienced anxiety, while a higher threshold for passing induces need for better feedback underway, as summarised in [7].

4 TWO EXAMPLES OF PROBLEMS

Pedagogical considerations on designing the problems for STACK are described in [5] and summarized in [6], so in this paper we just give two specific examples of STACK exercises. The combinatorial problem formulated in the first line (Figure 1) is too complex for most of our students. Thus, we have decomposed a complete solution into smaller pieces and asked several intermediate questions. When students enter a wrong answer into an input field, they get a hint. We call such setup for a minimal working solution.

The next problem (Figure 2) is taken from a STACK demonstration website and goes deeper in feedback, also taking one more representation form into consideration. Here the concept of an odd function is

Question 8
Not complete
Mark 0.00 out of 1.00

⚑ Flag question
⚙ Edit question

Hvor mange løsninger har heltallslikningen $x_1 + x_2 + x_3 + x_4 = 12$, om alle variablene skal være større enn eller lik 2?

Vi bruker bijeksjon fra (x_1, x_2, x_3, x_4) til binære strenger. Hver streng må da inneholde 12 nuller, hvor sammenhengende nuller representerer et tall og skilles med enere.

1) Hva må en streng starte med? `1`

1.1) Hva må det andre tegnet være? `2`

2) Hva må hver ener følges av (skriv inn alle symboler du veit om)? `3`

3) Hvor mange nuller og skilletegn kan man plassere etter man har funnet restriksjonene? `4`

4) Hvor mange skilletegn kan man plassere? `5`

5) Hva blir da antallet løsninger for likningen? `6`

Check

✖ Incorrect answer.

1) Feil, prøv igjen! Husk at hvert tall må minst være 2 eller høyere.

1.1) Feil, prøv igjen! Husk at hvert tall må minst være 2 eller høyere.

2) Feil, prøv igjen! Husk at hver variabel må minst være 2.

3) Feil, prøv igjen! Etter man har funnet riktig svar på de 2 forrige spørsmålene, hvor mange nuller kan man plassere ut da?

4) Feil, prøv igjen!

5) Feil, prøv igjen!

Marks for this submission: 0.00/1.00.

Figure 1: A problem with some hints

reformulated geometrically, and an actual graph of a student's suggestion is plotted, thus providing better insight to the problem. The white frame shows results of the validation process, namely, that student's input is considered as an expression with variable x.

Question 3
Not complete
Mark 0.00 out of 1.00

⚑ Flag question

Give an example of an even function.
$f(x) = $ `sin(x)`

Your last answer was interpreted as follows:

$$\sin(x)$$

The variables found in your answer were: $[x]$

Check

✖ Incorrect answer.
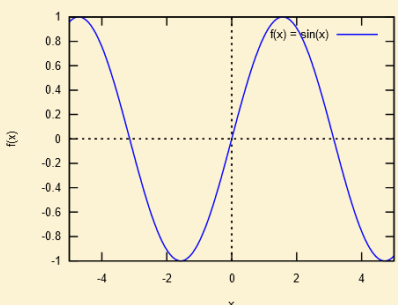Your function should be symmetric about the y-axis, but is not.

Figure 2: Problem involving graphics

5 RESULTS FROM INTERVIEWS AND OBSERVATIONS

5.1 Student panel reviews

As a part of NTNU's education quality management system, regular meetings with student panels were held. Summaries of the meetings have been systematically archived, and the following aspects have been indicated:

Immediate feedback Everybody appreciates instant feedback. This aspect is so strong that many waited for the STACK-variant before starting to work with a new problem set.

Detailed feedback As we were developing and testing at the same time, not all questions were programmed with constructive feedback. Students explicitly demanded detailed feedback for all exercises, and this has been emphasized rather strongly.

Number of attempts The unlimited number of attempts takes away the stress of "suddenly writing something in the wrong way". This also encouraged some students to try and get assessed different solutions to the same problem.

Technical challenges with input STACK as of today does not have a graphical interface to input equations. Certain students meant that it was technically challenging to input larger formulas and/or lists and sets. Given that the pilot was aimed at Computer Science students, the issue is considered serious.

Technical errors Instant feedback has a strong positive effect on motivation. On the contrary, even small errors in programming of questions, leading to the correct answer not being recognised as a correct answer, have negative effect on motivation and may cause frustration and bad feelings about "this stupid system".

The data might be affected by the fact that it was collected by the lecturer itself. However, it was mentioned to the students that they are a part of a pilot project, and their feedback on the STACK system will be important for the future generations of students.

5.2 Teacher's observations

As STACK enables assessment of different mathematical objects, such as complex algebraic expressions, sets and lists of those, there was almost no need to use multiple choice questions. We tried to assess authentic production of students, and in most questions it was not possible to guess the correct answer, even with an unlimited number of attempts. This forced each colloquium group to produce its own solution.

At the start we were worried about "the free passenger problem", namely, that just one person per group would produce the solution and the others would copy and paste. While we do not know the working patterns of those students who were online-only, observations during offline sessions show that most students kept their own paper record of solution methods to each problem. As though some members in each group were obviously faster/more efficient in their work, the presence of *solution methods*, rather than just *answers*, indicates that colloquium groups to a large degree did not just literally copy and paste.

STACK separates validation from assessment. This means that the student-provided answer is first parsed and then presented to the student in a formatted way. Only when the student agrees that the answer has been interpreted correctly, the answer is being sent for actual assessment. This drastically reduces needs for technical assistance, as most students are able to troubleshoot by themselves. This also has a positive effect on motivation, as technical issues are mainly separated from mathematical issues. Reports given by TA's show that the majority of technical issues resulted from bugs in exercise programming, and the rest of the questions asked were about deeper mathematical concepts. Thus, a combination of detailed feedback by STACK and work in colloquium groups results in better troubleshooting skills among the students, and brings conversations with available human guides to a higher level than looking for typo's and counting errors.

Observations are based on a selection of students who met for exercise sessions. The number of those students comprised all from 70% to 10% of a total number of students, and covid-19 restrictions had a clear negative effect on attendance of offline sessions. So, the observations may not hold for the whole class, and further investigation is necessary to draw deeper conclusions.

In the context of the covid-19 crisis, when many offline learning activities were not allowed or not recommended, STACK-based exercises provided at least some feedback to the students. STACK also enabled the lecturers to monitor the progress, as the students had to work with assignments systematically from their homes. Without STACK, an only real alternative would be to use quizzes within the official learning management system (LMS) for NTNU, namely, Blackboard learning. Such quizzes

would necessarily use multiple choice questions, while there are many solid arguments against this type of questions, see e.g. [5].

6 CONCLUSIONS AND CHALLENGES

The STACK system received a rather positive response from the students, and it is clear that the most appreciated feature is interactivity based on deeper analysis of the provided answers. This is also the feature that takes most time to program, and the work on adding extra interactivity to existing problems will continue.

Further, several new software features are desired in STACK. One is aimed at higher adaptivity, such that different student would be allowed to follow different solutions paths and be automatically guided on the way. Another area of software development could be aimed at a two-directional interactive graphics engine that would facilitate both a) construction of graphical representations of algebraic concepts and b) construction of algebraic representations of geometric concepts. Such an engine would allow new types of questions, where several representation forms of complex mathematical concepts are coupled together. For 2D graphics, substantial work has been done within JSXGraph project, [8].

As of now, STACK is not integrated into any system for digital exams that is used in Norway. Those systems have limited support for mathematical questions, and it could be beneficial to develop integration such that STACK could also be used for exam purposes. However, such integration has to be done by private vendors, and joint efforts from the academic community might be necessary to convince them for the integration project.

ACKNOWLEDGMENTS

REFERENCES

[1] Frode Rønning. *Innovativ utdanning i matematikk* UNIPED. 2015, 38 (4), 319-326. `https://hdl.handle.net/11250/2657706`

[2] Greta K. Johansen and Maarten T. P. Beerepoot. *På vei mot en kollegial tilnærming til undervisning – seminarundervisning og pedagogisk opplæring av stipendiater* Nordic Journal of STEM Education, Vol. 3 No. 1 (2019), ss. 38-41. `https://doi.org/10.5324/njsteme.v3i1.2992`

[3] Lev Vygotsky. *Interaction between learning and development.* In Gauvain and Cole (eds), Readings on the development of children. New York: Scientific American books, 1978.

[4] LAOS - Opplæring for læringsassistenter at NTNU. `https://www.ntnu.no/documents/1266530207/1270423601/LAOS/c636c00b-25db-435a-bc90-a4b682fcfeb1`

[5] Chris Sangwin. *Computer Aided Assessment of Mathematics.* Oxford University Press, ISBN: 9780199660353

[6] Morten Brekke. *Utvikling og oversetting av selvrettende oppgaver fra STACK for å bedre studentenes læring i matematikk* Nordic Journal of STEM Education, Vol. 3 No. 1 (2019), ss. 195-199. `https://doi.org/10.5324/njsteme.v3i1.2992`

[7] Christian Jørgensen and Helen Bråten. *Can 'Passed with Distinction' as a new grading scale favour the transition towards formative asssessment?* Nordic Journal of STEM Education, Vol. 3 No. 1 (2019), ss. 7-11. `https://doi.org/10.5324/njsteme.v3i1.2992`

[8] Alfred Wassermann et al. *JSXGraph – Dynamic mathematics running on (nearly) every device*, Electronic Journal of Mathematics and Technology, 2011, 5/1