Peer-Reviewing and Life Long Learning.

Jochen Koubek

Humboldt-Universität zu Berlin
Institut für Informatik
Unter den Linden 6
10099 Berlin
jochen.koubek@hu-berlin.de

Abstract: The review system is a web based learn-management-system that let students review the works of their peers. It has been proven successful in formal learning situations. In this paper a scenario is proposed to use it in non-formal or informal learning which is typical for adult training and life long learning.

1 Adult training

Adult training often does not require the students to participate in a given course but to prove their skills in a written or oral exam. The certification compares the skills and problem solving capabilities with principles and standards worked out and published before the exam takes place. The individual result does not depend of the average of the class but only of the student's achievements.

But how do students develop these skills? A common approach distinguishes three forms of learning [CO02]:

1. **Formal learning** is organized learning in the educational system and leads to a certificate opening doors to jobs or higher education.
2. **Non-formal learning** is the organization of learning outside the formal educational system e.g. in further education or training-on-the-job.
3. **Informal learning** is an intentional but less organized form of learning like explaining a college how to use a digital camera.

During informal learning students develop knowledge, skills, and competencies at their own responsibility and their own pace. But the lack of extrinsic motivations often inhibits systematic efforts to overcome individual learning obstacles. A simple learning plateau can mark the end of the whole learning process.

Extrinsic motivations in form of assignments, exams and certificates are typical for formal and non-formal education. These structure characteristics come along with larger classes or courses which often are not adapted to individual learning for several reasons:

- Courses progress at the pace of the teacher who at best respects the average learning level with disrespect to the slower and the faster students.
- The rhythm of the course is dictated by the teacher with exercises and tests, acting as milestones and emphasizing the power of the teacher over time and space of the course.
- The teacher's feedback capacity for the students' work must be distributed among all students. More students mean less feedback for everyone.

Qualitative exercises like written assignments or practical work are revised only by the teacher or his assistant. More often than not the feedback consists of some correction marks followed by a simple numerical value not explaining the problems and what should be done to progress.

Students learn how to master exercises but neither how to create them nor how to correct them. Creation and correction of exercises are important skills not only in the educational system. Creation means being able to break up a large subject into parts which are small but essential and to arrange them so that attending the parts enables the comprehension of the whole. Correction means developing personal judgment and expressing it in an interpersonal way.

2 The Review System

Given the right didactical setting some or all of the mentioned obstacles can be overcome. The *Review-System*, developed at the Humboldt-Universität zu Berlin is a tool to support this kind of settings. It is a web based learn-management-system, whereby students mutually peer-review the exercises they have participated. The system's technical and didactical basics has been described in other papers ([KO05], [KN08]) and will be briefly summarized:

Every student receives an anonymized random selection of works from other students for correction and review. A cycle for one exercise consists of the following steps:

1. The students get the exercise.
2. They have a limited amount of time to work out and upload a solution.
3. After the deadline each student is given five documents to review. A review contains both a numerical grade and a written report about the quality of the reviewed document. The deliverance of all five reviews are crucial to validate one's own points. A reviewer handing over less than the five demanded reviews would receive no points at all for his document. To prevent disadvantaging an author in case of missing reviews and points for his document, the overall score is filled up with the rounded average of the points of the delivered reviews.

4. Subsequent to the reviews the overall score for the exercise is calculated. More important, the qualitative reports are published at least to the author of the reviewed document.

The system assures that only those students who worked on an exercise participate in the review process. The reviewers therefore have a minimum of experience with the nuts and bolts of the particular problem and should be able to roughly judge the efforts of their peers. A more detailed judgment depends largely on the structure of the exercise.

3 The structure of exercises

Using the review system, course material should be transformed into exercises in seven steps:

1. Learning **objectives** should be precisely clarified. This can be done using the revised taxonomy of educational objectives as proposed by Anderson & Krathwohl [AN01]. The knowledge dimension covers factual, conceptual, procedural, and meta-cognitive knowledge. The cognitive process dimension distinguishes:
    i. Remember and retrieve knowledge from long-term memory
    ii. Understand and construct meaning in communication
    iii. Apply knowledge in a given problem
    iv. Analyze knowledge into its basic elements and discover the relations of these parts
    v. Evaluate and judge based on knowledge
    vi. Create new patterns or structures from these elements

    The objectives should be stated that they are easily understood by all students. This is crucial for the peers to know what to review.
2. For each objective one or more **criteria** should be given that clearly states if and when the objective is achieved. In most cases the verbal description of the objective includes the criteria.
3. Even many of the criteria may be qualitative and depend on the judgment of the reviewer, it is desirable to find a **metric** which measured a criteria and enables an objective feedback about achievement. As Tom DeMarco stated »You can't control what you can't measure« [DM86].
4. Sometimes the objectives are too general or too big for the students. In this case they must be divided into **sub-objectives** which seem to be in reach. In this case each sub-objective is treated as an objective and criteria and metrics must be added.

5. The objectives are split up into several related **groups of objectives**. Not all of the possible objectives mentioned above should be stated at once. Choosing to many objectives may discourage students from working on a given exercise.
6. **Problems** are to be created that cover each group of objectives. Sometimes these groups may have been chosen with a particular problem in mind, sometimes a good problem demands a re-grouping to match problems and objectives.
7. The problem is translated into an **exercise.** This exercise is given to the students together with the objectives, criteria and metrics.

The composition of exercises is a learning taxomony combined with a Factors-Criteria-Metrics (FCM) quality model. The use of a FCM-model is important because the reviewers must be able to judge the work of their peers without a developed pedagogical judgment which can be expected only from a professional teacher. The use of obejctive metrics relieves the judgment and assures comparable results among the reviews. Where objective metrics can not be found, at least good subjective criteria should be given to support the reviewers.

4 Peer Review in Virtual Communities

The review-system may be used in training settings where it is easy to find and pay a teacher offering lessons and designing exercises but who is unwilling to correct and review the students' works. The review-system is the tool of choice if

- the score attained in the exercises is not overall important for the students and
- feedback is important but can't be extensively offered by the teacher.

These are typical criteria for adult education and life long learning. These learning situations are primarily objective driven: students know what knowledge they want to learn or what skill they want to achieve. In informal or non-formal learning situations the final certificate depends not on the learning path and the problems and exercises solved on this way but on a final exam. Thus a feedback on the performance while exercising is more important than the score. If this feedback cannot be offered by the teacher it can be given by the peer students. The students do not necessarily know each other, complete virtual training sessions are possible.

An application may be courses for using a software, often necessitating intensive training. An online-community can form around a given software, getting their teaching material online in form of documents, examples and video-tutorials. This material can be supported by computer mediated communication such as chats, forums or newsletters. Adobe's design and training center is a good example for this kind of repository of online-materials accessible both for public and for registered users [AD08]. An expert is filmed solving a given problem like using layers in Photoshop. The student trying to solve the same problem can replay the film over and over stop at every moment and go

back to a part he didn't understand. Additionally written texts are offered explaining every step.

But reading about or looking at the teacher's performance is only one aspect of learning. Watching an expert's performance can lead to the false impression that the problem is not so hard and does not need the effort to do the same thing here and now. It may become the entertainment part of edutainment.

Successful building of competencies requires practical exercises which should be corrected in order to receive a feedback about ones performance. With consistent use of the review-system is it possible to build completely autonomous virtual communities of people eager to learn and getting feedbacks on their works but not depending on grades or rankings. These groups can be maintained so that each student can choose the time and level of exercises according to his pace. Since the review process only includes students that delivered a document not to participate in an exercise will not provoke discouraging reviews.

Given enough participants, each exercise can be processed a multiple times. If a student is discontent with the results for his work or got a constructive review telling him how to do better next time, he can deliberately choose to retry the same exercise till he and the peers judge his work to be conclusive. Thus the learning rhythm entirely depends on the students.

Since students need to write reviews for the assigned documents, each author will get as many qualitative feedbacks for their work. The review system does not depend on the person creating the exercise. They can be developed by a teacher or by the students. In order to help corrections and reviews, possible solutions can be made public after the deadline for delivering the documents.

Space

The learning community can be completely virtual sharing no common space. The review system can of course be used in a blended learning environment assisting a course where real life presence is needed.

Time

The timing of the exercises might be structured in several ways

1. No time limit: Exercises are always online, everyone can participate at all time. Whenever six (or n>6) students worked on an exercise, a review task is created amongst them. Everyone delivering five reviews can see the reviews for his work.
2. Fixed length: The whole course has a fixed duration during which a fixed number of exercises must be attended or a minimum number of points must be gained.

3. Milestones: Every exercise is limited in time, e.g. two weeks till submission. After this period, no further submissions for the exercise are accepted. This is the classical scheme for institutional courses.

Organization

The overall organization of a course can be presented like a game, wherein the players can choose between several modes:

1. Free for all: Everyone takes part whenever he wants to. No deadline is set, no records are kept. If a student wants to check her skills, she participates in an exercise, gets her reviews after reviewing the works of others.

2. High Score: Everyone takes part whenever he wants to but records are kept. The final score is calculated as a sum of the results. The sum may be weighted although the exercises do not depend on each other.

3. Exercise levels: A minimum of points must be acquired in order to participate in certain exercises. These points can be accumulated either with the same exercises like harvesting gold in World of Warcraft or with more complicated exercises offering more points but asking higher competencies.

4. Competence model: The level structure can be more complicated, if the next level not only depends on the number of points but on the solution of a specific set of exercises. This organization is based on a competence model and approaches the constrains of formal education, where certificates decide over the possibility to enter the next grade or the more advanced institution.

Exercise design

The exercises can be designed by different persons:

1. The **teacher** knows best what competences are to be met. But since the design of balanced exercises and choosing the right level of difficulty is not the same task as teaching, it is not always possible or desirable to let the teacher create the exercises.
2. An **exercise designer** with close contact to the teacher can transform the subject into exercises respecting the competence level of the students. In university teaching this part is often taken by the assistant or member of the working group.
3. Exercises can be taken from a **book** in which case they will be well designed and adjusted according to the authors material which is not necessary the same as the course of the teacher.

4. Exercises can be designed by the **students**. This is not in all cases possible, but can lead to satisfying results if they are reviewed by a third person already familiar with the subject.

For different exercises or levels of exercises the responsibility may vary. Easy tasks can be created by students while the more difficult problems may be arranged by the teacher or assistant.

Participants

The size of the learning group is not limited, the bigger the better. Furthermore it is not necessary to have a stable group of subscripted students. In one exercise there may be 100 participants while in others there are only 40.

Privacy

Participation may be completely anonymous, only a valid email-address is needed. If the results of the reviews are used in a certification process like a final exam, there must be a mapping between the mail-address and the real name. This mapping can be done during subscription to the system and kept secret during the reviewing process.

5 Case Study

For several years the review-system has been used in a master module of computer science and some but not all of the mentioned possibilities has been successful realized. The principal constraint was the need for grade with all the disadvantages such a number presents.

Space
Although a weekly course was offered at a given place, participation was not mandatory. Every student could work wherever he or she wanted to.

Time
The timing of the course was entirely centralized by the tutor following the milestone-model mentioned above. Every two weeks a new exercise was assigned with two weeks time limit for delivering and uploading the final document.

Organization:
Since every exercise was dealing with a new subject, a high-score model was chosen. Every exercise was independent of the others but an overall ranking was calculated. A minimum score must be achieved in order to participate in the final exam. Although this meant, that the peer reviews actually did have an effect on the grade, the score was kept so low, that no one complained about bad or unjust judgments of peers preventing him from the final exam.

Exercise Design
The exercises were exclusively designed by the teacher. Every document had to be reviewed by 5 peers and 0 to 5 points could be distributed. So each document could gain as

much as 25 points, a score rarely achieved for it has to convince 5 reviewers of its top quality.

Participants
During the years, group sizes from 30 to 100 students have been successfully tested.

Privacy
Anonymity is a key concept in peer review but for the sake of certifications each pseudonym was mapped to a real name outside the system. Thus we chose to use the matriculation number as pseudonyms.

A common objection against peer reviewing in classes is the assumption that the solidarity amongst students is too high to give each other bad grades. This kind of solidarity never occurred in our classes. The points were distributed around an average (Fig. 1). Few documents were judged to be very bad, few to be very good.

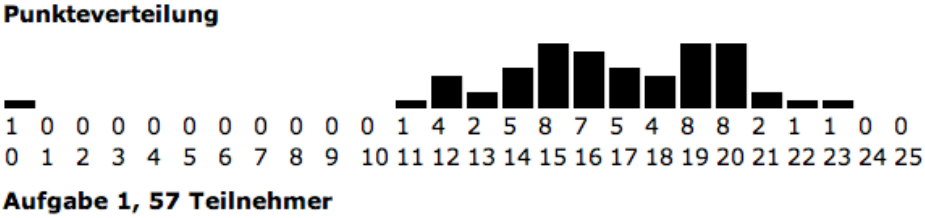

Fig. 1: Distribution of points for exercise 1 with 57 participants.

In a qualitative poll, students report, that it is helpful and motivating

- to know that their document will be reviewed.
- to see the documents of others.
- to have access to the documents which gained the highest score. For this purpose we offer a page *top* linking to all documents that gained more than 22 points.
- to get meaningful reviews, especially when they contain constructive critics.

Since the task changed with every exercise, nothing can be said about the learning effect before and after the reviews. But the review competence was enhanced during the course, the average review became more precise and concrete. Students who benefit from good reviews are willing to give this to others. There are of course always reviewers who content themselves with a review like »not too bad, 3 points«, but they rest a minority.

The review system has been successfully used since 2004 and will be in the future, for it is a win-win-situation for both teachers and students. Teachers have less work and are

not obliged to correct the same exercise over and over. Students have better feedback and learn how to analyze the work of others and how to express their impressions.

The review system will be refactored and published under an Open Source License by the end of 2008.

Bibliography

All URLs have been tested in May 2008.

[AD08] Adobe Design Center 2008: Internet http://www.adobe.com/designcenter/

[AN01] Anderson, Lorin W.; Krathwohl, David R., Ed.: *A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives.* New York, Addison-Wesley, 2001.

[CO02] Colley, Helen; Hodkinson, Phil; Malcolm, Janice: *non-formal learning: mapping the conceptual terrain. a consultation report.* 2002. Internet: http://www.infed.org/archives/e-texts/colley_informal_learning.htm

[DM86] DeMarco, Tom: *Controlling Software Projects: Management, Measurement, and Estimates*. Prentice Hall, 1986.

[KO05] Koubek, Jochen: *The Review-System.* Proceedings TESI 2005 (Training Education & Simulation International), Maastricht (Netherlands), March 22-24, 2005).

[KN08] Knaut, Andrea: *Open Source-Revision einer Software am Beispiel des Gutachtersystems.* Studienarbeit (not published), 2008.