Abstract— Vygotskij’s educational concept of the Zone of Proximal Development (ZPD) calls for an adaptation of learning activities to the present state of learner’s knowledge and abilities. At the same time, Vygotskij’s educational model includes a strong bent towards social and collaborative learning. In practice, a tight integration between personalized and collaborative learning is called for. Along this line, we exploit two previously implemented prototypes of systems for web-based e-learning, to investigate the integration of adaptive e-learning with social learning activities, both at individual and group level. LEcomPS is a web-based e-learning environment for the automated construction of adaptive learning paths. SocialX is a web-based system for shared e-learning activities, which implements a reputation system to provide feedback to its participants. We propose a two-way tunneling strategy to integrate the above prototypes: the LEcomPS student model is used to select the set of social activities (met in SocialX) according to the present individual learner state of knowledge; on the other hand, the solution of exercises, and the associated reputation gained in SocialX, are used to update the LEcomPS student model. Under the social perspective induced by the integration, we present a mapping between the student model and the definition of Vygotskij’s Autonomous Problem Solving and Proximal Development regions, with the aim to provide the learner with better guidance, especially in the selection of available social learning activities.

Keywords: zone of proximal development; adaptive e-learning; social collaborative e-learning; reputation system; group modeling

I. INTRODUCTION

The aim of the design strategy presented in this paper is to merge personalized learning paths and acquired learning reputation under the auspices of Vygotskij’s theory of proximal development [1][2], while further strengthening the sense of community in a class. In order to achieve these goals, we rely on the integration and exchange of data, about students’ performances and achievements, between two formerly developed web applications: LEcomPS, an environment for personalized e-learning, and SocialX, supporting collaborative and social learning activities.

Personalized learning paths allow to better respond to learners’ needs and preferences, while reputation based techniques tend to increase the student’s motivation. In our model, the learner’s activity in the socio-collaborative reputation system (in SocialX), and the reputation itself, are used to refine the individual student model (in LEcomPS). In the other direction, the learner’s overall profile (of LEcomPS) is used to enrich a course with social-learning activities. This strategy fully agrees with principles and guidelines of Learner Centered Design [3][4].

II. MOTIVATIONS AND BACKGROUND

A great deal of interest and research is presently dedicated to supporting automated construction, maintenance and delivery of adaptive e-learning courses. On the other hand, collaborative learning is considered a winning methodology to allow the development of meta-cognitive abilities, such as critical thinking in learners, and to foster the acquisition of new knowledge. Moreover, it appears to support better retention and deepening of knowledge over time [5]. In a collaborative environment the learners are prepared for team-based working activity, both by sharing their common experience and combining their skills [6]. Thus it becomes an aspect of interest the social dimension of e-learning, viewed as a community in which social activities take place and social interaction skills are developed by the participants. Along this line, in connectivism perspective [7] as well as in situated learning model [8], knowledge is a system which is accessed through people participating in (ecologically significant) activities. An example of e-learning environment implementing Organizational Learning is the environment E-MEMORAe2.0 [9]. Its aim is to prepare the students to work in team, and above all to share and exchange knowledge. Such skills are important both in formal and informal learning, and the same holds in the perspective of lifelong learning [10]. Moreover, while any organization, including a class of students, cannot learn without its members being continuously learning [11], it seems crucial for the organization to act as a Community of Practice (CoP), in order to better enable its members to share experiences, knowledge and skills, in a true peer-to-peer learning. So, development of CoPs is surging ahead. A CoP is “a group of people who share an interest in a domain of human endeavor and engage in a process of collective learning that creates bonds between them: a tribe, a garage band, a group of engineers working on similar problems.” [12]. CoPs develop along a number of dimensions. Wenger [13] lists eight dimensions, classifying technological products which may support CoPs to a different extent: ongoing integration of work and knowledge, (collaborative) work, social structures, conversation, fleeting interactions, instruction, knowledge exchange, documents. As we can
notice, these are the same core components of the collaborative side of an advanced e-learning system. The main difference is that, while CoPs rise and grow in a completely spontaneous and quite unsupervised way, e-learning systems still need the role of the teacher as main inspirer, supervisor and guiding figure. The integration of SOCIALX for collaborative aspects and LECOMPS for supervised ones is a first attempt towards the long term goal of fully developing an e-learning system including all required CoP-like features, in other words, a Learning Community of Practice (LCoP). A reputation system is a natural component for building a LCoP, suitable in particular to motivating and transforming a class (or some classes) of learners into a real community. The reputation captures the contribution of each learner, and makes it apparent to the community (the group, the class and the course). The reputation is both a motivational tool and a way to evaluate and understand learner's psychological characteristics (learning and communication style) as well as preferences, relations with others, and ability to analyze/judge others' work (meta-cognitive achievements).

III. VYGOTSKIJ’S COOPERATIVE MODEL

One of the main achievements of Vygotskij’s research [1][2] is a model aiming to demonstrate that cooperation provides the basis for the individual development. Even during childhood, targeted as well as casual interactions activate and spur the cognitive processes. The presence of people in the same environment, and the cooperation with peers, induces a reflection and an auto-regulation of one’s own behavior. Once such processes are interiorized, they become part of the child’s autonomous evolution. Social learning therefore precedes individual competencies and determines and prepares cognitive development. As a matter of fact, according to Vygotskij, the typical direction of learning is from outside to inside: the knowledge interiorization process mainly happens through the social “co-construction” (social learning), and then proceeds through a progressive transfer of the exterior social activity to the interior control. Once knowledge and processes are interiorized, the learner will be able to proceed in an autonomous and independent way. It follows that concrete growth can occur only in the Zone of Proximal Development (ZPD), which is characterized as [14] “the distance between the actual development level, as it is determined by the autonomous problem-solving, and the level of potential development, as determined by the problem-solving under an adult’s guidance or in collaboration with one’s own more capable peers”. In practice, when the learner has consolidated a region of Autonomous and independent Problem Solving (APS), it is useless to further suggest exercises related to the same level of difficulty. On the other hand, it is even more useless to suggest exercises which are completely out of reach for the learner (Unreachable Problem Solving – UPS). The right zone of complexity, is the one including exercises that the learner can solve, according to her own competence and/or with a moderate support and/or in collaboration with companions (Zone of Proximal Development – ZPD) (See Fig. 1).

IV. TWO WEB-APPLICATIONS FOR PERSONALIZED SOCIAL LEARNING

A. LECOMPS

LECOMPS [15] (Fig. 2) is a web-based e-learning environment, supporting learners and teachers in a variety of activities:
- authoring of learning objects (Learning Components – LCs) and their organization in pools defining “subject matters”;
- enrolment of learners and evaluation/maintenance of their personal Student Model (SM);
- automated construction and adaptive delivery of personalized courses, tailored over the learning goals (Target Knowledge – TK), stated for the subject matter, and the (evolving) individual SM.

The whole set of LCs dedicated to the subject matter composes the Learning Domain (LD) for the courses. A course $C$ is a sequence of LCs, a subset of LD, built to let the learner bridge the gap between her initial state of knowledge and the TK.

The knowledge associated to a learning component (required to study its content, or acquired during its study) is represented through the conceptual device of the Learning Objective (LO) [16], that is a predicate such as $LO(\text{level}, \text{keyword}, \{\text{concepts}\}, \text{context})$

where $\text{level}$ and $\text{keyword}$ are cognitive characteristics of the $\text{concepts}$ (topics about which the LO does express a skill), and $\text{context}$ designates the learning context of the
concept(s). (Keywords and level are grounded on the taxonomy of Bloom [17]).

In detail, a learning component, lc, includes:
- **learning content** (an XHTML-formatted resource, that can be given in different, LS-wise, versions;
- two sets of LOs: lc.RK (required knowledge), and lc.AK (acquired knowledge, see earlier);
- **questions**, related to the LOs in AK, used in questionnaires to assess knowledge acquisition;
- **effort**, informal measure of the LC content.

Among LOs, inference rules hold, such that the possession of certain LOs (such as a set \\{\textit{l}_i\}_{i \in \text{I}}\) can imply that some other LOs are possessed too: \\{\textit{l}_i\}_{i \in \text{I}} \rightarrow \\{\textit{l}_j\}_{j \in \text{J}}\). For instance, a skill possessed at a certain cognitive level implies the possession of the same skill at lower level:
\[ \text{LO}(3, \text{apply, cpt, cxt}) \rightarrow \text{LO}(2, \text{describe, cxt}). \]

The set of all LOs associated to LCs in the LD is the **Knowledge Domain** (KD) for the subject matter.

The student model SM denotes the state of knowledge (CS – cognitive state) of the learner and her Learning Style (LS) preferences, as couple <\textit{CS,LS}> where
- **CS** is a set of pairs \\{<\textit{lo}, cert(lo)>\}: LOs presently “owned” by the learner are listed and labelled by their estimated certainty (see later).
- **LS** is a 4-tuple <\textit{d}_1,\textit{v}_1; \textit{d}_2,\textit{v}_2; \textit{d}_3,\textit{v}_3; \textit{d}_4,\textit{v}_4>, where \textit{v}_i are in the range [0,1], and \textit{d}_i are values in the dimensions of Felder-Silverman’s model [18] (active/reflexive, sensing/intuitive, visual/verbal, sequential/global.).

The SM possibly changes during the study, allowing for an adaptive retuning of the course.

The acquisition and continuing possession of a LO in CS, and its **certainty**, are determined basing on learner’s answers to end-lesson tests (where a lesson is a segment in the sequence C). The system manages a set of parameters (configurable by the teacher on each subject matter) to drive CS updates during the course taking. In particular, when a \textit{lo} is added to CS after a test, it is assigned certainty \textit{c}_\text{IN}; then, after each further test, cert(lo) is decreased/increased, by \textit{c}_\text{KD}/\textit{c}_\text{OK}, depending on answers; should cert(lo) eventually be under/over thresholds \textit{c}_\text{promote}/\textit{c}_\text{demote}, though, it would be respectively extracted from CS or permanently included in it (with no further tests).

**B. SOCIALX**

SOCIALX [19] (Fig. 3) is a web-based system designed for collaborative and social aspects of learning. It supports the practical/exercise experiences of a course, through the management of socio-collaborative learning activities, in the framework of a reputation system [5] [6][20] [21].

In particular, in SOCIALX the learners can contribute by:
- **providing solutions** to available exercises;
- using others’ solutions to develop their own;
- evaluating others’ and one’s own work;
- discussing exercises in contextual micro-forums, in which direct reward (tokens) is earned for the perceived usefulness of one’s contributions (questions & answers);
- participating to group-based projects, with a “social” bent.

Several projects are available, structured in common stages; at each stage a group is expected to produce a deliverable, also by exploiting deliverables from previous stages of the same project; after stage completion, the group moves to the following stage of another project.

The reputation in SOCIALX is a representation of the following learner’s characteristics and qualities demonstrated during the interactions with the system:
- **involvement**, i.e. the active participation in the system (through the abovementioned contributions);
- **usefulness**: how a learner’s work is beneficial for others (e.g. her/his solutions are taken for extension and reuse, her/his groups deliverables are not harmful for the further advancements of other groups in the project);
- **competence**: measured on the basis of grades and judgments coming from other learners and, especially, from the teacher;
- **judgment and self-judgment**, i.e. the consistence with others’ (and especially teacher’s) judgments;
- **critical appraisal**: the ability to select others’ contributions to be extended or corrected;
- **group_reputation**, which measures aspects of the learner’s activity performed during group work (number of products delivered by the group, number of marks given by the members of the group to products received from other groups, average teacher’s marks on the group’s products).

[Image: SOCIALX functionalities.]

**V. MODEL OF INTEGRATION**

It is worth pointing out that, though Vygotskij’s model is frequently cited in literature about e-learning, at the best of our knowledge a concrete attempt to formalize and put into practice its features in a working software framework seems to be still to come. A first step in the integration of the web-based systems, mentioned in Sec. IV, along the lines of Vygotskij’s research, is the extension of the definition of SOCIALX exercises according to the structure of an LC, by RK/AK/effort. Quite straightforwardly, this allows admitting the SOCIALX learning assets (LCs that are actually mirrors of the exercises) into the LECOMPS pools and courses. They have no questions embedded, though, so we have to allow for a dedicated method of assessment of their AK; in brief, it is the reputation gained by the learner in SOCIALX, as related to the given exercise, which allows to state that its LOs are to
be included, with according certainty, in the CS component of the student model.

Assuming that a course \( C \) can now contain both LCs and exercise-LCs, we thought it essential to let the learner have a more personal support to the navigation of the LCs in the course, besides the default sequencing provided by LECOMPS: that is important due to the presence of exercise-LCs, where prolonged, interactive and collaborative activities are supposed to be met during the course span.

In order to provide such a support, a classification of the course LCs (or, rather, of their Learning Objectives) under the “zones” discussed in Sec. III is helpful.

In the following we assume that \( C=\{c_1, c_2, \ldots, c_n\} \subseteq LD \) is the course personalized for a given student, in the learning domain, according to stated learning goals TK. CS will be the personal Cognitive State of the student. Denoting the knowledge acquired through a LC \( c_i \) as \( c_i.AK \), the overall knowledge provided by \( C \) is \( C.AK = \bigcup_{c_i \in C} c_i.AK \). Accordingly, we’ll call projection of the knowledge domain over the course the set of all LOs relevant to the course:

\[ \prod^{KD,C} = C.RK \cup C.AK \]

(notice that \( C.RK \) and \( C.AK \) are not necessarily disjoint).

First we define the Autonomous Problem Solving region as the set of all the LOs that are in CS with maximal certainty

\[ \text{APS} = \{lo \in CS / cert(lo) = C_{promote} \} \]

As of the zone of proximal knowledge, we have to define which LOs of the KD (projected over \( C \)) are not “too distant” from the SM, i.e., from the knowledge in the CS. Those LOs are the ZPD, that can be shown the learner, to help her in selecting next LCs to take in \( C \).

In this we have to rely on two further definitions:

- a distance metrics: \( d(CS, lo) \) measures how far is a given LO, \( lo \), from the learner’s grasp (that is from the CS set);
- a “daring threshold” \( p() \) on whose respect declaring that a given LO is “not too distant” from CS (and so be in ZPD).

In order to acquire a given \( lo \in \prod^{KD,C} \), the learner is supposed to study along a path of LCs in \( C \). Each LC, \( c \), demands an effort, \( c.effort \), and the distance between CS and \( lo \) is computed as the (minimal) sum of such efforts. So, let \( G(CS,lo) \) be the minimal (wrt effort) subset of LCs in \( C \), to be studied in order to have CS be extended so to comprise \( lo \):

\[
G(CS,lo) = \{c_{lo} \ldots, c_{nk}\} \subseteq C
\]

with \( lo \in c_{nk}.AK \)

and \( \bigcup_{c \in G(CS,lo)} c.RK \subseteq CS \cup \bigcup_{c \in G(CS,lo)} c.AK \)

then it is

\[
d(CS, lo) = \sum_{c \in G(CS,lo)} c.effort
\]

(In Fig. 4, \( G(CS,lo) = \{\text{(red)-circled LCs}\} \) and \( d(CS, lo)=9 \).

The daring threshold should express a measure of how far from the APS the learner is to be reasonably allowed to reach, while undertaking the acquisition of new knowledge. Since such definition appears to be quite restrictive, we extend it, by possibly considering all the LOs in CS. In that, we weight such LOs by their certainty. Then the daring threshold for a given \( lo \) has a dependence on the overall (average) certainty of the LOs in CS that are used to reach it (i.e., those in \( G(CS,lo) \)).

So let us call support set in CS for \( lo \), the minimal subset of CS made of LOs involved in \( G(CS,lo) \) definition,

\[ \text{Supp}(CS,lo) = CS \cap \bigcup_{c \in G(CS,lo)} c.RK \]

(In Fig. 4, \( \text{Supp}(CS,lo) \) is the lighter (yellow) subset of CS, comprising the three LOs, with indicated certainty). An average certainty (\( \text{avgCert()}) \) for a set of LOs can be defined straightforwardly.

Then, the daring threshold is to be defined as a function \( p() \), expecting the average certainty parameter and monotonically increasing with it.

So, the zone of proximal development is defined as the set of those LOs in the course (excluding the APS) that are within the maximum distance \( p \) from CS, as determined by the average certainty of the support set:

\[ \text{ZPD} = \{lo \in \prod^{KD,C} \backslash \text{APS} / d(CS,lo) \leq p(\text{avgCert(Supp}(CS,lo))) \} \]

VI. SOME IMPLEMENTATION DETAILS

The Pool of LCs and LOs defines an acyclic directed layered vertex-weighted AND-OR graph, with the LOs that correspond to OR nodes (as they can be acquired by more than one LC) and the LCs that correspond to AND nodes (as they demand their RK LOs). To compute the distance of an LO, \( lo \), from the CS we compute the minimum subgraph for which \( lo \) is root and all leaves are in CS.

In Fig. 4, the distance of the rightmost darker (red) LO

\[
\text{Figure 4 Computation of the distance } d \text{ of an LO from the CS. (boxes represent LOs with their distance, circles represent LCs with their effort, arrows show the "acquired" relation, the highlighted set of LOs in CS is the support set of the red subgraph)}
\]
(d=9) is obtained by summing the efforts (3, 4, 2) of the minimum set of LC (the (red)-circled ones) that are necessary to reach the LO. In this process we consider only once the weight of any subgraph appearing on multiple paths.

Notice that, if the graph were a tree, the complexity of computing the distance of all LOs would be linear in the number of nodes, as we could apply bottom-up the following three cases: 1) the distance of an LO in CS is 0; 2) the distance of an LC is the sum of the distances of its required LOs plus its effort; 3) the distance of an LO outside CS is the minimum distance of the LCs that acquire it. In the case of a more general AND-OR graph, two different LOs could share part of their acquiring subgraphs, and thus the actual distance of an LC requiring them will be less than the one computed above. Solving this problem is equivalent to finding the minimum-cost abduction proof [22] of the LO, when we map the LOs in CS to unweighted assumable facts and we map LCs to weighted deduction rules.

VII. Group activities

In SocialX group activities are supported, and so they are going to be in the integration we are describing. A problem to address when selecting activities for a group is the determination of the overall group’s CS (and accordingly the overall group’s ZPD). Ideally, the group’s common zone of proximal development should be as large as it is possible, in order to maximize the group members’ gain coming from the common activities; on the other hand it should also avoid to present group members with activities that are, if not in their ZPD, at least not too far away, just in order to have that nobody is left behind. A first possible choice is to use the maximum common core of knowledge, of the group members (e.g. the intersection of their CSs). Unfortunately, this would limit each learner in the group to a ZPD basically defined over the weakest members: this would maximize the effect of outliers on the group ZPD, and probably reduce others’ motivation; moreover this would deprive us of the possibility to leverage the support that could come from more experienced peers, which is a key feature in Vygotskij’s model. Let us note that the dual choice, i.e. building the group’s CS as the union of all members’ CSs, would obtain similar negative results, mainly satisfying the brightest group members (them too being outliers), and leaving the others behind.

To satisfy both weaker and brighter students we may choose, therefore, an intermediate construction where we compute the group’s CS as the union of the members’ CSs, and where each LO has group-certainty assigned as the mean value of the certainty of that LO as in the members’ CSs. This choice reflects more precisely the level of confidence in the possession of the LO by the group members. As a side effect, this model could motivate brighter students to help their weaker peers, in order to both improving the overall group’s CS and letting the group ZPD grow, to comprise further, possibly more interesting, activities.

An alternative to such a direct construction of the group’s CS is in following a reverse strategy and just defining implicitly the group ZPD, by establishing criteria of admissibility of activities for the group of learners. In this case two conditions are defined, by working on the group APs (considered as “firm knowledge”), the group ZPDs, and the cognitive states of the group members.

For the first condition, we want to express requirements both on the group composition and on the selected activities. First, the group’s member’s (firm) starting points (APs) must have some common intersection, and, second, at the same time such common starting point must allow the group to fulfill the activities prerequisites. So, given a group of students ST={s₁, ..., sₙstud} and a set of activities AC={lc₁, ..., lcₙact}, 1) the group’s members must share a common APS, and 2) the activities’ prerequisites are well known by at least one of the members: (APSᵢ is the APS of the i-th learner in ST)

$$(\bigcap_{s_i \in ST} APS_i \neq \emptyset) \land$$

$$(\bigcap_{l_c \in AC} l_c.RK \subseteq \bigcup_{s_i \in ST} APS_i)$$

The second condition states that 3) students in a group ST must have some common proximal development, and 4) an activity lcᵢ in the group AC is admissible for ST iff, although it might be outside of the ZPDs of some members, it is not too distant from them, and it is comprised in the ZPD of at least one member (that could help the others reaching it): (ZPDᵢ is the ZPD of the i-th learner in ST; τ is a threshold to establish admissibility, for a learner, of an activity outside the learner’s ZPD)

$$(\bigcap_{s_i \in ST} ZPD_i \neq \emptyset) \land$$

$$(\forall l_c \in AC, \forall s_i \in ST : d(lc_i, ZPD_i) < \tau) \land$$

$$(\forall l_c \in AC, \exists s_i \in ST : l_c.RK \subseteq \bigcup_{s_i \in ST} APS_i)$$

Notice that if we used APS in place of ZPD, in the definition of the second condition, we would have some of the brighter members of the group left without anything new to learn in each one of the admissible activities, which could diminish their motivation.

VIII. Conclusions

Our long-term goal is the design of a Learning Community of Practice (LCoP) where the teacher plays the role of organizer and supervisor. Along this way, this work supports an adaptive design of the path of exercises that the learner follows during an adaptive course, and the management of an extended student model, gaining the feedbacks coming from the socio-collaborative activities supported by a reputation system. We addressed the problem of defining both an individual Zone of Proximal Development and a group one, in terms of the Learning Objectives stored in each individual student model. The determination of ZPD is done adaptively, according to the
evolving student model; for that we have defined a cognitive distance between the state of knowledge of the learner and the learning objectives to acquire, and a (“daring”) threshold function, dynamically stating the maximum distance, from the learner’s CS, of the learning objectives in the ZPD, according to the certainty factors of the LO already in the student model. Cognitive distance and threshold function may deserve some further work, e.g. to consider meta-cognitive traits. Additional efforts will be directed at evaluating the correctness of the underlying assumptions, and longitudinal behaviour of the overall resulting framework.

ACKNOWLEDGMENTS
This work was partially funded by the Understand IT EU Life Long Learning Program project num. 2010-1-NO1-LEO05-01839.

REFERENCES