Design of Erroneous Examples for ActiveMath

Erica Melis
German Research Institute for Artificial Intelligence (DFKI)
66123 Saarbruecken, Germany
melis@dfki.de

Abstract. The behaviorist view of learning that informs much of traditional schooling is not likely to invite students and teachers to see errors in a positive light. This is particularly true for mathematics. Our goal is to change this situation by including erroneous examples and other error-related learning opportunities in ActiveMath.

This paper investigates the systematic design of erroneous examples. For this, it analyzes the potential benefits that erroneous examples can have and distinguishes different presentation patterns. This analysis together with first experiences from school and from a university course with ActiveMath informs further research on effects, adaptive choice and presentation of erroneous examples in ActiveMath.

1. Introduction

The behaviorist view of learning that informs much of traditional schooling is not likely to invite students and teachers to see errors in a positive light. Behaviorism assumes that learning is enhanced when correct responses are rewarded (positive reinforcement) and incorrect ones are either punished or extinguished through lack of attention (withholding of positive reinforcement) [8]. Approaches to use errors as learning opportunities may help to overcome the traditional transmission view of mathematics teaching and learning.

Within the traditional framework, paying explicit attention to (mathematical) errors in class is even considered by many as dangerous since it could interfere with fixing the correct result in the student's mind. Indeed, the effectiveness of erroneous examples for different kinds of learners is an open issue and may depend on the individual learner. [15] investigated teachers' point of view on this and other issues with no conclusive results.

We know only of little research in psychology [12,5] which targets learning with erroneous examples. Some research in maths education addresses learning from errors that others made or that are deliberately introduced [1,9,14]. Mostly, these describe positive and creative reactions of teachers to student errors in the classroom which may be hard to implement in a learning environment. Hart [6] addresses the need to diagnose the learner's misconception (rather than the teacher's conceptions) for a proper reaction.

An intelligent system should use its potential to work with errors productively. One way to do this is through providing feedback on errors the student made. Another way is to include erroneous examples – a rather unusual type of exercises – into the learning experience.

This paper reports first steps and experiences with erroneous examples in the adaptive learning environment ActiveMath [7]. This sets the stage for other computational issues such as generaltion of erroneous examples and adaptive choices. It investigates dimensions for the systematic design of erroneous examples. For illustration, the paper includes examples from our fraction course (school) and the derivatives course (university) which both are available online.

We would like to stress that the described design of erroneous examples does not primarily target the design of erroneous examples for lab experiments. Presumably, for this a more fine-grained tweaking is needed to obtain statistically significant results in a limited time-on-system.

2. Targeted Dimensions of the Learning Process

Including erroneous examples as exercises into a learning experience can serve several purposes:
(1) improvement of learner's motivation [14] and influence on students' attitudes towards failure and success.
(2) Proper understanding of concepts which includes conceptual change in case of a misconception [13] and understanding concept's boundaries. For concept learning, previous research indicates that people tend to use positive instances and ignore negative instances, see, e.g., [2]. This is an inefficient strategy. One measure to push students to look at negative instances is to require an explicit work on erroneous examples.
(3) Improve reasoning capabilities, e.g., the correct application of rules and the application of correct rules as well as hierarchical/structured problem solving.
(4) Train meta-reasoning including critical thinking, self-monitoring, and enforce self-explanation [12] to judge solution steps as correct or faulty. Meta-cognitive skills are required to overcome the barriers imposed by the student's prior knowledge and conceptions [10], and finding and correcting errors in an example can stimulate and prompt meta-cognitive activities. Critical thinking is sometimes neglected for mathematics and its applications. However, in real life people have to be able to judge whether a mathematical result is acceptable or to discover the conditions under which it is correct. They have to be able to find out the reason for an error. Learning should therefore, target this capability.
(5) Encourage exploration. Borasi [1] reports striking experiences on how even below-average students start questioning and exploring mathematics, when confronted with an error and encouraged to dwell on it.
(6) Change attitudes. In the traditional classroom there is not much room for being wrong, not even temporarily. Schoenfeld [11] reports that most students believe that if you can't solve a problem in a few minutes, you can't solve it at all. A mistake is interpreted as an ultimate failure and there is little room for experimentation (and debugging). When guessing, experimenting and playing

with partially correct conjectures are discouraged, the only remaining alternative for many students is getting stuck. Schoenfeld concludes that this attitude is an important factor in students' inability to cope with non-routine problems.

3. Design of Erroneous Examples

It is an art to design erroneous examples that include an obvious inconsistency and provoke conflicts. The most relevant variables for the design of an erroneous example are the actual error/misconception addressed and the example's actual presentation. The first is addressed implicitly in the examples below because it depends on the domain and on the typical errors that occur, the second is explicitly addressed.

There are several types of (typical) errors including buggy rules, misconceptions, and frequent slips such as wrong labels for quantities. A vast pedagogical literature about typical errors exists for school mathematics, e.g., for computation with fractions [4]. They collect and analyze procedural errors as well as misconceptions, e.g., [13].

As for the presentation, alternatives of the following Derivation Erroneous Example are described below.[1] In section 5 we summarize observations on when which presentation seems appropriate.

Eve wants to compute the derivative of the function: $y = \frac{1}{(1-2\cdot x)^2}$ *for* $x \neq \frac{1}{2}$. *Her solution contains one or more errors. Please find the first error.* [2]

Eve's solution: since $x \neq \frac{1}{2}$ *holds, the function is differentiable in its domain. She uses the Chain Rule for computing the derivative.*
The Chain Rule states that the derivative of a composite function $f \circ g$ *can be calculated as follows* $(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$.
Eve chooses $f = \frac{1}{g^2}$ *and* $g = 1 - 2 \cdot x$
Now, Eve calculates the first factor $(\frac{1}{g^2})'$.
She begins with rewriting $f = \frac{1}{g^2}$ *as* $f = g^{-2}$ *which leads to* $f'(g) = (-2)g^3$
Then she calculates the second factor: $g'(x) = -2$.
Finally, she combines the factors as follows:
$(f(g(x))' = (f \circ g)'(x) = (-2) \cdot (1 - 2 \cdot x)^3 \cdot (-2) = 4 \cdot (1 - 2 \cdot x)^3$.

Erroneous Results vs Erroneous Worked Solution The Derivation Example shows an erroneous worked solution. An alternative presentation that can be generated consists of the erroneous result only.

Eve wants to compute the derivative of the function $y = \frac{1}{(1-2\cdot x)^2}$ *for* $x \neq \frac{1}{2}$. *Her solution is* $f'(x) = (1 - 2 \cdot x)^3$. *Please find the error.*

Correcting Errors vs Finding and Correcting In the first version, the errors are marked in the presentation of the erroneous example and the student is asked to correct them. In the second, the learner has to find the errors first and then correct. These alternatives can be produced automatically.

[1] This example is one from a set of erroneous examples we used in ActiveMath
[2] correcting the errors is requested subsequently

High-Level vs. Low-Level Questions Low-level questions ask for a particular step in the worked solution. For the Derivation Example, a multiple choice question (MCQ) with low-level choices asks to decide which of the following alternatives did actually occur in the erroneous example:

- the Chain Rule is not applicable here
- Eve differentiated $\frac{1}{g^2}$ wrongly
- Eve differentiated $1 - 2 \cdot x$ wrongly
- the computation of $(f \circ g)(x)$ is wrong
- a condition is missing.

A high-level question may cover several occurrences in a worked solution or ask for violated principles. An MCQ with high-level questions for the Derivation Example asks which type of error occures (first):

- a wrong derivation rule was chosen
- a rule was applied incorrectly
- an algebraic transformation was wrong
- the solution is correct only under certain conditions

MCQ vs Marking Both, MCQ and Marking exercises are *choice* exercises. Therefore, they can have the same representation from which either a low-level MCQ- or a Marking-interaction can be generated.

Describing as Erroneous vs Asking Student for Decision. The above Derivation Example indicates that Eve's solution is erroneous. Alternatively, the student is asked whether this solution is correct or not and why. If we decided for the second strategy, then it needs to include similar prompts for correct examples. A special case of 'Asking' addresses (missing) conditions (as for $x = \frac{1}{2}$ in the Derivation Example) and asks "in what circumstances could this result be considered correct?". Another special case of 'Asking' is the presentation of two solutions of the same problem for which one of them is flawed.

Feedback vs no Feedback In their study Grosse and Renkl [5] do not provide feedback to students. We think that feedback is crucial.

4. Adaptation wrt. Concept and Presentation

A user-adaptive system will choose erroneous examples (1) according to a meta-goal of learning, (2) according to a particular concept or rule the learner needs to understand and (3) appropriate wrt. difficulty. That is, the choice will depend on what the student model exhibits about the learner's misconceptions, buggy rules and attention, about his learning goals, and general capability.

For one and the same erroneous example there could be different reasons to choose it for different students. For instance, the Proof Example below can target the fringe conditions of division for one learner and target better attention and monitoring of his problem solving process for another student.

The learning goal and concepts can be served by the choice of an particular erroneous example and by the level (and content) of the questions/tasks for the learner. The difficulty is greatly influenced by the tasks and form.

5. Hypotheses and Observations for Erroneous Examples in Tests

This section summarizes first observations from two formative tests we have been running with erroneous examples in ActiveMath. The study with about 120 students at an under-privileged school (6th grade) did not allow for controlled conditions. For now, we can report observations only. Another study was performed in a seminar with 17 second to fourth year computer science students at the University of Saarland and we tested the acceptance and problems of working on erroneous proofs and erroneous derivation examples. In addition, a very mixed population (academics, non-academic adults and children) with 53 subjects was tested with erroneous proof of $2 = 1$ given below. The conditions were not controlled.

For the school test with ActiveMath, we interviewed teachers on the errors they would target for fractions. The resulting most frequent errors concern buggy addition procedure. These errors are addressed in erroneous examples of the current ActiveMath fraction course, for instance

Eve made a mistake when computing the sum of $\frac{1}{8}$ *and* $\frac{3}{8}$.
She computed $\frac{1}{8} + \frac{3}{8} = \frac{4}{16}$
Find her mistake! (and later: compute the sum of $\frac{1}{8}$ *and* $\frac{3}{8}$ *correctly).*

For the university test with ActiveMath, we employed the Derivation Erroneous Example and other examples with the following frequent errors for computing derivatives in terms of misconceptions and buggy rules

- wrong derivation rule used
- wrong application of a rule
- misconception of composite function, e.g., wrongly assumed commutativity
- misconception about variables or about dependency of variables
- misconception about fringe elements. No restriction of function domain
- wrong interpretation of the derivative in word problems

Moreover, we tested subjects with the erroneous Proof Example:

Let	a	$=$	b
multiply both sides of equation with a	a^2	$=$	ab
add $(a^2 - ab)$ *on both sides*	$a^2 + a^2 - 2ab$	$=$	$ab - a^2 - 2ab$
take out $(a^2 - ab)$	$2(a^2 - ab)$	$=$	$1(a^2 - ab)$
division by $(a^2 - ab)$ *on both sides*	2	$=$	1

To summarize, observations at *school* indicate that
(1) replacing examples by erroneous examples increased the motivation of almost all students
(2) students read/studied the erroneous examples more carefully than normal examples (which they obviously did not self-explain). That is, erroneous examples fought the problem that in maths classroom many students do not read instructions, definitions, examples carefully and do not spontaneously self-explain but immediately go to the exercises (performance-orientation)
(3) working with erroneous examples took longer than working with material that included examples instead. This indicates a conflict with the 'economy of learning' that prefers performance-oriented ways of learning.

Observations in the *university* experiment indicate that those students who were well-trained in logic and knowledgeable about epsilon-delta proofs, judged the task of finding and correcting some of the errors as 'too easy', even for a mistake for which other students struggle to discover it. Not surprisingly, this indicates that the choice of erroneous examples needs to be adapted to the learner's prerequisites and capabilities.

The test in which the above Erroneous Proof Example $2 = 1$ was used, was performed with a mixed population of 53 subjects 38 found the error and 15 did not. The Erroneous Proof gave rise to an unusually high attention (between 1 minute (for quick solvers only, the lowest dropout time was 5 minutes) and 20 minutes (one non-solver took even 45 minutes)!). 10 non-solvers rated erroneous examples exercises as a "useful way to learn mathematics". 5 non-solvers rated erroneous examples exercises as not useful. 28 solvers rated erroneous examples exercises as useful. 10 non-solvers rated erroneous examples exercises as not useful. A possible reason for this relative high attention and acceptance rate may be the obvious conflict $2 = 1$ which can be thought provoking.

For the different ways to present erroneous examples in §3 the following hypotheses were (partially) supported in the tests.

Erroneous Results vs Erroneous Worked Solution When given only an erroneous results, the task was more challenging. Students had to build possible solutions paths themselves. On the one hand, this seems to be more difficult than judging an erroneous worked example (and low-achieving students give up more easily). On the other hand, constructing alternative solution paths provides precious training. If a student is not able to find the error when given the result only, then presenting the erroneous worked solution can be the next choice.

We hypothesize that similar to the setting of self-explaining worked examples, the parts of the (erroneous) worked solution provides reminders and more support to a student than a full problem solving exercise.

Correcting Errors vs Finding and Correcting Finding and correcting errors was more difficult for (weak) students than only correcting errors with feedback. Finding and correcting involves two types of activities, the first one for reasoning and explaining and the second one for problem solving. That is, 'finding' required reasoning, self-explaining and/or careful watching each step in the example. This first interaction provides good learning opportunities. Therefore, only if a student cannot 'find' the error, she should obtain a correction-only presentation of the erroneous example.

High-Level vs. Low-Level Questions. Sometimes it is difficult to ask reasonable high-level questions other than 'is the result correct or incorrect?' To answer abstract questions, the student has to understand what the principles are and where they occur in the worked solution. Since high-level questions can be followed by lower-level questions or marking, the guidance itself is structured and thus, can support a more structured reasoning. This was observed in the university course. In the school test, this situation was observed too for tutor interventions but not yet tested with ActiveMath.

Low-level MCQ vs Marking. We observed that marking seems to be more difficult for low-achieving students and can be somewhat more confusing at places (should a formula/result be marked or the reasoning/text that led to it?). We hypothesize that this is due to the smaller number and the explicit choice in case of MCQ.

Describing as Erroneous vs Asking Student for a Decision. Especially, the knowledgeable university students judged the more open format (in which they had to decide themselves about correctness) more interesting than a design stating that the solution is erroneous. We hypothesize that such a presentation will be more motivating for capable students. Moreover, the student has to be able to checking solutions and to inspect the problem solving space in order to be able to succeed with the problem.

Feedback As opposed to erroneous examples in the experiments of [5] ActiveMath provided orienting feedback for the finding phase as well as for the correction phase of erroneous examples. More detailed feedback is still under construction for the school course. For school students, the observations suggest that visualizations of the consequences of a learner's response may be needed in order to provoke cognitive conflicts.

6. Conclusion

Currently, erroneous examples are a rather unusual type of exercises in schools and in learning systems. However, they offer an interactivity that is primarily learning-oriented rather than performance-oriented.

We designed erroneous examples in ActiveMath with the long-term goal to improve the quality of learning at the cognitive and meta-cognitive level. This paper discussed several potential benefits of erroneous examples and different ways to design erroneous examples.

We reported the (informal) experiences from tests of ActiveMath with erroneous examples in a school and at the university.

Future Work

This work provides a basis for adapting to learning goals and students' capabilities. Future work will investigate in which situations erroneous examples are beneficial, how they have to be adapted for which learners, and how to generate useful feedback. Another problem is how to measure the learning effects that differ from performance improvement. This is important because performance is not the only dimension and may not even be the most important dimension of growth as discussed in section 2.

Acknowledgment

I thank Aiping Chen and Michael Dietrich for encoding erroneous examples on derivatives and on epsilon-delta proofs, Edgar Kessler for encoding the first erroneous examples for fractions and Martin Homik for his devoted activities in the university seminar.

This publication is partly a result of work in the context of the LeActiveMath project, funded under the 6th Framework Program of the European Community – (Contract IST-2003-507826). The author is solely responsible for its content.

References

[1] R. Borasi. Capitalizing on errors as "springboards for inquiry": A teaching experiment. *Journal for Research in Mathematics Education*, 25(2):166–208, 1994.

[2] L.E. Bourne, B.R. Ekstrand and B. Montgomery. Concept Learning as a Function of the Conceptual Rule and the Availability of Postive and Negative Instances. *Journal of Experimental Psychology*, 82(3): 538–544, 1969.

[3] A.T. Corbett and J.R. Anderson. Locus of feedback control in computer-based tutoring: Impact on learning rate, achievement and attitudes. In *Proceedings of CHI 2001*, pages 245–252. ACM, 2001.

[4] H.D. Gerster and U. Grevsmuehl. Diagnose individueller Schülerfehler beim Rechnen mit Brüchen. *Pädagogische Welt*, pages 654–660, 1983.

[5] C.S. Grosse and A. Renkl. Learning from worked examples: What happens if errors are included? In P. Gerjets, J. Elen, R. Joiner, and P. Kirschner (eds) *Instructional design for effective and enjoyable computer-supported learning*, pages 356–364. Knowledge Media Research Center, Tuebingen, 2004.

[6] K. Hart. Ratio and Proportion. In K. Hart (ed), *Children's Understanding of Mathematics*, pages 88–101, John Murray, London, 1981.

[7] E. Melis, E. Andrès, J. Büdenbender, A. Frischauf, G. Goguadze, P. Libbrecht, M. Pollet, and C. Ullrich. ActiveMath: A generic and adaptive web-based learning environment. *International Journal of Artificial Intelligence in Education*, 1002(4):385–407, 2001.

[8] P.H. Miller. *Theories of Developmental Psychology.* Freeman, San Francisco, 1983.

[9] F. Oser and T. Hascher. Lernen aus Fehlern - Zur Psychologie des negativen Wissens. Schriftenreihe zum Projekt: Lernen Menschen aus Fehlern? Zur Entwicklung einer Fehlerkultur in der Schule, Pädagogisches Institut der Universität Freiburg, Schweiz, 1997.

[10] A.H. Schoenfeld. Explorations of students' mathematical beliefs and behavior. *Journal for Research in Mathematics Education*, 20(4):338–355, 1987.

[11] A.H. Schoenfeld. Explorations of students' mathematical beliefs and behavior. *Journal for Research in Mathematics Education*, 20(4):338–355, 1989.

[12] R.S. Siegler. Microgenetic studies of self-explanation. In N. Granott and J. Parziale (eds), *Microdevelopment, Transition Processes in Development and Learning*, pages 31–58. Cambridge University Press, 2002.

[13] S. Stafylidou and S. Vosniadou. The development of students' understanding of the numerical value of fractions. *Learning and Instruction*, 14:503–518, 2004.

[14] Ch. Strecker. Aus Fehlern lernen und verwandte Themen. http://www.blk.mat.uni-bayreuth.de/material/db/33/fehler.pdf, March 1999.

[15] P. Tsamir and D. Tirosh. In-Service Mathematics Teachers' Views or Errors in the Classroom. *International Symposium : Elementary Mathematics Teaching*, Prague, August, 2003.