Learning from peers: motivating students through reputation systems

Marco Temperini and Andrea Sterbini
Sapienza University of Rome, Via Salaria 113, Rome, Italy
marte@dis.uniroma1.it, sterbini@di.uniroma1.it

Abstract

Our on-line students, being mainly busy worker-students, study almost alone. To improve their interaction we use asynchronous tools (Wiki or forums), but we notice that interaction becomes high mainly when the discussion is focused on a task to be graded for the exam or when the teacher/tutor is very active in the community.

We present SOCIALX, our exercise sharing tool, an application to e-learning of a simple reputation system to increase the student motivation and interaction, and to let them learning from each other, either by reusing other's solutions or by correcting other's mistakes. Moreover, students gain reputation from others reusing their solutions. In this we want to engage students in learning activities at the highest cognitive levels of the Bloom taxonomy [1].

1. Introduction

We present an application to e-learning of basic reputation system techniques, developed with the aims of increasing the motivation and level of interaction among students, and helping the students to learn from each other, such that each one can reuse (at different levels) a solution proposed by another, and possibly spot some mistakes found there.

We have implemented SOCIALX, an exercise sharing tool where students gain reputation in front of the teacher and (especially) of the other students by submitting solutions to exercises and by collecting endorsements by other students inspired to reused solutions. In this we want to engage students in learning activities at the highest cognitive levels of the Bloom taxonomy [1].

2. The “SOCIALX” system

The SOCIALX system allows three types of users:
• teachers create courses, add exercises to topics of their courses and “endorse” a solution by stating if it's correct or wrong,
• the administrator enables new teachers,
• students browse courses, exercises and solutions and add votes and new solutions,

3. Exercises and solutions

Exercises are associated to topics inside one of the courses of a teacher. E.g. The course “Linguaggi per il Web” (Web languages) contains 50 exercises on the topics:
• XML (6 exercises)
• XHTML (11 exercises)
• CSS (10 exercises)
• XHTML-CSS (6 exercises)
• CGI (4 exercises)
• PHP (10 exercises)
• PHP-MySQL (3 exercises)

An exercise is added to the system by uploading a file containing all the material needed to describe and solve it (in the simplest case, a PDF file describing the problem, or a compressed archive containing all the required files), and associating the exercise to a specific course and topic.

A student can try to solve an exercise from scratch or can download and examine other's solutions and get inspiration from them, and later add his/her new solution. When a new solution is added the student should state what is the level of reuse of the other solutions (in a range going from “simple inspiration” to “almost total reuse”). In doing this s/he increase the reputation of the author of the reused solution.

A student can vote for a solution that he likes/dislikes with a mark ranging from 0 (worst) to 10 (best), thus improving/reducing the other's reputation.

The teacher can mark a solution as “good” (i.e. correct), so that the author will gain reputation from its reuse, or “bad” (i.e. incorrect), so that other students will gain
reputation by finding errors and proposing their correct solution.

The system is “forgetful” with respect to bad solutions i.e. a bad solution doesn’t reduce the reputation (we want to encourage participation). Yet, one can improve his reputation by correcting a bad solution (we want to encourage critical analysis of each other's work).

The solutions not examined by the teacher are left unmarked, thus showing that nothing is known on their correctness.

Both the student and the teacher can analyse the reputation earned by the students either at global level (over all the courses) or respect to the selected course or to the selected topic.

4. The reputation system

We base our system on the exploitation of the social interaction of the group of students. We engage them in a “social game” where they are rewarded for good solutions and for good social practices. In this we get inspiration from well known reputation-based systems for big communities of users [4, 5, 6, 7] and in other reputation systems applied to education [3].

As the class community is very small (in respect to the above mentioned societies) we motivate the participation in the system by awarding part of the final course grades depending on the reputation collected in the system.

A reputation-based system is characterized by two facets:

- each user collects a personal reputation reflecting his perceived utility to the community,
- the effect/weight of his actions in the community depends on his reputation.

The second part could imply that the reputation of a user depends itself on the reputations of the others grading him. This sometimes implies a recursive schema, which adds the constraint that the algorithm to compute the reputation of users should converge to a fixed point within reasonable time. In our current simple implementation we compute the reputation of a user independently from the reputations of other users, thus the convergence of the computation is not an issue.

Our previous implementation

In a previous implementation [2], the reputation of a participant was quite elementary built from two aspects of each submitted solution: the average mark of the submission and the amount of reuse that the solution had enjoyed. In particular, the reputation of a student was:

\[
\text{REPUTATION} = \text{ACTIVITY} \times \text{REUSE}
\]

where

\[
\text{ACTIVITY} = \text{avg. votes} \times \text{n° of sol. Submitted}
\]

\[
\text{REUSE} = \text{weighted sum of the n° of inspired sol.}
\]

Moreover, an incentive was introduced to encourage students both to solve exercises by reworking on an already present solution, or by correcting other's mistakes.

5. The new reputation system

We have updated our initial simple reputation system to evaluate and encourage several facets of the student participation:

- Involvement
- Usefulness to others
- Competency on the topic
- Ability to judge other's solutions
- Critical thinking (while correcting other's work)

In this, we try to apply the same social/psychological mechanisms used in other reputation systems applied to education (e.g. Comtella) to encourage the student's behaviour towards the learning goals [3].

In his personal page, a student can examine his reputation with all the details of the different factors that contributed to it. This is particularly important to give a clear description of what are the behaviours that can be improved in order to obtain a higher mark.

![Fig. 1: Personal reputation profile](image)

The total reputation is obtained by weighting the 5 facets as follows:

\[
\text{reputation} = \left( 2 \times \text{involvement} + 3 \times \text{usefulness} + 5 \times \text{competency} + 1 \times \text{judgement} + 4 \times \text{critical_thinking} \right) / 5
\]

The teacher can change the weights of the facets to stress the one s/he thinks is more important.
The five facets are evaluated by analysing the interaction of the student with the system, the votes collected from fellow students and from the teacher and the reuse factor of the submitted solutions.

**Involvement factor**
To engage students both on solving exercises and on participating to the reciprocal evaluation of solutions we estimate the involvement factor by weighting both the number of solutions submitted and the number of votes given to other's solutions:

\[
\text{involvement} = 3 \times \#\text{submitted\_solutions} + \#\text{voted\_solutions}
\]

**Usefulness factor**
The overall usefulness factor of the student's solutions (corresponding to the previous REUSE factor) is computed by weighting the number of solutions submitted that somebody else has reused in their work (declaring how much she reused it, in a scale from “simple inspiration” to “almost total reuse”). Solutions that have been almost totally reused are given the highest weight, thus giving higher reputation to the student.

\[
\text{usefulness} = 1 \times \#\text{simple\_inspiration} + 2 \times \#\text{medium\_reuse} + 3 \times \#\text{extensive\_reuse} + 4 \times \#\text{almost\_total\_reuse}
\]

**Competency factor**
The competency factor of the student is obtained from the votes s/he has collected both from the teacher and from the fellows students (votes range from 0 to 10). The teacher's votes are more important.

\[
\text{competency} = \frac{3 \times \text{avg\_teacher\_votes} + \text{avg\_student\_votes}}{10}
\]

Notice that we separately compute the two average values. We do this for three reasons:
- because the votes come from people with very different expertise
- to give a clearer feedback to the student
- to lower the impact of the “asking friends for high marks” strategy

Notice that we have expanded the way votes are given to solutions. Previously, only one vote was given to a solution, now the teacher can define several properties to be voted (e.g. clarity, correctness, synthesis, difficulty), some of which can be marked as critical factors for the course, that will give bonuses to the students.

**Bonus points**
We add a set of bonus points to the competency to handle cases that imply more effort from the student or to endorse particularly good contributions:
- +6 for a high mark (>7) from the teacher on the “most important property” of the course
- +3 for a high mark (>7) from a student on the “most important property” of the course
- +10 if this is the first solution and it's accepted by the teacher
- +1 for each solution endorsed (good) by the teacher
- -3 for a very low mark (<4) from the teacher on the “most important property” of the course

**Judgement factor**
To evaluate the ability to judge other's work (the judgement factor) we compare the votes given by the student to the ones given by the teacher. We count how many are exactly equal and how many differ just by one. Exact votes are weighted more than “similar” ones:

\[
\text{judgement} = 5 \times \#\text{same\_as\_teacher} + 2 \times \#\text{similar\_to\_teacher}
\]

**Critical thinking factor**
Finally, we try to evaluate (and thus to encourage) the high-level conceptual work (as for the Bloom's taxonomy) needed to understand and critically judge the others' solutions in two particular cases:
- looking for a correction for a solution marked “bad” by the teacher,
- using a not-yet-approved solution (that the student cannot safely trust) as starting point.

Thus we compute the critical thinking factor by counting both how many solutions are inspired to incorrect ones and how many are inspired to unknown (at the time of submission) ones. In the former case the student already knows that there is something wrong, thus the job is considered somewhat easier.

\[
\text{critical\_thinking} = 3 \times \#\text{sols\_inspired\_to\_bad\_sols} + 6 \times \#\text{sols\_inspired\_to\_unknown}
\]

6. Global ranking

Reputations can be compared with a series of synthetic tables both globally (over several courses), at the course level, and at the topic level.

In the current system the reputation values are not normalized, thus, to give the student a clear definition of the teacher's evaluation criteria, the reputation value is translated to a set of verbal descriptions. E.g.
The definition of levels can be tuned by the teacher with an appropriate interface.

7. First experiments

We are going to start two pilot tests using the SOCIALX system in Rome in two different settings:

• with the students of the course “Linguaggi per il Web” (Web languages), that is delivered face-to-face to Master-level students in the Computer Engineering Dept. in the “La Sapienza” university, in Rome, Italy. The course contains 50 exercises distributed among its 7 main topics. Circa 40 students are enrolled in the course. Being face-to-face all students study (more or less) at the same time, thus a higher cooperation is expected. As we said earlier, we motivate such a small number of students by awarding part of the final grade through the collected reputation.

• with the on-line students of the course “Fondamenti di Informatica 2” (Computer Fundamentals 2), Laurea degree in Computer Science Engineering at “La Sapienza”, in the NETTUNO Consortium Programme, which is a video-course delivered through TV with on-line tutorship. The course has normally 50 students/year. Here the students enrol and study the course when they like, thus the cooperation between students is expected to be very desynchronized across the whole year, and mainly based on reuse of older solutions. Even if the “community” cannot really work well in this setting, the system should be very helpful because the students are really alone and need good examples of solving exercises.

8. Conclusions and future work

We have presented our SOCIALX exercise sharing tool. Beside the pilot tests on its usage and on its impact on on-line students, we are planning some improvements, and integration with other e-learning tools under development in our group [8, 9].

Cheating the reputation system: Our simple reputation system is somewhat weak respect to misuse of the system by the student. High (untrue) values both of the votes and of the reuse factor of a solution increase the reputation of the fellow students, thus creating distortions in the discrimination power of the reputation value. To discourage such misuse we track the authors of each action, so that a student knows to be always responsible for his own behaviour. Moreover we plan to implement a set of tools to easily expose malpractices, to avoid playing the “cop and thieves” game with the students we will try to introduce other reputation factors that encourage in correct behaviour.

Metadata description: In the current system exercises are associated to the Teacher/Course/Topic triple, thus there is a low level of reuse between parallel courses or different teachers. A better metadata description of the exercises would allow wider reuse among several courses and automatic selection of appropriate exercises to be associated to specific points in learning paths [10].

9. Acknowledgements

We deeply thank the anonymous referees for their very useful comments and suggestions.

10. References