Research Paper

Automatic assessment of mathematics exercises: Experiences and future prospects

Antti Rasila
Matti Harjula
Institute of Mathematics, Helsinki University of Technology, Finland
Kai Zenger
Control Engineering Laboratory, Helsinki University of Technology, Finland
antti.rasila@tkk.fi, matti.harjula@tkk.fi, kai.zenger@tkk.fi

Abstract

Our experiences and future plans for computer aided assessment of mathematics exercise assignments are outlined. A summary of the results from a pilot course arranged in autumn 2006 is given. While the pilot course concerned basic engineering mathematics, computer aided assessment technology can be used in any related field as well. Integration of the technology to existing control engineering web courses and future applications in engineering are also discussed.

Introduction

Mathematics and physics constitute the ground of all engineering sciences, and a good knowledge on them is a prerequisite for mastering sophisticated applications in different engineering fields. One example area is automation technology and especially control engineering, which use a lot of advanced mathematics to solve problems e.g. in process technology, machine design, power electronics and electric drives, telecommunication applications, micro- and nanotechnology etc. [18]. In order to study control engineering efficiently and successfully, a firm mathematical ground and practical routine to apply mathematics is needed of the student.

However, in modern age the amount of new information and new technological tools has grown so much that learning and training the basic skills by "pen and paper" is not so popular any more among the students. This claim, although not proved, is believed to be true among the majority of students in technical universities. In the new educational system in EU a lot of emphasis is put on planning the quality of university level teaching and learning, and new activating learning methods have been introduced and tested. However, the fundamental difficulty of motivating such students, who feel that the subject is too difficult, too theoretical or purely not useful for his/her further studies and work, remains.

In this article we outline our experiences and future plans for Computer Aided Assessment (CAA) of mathematics exercise assignments. CAA has been used in computer science with good results, see e.g. [1, 4], but it is believed to be new in mathematics. This is a joint research and development project between the Control Engineering Laboratory and the Institute of Mathematics in Helsinki University of Technology. The basis of our project is the experience in interactive learning environments and digital study material that has been gained within the long-running MatTaFi project [7], and expertise in CAA systems and web based study materials in the Control Engineering Laboratory from earlier projects in the field.

Our main motivation is to provide new opportunities for learning mathematics, with a specific goal of improving the quality of teaching in mathematics and related fields [10]. The technology has shown promise in student activation in particular, but the other benefits from automatic assessment include continuous diagnostic information and improved use of resources. The technology was tested on a pilot course in Helsinki University of Technology in 2006. Results from the pilot experiment were promising. The student reaction to web based mathematics exercises was positive, and we found the technology to be sufficiently mature for large scale testing. The tested exercise assessment framework was based on free open source software (FOSS) solutions.

Anatomy of a CAA system

A CAA system consists of the following parts:

1. Access and identity control for identifying students and teachers.
2. User interface for the student (for taking exercise assignments).
3. User interface for the teacher (for making new assignments).
4. A database where the assignments are stored.
5. Computer Algebra System CAS (e.g. Maxima, Maple, Mathematica) that is used to grade the student answers and give relevant feedback.
6. A grade book where the results are stored (optional).

For practical reasons, the user interface is usually accessed with a web browser. There are several CAA systems suitable for assessment of mathematics exercises, including AIM [3, 12, 13], Maple T.A. [6] and STACK [14]. In case of STACK the CAS behind the system is open source Maxima.

Before STACK was chosen for the pilot course several free/proprietary solutions were evaluated. None of the available solutions was found to be entirely satisfactory for the projected usage. The main issues were related to user management in large courses, support for Finnish language and cross-platform compatibility. The most important advantage with open source software was that any modifications and improvements deemed necessary could be developed locally. Other benefits of using open source software include lack of licensing fees. In Figures 1 and 2 a basic example illustrating how STACK can be used is given.

Figure 1: Assignment editor (teacher) view of an exercise. Dimensions and coefficients of the matrices are random numbers.

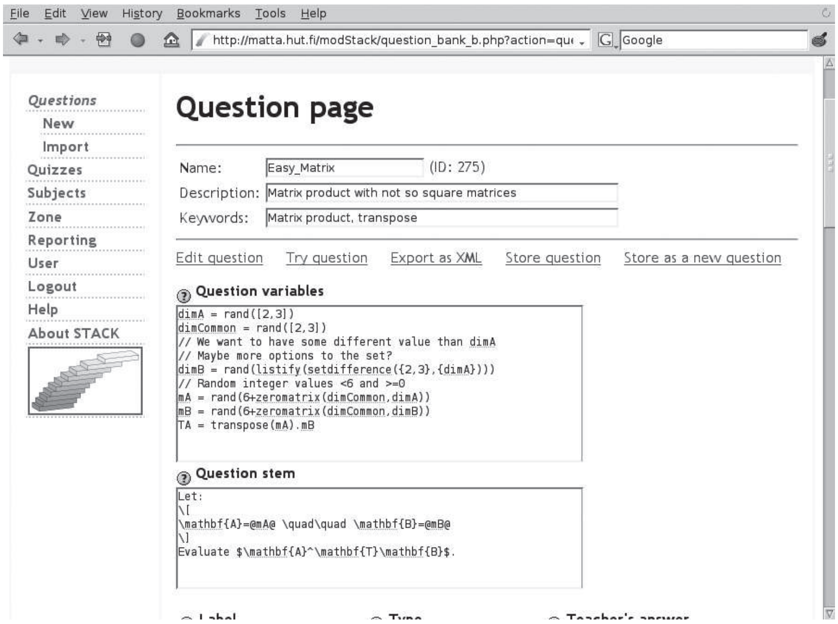
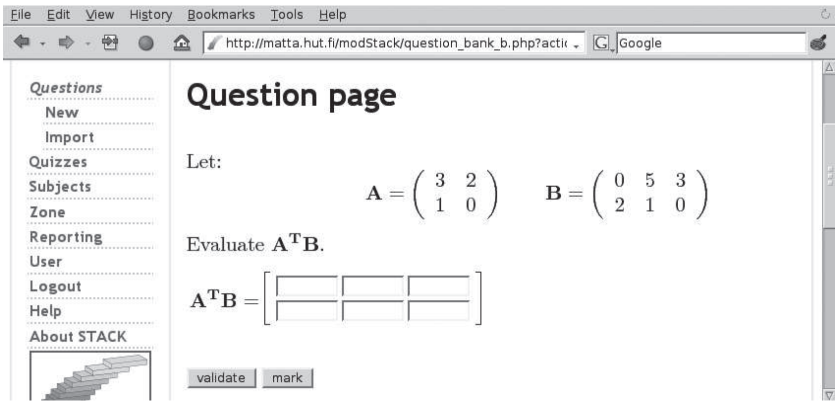


Figure 2: Student view of the same exercise. Because randomized elements are used, each student will have a slightly different exercise assignment.



Pilot course KP3-I in autumn 2006

The basic course in mathematics KP3-I introduces second year students to certain new concepts in engineering mathematics. The students in the course are majoring in e.g. construction, chemical and mechanical engineering. Contents of the course include:

- basics of complex analysis,
- integral transformations (i.e. the Laplace transformation),
- and the Fourier series with applications to differential equations.

In autumn 2006 a modified version (see [5]) of the STACK software was tested as a part of (voluntary) homework exercises. Traditionally, students have been given extra points for those exercises that the students said they were prepared to present in the exercise sessions. These points are then added to the score from the midterm examinations, and the sum is used to determine the final grade of a student. On the pilot course STACK exercises were used along traditional exercises (four traditional and four STACK exercises/week). There where 207 students that were enrolled in some way to the course. Some of them did not participate in exercises of any type and came just for the exam.

Each week a student had an option to participate in traditional exercise sessions, STACK exercises or both. Students were given bonus points for participating exercises of any type, but the number of points for each web-exercise was one half of that given for solving a traditional one. The reason for this was that web-assignments were considered to be narrowly focused when compared to traditional ones, e.g. there was only one question in each assignment while traditional assignments of often included separate questions (a) and (b). Students could earn maximum of six bonus points from exercises which were added to the points from exam (max. 30) for determining the final grade.

The goal of the pilot course was to test computer aided assessment technology in a real course environment. In particular we were interested in following questions:

1. Is the technology mature enough to be used in a large course?
2. What is the student reaction to computer aided assessment (student activity when compared to the traditional exercises)?
3. Can the students use the system? If not, how much training is required?
4. What are the most common problems with the system?
5. Does automatic assessment somehow change the way how the students work with their exercises? For example, how much do the students use the option to work outside the office hours?
6. How much plagiarism will occur if exercises are not randomized?
7. How do the students learn?

Figure 3: Percentage of students who actively participated either STACK or traditional exercise sessions (the groups are not mutually exclusive). Here total is the number of students who attended the exam. The number of participants is higher in STACK exercises through the course. The decrease of student activity is also steeper in traditional exercises

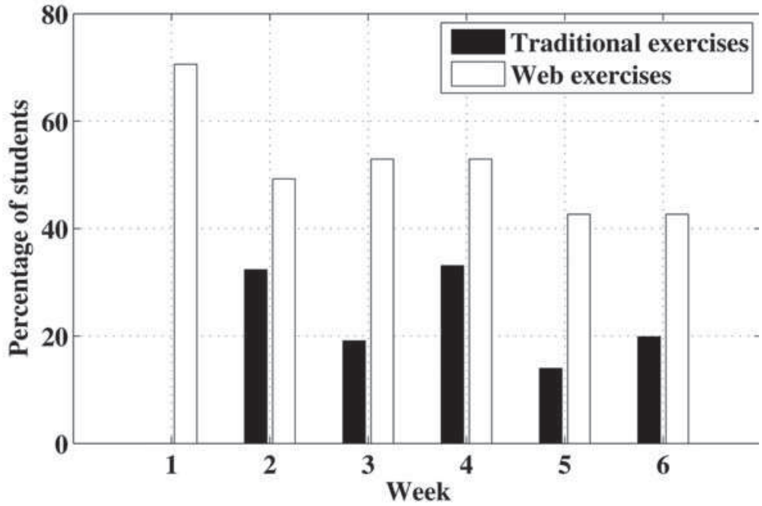


Figure 4: Student activity in as time approaches the deadline. It is not surprising to see that most of the activity concentrates to just before the deadline (Sunday 24:00). The last hour counts as 6 % of the total activity. The dashed line is the time of day 00:00–24:00 (scaled).

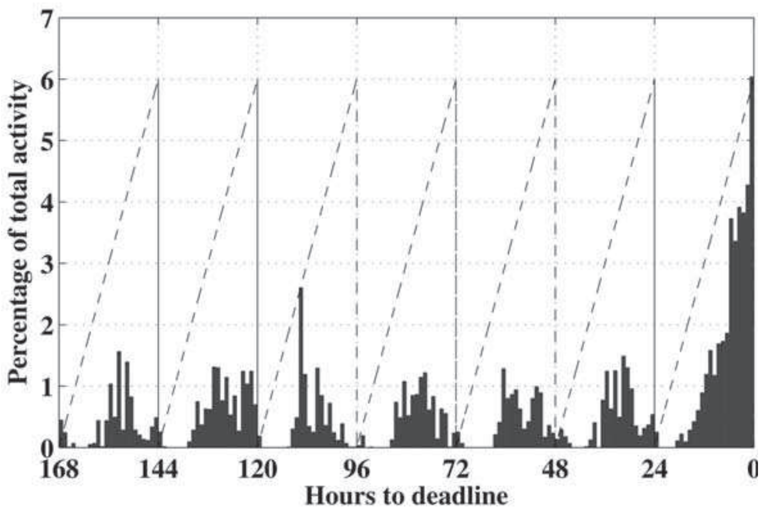
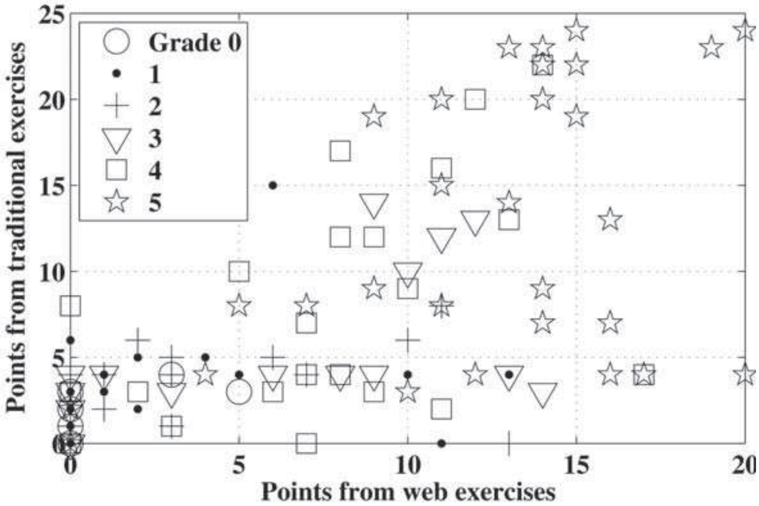


Figure 5: Participation in traditional and web-based exercises. The figure indicates that mostly the same students participated the both. The best grades were, in general, given to the most active students. Symbols in the figure indicate the final grades.



Conclusions

The most important result from this experiment is that, in general, the answer to the first question is affirmative. The technology is not perfect but it works well enough to justify wider testing and future development efforts. The student reaction to the system was also positive as demonstrated by the activity data (Figure 3). The basic use of the system required very little instruction or training. There were some problems with entering more complicated answers. In particular, problems arose with question types requiring students to enter long algebraic formulas. Many improvements to the user interface have since been made in order to make this kind of problems as rare as possible [5].

Also, questions have to be formulated very carefully and explicitly to avoid confusion. Interestingly, some students found web-based problems easier than similar problems given in the traditional way. This could be related to difficulties in reading mathematical text, making it easier to answer a problem that is stated very explicitly, even if the mathematics actually needed for solving the problem remains the same.

The student activity graph (Figure 4) shows that many students like to work outside the office hours. Unsurprisingly, much of the activity is concentrated just before the deadline. Many students spent a long time with each exercise [5]. This probably indicates that students mostly do not solve the exercises beforehand but work with mathematics while the web session is open. Thus, printed exercise assignments may not be needed in the future.

Participation to traditional and web-based STACK exercises is presented in Figure 5 along with the final grades from the course. The figure seems to indicate that mostly the same students were active both in traditional and web-based exercises. The figure also shows that the situation with plagiarism is probably not worse with web-based assessment when compared to traditional exercises, even if no randomization is used, although this hypothesis is difficult to test statistically. Obviously, the situation may change in the future as the technology becomes more widely used.

Table 1: Correlations between final grades, points from the exam, bonus points earned from exercises (web/traditional) and participation in exercises. A student who had tried to solve at least one problem of the given type is considered a participant.

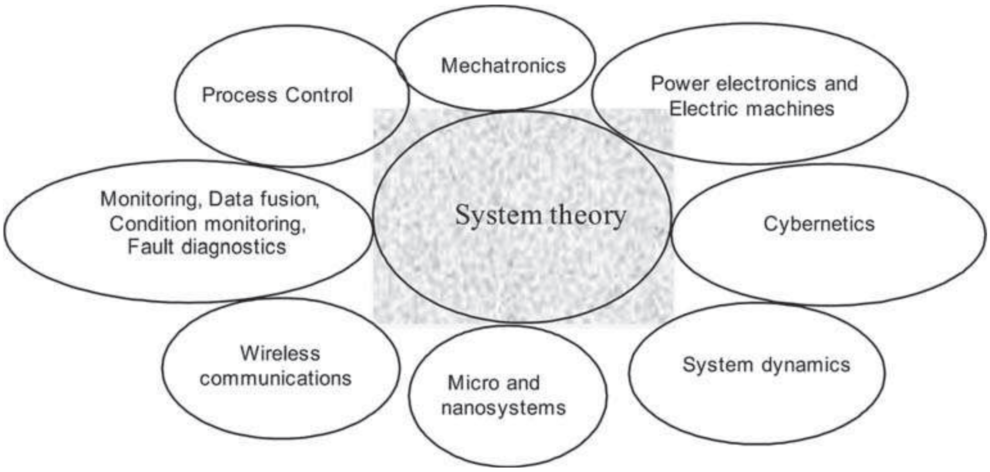
	Final grade	Exam	Bonus/ web	Bonus/ trad.	Partic./ web	Partic./ trad.
Final grade	1.0000	0.9720	0.6416	0.5842	0.6259	0.5897
Exam	0.9720	1.0000	0.5996	0.5466	0.5605	0.5515
Bonus/web	0.6416	0.5996	1.0000	0.6440	0.9066	0.6201
Bonus/trad.	0.5842	0.5466	0.6440	1.0000	0.6157	0.9669
Partic./web	0.6259	0.5605	0.9066	0.6157	1.0000	0.6202
Partic./trad.	0.5897	0.5515	0.6201	0.9669	0.6202	1.0000

Final grades given to the students seem to be closely connected to the activity in exercise sessions (see also Table 1). On the other hand, this result may be somewhat distorted by the fact that questions in the exam were very similar to the exercise assignments. Students were also given some bonus points for participating exercises, although this was not significant for most students. The last question, how do the students learn, is deep and philosophical. It would be too simplistic to claim that e.g. student answers to given exercise assignments accurately measure their learning. But obviously learning is somehow reflected in the answers the students enter to the system. The long term goal is to use this information to obtain understanding of the learning processes involved. How this could be made in practise is discussed next by using control engineering as an example case.

Case Control Engineering

Control engineering is a general discipline, where basic skills especially in mathematics, physics, chemistry and computer science are continuously applied and needed. The main difficulty in learning it is that control engineering – or system theory – is generally considered to be a theoretical discipline. The reason for this can be described by looking at Figure 6.

Figure 6: Research clusters in one control engineering unit

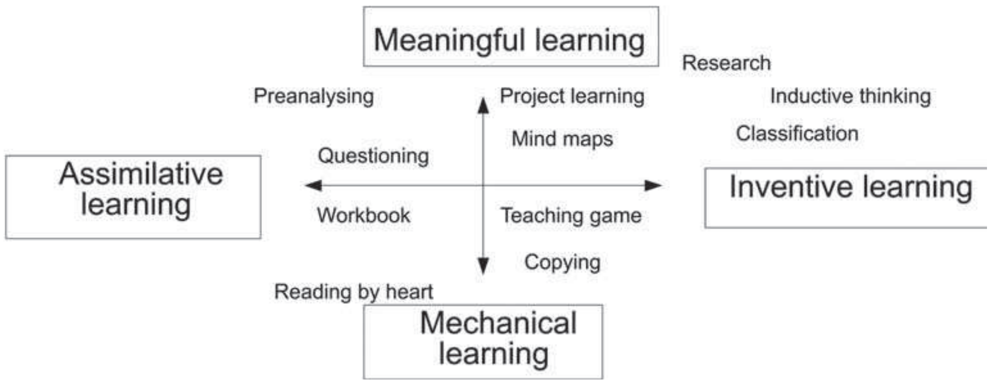


In order to learn control engineering on a wide variety of application areas a system-theoretic framework must be learned and appreciated. System theory is based on the basic sciences – especially mathematics and physics – and therefore the students must learn to apply them on a variety of different contexts. However, it has turned out in practical teaching work that many of the students do not want to learn topics, which they consider to be too theoretical and difficult to grasp. The gap between theory and practice seems to exist year after year.

The dilemma becomes even worse, when dynamic system modelling, control and simulation are needed in new "untraditional" application fields in engineering like biology, medical sciences, business processes etc. Now researchers and students do not form a traditional audience with an engineering background [15, 17]. How to teach control engineering, systems thinking, simulation of dynamical systems, principles of feedback etc. to them is a big and especially important question today [8].

The pedagogical challenge can be analyzed by using e.g. the taxonomy by Ausubel [2] shown in Figure 7. The idea is that the student in the beginning of his/her studies proceeds from mechanical and assimilative learning methods to more meaningful and inventive ones. At the same time the cognitive skills develop and make deep learning possible. This idea can be used in the planning of study programs for advanced disciplines like control engineering. To that end, in the Control Engineering Laboratory at Helsinki University of Technology (TKK) new teaching methods have been developed. These consist of interactive web courses and laboratory exercises over the net. The new idea to be used for training of basic control mathematics and possibly also for carrying out examinations is the STACK-based system described in the paper, now adapted for control engineering purposes.

Figure 7: Learning methodologies.



At the moment there are three web-courses available for control engineering studies at TKK: *Basics in control engineering*, *Analog control* and *Digital control*. The web-courses contain a full set of learning material of the course: textbook divided into learning units, short interactive drill exercises and normal exercises to be done with pen and paper and with a computer. Full solutions are given to each exercise. The material from the web-courses is available at [19, 20].

For more details on the control engineering courses, see [11]. As for the laboratory exercises carried out over the net, see e.g. [9, 16].

The next step in the development of control engineering courses is a training or examination system like STACK. The idea is to teach the basic mathematical tools needed in control engineering studies by using short problems like in the mathematics courses, where the system has already been tested. Suitable simple problems have already been implemented in STACK, and the tests for adapting STACK in control engineering studies is currently in the development stage. Also, including problem solving sessions or examinations with a tool like STACK make it possible to collect large amounts of data of student performance in a course, with the ultimate goal to answer the fundamental question: How do the students learn and how to teach better to motivate for deeper learning. In particular, it is possible to obtain information about long term effects of changes in teaching methodologies.

Future prospects

The plan is to expand usage and testing of the technology to other mathematics courses. As noted above, the pilot course in autumn 2006 concentrated around the technical aspects of the system as well as general questions about how the system should be used. A more ambitious experiment will be held this autumn in a basic course in mathematics for electrical engineering students S1. New features that have been developed since the KP3 experiment will be tested there. We also plan to use this opportunity to gather thorough data about student attitudes and impact to learning results.

In control engineering education at Helsinki University of Technology there is a course "Basic mathematics for control theory", which gives the background of mathematical tools specially needed in the control engineering curricula [18]. Although the course is taught in the classical way (lectures, exercise hours, written exam) there is an independent web-course on the topic also. That included lecture sections, interactive easy problems, problems to do by pen and paper, and problems to do by using computer. All problems have detailed answers. It is known that the web-version is used much by students in different technical universities and universities of applied sciences in Finland. However, as noted above, the course examination is a traditional one, where the students solve given problems to pass the course. Such an exam is arranged four times in a year.

The proposed STACK-assessment system offers totally new views to the teaching of such courses like "Basic mathematics for control theory". It would be possible to drop lectures totally from the course and substitute them by an efficient use of the web-course and a possibility to have lots of practical "drill" problems to the students for practising purposes and eventually for passing the course by STACK-based examination.

REFERENCES

[1] Ala-Mutka, K. A Survey of Automated Assessment Approaches for Programming Assignments, *Computer Science Education*, 15 (2) (2005), 83–102.

[2] Ausubel, D. *Educational psychology: A cognitive view.* (New York: Holt, Rinehart and Winston), 1968.

[3] Delius, G.W. Conservative approach to computerised marking of mathematics assignments. *MSOR Connections* Vol. 4 No. 3 (2004), 42–37.

[4] Douce, C., Livingstone, D. and Orwell J. Automatic test-based assessment of programming: a review, *ACM Journal of Educational Resources in Computing*, 5 (3) (2005) http://doi.acm.org/10.1145/1163405.1163409

[5] Harjula, M. *Mathematics exercise system with automatic assessment*. Master's thesis, Helsinki University of Technology, Department of Automation and Systems Technology. In preparation.

[6] Heck, A. *Assessment with Maple T.A.: creation of test items.* Technical report. URL (referenced 19.8.2007): http://www.adeptscience.co.uk/products/mathsim/mapleta/MapleTA_whitepaper.pdf

[7] MatTaFi project home page http://matta.hut.fi/mattafi/

[8] Murray, R. M. Future Directions in Control, Dynamics, and Systems: Overview, Grand Challenges, and New Courses, *European Journal of Control*, Vol 9, No 2-3, 2003, pp. 144–158.

[9] Pohjola, M., Eriksson, L., Hölttä, V. and Oksanen, T. *Platform for monitoring and controlling educational laboratory processes over Internet*, 16th IFAC World Congress, Prague, Czech Republic, CD Rom format, 6pp, 2005.

[10] Rasila, A. Automaattisesti tarkastettavat tehtävät matematiikan opetuksessa (in Finnish). To appear in *Proceedings of the Researcher Meeting of the Conference ITK – Interaktiivinen Tekniikka Koulutuksessa* 2007.

[11] Riihimäki, V., Ylöstalo, T., Zenger K. and Maasalo, V. A new self-study course on the web: Basic mathematics in control, In (J. Lindfors, Ed.): *Preprints of the IFAC Symposium ACE 2003*, 149–153, Oulu, Finland, 2003.

[12] Sangwin, C.J. New opportunities for encouraging higher level mathematical learning by creative use of emerging computer aided assessment. *International Journal for Mathematical Education in Science and Technology*, 34 (6) (2003), 813–829.

[13] Sangwin, C.J. Assessing mathematics automatically using computer algebra. Teaching Mathematics and its Applications, 23 (1) (2004), 1–14.

[14] Sangwin, C.J. Assessing elementary algebra with STACK. To appear in International Journal of Mathematical Education in Science and Technology.

[15] Sterman, J. D. Business Dynamics, (McGraw-Hill), 2000.

[16] Varso, J. and Koivo, H. Multivariable PI-control of a lamp system, In (J. Lindfors, Ed.): Preprints of the IFAC Symposium ACE 2003, 299–304, Oulu, Finland, 2003.

[17] Wolfram, S. A New Kind of Science, Wolfram Media, 2002.

[18] Zenger, K. Control engineering, system theory and mathematics: the teacher's challenge. To appear in the European Journal of Engineering Education.

[19] Zenger, K., Ajanki A., Maasalo V. and Riihimäki, P. Analogisen säädön *verkkokurssi* (digital self-study material) http://www.control.hut.fi/Kurssit/AS-74.2111/verkkokurssi/

[20] Zenger, K., Härkönen, A. and Maasalo V. *Digitaalisen säädön verkkokurssi* (digital self-study material) http://www.control.hut.fi/Kurssit/AS-74.2112/verkkokurssi/