Intelligent computer-aided assessment in math classrooms: state-of-the-art and perspectives

Wolfgang Müller, *muellerw@ph-weingarten.de*
University of Education Weingarten, Media Education and Visualization Group, Leibnizstr. 3, D-88250 Weingarten, Germany

Christine Bescherer, *christine.bescherer@uni-flensburg.de*
University Flensburg, Institute for Mathematics and Mathematics Education, Auf dem Campus 1, D-24943 Flensburg, Germany

Ulrich Kortenkamp, *ulrich.kortenkamp@ph-gmuend.de*
University of Education Schwäbisch Gmünd, Computer Science Department, Oberbettringer Str. 200, D-73525 Schwäbisch Gmünd, Germany

Christian Spannagel, *spannagel@ph-ludwigsburg.de*
University of Education Ludwigsburg, Institute for Mathematics and Computer Science, Reuteallee 46, D-71634 Ludwigsburg, Germany

Abstract
Research in math education emphasises – since TIMSS and PISA – on extended problem solving skills and mathematical processes. But, to solve complex and real life problems students also need to be proficient in the "basic" skills like solving systems of equations or constructing simple geometrical figures. These more or less technical "routine skills" have to be trained individually and then evaluated by the teacher. Limited resources, however, often prohibit an adequate evaluation of such skills. A possible solution to this problem is a wider application of Intelligent Assessment tools.

In this paper we give a brief overview on of the state-of-the-art in Intelligent Assessment. We present a detailed discussion on the potentials of Intelligent Assessment in the math classroom. A central paradigm we propose is that the use of these tools does not replace a teacher but saves the teacher's time in the necessary training of routine skills, leaving therefore more time to work and assess non-routine problems and projects. From a learner's perspective, Intelligent Assessment would mean that they do not have to follow one of the few "correct" paths to the solution allowed by the software, but can solve the problem following their own approaches. We demonstrate how most of today's approaches to Intelligent Assessment do not fulfil these requirements from a pedagogical vision.

We present three innovative approaches for Intelligent Assessment in the math classroom, coming closer to the presented vision for Intelligent Assessment. These tools support teachers i.e. in checking homework or even part of tests in "sorting out" correct answers or "standard mistakes" like erroneous addition or multiplication while solving equations or estimating instead of constructing a perpendicular line. Based on the pedagogical vision provided and the discussion of the technical state-of-the-art we will formulate open questions from a research point of view. This could be seen as a first step to a research agenda for Intelligent Assessment.

Keywords
Intelligent assessment, routine math problems, inference techniques, computer algebra system, dynamic geometry system

INTRODUCTION AND MOTIVATION

The acceptance of information technology in the classroom is increasing slowly but steadily. There is a growing number of teachers who believe that a computer may not only aid in preparing teaching material, but also represents a valuable teaching and learning tool: computers help to access the world-wide information space via WebQuests, allow a dynamic approach to geometry, enable the modelling and solving of real life problems using spreadsheets or digital image or audio processing or sometimes just make classes simply more interesting. Acceptance will increase further with the understanding that the computer does not restrict the teacher's decisions but assists in deciding, or that using computers does not entail a replacement of the teacher, but helps teachers to focus their attention on situations where it is needed most.[1]

A frequent and usually time-consuming task for teachers is the assessment of students' work. Assessments do not only represent the basis for grading students' work. On a regular basis they are important to the teacher to motivate students, to give feedback to students' learning products and processes and to evaluate the effectiveness of teaching. Assessments can be used diagnostically to identify areas within the course where students still have difficulties. Only if the teacher knows the students' individual problems, he may provide adequate help and support. From this viewpoint, individual assessment is the first step to student-centered teaching. In general, however, a teacher cannot afford such an assessment of students' performances on an adequate level due to limited time resources.

Computer Aided Assessment (CAA) has been proposed as a solution to this problem. CAA refers to a number of approaches to assess students' performance using a computer. CAA promises a big advantage: test results may be analyzed and compared with minimal effort in minimal time. The time and resource savings allow more regular assessments than otherwise possible. As a result, teachers may gain more detailed knowledge of students' progress and may quicker identify problems. Last not least, tests can be tailored to match students' abilities. Using adaptive approaches (i.e. Computer Adaptive Testing) it is also possible to match students' weaknesses as they emerge during the test and to adapt test content correspondingly on the fly (Chalmers and McAusland 2002).

CAA is widely used in distance learning and in higher education, where the number of students is much bigger than in schools, and assessments require even more resources. CAA is also being used for comparative assessments between schools and institutions. Still, CAA hardly made it into the classroom as a teacher's daily tool. There are several reasons for that (Chalmers and McAusland 2002):

- CAA techniques are usually restricted to assess factual knowledge based on objective tests and multiple-choice questions.
- Even when restricting to objective tests, the construction of adequate assessments is very time consuming and requires some specific knowledge. Both time and corresponding knowledge are typically not available at the schoolteachers' level.
- CAA requires the availability of adequate technologies for authoring the tests and performing the assessment. Frequently, this represents another problem for teachers.

In addition, experience from the first phase of an internal project at the Technical University of Berlin, where CAA is used in mathematics courses for students of

[1] We focus on using ICT in a "real classroom" in face-to-face settings. In this paper we are not concerned with online courses.

engineering, has shown that CAA might need even more personal resources for assisting students than paper-based testing due to technical overhead.

While assessment is a problem for all subjects in this paper we will concentrate on mathematics[2]. In mathematics the correctness of a solution is usually much more easily agreed on than in other subjects. On the other hand the focus on assessment in mathematics classroom is shifting rapidly. Especially in view of the shifting emphasis to process oriented mathematics education – contrary to math problems where only the numerical solution is relevant – there are currently two main lines of action. One area of research is based on the approach to further the use of open-ended problems. Assessing such open-ended problems however is not easy and often not practical: if students are supposed to find their own and maybe unique way to solve a problem and communicate their ideas and reasoning, a multiple choice test can hardly capture the full range of these competencies. Unfortunately, teachers do not have the time to read and grade maybe 50 extended essays a week. The other line of action concerns the resource problem of the math lessons: open-ended and complex problem solving takes a lot of time. But still the basic technical math skills have to be trained, because without them process-oriented skills cannot be applied. This training can be done in self-directed learning settings if adequate feedback is provided.

Intelligent Assessment refers to approaches to overcome the restrictions of CAA. In the meantime, some work has been published claiming to provide intelligent solutions in this field, for instance:

- Various approaches have been proposed to allow for free-text assessment of writings, such as the Project Essay Grade (Page 1966) or the Intelligent Essay Assessor (Foltz et. al. 1998). A detailed description of such approaches can be found, for instance, in Whittington and Hunt, 1999.
- In the context of math education there have been some approaches to adapt the "fill in the blanks" metaphor to mathematics. An example for such an approach represents (Patel et al. 1998).
- There are various approaches to utilize Computer Algebra Systems (CAS) for math assessment in a Web environment (e.g., Klai et al. 2000, Maplesoft 2006). In general, the integration of CAS corresponds to the integration of inference techniques in assessment, and typically this is being used to allow specifying possible solutions in a more general and more abstract way. There are only few examples going beyond such an application of CAS.
- Another approach to Intelligent Assessment is the integration of simulators into assessment. Simulators are typically designed for very specific application fields. As a result, the utilization of simulators for assessment requires typically complex integration steps, and the application of such assessment tools remains limited. There are, however, successful examples, such as the Java eLearning Simulations (JeLSIM, Thomas et al. 2004), a set of Java-based simulation components, which can be integrated for assessment into the PASS-IT Assessment Engine. Another very specific example is an interactive computer graphics course where the results of graphical algorithms specified terms of Java Beans are matched against expected prototypes (Klein et al. 1998).[3]

In summary, the number of approaches to Intelligent Assessment, especially in mathematics, is – despite its importance for the new process-oriented teaching and

[2] One of the reasons is definitely that three of the authors work in mathematics education.
[3] Apparently, this is closely related to the test-driven approach of software engineering (cf. Beck & Gamma 2006)

learning styles – limited. In general, most of these approaches do not consider process knowledge at all during assessment. The exceptions are simulators, but their application in school mathematics is limited and their integration in assessment is complex and expensive.

In the following chapters we will discuss requirements, approaches, and problems to Intelligent Assessment in mathematics education in more detail.

REQUIREMENTS FOR INTELLIGENT ASSESSMENT IN MATHEMATICS EDUCATION

Basic mathematical skills like mental arithmetic, solving linear equations, or basic geometric constructions are absolutely necessary for solving complex mathematical problems. But working on complex open-ended problems takes a lot of time and effort. Often there is not enough time for training these basic skills and problem-solving skills. Intelligent Assessment allows students to practice these techniques while getting "helpful" feedback, for instance, by feedback like "routine mistake while doing multiplication", "no progress in solving the equation system detected", and so on.

Students can use this feedback also to practice basic skills. Learning by trial and error is one of the first ways small children learn when they start to crawl, walk, or eat with fork and knife. It is obviously still used while working with online tutorials, which give automatic feedback (Baruah et al., 2005). The effects of trial and error in the development of mental models in i.e. mathematics are not fully understood yet, but trial and error is undeniably used quite often in learning systems with automatic feedback features.

Another aspect is that the complete solution to a mathematical exercise is not described by its final result alone. All the steps that lead to the final result are important. This is true for complex exercises, but also for routine skills like solving linear equations or, much more basic, the mental arithmetic of "288 – 154". If one's aim is to encourage students to find their own "best strategy" for solving problems like these, then *all* the different possible – and correct – solutions have to be evaluated as such. In a lot of training programs for mathematics it is still quite common that correct answers obtained by alternate solution strategies are evaluated as "incorrect".

If a solution of a complex open-ended math problem is correct then it's assumed that the student has the adequate process skills und techniques of solving this kind of problems[4]. And it is not important – and nearly impossible to find out without i.e. extended interviews – whether all the mathematical concepts are fully understood by this student. Only if there are faulty solutions there is a need to work on the skills (process and otherwise) of the learner.

In teaching mathematics, students' mistakes are very good indicators of misconceptions in mathematical thinking. "Only by examining misconceptions and errors can students [and teachers] deal with them appropriately." (NCTM 2000, p. 272). And "the teacher could look at students' incorrect observations and design a lesson to address those misconceptions. In this way, the students' knowledge becomes a starting point for instruction [...]." (NCTM 2000, p. 351). If mistakes are

[4] Of course only if he can solve similar problems as well he'll be considered a good math student.

used in this way the correct answers are not interesting – except for the number of correct answers.[5]

A problem with recording mistakes for feedback or assessment is the sheer number of possible "wrong turns." "Mistakes are of an individual nature and in their possible number unlimited."[6] These are exactly some of the reasons not to use *automatic* but *Intelligent* Assessment. If the focus is on the individual processes of problem solving a lot of different solutions will have to be recognized as correct (even if the teacher has not thought up this specific one before).

Also, there are different levels of seriousness of mistakes. In solving linear equations or linear equation systems there are some recurring errors, for instance mixing-up of addition and multiplication, non-application of the distributive law, or wrong elimination of parenthesis, which do not have anything to do with the solving of linear equation systems in itself. If some students have special problems with doing the multiplication they should practice their multiplication tables – using intelligent assessment – before returning to equation solving.

In summary, intelligent assessment has a wide field of possible applications in math education. As seen below there are some first realizations using intelligent assessment in schools but it is just coming up as a topic for development and research. On the other hand the use of specific software in the math classroom like spreadsheets, Dynamic Geometry Software or Computer Algebra Systems is increasing steadily.

TECHNICAL APPROACHES FOR INTELLIGENT ASSESSMENT IN MATHEMATICS EDUCATION

The previous discussion shows that there are a number of requirements for Intelligent Assessment from a pedagogical point of view. Technical feasibility presents the other side of the coin. While multiple-choice questions represent the state-of-the-art in today's e-learning systems, concepts such as intelligent tutoring are not as easy to implement. Nevertheless, a number of promising approaches have been presented lately. We will show three interesting examples from different areas of school mathematics for such Intelligent Assessment systems. The central ideas of these approaches will lead us to some open research questions in the final section.

Example 1: Cinderella

There is the intelligent tutoring system[7] of the interactive geometry software Cinderella (Richter-Gebert et al. 1999) using an automatic geometric theorem checking technique. The Dynamic Geometry System (DGS) acts as an authoring tool for geometric construction exercises. The teacher is required to do a sample solution, and then one or several checkpoints in the construction sequence may be defined, together with individual help and comments for students that may be stuck. Later, when doing the exercise, students can ask for a tip. The software then decides, based on the geometric objects that were already constructed and have certain geometric properties, which tip might be appropriate and either shows a text, opens a webpage, or adds missing elements to the construction. The same data is

---

[5] If 90 % of the students gave correct answers there is no need to spent more time on the topic. This is different if only 10 % gave correct answers.

[6] Translated by the authors from a study on "Increasing the Efficiency of the Mathematics and the Science Classrooms", BLK 1997, Chap. 3.6

[7] For a more detailed description we refer to (Kortenkamp & Richter-Gebert 1998, Richter-Gebert & Kortenkamp 1999, Kortenkamp 1999).

also used without request all the time during the solution process to monitor the students' progress, give additional information, and finally to check whether the desired result has been reached.

Despite the fact that the geometric construction exercises are much more advanced than, say, a multiple-choice solution, they still leave much room for improvement. The software does not have a mental model of the learner, but it can only use a single reference solution for comparison of the *output* the learner generated. This is enough to get students "back on track", but it does not support individual help "off track". Thus it helps to allow for individual solutions, but it does not *encourage* them. This drawback has been removed with the new release *Cinderella.2*, which supports multiple sample solutions, and can also give automatic feedback for common errors that were anticipated by the exercise author.

Although it is now possible to create even more helpful guided exercises, their use is mainly targeted at homework or other unsupervised learning situations. Students can get help for their specific needs, but the teacher does not get much information about the students' progress. The final product of a student with severe difficulties and the product of a student who easily handled the exercise are indistinguishable. In other words: There is no feedback channel – no information about the learners' progress, misconceptions, etc.

Another new and experimental feature of Cinderella was designed to remedy this situation. The CINErella module (Kortenkamp 2005) uses semantic event stream recording in order to provide record and playback features for the whole solution process. The most basic use of this module is to prepare animated presentations of evolving geometric constructions, in order to demonstrate solutions or problems. Both the teacher and the learner can do this easily and save their solutions to the web, where they can be accessed both in classroom situations and virtual configurations over the network.

For the teacher, it does not scale to use this mode of operation with all students. If in a class of 30 students everybody prepares a 3-minute demonstration, it will take at least 1.5 hours only to have a look at each of them, and this does not take into account that the analysis of the problematic cases that will take more time. The time necessary for analysis can hardly be reduced, but we can use the fact that we are recording *semantic* events to avoid the manual (or visual) scan through all the solutions.

Using the information about the construction, which is now available for the whole sequence of construction steps, and statistic information (for example about the movements of the mouse, the creation and deletion of elements, undo/redo operations, or the time it takes to perform a certain subtask) derived from the semantic event stream, we can

- Cluster demonstrations that share common patterns, and thus might give a hint on misconceptions,
- Find extreme – good or bad – solutions, that might be used in explanations or serve as examples for the class, or
- Quickly identify students who need individual help.[8]

---

[8] The data could also be used to grade students, but extreme care has to be taken here. For example, it is not at all clear that a quicker solution of a task shows that one student is better than the other. Actually, it only shows that he or she found the solution faster than the other.

In conclusion, CINErella allows for monitoring the complete learning process while using a DGS, which was only possible using video recording (which does not provide the semantics und thus is far inferior) or in person (which lacks the persistence of a recording, interferes with the students and influences their performance, and simple is not possible for a single teacher in a classroom). If we combine this with the automatic theorem checking approach, we can build a system that is capable of semi-automatic assessment in geometry education.

Example 2: Saraswati

Saraswati (Bescherer et al. 2004) is an example for an Intelligent Assessment system in the field of algebra, more specifically for finding solutions of linear systems of equations. Similar to CINErella, Saraswati targets to grasp the process of finding a solution, not only the correctness of the final solution. In this case, the correctness of individual transformation steps is also validated and in case of errors specific information is made available on the type of error and possible corrections. The Saraswati approach has certain similarities to the APSLUSIX system (Chaachoua et al. 2004) or Mathlantis (Cornelsen 2005). The Saraswati use case, however, resembles more an assessment system than these two, allowing for continued transformations even in case of an error and providing detailed statistics for a class and individual students.

The Saraswati system itself provides more than an assessment system, but a complete framework for homework authoring, distribution, solving, assessment, and analysis. Three different components provide this functionality:

- The authoring component, where the teacher can compile exercises,
- The assessment client, in which students may interactively solve and edit the exercises (see Fig. 1), and
- The analysis component, which provides a detailed analysis of the collected students' solutions.

As mentioned before, Saraswati's assessment component does not only check the correctness of the final solution, but also verifies the correctness of the transformations leading to the solution. The general principle applied here is the assertion of the equivalence of the solution space for a transformed linear system and the untransformed, previous one. This is controlled with the help of an integrated computer algebra system. In case the solution spaces are found to be different in the path of solving a linear system of equations, an additional analysis is triggered to classify the error. This is achieved by taking a heuristic approach, where the identified error is tested against a set of most relevant error classes.

For instance, a student might introduce an error in one of the equations of the system when dividing both sides of the equation by invalid cancelling of a single factor. In this case the solution spaces of the linear system of equation before and after the transformation will differ. Iterating through the equations of the new system and exchanging the equation with the corresponding equation of the system before the transformation, it is possible to test whether the solution space can be corrected. Thus, the equation line where the error occurs first can be identified. For the equation introducing the error we can try whether a correcting factor for one of the variables will re-establish the solution space. If this is true, we can infer a simple error in a multiplication or division of a factor. Similarly, typical errors in solving linear systems of equations, such as wrong additive components, loosing an equation, or loosing a variable can be identified. Saraswati also identifies in how far the student succeeded to solve the exercise, that is if the system of equations was completely solved, whether a variable was isolated, or whether no progress could be detected.

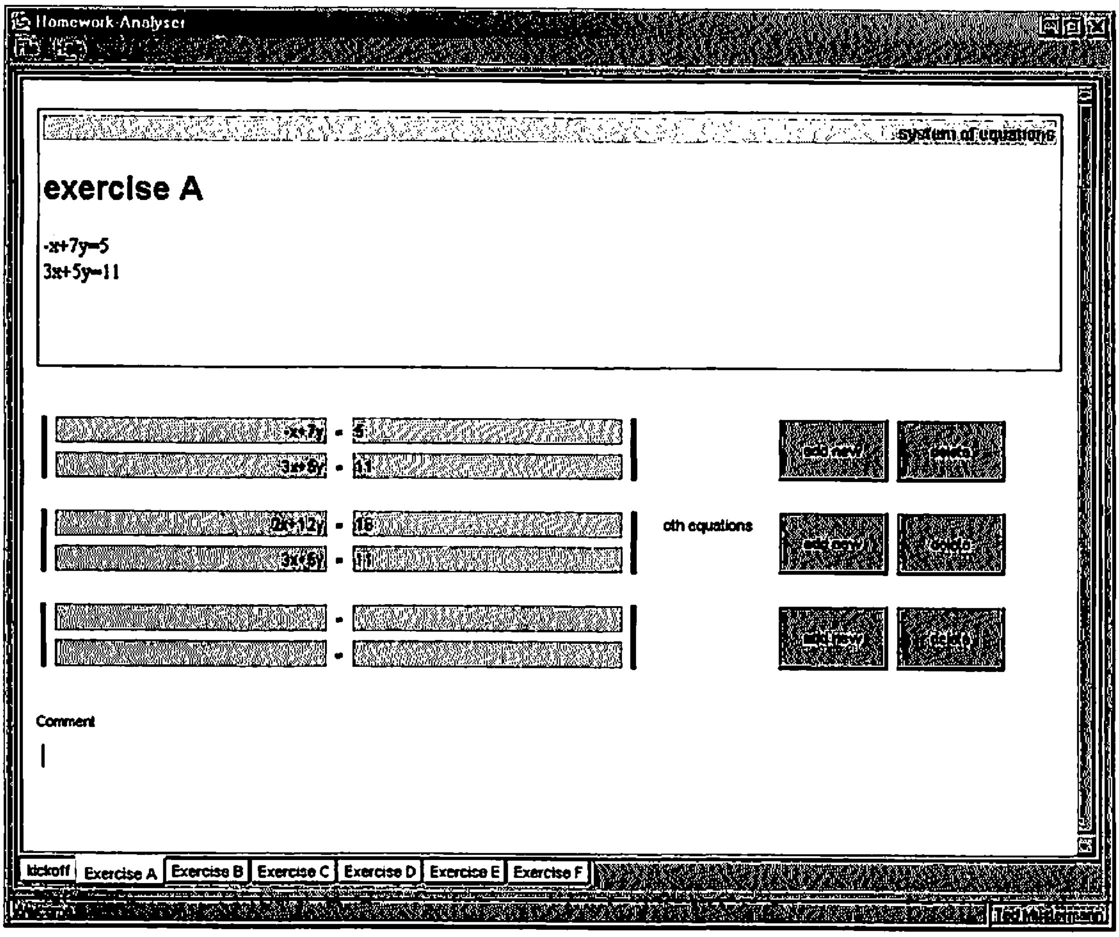


Figure1: The Saraswati assessment client presenting an exercise and text fields to enter the equation system, its transformations and the final result.

A very important contribution of Saraswati is the paradigm of dividing a student's solution process into discrete steps and asserting the correctness of the path by comparing the solution space after a transformation step to the initial and expected solution space. This paradigm can be easily mapped to other domains, even outside the field of mathematics. Consequently, it can be seen as general approach to acquire and assess process knowledge.

Example 3: Explorations on the number line

Yet another approach to Intelligent Assessment is the use of an augmented capture and replay tool (CleverPHL[9]) to obtain data on on-line behaviour of students working in microworlds (cf. Klaudt 2003). The objective of this thesis project was to infer the students' mental representations of numbers by observing how they directed the mouse pointer to the correct spot. The mental model of the number line and solving problems using the number line are very important in the development of a number sense. Therefore the knowledge of the students' mental models in a class is essential to math teachers in primary schools. Figure 2 shows a typical task for the first graders.

[9] CleverPHL is part of the Jacareto capture and replay toolkit and can be downloaded at http://jacareto.sourceforge.net/, last visited: 15.1.2006

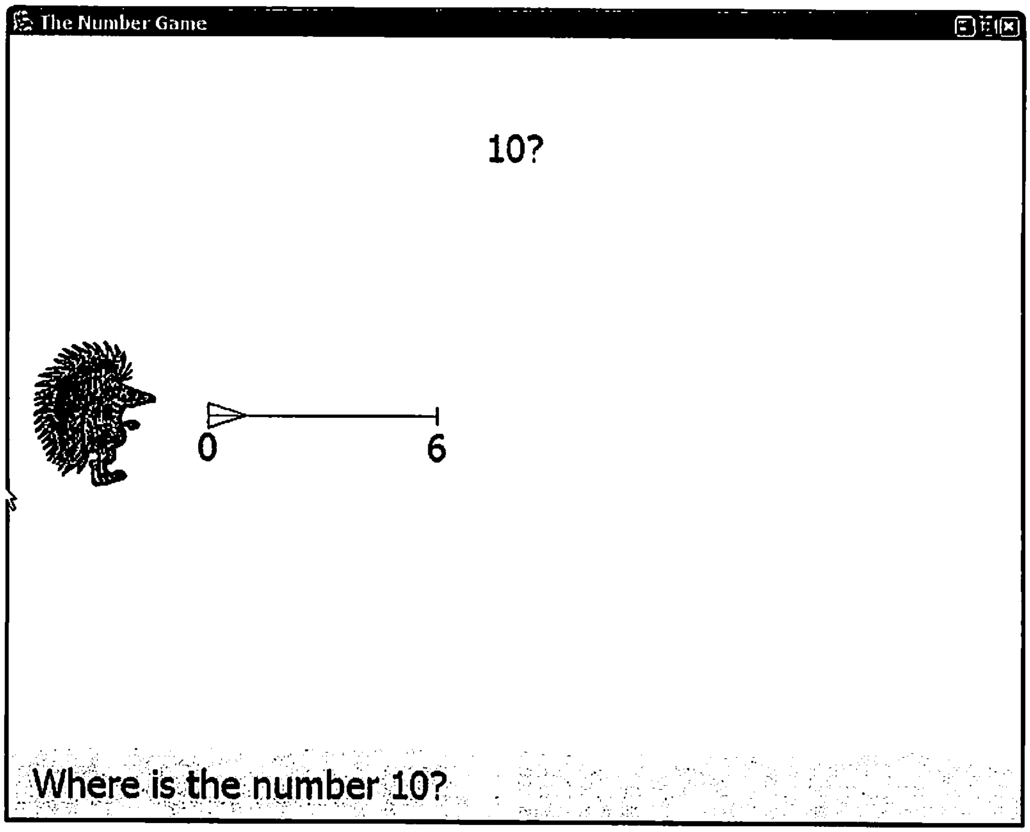


Figure 2: Typical task of the LOGO microworld

The first graders were set to find the "treasure" hidden at the place where the "10" is. Some students just "counted" from 0 onwards – they just moved the mouse pointer in six nearly equidistant steps from 0 to 6 and then took four more steps to the 10. Others used also the counting strategy but started at the 6. Still others used a doubling strategy – they doubled the distance from 0 to 6 and then moved a little bit to the left. Using a capture and replay tool the teacher would be able to replay all the students mouse moves and mouse clicks. But this would take a long time and simply can not be done in a normal school day.

For this reason – among other things – the capture and replay tool CleverPHL has been developed which offers methods to record user actions, to replay and to automatically analyze them (Spannagel, Gläser-Zikuda and Schroeder 2005). With CleverPHL, user actions performed in the microworld can be captured and stored in chronological order. The list of captured actions (also called "interaction record") can be replayed in order to watch the recorded behavior again. In addition, interaction records can be automatically structured by categorizing sub-sequences of actions. For example, given an interaction record created in the microworld, the list can be split up between actions belonging to successive tasks. The resulting structure can be refined by grouping actions together which belong to a single click on the number line. The result is a hierarchical representation of the interaction record, containing information about the user behavior at different levels of detail. Figure 3 shows the first part of an unstructured interaction record which goes on quite some more, and the same record automatically structured (right).

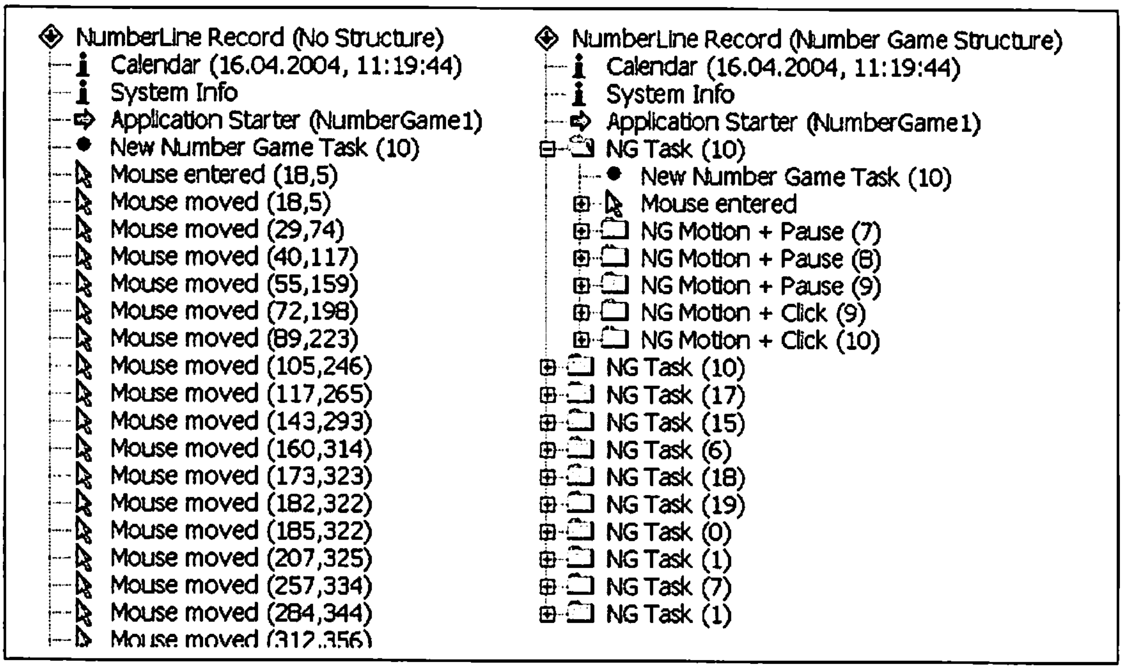


Figure 3: (a) The first part of an unstructured interaction record and (b) the same record automatically structured

Furthermore, CleverPHL allows for automatically categorizing action sequences. For example, given the sequence of all actions belonging to one task in the microworld, CleverPHL can automatically detect the strategy which has been used to find the number. At the moment, the "direct hit" and "sequential search" strategies are recognized automatically. The detection of other operational strategies ("double the distance and count") are not yet implemented. However, the necessary software infrastructure and the important possibility of using special events defined by the teacher as a marker for a new problem or some more help or whatever are present and could be exploited further.

Given a set of structured interaction records, CleverPHL allows for extracting quantitative data. With this feature, information about a whole group of students can be gathered and processed in statistical software. The record structure may further be used as index to replay only one part of the captured user behaviour. The possibility to selectively replay sub-sequences of interaction records saves time, because the teacher must only screen the actions s/he is interested in.

CleverPHL can be used in conjunction with Java applications and applets. New categories for special user behavior can be added by implementing detection algorithms which can be integrated into CleverPHL's framework. The combination of Java-based software with specific capture and replay tools allows the use of a wide range of subject-specific software, which is in itself a merit. The symbolic and therefore automatically processable representation of interaction records is the prerequisite to all kinds of analysis and assessment.

DISCUSSION

The examples above show that Intelligent Assessment can be much more than just basic multiple-choice checkers or simple free text comparison. The examples differ from other approaches to assessment frequently referred to as "intelligent". In our understanding Intelligent Assessment in mathematics education means:

- Mathematical processes are being made detectible and visible. Without this, a discussion about improving competencies in mathematics is fruitless.

- The workload of teachers in assessing solutions of open-ended problems is reduced, and therefore teaching and learning scenarios that foster mathematical process skills are supported.
- Students are permitted to follow their own strategy of solving problems rather than following a given recipe blindly.

Moreover, we are able to specify the characteristics that lead to these enhancements of Intelligent Assessment. Although our three examples are coming from very different areas of K-12 mathematics education, they share four common principles:

- Students do not choose from a set of answers, but can give free answers by either entering text or performing actions with the mouse.
- The process of compiling a solution to an exercise is recorded and validated.
- The process of finding a solution is divided into discrete entities that make up a series of semantic events.
- The correctness of each step is not checked literally, but by testing each intermediate result for mathematical equivalence with the desired solution space.

We see the above principles as a starting point for a *new paradigm* for the design of Intelligent Assessment systems that take into account process knowledge. Still, for a broader application of this paradigm there remain a number of open research questions. We consider the following problems crucial from a pedagogical point of view, and they should be addressed in future research:

- We need a "theory" of Intelligent Assessment in mathematics education, which describes categories of mistakes in general and in special mathematical fields like algebra or geometry. A similar classification is necessary for process skills.
- This theory needs to be extended to the level where processes are divided into discrete steps. This involves, for instance, a distinction of stages in the solution process, such as generating solutions, evaluating and selecting the appropriate strategy, planning and then undertaking an activity that solves the problem (Thomas et al. 2004). For specific solution processes a further split-up to individual solution steps is necessary, as in the case of solving linear systems of equations.
- The single problem solving steps have to be related to mechanisms for identifying possible errors. A somehow general approach would be the description of expected solution spaces that could be exploited using automatic theorem proving or unit tests.
- Error classes and patterns have to be identified on this discrete step level. If possible, these error classes should be independent of the specific subject and problem. This implies that we need a meta-description of semantic events. It remains an open question, however, whether we can identify problem solving strategies across different subjects.

Finally, from a technical point of view there are still a number of problems to be solved, for instance: Technologies for process recording need to be developed (or improved) and applied in different application contexts; the easy interfacing to statistical treatment has to be ensured. Yet, the recording of the processes alone does not solve the problem. The detection of meaningful events in the recorded process path is still a challenging problem. In addition, system events have to be mapped to accomplished steps in solving the problem.

CONCLUSION

Undoubtedly there is a need for computer aided assessment for quality improvements in teaching. The recent approaches of Intelligent Assessment are promising, but they lack a generality that could be used to transfer the results to other subjects. In particular, process orientation and its implications on individual assessment can help to overcome this situation. Based on the three examples from math education presented in this article, we can derive general guidelines for creating Intelligent Assessment solutions. The separation of the solution process in series of discrete steps, the robustness against minor mistakes in intermediate steps, and the openness of the assessment environment are key ingredients for assessments that give proper feedback about the concepts and misconceptions of the learner while taking away routine and time-consuming tasks from the teacher.

REFERENCES

Baruah, N., Ellis, E., Gill, M. & Greenhow, M. (2005) Recent Developments in Setting and using Objective Tests in Mathematics using QM Perception, in Proceedings of CAA Conference 2005, online at http://www.caaconference.com/pastConferences/2005/proceedings/BaruahN_EllisE_GillM_GreenhowM.pdf, retrieved 4-6-2006.

Beck, K.; Gamma, E. (2006) Test Infected: Programmers love writing Tests. Junit documentation, online at http://junit.sourceforge.net/doc/testinfected/testing.htm, retrieved 4-6-2006.

Bescherer, C.; Müller, W.; Heinrich, F; und Mettenheimer, S. (2004) Assessment and Semi-Automatic Analysis of Test Results in Mathematical Education, in Proceedings of ED-MEDIA 2004 - World Conference on Educational Multimedia, Hypermedia & Telecommunications, Lugano.

BLK Bund–Länder–Kommission für Bildungsplanung und Forschungsförderung (1997) Gutachten zur Vorbereitung des Programms "Steigerung der Effizienz des mathematisch–naturwissenschaftlichen Unterrichts". Materialien zur Bildungsplanung und zur Forschungsförderung, Heft 60. Bonn online at: http://www.blk-bonn.de/papers/heft60.pdf, retrieved 4-6-2006.

Bull, J.; and McKenna, C. (2003): Blueprint for computer-assisted assessment, Routledge Falmer Pub.,

Chaachoua, H.; Nicaud, J.F.; Bronner, A.; and Bouhineau, D. (2004) APLUSIX, a learning environment for algebra, actual use and benefits, Proccedings of 10th International Congress on Mathematical Education (ICME 10), Copenhagen, Denmark.

Chalmers, D., and McAusland, W.D.M. (2002) Computer-assisted Assessment. In: The Handbook for Economics Lecturers: Assessment. Edited by J. Houston and D. Whigham. Economics LTSN, online at www.economics.ltsn.ac.uk/handbook/, retrieved 4-6-2006.

Cornelsen (2005) Mathlantis, Cornelsen Verlag. http://www.cornelsen.de, retrieved 4-6-2006.

Foltz, P. W.; Landauer, T.K.;and Laham, D. (2004) The Intelligent Essay Assessor: Applications to Educational Technology, Interactive Multimedia Electronic Journal of Computer-Enhanced Learning, Vol. 2, No. 4, online at http://imej.wfu.edu/articles/1999/2/04/, retrieved 4-6-2006

Klai, S.; Kolokolnikov, T.; van den Bergh, N. (2000) Using Maple and the Web to grade Mathematics Tests, Proc. International Workshop on Advanced Learning Technologies.

Klaudt, D. (2003) Exploring number sense—a research project in Primary School using Logo. In Gerald Futschek (ed), *EUROLOGO 2003*, Be creative ... re-inventing technology on education, Proceedings of the 9th European Logo Conference, Porto, Portugal, pp. 62--68. Coimbra, Portugal: CNOTINFOR.

Klein, R.; Hanisch, F. (1998) Using a modular construction kit for the realization of an interactive Computer Graphics course, in: Th. Ottmann, I. Tomek (eds) Proceedings of The World Conference on Educational Multimedia and Hypermedia (ED-MEDIA) and the World Conference on Educational Telecommunications (ED-TELECOM), Association for the Advancement of Computing in Education (AACE), USA, June 1998.

Kortenkamp, U. (2005) Experimentieren und Publizieren. In: Peter Bender, Wilfried Herget, Hans-Georg Weigand, Thomas Weth (eds) "WWW und Mathematik - Lehren und Lernen im Internet?" Proceedings 21. Arbeitstagung des Arbeitskreis Mathematikunterricht und Informatik in Dillingen/Donau, Franzbecker, Hildesheim.

Kortenkamp, U. (1999) Foundations of Dynamic Geometry, Dissertation, ETH Zurich, Department of Theoretical Computer Science, November 1999, http://kortenkamps.net/papers/1999/diss.pdf, retrieved 4-6-2006.

Kortenkamp, U.; Richter-Gebert, J. (2004) Using Automatic Theorem Proving to Improve the Usability of Geometry Software, Proceedings of MathUI 2004, online at http://kortenkamps.net/papers/2004/ATP-UI-article.pdf, retrieved 4-6-2006.

Kortenkamp, U.; Richter-Gebert, J. (1998) Geometry and Education in the Internet Age, In: Proceedings of ED-MEDIA 98, Freiburg, Germany, AACE, 1998, http://www.cinderella.de/papers/geo-i.pdf.gz, retrieved 4-6-2006.

Maplesoft (2006) MapleTA - Automated Online Testing & Assessment, online at http://www.maplesoft.com/products/mapleta/, retrieved 4-6-2006.

Müller, W.; and Bescherer, C. (2005) Saraswati – Ein System zur elektronischen Leistungserfassung und semi–automatischen Fehleranalyse im Mathematikunterricht, in Engel, J., et al. (eds): Strukturieren - Modellieren - Kommunizieren. Leitbilder mathematischer und informatischer Aktivitäten, Festschrift zur Verabschiedung von Prof. Löthe, Prof. Krauter, Prof. Klose und Prof. Dr. Wölpert, pp. 293–300. Franzbecker Verlag, Hildesheim.

NCTM (2000) National Council of Teachers of mathematics: Principles and Standards for School mathematics. Reston, Virginia, USA.

Page, E. B. (1966) The imminence of grading essays by computer. Phi Delta Kappan, January, pp. 238--243.

Patel, A.: Kinshuk, K.; Russell, D. (1998) A computer based intelligent assessment system for numeric disciplines, Information Services & Use, Vol 18, No. 1-2, pp. 53–63, Amsterdam.

Richter-Gebert, J.; Kortenkamp, U. (1999) The Interactive Geometry Software Cinderella. Springer-Verlag, Heidelberg, 1999. http://cinderella.de retrieved 4-6-2006.

Spannagel, C; Gläser-Zikuda, M & Schroeder, U. (2005) Application of Qualitative Content Analysis in User-Program Interaction Research [39 paragraphs],. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research [On-line Journal], *6*(2), Art. 29. Available at: http://www.qualitative-research.net/fqs-texte/2-05/05-2-29-e.htm, retrieved 4-6-2006.

Thomas, R.; Ashton, H.; Austin, B.; Beevers, C.; Edwards, D.; and Milligan, C. (2004) Assessing Higher Order Skills Using Simulations, Proc of CAA Conference 2004, online at http://www.caaconference.com/pastConferences/2004/proceedings/Thomas_R.pdf, retrieved 4-6-2006.

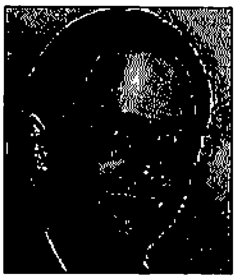
Whittington, D.; and Hunt, H. (1999) Approaches to the computerized assessment of free text responses, Proceedings of the Third Annual Computer Assisted Assessment Conference, pp. 207--219, online at: http://citeseer.ist.psu.edu/697625.html, retrieved 4-6-2006.

Biographies

Wolfgang Müller is a Professor for Media Education and Visualization at the University of Education Weingarten, Germany. His special interest is in E-learning Technologies, Game-based Learning, and Information Visualization.
Wolfgang Müller holds a Diploma in Computer Science and a doctorate in Computer Science (Dr.-Ing.) from Darmstadt University of Technology.

Christine Bescherer is a Junior Professor for Mathematics Education at the University of Flensburg, Germany. She holds a doctorate in Mathematics Education. Her special interest is in computers in mathematics education.

Ulrich Kortenkamp is Professor of Media and Computing at the University of Education Schwäbisch Gmünd. Before that, he held a position in Mathematics Education at the Technical University of Berlin, and was scientist in charge for the application area G (Education) at the DFG Research Center Matheon. He is also co-author of the Interactive Geometry Software Cinderella.



Christian Spannagel studied computer science and is now Ph.D. student at the Institute for Mathematics and Computer Science, University of Education Ludwigsburg. He is co-author of the software Jacareto/CleverPHL.

Copyright Statement
This work is licenced under the Creative Commons Attribution- NonCommercial-NoDerivs2.5 License. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/2.5/ or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.