Towards a Dynamic Visualization of Online Collaborative Learning

Malik Koné1,2, Madeth May1 and Sébastien Iksal1

1LIUM, Le Mans Université, Le Mans, France
2LARIT, INP-HB, Yamoussoukro, Cote d’Ivoire

Keywords: Collaborative Learning, Visualization, MOOC.

Abstract: Socio-constructivism and connectivism theories pinpoint the importance of collaboration for learning. Nevertheless, the online social interactions underlying the collaboration processes are still not well understood. As a result, learning designers have difficulties creating effective collaborative activities in Massive Open Online Courses (MOOCs). As for online learners, they are often isolated and require a lot of self-regulation to succeed. The research effort presented in this paper covers a review of visualization techniques supporting the online collaborative learning process. Our findings show that some visualizations have the potential to develop the learners’ reflexivity. Therefore, we give an overview of collaboration importance and how it could be enhanced with such visualizations. Our goal is to identify a new approach to visualize learners’ activities in MOOCs, while supporting collaboration and self-regulation.

1 INTRODUCTION

For socio-constructivism and connectivism theories, people learn within a group when they exchange and confront their ideas (Bandura and McClelland, 1977; Siemens, 2014). Discussing opinions, arguing, expressing ideas help learners to evaluate and consolidate their knowledge. While doing so, they explore their Zone of Proximal Development (ZPD) and can gradually shift it towards new knowledge (Vygotskij et al., 1985).

Online collaboration analysis is based on observable electronic interactions between learners. Graphs are often used to visualize these interactions. A set of nodes represents the agents and links between nodes idealize the interactions, but collaboration is also a process and it cannot be fully captured with a static image. Most studies on collaborative learning give an overview of the collaboration at the end of the learning session. They are intended for instructors or researchers, to help them reflect on the course design and to support future pedagogical interventions. In our case, we also want to provide learners with feedback. Yet, if the latter comes at the end of the lesson, it will be too late. Hence, we need to visualize the collaboration as the course unfolds. Also, we will look at visualizations to ensure that temporal information, from the collaborative activities, is not lost. Animations, timelines or interactive visualizations are among the most effective ways of achieving this. We refer to them as dynamic visualizations. In the following section, we define collaboration and discuss the challenges to model collaborative learning. Then we illustrate the change in the Learning Analytics (LA) goal, moving from modelling to supporting learning. In the third section, we will detail the visualization concept and why it is important in the new LA paradigm. Finally, a review of dynamical visualizations techniques, supporting learner’s collaboration, is proposed in the fourth section.

2 COLLABORATIVE LEARNING

Roschelle et Teasley (1995) cited in Dillenbourg (1999) define collaboration as “a coordinated synchronous activity born from the persistent will to share a common perception of a problem”. Dillenbourg perfects this definition by specifying that collaboration occurs between people roughly equal in social status. There is no collaboration if one is an instructor and the other his student.

We call online collaborative learning or collaborative learning the work and process of learners interacting online, supporting each other to achieve a precise common learning goal. For example, in a MOOC, it could be the grouped activities of online
learners interacting and coordinating over a forum to help each other understand the subject matter or to get good grades at the end of course.

Online collaborative learning is based on networks, which we call online learning networks. We define an online learning network as the set of individuals and their interactions in a MOOC. Mathematical tools from the network science provide models to study online learning networks. Graphs for example, can have nodes representing individuals (or social entities), and links between them can symbolize the social interactions. But graphs are not enough to model online collaborative learning. This endeavour has challenges that we will present now.

2.1 Challenges in Modelling Collaborative Learning

Human have a unique talent to learn. We infer efficiently generalizable knowledge from a few specific examples (Tenenbaum et al., 2006). Artificial Intelligences (AIs) based on neural networks, despite their biological inspiration and success – driving cars, winning world class go players– cannot explain how humans learn. Lake et al. (2016) argue that human learning models should be causal models of the world able to support explanation and understanding. These models should learn-to-learn and be reflexive. Bayesian models of cognition may go in that direction but they are not yet able to implement the subtle and intuitive psycho-sociological phenomena described by the theory of mind and at play in collaborative learning.

On the other hand, traditional sociological approaches of collaboration base themselves on field observations and questionnaires, but contingencies in time and human resources usually limit the experiments to small sets of individuals. In contrast, online collaborative learning studies are cheap. It is an opportunity to gather data about behaviours of humans learning. It could help elaborate a model of online collaborative learning. But, with this abundance of information, data scientists need to carefully ground their approaches in learning theories. The exploration of correlations and predictive models may dominate the search for causal relationships and explanations. Kirschner (2016) alerts that the latter should be the essence advancing learning sciences.

Learning Analytics, as the practice of informing learning with statistical analysis (Kirschner, 2016), may not yet be able to produce stand alone collaborative learning models but it could effectively support it, says Baker (2016).

2.2 Supporting Collaborative Learning with LA

In distant education, learners and instructors are isolated from the cohort. They hardly know each other. Collaborating is therefore difficult.

Collaboration does not spur on its own says Dillenbourg (2002). Grouping learners and letting them loose will likely fail to induce a successful collaboration. Agreeing on a common goal, coordinating actions during a substantial time need a coercive force that is usually the role of the instructor. In MOOCs, the number of learners is overwhelming. They are often geographically distant, may live in different time zones and have different learning objectives. In this context, it is not possible for an instructor or even a team of instructors to manage the collaborative learning efficiently.

One option is to help automate the collaboration process with scripts. Scripts are sets of instructions, which in this context tell learners how to collaborate. They may define criteria for team building, specify the role of each member and the means offered for collaboration. Scripts have been LA scientists’ attempts to model collaboration, but over-scripting sometimes lead learners in clumsy situation (Dillenbourg, 2002). This is partly why Baker (2016) argues for a paradigm change. Instead of seeking to create “intelligent tutors” (or AI), which fail to grasp the subtle psychological complexity of human learning, he says provocatively that the community should aim for “stupid systems”... but stupid systems expert in intelligent human support. Following his logic, let’s review some LA tools supporting collaborative learning.

2.2.1 LA Tools

Soller et al. (2005) compile a list of articles evaluating the tools effect on group learning and categorize the different approaches relative to the degree of control devoted to the Learning Management System (LMS).

Mirroring Tools: display gathered data almost in raw format. For example, connections’ numbers, posts’ percentage by such or such user.

Meta-cognitive Tools: present elaborated measures of higher order cognitive functions. Motivation, mastery level are often compound measures built on top of lower level indicators, such as grades, post, # of friend.

Guiding Tools: propose actions to take. This can be a learner to help, a keyword to define a post, the next learning resource to access.
Table 1: We classify collaborative learning articles in two groups: visualization and indicator based studies. Visualization based studies, focus on the impact of visualizations in online collaborative setting. Indicator based studies focus on ways to measure online collaboration and its relation with performance. Different social contexts are covered, small groups (2 or 3 persons), medium (a few tens), big (a few hundreds) and massive (a few thousands).

<table>
<thead>
<tr>
<th>Papers</th>
<th>Social context</th>
<th>Audience</th>
<th>What is measured</th>
<th>Expected output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duque et al. (2015)</td>
<td>Small</td>
<td>Instructors</td>
<td>Interactions</td>
<td>Collective work</td>
</tr>
<tr>
<td>Rehm et al. (2015)</td>
<td>Medium</td>
<td>Researchers</td>
<td>Hierarchical position</td>
<td>Activity &amp; Performance</td>
</tr>
<tr>
<td>Hommes et al. (2012)</td>
<td>Big</td>
<td>Researchers</td>
<td>Social network structure, learners’ motivation &amp; proio-performance</td>
<td>Performance</td>
</tr>
<tr>
<td>Wang et al. (2016)</td>
<td>Massive</td>
<td>Researchers</td>
<td>Interactions</td>
<td>Performance</td>
</tr>
<tr>
<td>Tempelar et al. (2015)</td>
<td>–</td>
<td>Researchers</td>
<td>Online activity &amp; Survey</td>
<td>Performance</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Papers</th>
<th>Social context</th>
<th>Audience</th>
<th>What is visualized</th>
<th>Expected output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anaya et al. (2016)</td>
<td>Small</td>
<td>Instructors &amp; learners</td>
<td>Decision Tree &amp; Collaboration</td>
<td>Learners’ activity</td>
</tr>
<tr>
<td>van Leeuwen et al. (2014)</td>
<td>Small</td>
<td>Instructors</td>
<td>Agreement level &amp; Contributions</td>
<td>Teacher’s activity</td>
</tr>
<tr>
<td>Medina et al. (2016)</td>
<td>Medium</td>
<td>Learners</td>
<td>Activity, Grades &amp; Usability</td>
<td>Users’ satisfaction</td>
</tr>
<tr>
<td>Yousuf and Conlan (2015)</td>
<td>–</td>
<td>Learners</td>
<td>Engagement</td>
<td>Learner’s satisfaction</td>
</tr>
<tr>
<td>Lonn et al. (2015)</td>
<td>–</td>
<td>Instructors &amp; Learners</td>
<td>Activity</td>
<td>Motivation &amp; Performance impact</td>
</tr>
</tbody>
</table>

AI Tools: act on behalf of a human. They may automatically group users depending on a predefined model. The AI model could use a learner’s past performances to automatically adapt quiz difficulty.

Meta-cognitive and guiding tools seem the most promising tools to help user gain insight on their collaborative behaviours, without been overwhelmed by data or proned to over-scripting.

Duque et al. (2015) apply a multidimensional generalization of the median mathematical concept to identify central and peripheral learners based on the activity. The authors use dimensions such as the level of work, interaction, coordination, speed, syntactic correctness and quality to group learners heterogeneous; resting on Wang et al. (2007)’s suggestion that heterogeneous groups induce more collaboration than homogeneous ones. The authors emphasise the need to give more control to the instructors over the heterogeneity degree used to group the learners.

In our review (Table 1) we notice that few LA researchs focus consider the learners as the final tool users. The reason is that learners may not have sufficient knowledge to effectively interpret recommendations made by the LA tools. But Anaya et al. (2016) investigate the students’ reactions to a semi-automated recommendation about their collaborative behaviour. The authors provide a decision tree visualization, giving students insight on the recommendation machinery (Figure 1). The recommendation is based on six meta-cognitive indicators: reputation, leadership, initiative (I-level, I-regularity), activity (A-level, A-regularity). Thus students better understand why they receive the recommendation, this makes it more effective. The decision tree also reinforces the learners’ meta-cognitive abilities by helping the learners gain insight on the way to improve their collaboration.

These studies points to visualizations as promising tools to support collaborative learning from an instructor’s point of view but also from the learners’.
3 VISUALIZATIONS

Kosslyn, cited in (Twissell, 2014), break down the visualization concept in two parts, which he terms visual perception and visual mental imagery. Visual perception points to the bio-physical process. We perceive an object when it reflects light and stimulate the retina’s cones cells. This process is an efficient high bandwidth way to receive information about the environment. It can provide rapid and precise feedback about a specific situation and it is therefore a useful awareness mechanisms. Visual mental imagery identifies the process of representing, manipulating and transforming mental images. It is the cognitive process, the persistent image one can have of a diagram, an animation or a photo when the external stimuli cease. This is what we call visualization.

Visualizations enable creative and holistic thinking. They improve the ability to make effective inferences while translating or making visual analogies reinforces conceptual development. They impact cognition, help sense making and understanding (Twissell, 2014; Klerkx et al., 2014). The designer’s ingenuity is his ability to convey a specific information or an emotion by arranging judiciously symbolic elements (Figure 2).

Animations alleviate visual complexity by matching the temporal dimension with that of the concept to analyse. “A meta-study of algorithm visualization effectiveness” (Hundhausen et al., 2002) describes how animations grounded in cognitive constructivism are an efficient ways to teach algorithms. Indeed, since algorithms are processes, animations facilitate their exploration and understanding.

Navigation with interactive visualizations can optimize topic exploration. INSIGHT, a web application uses Latent Dirichlet Allocation (LDA) to automatically extract major topics from the stackexchange.com network and displays them using bubbles of different sizes and colours. The bubbles’ attributes depend on the semantic proximity of the posts’ keyword and the searched concept.

In the following subsection, we look at other visualizations specifically supporting collaborative learning.

3.1 Supporting Collaborative Learning with Visualizations

Many studies proved visualizations useful from the instructors’ point of view. Theses researches are closely related to the field of Computer Supported Collaborative Work (CSCW). There, users are not necessarily learners. They can be experts in their field and therefore have less difficulties visualizing correctly the information. Gilbert and Karahalios (2009) give another example of a study supporting collaborative users with, CodeSaw, which let users infer knowledge from the logged data generated while working on an open source project. The authors advocate the use of simple visualizations, leaving most of the reasoning on the users’ side.

In Computer Supported Collaborative Learning (CSCL), collaboration visualizations mostly deal with supporting exploration of the user’s social context (Heer and Boyd, 2005; Klamma et al., 2006), or to help them collaborate by enhancing their reflexivity. The difficulty prevails with visualizations aimed at informing the learners but a few examples exist.

3.1.1 Visualizations for Learners

In cognitive science, reflexivity is the meta-cognitive ability to visualize one’s own cognitive abilities. It is, for example, the ability to answer the following questions: do I think that I am able to play tennis? How should I change my behaviour to succeed in that online course? It is closely related to Zimmerman (1990) Self-Regulated Learning concept. Reflexivity

![Figure 2: Dmitry Mendeleev’s periodic table. The clever organization of information provides a deep understanding and facilitates memorization of the chemical elements’ properties. It illustrates “how a diagram is (sometimes) worth ten thousand words” (Larkin and Simon, 1987).]
helps learners build a model of their own learning behaviour, which they can confront to the learner model held by the LMS. Learner Models (LMs) can be as simple as an ID number. It can also have thousands of dimensions trying to model a learner’s motivation, knowledge level. Accessing and visualizing this information is capital for the learner say Bull and Kay (2016). Therefore, this makes it an important element for learners to visualize.

Ideally, learners should also be able to interact with an Inspectable learner model. Govaerts et al. (2012) use learning data to generate feedback for the learners, creating activity awareness and fostering reflexivity. More recently Yousuf and Conlan (2015) found that it indeed helped their learning. But as we mentioned in Section 2.2.1 on page 2, visualizations may not be easily interpreted by learners.

### 3.2 Visualizations Limits and Bias

Klerkx et al. (2014) review visualizations applied to learning and find that they have a positive impact on cognitive abilities but Twissell (2014) identifies the following limits: a) different learning styles, natural differences in learners have a significant impact on the way diagrams are perceived, visualized and understood b) visualizations do not equally affect all types of learning activities.

Kahneman (2011) and Dehaene (2007) neuro-psychological findings, pinpoint to our visual and cognitive biases, an anchoring mechanism, which succinctly means that we cannot refrain ourselves from using our personal history to infer meanings. Therefore, there is no objective way of representing an information. Even a “simple” tasks such as evaluating a length or an area is bias. We are more precise at evaluating vertical length than oblique ones because statistically we encounter the first in our environment more often than the latter and developed more neurons to recognize it (Girshick et al., 2011).

Thus far, the choice for the collaborative learning’s representations aimed to learners, should carefully be thought and rest on works from Bertin (1983).

Visualizations can easily be misunderstood, as in this study from Lonn et al. (2015), where they find a modification of the students’ goals related to how often the students visualize their progress. Mastery goal orientated students, which are those seeking a deep understanding, became competency based students (Pintrich and De Groot, 1990), only caring to show external proof of performance.

In the last section we present visualizations which believe could support online collaborative learning from a student point of view.

### 4 Dynamic Graphs

We propose to build a visualization considering the two pedagogical theories: socio-constructivism and cognitivism. Socio-constructivism because the visualization shows the learners’ social online interactions in the MOOC. It emphasizes the collaboration rather than the competition. Learners will not be able to derive each others exact participation from the visualization. The self-regulated learning or cognitivism aspect comes from the feedback provided through the visualization to each learner about his behavioural pattern.

Our visualization should support cognitive competence, such as improved performance, and goal achievement. It should provide some means for comparison but only with cohort of similar students or with one’s future goals and past progression.

We envision that a dynamic network visualization could satisfy those characteristics. The network should be content focused. Nodes would represent forums and not users. Two forums would be linked if the same user posted in both. Also, the size of the forum-nodes, the links’ width would be a function of the user count. The keywords extracted from the forum would help build a conceptual map from the users’ interactions.

Another reason not to represent learners as nodes is a readability issue. There are far more learners than forums discussions. Therefore a network showing forums as nodes should be more comprehensible that one showing learners. Meanwhile, it would still be possible for instructors to transform a forum-network in its dual, that is the a learner-network.

Boroujeni et al. (2017) propose an integrated approach to analyse the dynamic of MOOCs discussion forums. They describe the interplay of temporal patterns, discussion content and social structure emerging from the learners’ interactions. Our objective is to find a visualization carrying this information to learners. In the following subsection we summarize the different visualization techniques for dynamical graphs.

#### 4.1 Dynamic Graph Visualizations

A collaborative learning visualization could build upon dynamic graph visualizations. Directed graphs (whose links have directions) can be represented in four ways (see Figure 3 on the next page). 1) Force-directed, strong links are the shortest and nodes are correspondingly placed. 2) Orthogonal, nodes are placed on an orthogonal grid. 3) Hierarchical, links define node hierarchies. They tend to always go in
the same direction, for example from top to bottom. Nodes can therefore be grouped based on their vertical position. 4) Matrix, nodes are represented by the matrix’s lines and columns. The line-column intersections correspond to the interaction between two nodes. A “1” value at an intersection code for the existence of a link, a “0” for its absence. Intermediaries values can quantify the link’s strength.

Beck et al. (2017) further identify two major categories of dynamic visualizations: animations and timelines. By dynamic visualization we consider all visualizations evolving in time, and because collaborative learning is a process, therefore evolving in time, we believe that a dynamic visualization will reduce the learners cognitive load.

Animations integrate time changes in unique moving representations. They are general-purpose, special-purpose or matrix based representations.

General purpose methods are further subdivided in online and offline approaches. In offline approaches, the full sequence of the interactions is known. For example in a playback situation the user sees how the graph evolved up to present. In that case, the graph layout can be optimized to smooth transitions from the first to the last layout. For online, or real-time, approaches. The full set of interactions is not known. The graph layouts are arranged without knowing the future. The transition problem, here, is salient. How should nodes and links’ addition deletion be displayed to keep a visually comprehensible graph? One of the approaches proposed by Beck et al. (2017) is that of Friedrich and Eades (2002): Marel. It consists of four phases. 1) nodes and links are removed; 2) the whole graph is translated to a new optimal position; 3) each node takes its new position; 4) new nodes and edges are placed.

Special purpose techniques are particularly useful for visualizing subcommunities, if we know for example that the graph has a hierarchical structure.

In general, node-link animations are well suited for medium sized graphs (with a few hundred nodes). Bigger graphs rapidly become unreadable even for experts. A common option, is the to use colours to mark different communities. But structural changes may always be difficult to track on big scales, and if we were to represent learners as nodes we would need other techniques to provide a collaborative learning overview.

Other techniques include timeline, matrix and hybrid visualizations.

Timelines map different timesteps in space. Each timestep requires a graphical representation. Therefore changes in the interactions give rise to several graphs combined in one visualization. Such techniques shrinks each node-link graphs, but it facilitates comparison. Timelines make it easier to analyse structural changes when nodes are at fixed position (see Figure. 4)

Matrix visualizations are found both with animations and timeline techniques. Due to its concise size, matrices scale better than other approaches. It can conveniently represent data with several thousand nodes.

A few hybrid methods exist. They combine timelines and animations. For example, a selection of timesteps could animated on request by the user. This brings up our discussion on the interactions.

As Yi et al. (2007) recall, information visualization has two parts: representation and interaction.

4.2 Interactions to Support Visual Analytics

Heer and Shneiderman (2012) emphasize the importance of interaction in information visualization. With Yi et al. (2007), they provide two overlapping user centred taxonomies listing the tools to support a fluent use of visualizations. a) Data & View specification tools filter out, sort and reconfigure the data

Figure 4: Three timesteps representation of the same graph. In Δ₁, some users posting Forum 3 (F₃) also posted in F₂. Some from F₂ also posted in F₇. Some users posted in F₁ but did not post elsewhere during this time interval. In Δ₂, a new forum (F₅) and a new link (F₁ → F₇) are created. So people are still posting in the previous forums. Finally in Δ₃ since nobody posted in F₆ it is removed.
to give it meaning, b) View manipulation tools select items, navigate, explore, abstract/elaborate, encode the information, c) Process & Provenance tools record, annotate, share visual configurations.

We propose to integrate the following interactions in our collaborative learning visualization: a) filter forums based on keywords, time stamps, posting/reading users, b) filter links based on quantitative threshold, time stamps, users’ characteristics, c) sort the nodes’ dispositions based on forum creation date, most recent post, topics, or a hierarchy derived from the users’ interactions, d) get detailed information about a selected forum, e) see the path from forums to forums for a selected user or cohort, f) zoom, pan, scroll the forum network map, g) enable an overlay instructor based concept map, h) annotate, comment and share visualizations.

We understand that, as shown by Jivet et al. (2017), awareness is not enough. With our interactive visualization, we hypothesize that the users will engage in collaborative activities to find topics of interest and collaborative partners, supporting their learning goals without fostering competition. They develop self-regulated competences.

5 CONCLUSION
In this paper we presented the challenges related to modelling learning and collaborative learning. We explained why the LA community focuses on designing and evaluating tools supporting and not replacing human learning. We argued that visualizations can be used to support some parts of learning if they are well grounded in learning theories. We found that static graph images of collaboration are not suitable to support learners because they are only available at the end of a course. Instead, we proposed a dynamic visualization to support collaborative learning in a self-regulated learning and socio-constructivist framework. Hence, our future challenge is to represent in a real-time and intuitive way to the learners, the collaboration dynamics. We plan to achieve this using a dynamic and interactive graph, which learners could explore and derive insights from, about their learning behaviour.

We are in the process of building a prototype to evaluate how this will scale up to our final objective, supporting MOOCs’ collaborative learning.

REFERENCES


