Software Agents in Large Scale Open E-learning: A Critical Component for the Future of Massive Online Courses (MOOCs)

Roxana Bassi [1], Thanasis Daradoumis [1,2], Fatos Xhafa [3], Santi Caballé[1], Ardiana Sula[4]
[1]Department of Computer Science, Multimedia, and Telecommunications
Open University of Catalonia, Barcelona, Spain
rbassi0@uoc.edu, adaradoumis@uoc.edu, scaballe@uoc.edu
[2]Department of Cultural Technology and Communication (D.C.T.C)
University of Aegean, Mytilini, Greece
[3]Department of Languages and Informatics Systems
Technical University of Catalonia, Barcelona, Spain
fxhafa@uoc.edu
[4]Sacred Heart University, USA
ardianasula@gmail.com

***Abstract* — MOOCs or massive open online courses are a recent trend in online education. They combine online resources with social tools and have unique challenges due to the large number of simultaneous participants. This paper analyzes some of the challenges in the areas of MOOC design, delivery and assessment. Then the authors present an approach using software agents to overcome some of the challenges that have been identified, as well as optimize efficiency, reduce costs, and ensure the pedagogical effectiveness and educational quality of large scale online learning courses. This paper is a first step towards research in the usage of software agents in massive online courses that we hope will shed more light on potential real life applications.**

Keywords— MOOC, OER, Software Agents, E-learning, Virtual Learning Environment, Artificial Intelligence.

I. INTRODUCTION TO MOOCS

MOOCs or massive open online courses have become a new popular theme of discussion in debates about online education. MOOCs can be defined as online courses based on open educational resources (OER), with a large number of simultaneous participants, and including interaction among participants using social tools.

According to Siemens, MOOCs are a continuation of the trend in innovation, experimentation and the use of technology initiated by distance and online learning, in order to provide learning opportunities for large numbers of learners [15]. Supported by many, feared by others, cautiously treated by the traditional academic community, MOOCs seem to offer the chance for millions to take part in true "education for all". As of December 2013 [6] there are more than 1,100 active MOOC courses worldwide, and the trend continues to grow as new institutions launch further courses in more languages. As summarized by the objective of one of the MOOC platform sites (Coursera, [3]) "We envision a future where everyone has access to a world-class education that has so far been available to a select few".

Although many MOOC formats exist, according to a sample made in 2012 [8] most courses exhibit common defining characteristics that include massive participation, online and open access, lectures formatted as short videos combined with formative quizzes, automated assessment and/or peer and self–assessment, and online forums and social applications for peer support and discussion.

MOOCS have distinctive characteristics, including supporting thousands of simultaneous remote students during a single course and adding the power of online community interaction to the learning process. What emerges from existing experiences is that it is not possible to design, deliver, manage or assess a MOOC the way a traditional e-learning human-tutored course is. In a critical analysis of the MOOC model [2], it has been observed that all MOOC initiatives are still delivering a Learning 1.0 product in a Web 2.0 world. He indicates that according to his view the platforms have replicated all of the problems of the traditional industrial-age model of lecture-based teaching and testing that has minimal linkage to student outcomes.

The fact that MOOCs are different from traditional e-learning became clear when a Coursera MOOC was abruptly closed leaving 40,000 thousand students out. Siemens reflected on the source of frustration for the participants "Faculty own the content, Coursera owns the platform. But neither should own the conversation. That belongs to the learners" [16]. It looks like MOOCs are challenging the very essence of traditional education, and as all changing forces, they still need to evolve into more practical and realistic applications. MOOC models are evolving quickly [15] and there is the need to use imaginative new solutions to overcome some of its disadvantages as well as making them sustainable.

In 2012 and 2013 several new organizations have been born in order to provide MOOC infrastructure and services, mostly for free and some including paid value added services, like certification or identity validation,

978-1-4799-6387-4/14 $31.00 © 2014 IEEE
DOI 10.1109/INCoS.2014.15

for a small fee [17]. Though MOOCs may be free to learners, they are not free for institutions that design, host and manage them. One source indicates that the average cost of just developing an xMOOC can be between US$50,000 – US$100,000 [1]. Therefore for MOOC owners, these types of courses are expensive to create and manage, and many different business and educational models are currently being tried out in this very competitive arena. Though many options for economic sustainability have been suggested, the commercial model to make them sustainable is not clearly visible – yet.

Certainly, there is a constant evolution in MOOCs design, delivery and assessment, but MOOCs unique characteristics in an evolving environment still pose specific challenges and several issues that need improvement: better instructional design, improved training of MOOC teachers and moderators, improvements in the platforms to monitor student progress (e.g., more effective software modules for learning analytics), and improved assessment methodologies.

We do not disregard that these issues are very important and should be taken into account. However, together with these aspects we also need a non-intrusive mechanism to add new features and capabilities into MOOCs management. In the following sections we will discuss some of these challenges in more detail and explain why we believe software tools like agents can be used to overcome some of the drawbacks that have been identified, as well as optimize efficiency, reduce costs and improve the learning process in large scale online learning environments.

II. SOFTWARE AGENTS

According to MIT software agents group [12], agents are "computer systems to which one can delegate tasks. Software agents differ from conventional software in that they are long-lived, semi-autonomous, proactive, and adaptive". Software agents can use artificial intelligence methods in order to analyze information and react to it.

The opportunities for using agents in e-learning courses are enormous. Agent characteristics like autonomy, abilities to perceive, reason and act in specialized domains, as well as their capability to cooperate with other agents makes them ideal for e-learning applications [13].

Research suggests that there are also simpler agents that are not necessarily intelligent, and the ability to incorporate a non-intelligent agent into a multimedia learning environment with relative ease may increase the effectiveness of the environment at minimal cost [14].

Agents have been used for years in educational environments, first in Intelligent Tutoring Systems (ITS) and lately in Virtual Learning Environments (VLE). The potential use of several agents in e-learning environments has been researched many years before the advent of MOOCs, defining some potential roles for them like for example Pedagogical Agents (tutor, mentor, assistant), Web Agents (working with Internet applications and social networking tools), Learner's agents and mixed agents which could teach and learn [10]. Though the very first MOOC "AI Class" at Stanford used a rather sophisticated software to assess student responses and provide feedback, no research has reported so far evaluation of the use of agents in the management or delivery of MOOCs.

In MOOC environments, agents could be used to analyze data produced by the MOOC platform, systems and participants, and use it intelligently or mechanically to improve design, delivery and assessment. This article is a continuation of a framework already proposed to integrate agents in MOOCs [4].

III. KEY AREAS OF POTENTIAL AGENT USE IN MOOC DESIGN, DELIVERY AND ASSESSMENT

In the following sections we will further discuss some of the current key limitations and challenges of MOOCs, and present some ideas on how software agents could be used to improve design, delivery and assessment of large scale e-learning systems.

A. Design.

Currently the average MOOCs consist of a lesson plan lasting a few weeks, with online content (text, presentations, video, animations, and additional resources), interactive applications, social networking among students and automated or peer-reviewed evaluations.

One of the most critical challenges is that massive attendance to the same course content by thousand of participants with different cultural backgrounds, educational levels, languages and objectives makes it difficult to ensure that all participants are actually learning. In a face to face course this could be managed by the teacher adapting content according to feedback collected, physical clues from participants, answering questions, providing additional reading and explanations as needed, etc. But this adaptive and interactive response to needs is impossible to emulate in a massive online course. What can be done in order to minimize learning problems is invest heavily during the content planning, design and review stages [9]. That way we can make sure content is easy to understand, politically correct and comprehensible for a very diverse audience with individual learning objectives. If not enough attention is paid to course content, the learning environment could be perceived by participants as frustrating because it lacks adaptability to their specific needs and learning styles. Therefore a potential improvement could be the use of agents to provide personalized content and tutor support according to each participant's profile and needs.

1) Personalized Content Potential Improvements

Dynamic content customization could be planned taking into account for example some of the specific variables identified below:

- User profile: country of origin, location, socio-cultural background, language, religion, ethnicity, units of measure, customs and other cultural characteristics,

education level, prior background knowledge, disabilities, as well as specific needs, interests, preferences.

- Technical specifications: data display formats according to the user's preference, access equipment or technical limitations (i.e. formats, language, players, bandwidth, subtitles, predefined links, access speeds, browser).

-Learning style: identified preferred learning styles and tools (for example, visual, sounds, reading, animation), more practical or theoretical content, with references or additional readings according to each participant's profile, education level, interest and location.

- Time factor: including factors/restrictions that affect dedication to the course such as: employment, current studies, family obligations, hobbies, organizational capabilities, geographical location, time of the year and day.

- Engagement, motivation, responsibility and attitude: student's motivation for taking the course (including the reasons for which a student chooses a course), engagement level, responsibility, expectations of the course, emotional state, response to assessment and success rate per section, speed at which the participant submits the answers to quizzes, current motivation and satisfaction, number of times he/she accesses help pages, participation in forums, number of support requests, level of participation and interaction within the learning community, etc.

2) Tutor Support Potential Improvements.

Software agents integrated in the MOOC management tools could help tutors detect and plan improvements in the initial design of the course they are tutoring for future cohorts, do adjustments to ongoing courses or even "re-shape" a current course if critical problems show up. Options for tutor support are further discussed in the next section on delivery and management.

B. Delivery and Management.

As mentioned before, the current average MOOC has several thousand participants learning at the same time, participating in an online course lasting a few weeks. Massive participation raises concerns about isolation and overwhelming student-instructor ratios [16] and makes it impossible to attend to student's needs and request individually, as would be the case in a more traditional e-learning environment. In order to provide support, most MOOCs use forums and peer support tools, support documentation and tutor assistance through email addresses and social tools to detect and sort out potential problematic issues. Cultural sensitivity and languages also have to be considered in user support for such a diverse audience, as well as 24 by 7 support to cover all time zones. Constant content monitoring is also critical, in order to detect missing resources, broken links, potential cheating information and other technical inconveniences that can be preventing participants to accessing course content or affecting the learning process. In order to care for all participants adequately, some MOOCs employ an army of teaching assistants and volunteers to provide support, answer questions, encourage participants and clean up inadequate content in forums [5]. This complex support structure poses organizational challenges and increases costs.

Using information collected from platforms logs, forums, activity and user profiling, software agents could detect and alert designers, managers, participants and tutors in the areas described below.

1) Recommender System

Modern learners are interested in optimizing the time spent in learning activities and their effectiveness with respect to their individual capabilities, expertise, preferences and learning objectives. In this regard they would like the MOOC environment to understand their learning style and adapt the learning situation in terms of specific content, didactic approaches, the type of media to be used, the way concepts to learn are sequenced and so forth, while getting adaptive/personalized feedback for improving their performance and their motivation as well [4]. Besides enhancing the learning process, recommender systems can help course designers discover how certain course content given in a specific order, help some students more effectively at different points in a course.

2) Content Monitoring.

Standard e-learning courses store a large quantity of valuable data on course usage and access, though it might not always be easy for tutors to interpret it or access it when it is needed. MOOCs present a critical improvement from the small populations of most online courses: large populations can be used statistically to detect trends, make inferences more accurately and identify anomalies with more certainty. Agents could be used to provide statistical analysis of content access that could indicate content that is not accessed, or that is frequently accessed, as well as access errors. This can help in avoiding errors and also improving the course content. As summarized by the MOOC platform Edx [5] "By carefully assessing course data, from mouse clicks to time spent on tasks, to evaluating how students respond to various assessments, researchers hope to shed light on how learners access information and master materials, with the ultimate aim of improving course outcomes".

Agents could help identify potential problems, gaps and limitations of the initial course design, for example improper planning, improper distribution of course constituents, inadequate time assignment to the course different issues, errors in tests and evaluations, etc.

Course managers could also receive information collected by agents that would allow them to analyze the cost/effectiveness ratio of the courses, measure the quality of the learning offering provided, predict success/failure and drop-out rate of their learners and adjust their learning offer accordingly.

3) Student support and guidance (feedback).

In order to sustain course management for massive attendance, MOOC managers, tutors and assistants need tools that help them serve a large number of participants with fewer human resources, optimizing the use of their time.

Software agents could be used for real time analysis of content accessed by the participants, in order to detect potential problematic areas, solve some pre-defined detected issues automatically, enhance participants support and optimize the use of human support time by making them focus on the most critical issues.

Because of the high number of participants, the activity of all the participants becomes a database of information that can be statistically analyzed in order to detect deviations from the norm. In this way, for example, an exceptional performance could indicate an exceptionally gifted student or potential cheating, while on the other hand below average performance could indicate a demoralized students facing some problems.

As for participants, they could be automatically alerted about missing task completion, deadlines, deviations from the norm, encouraged, etc.

Automatic systems can be used to monitor support accounts and discussion forums for critical issues or problems and notify tutors, for example selecting the most adequate tutor according to time zone, expertise, language or culture, in order to optimize response times.

C. *Assessment.*

Assessment is a critical area for MOOCs because unless the new knowledge acquired can be evaluated in a valid, certified manner, online courses will have limited recognition in academic and professional environments.

Assessment has two main areas of concern: pedagogical and technical. The pedagogical part is about how to measure the actual knowledge acquired. In a recent course one of us attended (Egyptology in Coursera) one participant commented in the forum "Without an evaluation adapted to the learning objectives, what is the difference between doing a MOOC and watching a video on YouTube?" Butin [2] reflected on the scarce creativity demonstrated by one of the main MOOC platforms when it comes to improving assessment. He suggested that the platform should have created an adaptive testing model in which students are presented with a question about a lecture topic at an appropriate level of difficulty based on their previous correct or incorrect answers to previous segments. He also indicated that the platform could have taken all of its student profiles and usage data in order to create personalized feedback and evaluation.

The other, more practical side of assessment has to do with conducting the actual evaluations in an effective manner. MOOCs have unique characteristics that make conducting evaluations more difficult even than in traditional e-learning environments. The virtuality of the course makes it difficult to guarantee the identity of the student or ensure the adherence to "ethical codes" that reduce the possibility of fraud and cheating. The massivity of attendance makes it impossible to provide marks and feedback that is not either automated or peer assessed [8]. And the online community factor makes it even more difficult to reduce sharing of test results, cheating and required re-designing the tests for every new cohort.

A large body of research on automatic grading systems exists, which could be used in the context of MOOCs [11]. As a consequence current MOOCS are either evaluated using automatic applications (quizzes, auto-graders, robots) or peer evaluation systems, with other more creative systems, like project-based evaluations, being explored. To solve the identity validation issue, platforms like EdX and Coursera are launching systems in order to verify the participant's identity and to certify his or her presence during the online tests. These methods include the use of webcams, biometric like typing pattern and even human-monitored identity controls.

Agents could improve assessment further as we describe in the next sections.

1) *Personalization and customized testing.*

As mentioned before, MOOCS attract different type of participants with different educational objectives, background education and learning needs. Besides customizing course content like we described before, agents could help in defining evaluation parameters that are personalized according, for example, to the participant's educational level, previous performance, etc.

2) *Testing delivery.*

By analyzing usage parameters and content, agents can be used to detect potential cheating (like plagiarism) and alert tutors of sharing of critical evaluation content on social networks.

Software agents could be used as part of the methods to validate the identity of a person, using, for example, typing patters and face recognition.

Agents could also help in delivering randomized evaluations that are difficult to share on social networks, and make it simpler to design tests for next cohorts.

3) *Performance.*

Emerging student patterns in MOOCs suggest that different participants can have different objectives when participating in a MOOC [7]. Therefore just evaluating the new knowledge acquired may not exactly satisfy all needs, or represent the real value of a course. Agents can be used to improve testing methods in order to measure other achievement indicators, besides delivering automated tests. Agents could be used in designing evaluation indicators that consider other parameters like peer-evaluation, social participation, creative thinking, problem solving, application of knowledge to a local reality, etc.

Agents can, by providing more information on usage patterns, also be used to improve retention rates and reduce drop-outs.

Student satisfaction is another area where agents can be used. By collecting platform usage information and

profiles we can discover how to measure student satisfaction with the course, how the satisfaction evolves as the course progresses, and even act in order to stimulate discouraged participants before they abandon the course.

IV. CONCLUSIONS AND FUTURE WORK

As we described in this article, on one hand we believe the use of agents in MOOC, supported by data mining services and learning analytics to accomplish their tasks, can have a positive impact on course planning, management, delivery and assessment. On the other hand, we are concerned that the power of software agents has not been yet explored and applied to MOOCs, so it constitutes an interesting and challenging area of research. This gap in research is probably due to some of the drawbacks of agent usage, for example the fact that their introduction in a virtual leaning environment can be a quite complex and demanding task, as well as increase costs. From the design point of view, designing a course that can be personalized according to the participants' profile, patterns and other variables identified in the previous sections, adds a significant workload because it requires programming an intelligent agent that is able to provide adequate support to content planning and development. Agent design also requires employing highly qualified programmers that are able to program artificial intelligent adaptive algorithms, abilities that are scarce and expensive.

We believe that cost considerations are a key reason why agent usage has not been employed further in massive e-learning. In an environment where the business model of MOOC is still unclear, any further cost and time demands are difficult to justify. That is probably why we don't know of any MOOC platform that integrates any agent- supported functionalities yet. One potential solution to the cost issues would be to make use of the open source community to help develop software agents as add-ons to existing open platforms.

We are starting a research hoping to shed more light in the coming years on the potential use of agents in MOOC environments. An important step to this end is to explore first in depth all the approaches that have been undertaken in order to improve MOOCs design, delivery and assessment (including potential agent use in MOOCs). Doing so, we will be able to identify positive or negative examples and see why and how agents can fit better for MOOCs improvement.

The next step will be that of identifying the most suitable agent design characteristics that provide an optimal support on MOOCs development. This research work will be based on an in-depth analysis of student satisfaction and successful performance with MOOC courses that will lead to define and measure specific analysis indicators which in turn will guide the most adequate agent design for MOOC support at all three levels of course design, delivery and assessment. The aims of agent incorporation in all the above stages of a MOOC course are to optimize efficiency, reduce costs, and ensure the pedagogical effectiveness and educational quality of large scale online learning courses.

REFERENCES

[1] Bates T., 2013. Look back in anger? A review of online learning in 2013 (online). Retrieved from http://www.tonybates.ca/2013/12/15/look-back-in-anger-a-review-of-online-learning-in-2013/ on 12/12/2013

[2] Butin D., 2012. What MIT Should Have Done (online). Retrieved from http://elearnmag.acm.org/featured.cfm?aid=2263018 on 20/12/2013

[3] Coursera, 2013. Coursera website (online). Retrieved from https://www.coursera.org/about on 01/11/2013

[4] Daradoumis, T., Bassi, R., Xhafa, F., Caballé, S., 2013, A review on massive e-learning (MOOC) design, delivery and assessment. In proceedings of the 8th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC 2013), Compiegne, France. October 28-30, 2013. pp. 208-213.

[5] EDX, 2013. Exd website research section (online). Retrieved from https://www.edx.org/research on 01/11/2013

[6] European Union, 2013. Open Education Europa, European MOOCs Scoreboard (online). Retrieved from http://openeducationeuropa.eu/en/european_scoreboard_moocs on 17/12/2013

[7] Hill P., 2013. Emerging student patterns in MOOC. Retrieved from http://mfeldstein.com/emerging_student_patterns_in_moocs_graphical_view/ on 10/10/2013

[8] Glance D., Forsey M., Riley M., 2012.The pedagogical foundation of massive online courses, First Monday Journal (online). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/4350/3673 on 12/6/13

[9] Guàrdia L., Maina M., Sangrá A., 2013. MOOC Design Principles: A Pedagogical Approach from the Learner's Perspective. Retrieved from http://elearningeuropa.info/sites/default/files/asset/In-depth_33_4_0.pdf on 21/6/13

[10] Martins Giraffa L.M. and Vicari Rosa R.M., 1998. The Use of Agents Techniques on Intelligent Tutoring Systems. Retrieved from http://lsm.dei.uc.pt/ribie/docfiles/txt200342413856156.PDF on 3/5/2013

[11] Matthews, K., Janicki, T., He L., Patterson, L., 2012. Implementation of an Automated Grading System with an Adaptive Learning Component to Affect Student Feedback and Response Time. Journal of Information Systems Education, Vol. 23, No. 1. 71-83.

[12] MIT, 2013. MIT software agents group website. Retrieved from http://agents.media.mit.edu/ on 10/10/2013

[13] Papazoglou M.P., 2001.Agent-oriented technology in support of e-business, Communications of the ACM. Vol. 44 No. 4, pp. 71-7.

[14] Schroeder N., Adesope O., 2012. A Case for the use of Pedagogical Agents in Online Learning Environments. Retrieved from http://jotlt.indiana.edu/article/view/2126/3045 on 2/6/2013

[15] Siemens G., 2013. Massive Open Online Courses: Innovation in Education?. Commonwealth of learning, Perspectives on Open and Distance Learning: Open Educational Resources: Innovation, Research and Practice, p. 5. Retrieved from http://www.col.org/resources/publications/Pages/detail.aspx?PID=446 on 21/5/13

[16] Siemens, G., 2013. Negating the learner in the learning process (online). Retrieved from http://www.elearnspace.org/blog/2013/02/03/negating-the-learner-in-the-learning-process/ on 10/12/2013

[17] Watters A., 2013. Top Ed-Tech Trends of 2013: MOOCs and Anti-MOOCs (online). Retrieved from http://www.hackeducation.com/2013/11/29/top-ed-tech-trends-2013-moocs/ on 1/12/2013.