Tracing a little for big Improvements: Application of Learning Analytics and Videogames for Student Assessment

Ángel Serrano-Laguna, Javier Torrente, Pablo Moreno-Ger, Baltasar Fernández-Manjón
Department of Artificial Intelligence and Software Engineering
Complutense University
Madrid, Spain
{aserrano, jtorrente}@e-ucm.es, {pablom, balta}@fdi.ucm.es

Abstract—**Assessment is essential to establish the failure or success of any educational activity. To measure the acquisition of the knowledge covered by the activity and also to determine the effectiveness of the activity itself. The increasing adoption of new technologies is promoting the use of new types of activities in schools, like educational video games that in some cases are developed by the teachers themselves. In this kind of activity, interactivity increases compared to traditional activities (e.g. reading a document), which can be a powerful source of data to feed learning analytics systems that infer knowledge about the effectiveness of the educational process. In this paper, we discuss how a part of the students' assessment can be achieved semi-automatically by logging the interaction with educational video games. We conclude that even the application of rather simple tracking techniques means an advantage compared to other systems that are fed with less quality data.**

Keywords: Learning Analytics, educational games, data mining, assessment

I. INTRODUCTION

Increasingly, teachers of all education levels and knowledge branches are becoming enticed by the possibilities new technologies can offer to their daily work. Among other activities, teachers are starting to use educational videogames in order to explore new ways to educate their students [1]. Although there is evidence to support that characteristics of videogames such as high interactivity, supply of engagement and challenge can improve the educational processes [2], [3], most teachers are reluctant to use them as assessment tools and usually end up turning to traditional methods, like written exams. Videogames are eventually left as low-weight complements that have little or no impact on the final mark, even when games can support new assessment approaches [4], [5] as they foster problem-solving, critical thinking, observation and reasoning.

There are solid reasons that motivate this distrust. First, and probably most important, it is difficult to implement Q&A (Question and answer) structures with games. In games students are constantly solving problems at a certain pace that is designed to make the game challenging but not frustrating and the explicit introduction of a test or questionnaire in any form is seen as disruptive and breaks the immersive atmosphere [6]. Tracking students' interactions and extracting conclusions about the process seems a better approach for the assessment of learning through games. But to do so, teachers need tools that analyze data coming from the games and infer high-level knowledge that can be understood by a human (e.g. the student has problems with concept 'A' or skill 'B').

Most of commercial games used by educational institutions, developed by publishers or government education departments are distributed as black-boxes in order to protect intellectual property. As a consequence it is unfeasible to collect any kind of data from them, leaving teachers with the only choice of using traditional methods, i.e. written exams.

Teachers doing their own games are using easy-to-use game editors [7]. These tools give them more control of the whole game, which allows them in some cases to extract information for evaluation purposes. However, to maximize effectiveness it is necessary that these tools include features to track students' interactions and display high-level results in a teacher friendly way without requiring technical knowledge.

Assessment systems of this kind can be complex to develop as they require advance techniques of data mining and precise information about the game and study domain. Nevertheless in this paper we discuss that even simple systems tracking a few types of traces can produce a lot of assessment information with little teacher intervention.

Our approach is framed in the Learning Analytics discipline, an emerging field [8] that advocates for the analysis of students' interaction data with online educational resources to better understand their learning process. This field is directly related to other disciplines that apply data mining processes, like Business Intelligence or Web Analytics. While Business Intelligence provides advice to take action on certain alerts by analyzing numeric data, Web Analytics is centered on the analysis of users interaction with web pages, and its main purpose is to report the collected information, but no conclusions are drawn automatically (data must be interpreted by the webmaster).

The next organizations have partially supported this work: the Spanish Ministry of Science and Innovation (grant no. TIN2010-21735-C02-02); the European Commission, through the Lifelong Learning Programme (projects "SEGAN Network of Excellence in Serious Games" - 519332-LLP-1-2011-1-PT-KA3-KA3NW and "CHERMUG" - 519023-LLP-1-2011-1-UK-KA3-KA3MP) and the 7th Framework Programme (project "GALA - Network of Excellence in Serious Games" - FP7-ICT-2009-5-258169); the Complutense University of Madrid (research group number GR35/10-A-921340) and the Regional Government of Madrid (eMadrid Network - S2009/TIC-1650).

We propose a double approach. In first place we use a similar philosophy to Web Analytics: data collected from interaction is reported to teachers using convenient graphics. In this case data is treated in a game-independent manner and teachers are responsible to give meaning to those results, as no automatic inference is possible. In second instance we propose adding an extra layer that allows the automatic inference of conclusions by using game-specific data. In this case additional input from teachers is required, as they have to define assessment rules that help the system to infer conclusions.

This paper is structured as follows: first we present the types of traces that we propose to log to facilitate assessment, and what information can be extracted from them. Second, we discuss what kind of teacher-defined assessment rules are needed to reach a more fine-grained inference of conclusions and high-level generation of *knowledge* about the educational process. The last section discusses final remarks, restrictions and possible extensions.

II. TRACES LOGGED

Educational video games can be varied and embrace many genres. However, most games share a certain common characteristics and therefore a basic but fundamental set of interaction traces can be defined. The more concrete the game design is, the better these traces can be defined. However in this paper we elaborate on a general basis that could become the breeding ground for more elaborated models.

Our basic set of traces is presented below, with all the information associated that can be extracted from them.

A. Start game, end game, quit game

Generating a trace with a timestamp whenever a student begins to play some game provides information about when and who, if the students are uniquely identified, the game is being played. Adding a session identifier to the trace, we can track how many times any user tries to beat the game. With all this information we can also obtain group stats about the total number of users who played the game.

This information might seem trivial, but having the students uniquely identified is the first step towards automatic assessment.

Also, generating time stamped traces whenever a student finishes the game provides more relevant information. Firstly, this allows for determining if the student accomplished the game (i.e. s/he completed the main goal for which the game was designed). Secondly, by comparing the final and start timestamps the total time spent to fulfill the game can be calculated. Global statistics can be generated as well, like students' success rate and mean completion times.

Start traces are always generated but end traces can be logged only if the student finishes the game. To deal with situations where the student quits the game it is necessary to generate another time stamped trace with information about the state of the game to let teachers know the exact point where students stopped playing. Sometimes, due to game platforms restrictions, the quit trace could not be generated. In those cases, the game state can be logged periodically at a fixed time rate. Checking the quit trace against the newest time stamped game state would provide information about the last game point reached. Moreover, by processing the evolution of these game states relevant information about the students' game play can be inferred.

B. Phase changes

Most of video games structure the narrative in chapters, missions or phases. Every phase can be associated with a secondary or mid-term goal in the game. A typical game structure is to consider the game as fulfilled or completed after all the secondary goals are fulfilled. In some games second term goals must be achieved in a specific order, while in other phases' can be explored freely.

If possible, logging traces with timestamps whenever a student starts and ends every phase provides relevant data. We obtain more fine-grained information about how the student is distributing her/his game play time within the games. Moreover, if the phase exploration sequence is not linear, teachers can gather insight on how each part of the game is being accessed.

Time spent in every phase can be used to identify most time-consuming parts of the game and lead to revising the game design if the results are unexpected (e.g. readjust difficulty, adapt content, etc.).

C. Significant variables

All games use variables to keep some sort of state. By tracking these variables the game play for every student can be reproduced.

A game can contain thousands of variables, but only a few are actually significant for assessment purposes. These variables can contain game scores or opportunities used to complete the game. A trace can be generated to reflect final values for these significant variables or, if necessary, when their values change.

This type of trace depends on the game mechanics and also on the possibility of working with game variables. In some game platforms this information is just kept hidden from the user/teacher.

D. User interaction

Previous traces are based on in-game situations and they contain information about what is going on in the game. But there is also interaction data produced by the user interaction with the game that may also contain relevant assessment information. Low level events such as mouse clicks, screen touches (on smartphones or tablets) and keys pressed register how students are interacting with the game and if all of them were logged and reproduced, teachers could replay the whole game play of their students.

However, in most cases it results inefficient to analyze every single user interaction with the game. Then, it is necessary to filter interactions that are not relevant for assessment purposes.

In our general approach, mouse clicks can be logged in order to create heat maps marking hotspots distributed by phases or game scenarios for example. If the game control is based on keyboard interaction, pressed keys can be logged to check if students are using controls correctly.

When the educational videogame is multiplatform (having versions for desktop, mobile platforms, etc.) or have more than one input method (e.g. voice, mouse, keyboard, etc.) it may be necessary to have an abstract low level interaction model all different could be mapped onto to enable comparison.

III. Extracting information

Once all data are collected from the game, we can start to extract some information from them.

A. Derived and combined data

The aforementioned types of traces can be automatically collected and displayed in human-readable reports to the teachers. The only requisite is to have a proper logging system available in the game and visualization tools to represent the data.

A game-independent analysis of data collected is useful, as it allows teachers to identify, for example, how long it took students to reach a certain point B from certain point A. However, *knowledge* can only be inferred automatically if information about the game is taken into account (e.g. structure, images, etc.). Using the previous example, if game-specific information is considered time intervals like those mentioned could be linked to the game structure to generate statistics about phase completion times.

Having game-specific information can be used to improve the quality of the reports as well. For example, heat maps could use real images as backgrounds

Finally, combining different types of traces it is possible to obtain new information. For example, if in some phases users spent more time than expected, phase heat maps could be checked in order to find what the problem is and redesign the game accordingly. This would allow an in-depth analysis of the feedback provided by the game by detecting students' misconceptions.

For example, using heat maps it could be easy to identify situations where the player fails to interact with the right element or repeatedly interacts with an irrelevant character or object. These situations usually depict an incorrect or insufficient use of feedback, leading to an inappropriate level of guidance that prevents the student to reach an optimum decision on what is the best action to progress in the game. By automatically identifying these situations, which is not a great challenge from a technical perspective, teachers or game designers would be able to provide more explicit feedback when needed.

B. Assessment

To generate knowledge about the process it is necessary to build a rule-based system on top of the interaction logging system. This system takes as inputs the assessment rules defined by the teacher and game-specific information. Useful information about the game includes its structure, specific goals, phases, and information about characters and objects present, for example. The system uses the teacher-defined rules to analyze the interaction logs and infer conclusions which are reported on high-level terms (see Figure 1).

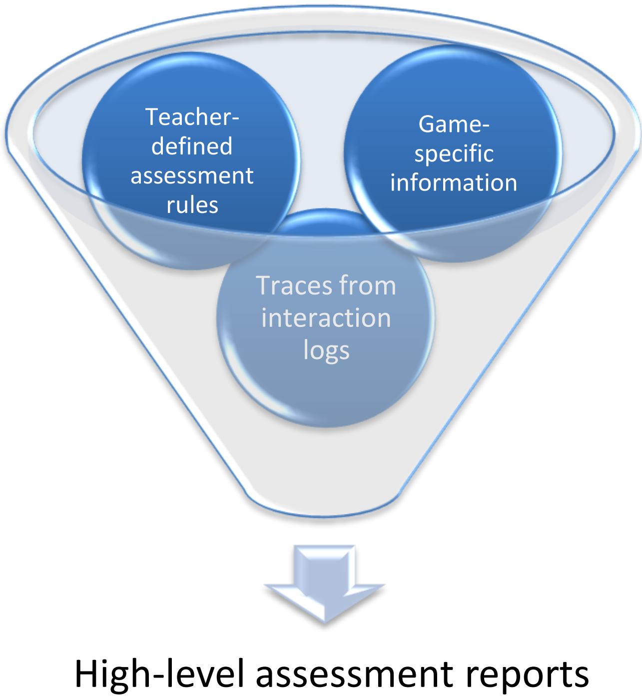

Figure 1. Diagram representing the process for generating high-level assessment reports.

To avoid an excessive increase of the system complexity, assessment rules must be based on quantifiable parameters. Some of these rules are briefly discussed below:

1) Measuring times

Time spent to reach a specific goal can be relevant to determine student success. With the presented traces, it is possible to measure different time intervals: time spent to complete every phase and total completion time.

Several assessment rules can be defined on these numbers. For example, a maximum time threshold can be set for each phase or the whole game, and use the comparison of the actual completion time with the threshold to automatically grade the student.

A minimum threshold could be set for games which have some repetitive actions that should keep students busy for a while. The maximum score would be given to students improving that threshold.

2) Variables values

All the significant variables chosen to be logged are a potential source of assessment information. These variables are usually numerical and assessment rules can be directly defined on the comparison between actual and expected values of these variables.

For example, from a variable that represents a score it could be extracted a numeric mark based on its division by a maximum score. Or some goal can be considered as fulfilled only if a set of variables are greater than values defined by the teachers.

3) *Group results*

Also, complementary assessment rules can be defined by the teachers at the student group level.

Teachers could define minimum success rates for every game and program alarms for when these rates were not achieved. They also could define the number of students assumed for the game and be alerted if these number is not fulfilled when the deadline to complete the task is approaching.

By analyzing group results teachers could gauge the assessment system as well. For example, if the number of students that receive good marks is too little, time thresholds can be relaxed and get new assessment reports automatically.

IV. FINAL REMARKS

Some considerations about the model proposed deserve discussion. The game platform used for running the games must comply with some technical requisites. First, it must allow collecting or generating traces as described in this paper. This issue could be addressed extending open source game engines or using one with a built-in tracking system. Second, the game platform must include an explicit model to represent the definition of a game. This model should be kept in separate files from the code that runs the games in a format that is easy to process by a machine. And third, data logged must be stored somewhere, normally in a database, which could increase the difficulty to deploy the games. But the game engine could also log locally (in files with some specified format, for example) and then pass the results to an external analyzer. For data visualization, one of the many open source libraries available could be used to generate graphics and reports.

In this paper we propose a double-step approach that applies Learning Analytics techniques in educational videogames. In first place it pretends to be general enough so to be applied to multiple game genres and game mechanics, treating information collected as game-independent to produce reports that teachers could analyze. In a second step, an additional assessment layer could be used to generate more depurated reports. These reports serve two purposes: assess and evaluate how students' progress within the game, and also debug the game design, spotting weak points in the game that could be fixed by the teacher if necessary.

We think it is a first step towards a new model of student assessment based on educational games that can complement other methods. These ideas will not substitute traditional assessment techniques, but they can provide more information about the educational process to the teachers in a rather automatic way. More sophisticated algorithms can be developed in order to expand and obtain more fine-grained information, narrowing the games genre or the game mechanics as needed.

V. REFERENCES

[1] L. Johnson, S. Adams, and M. Cummins, *NMC Horizon Report 2012 Higher Education Edition*, vol. 2012, no. 2 March 2012. The New Media Consortium, 2012, p. 42.

[2] L. A. Annetta, J. Minogue, S. Y. Holmes, and M. Cheng, “Investigating the impact of video games on high school students’ engagement and learning about genetics,” *Computers & Education*, vol. 53, pp. 74–85, 2009.

[3] H. Tuzun, M. Yilmazsoylu, T. Karakus, Y. Inal, and G. Kizilkaya, “The effects of computer games on primary school students’ achievement and motivation in geography learning,” *Computers & Education*, vol. 52, no. 1, pp. 68–77, Jan. 2009.

[4] A. Serrano, E. J. Marchiori, A. del Blanco, J. Torrente, and B. Fernandez-Manjon, “A framework to improve evaluation in educational games,” in *Proceedings of the 2012 IEEE Global Engineering Education Conference (EDUCON)*, 2012, pp. 1–8.

[5] Á. Del Blanco, J. Torrente, E. J. Marchiori, I. Martínez-Ortiz, P. Moreno-Ger, and B. Fernández-Manjón, “A framework for simplifying educator tasks related to the integration of games in the learning flow,” *Education Technology & Society (In Press, expected publication year: 2012)*, 2012.

[6] J. Chen, “Flow in games,” *Communications of the ACM*, vol. 50, no. 4, pp. 31–34, 2007.

[7] J. Torrente, Á. Del Blanco, E. J. Marchiori, P. Moreno-Ger, and B. Fernández-Manjón, “<e-Adventure>: Introducing Educational Games in the Learning Process,” in *IEEE Education Engineering (EDUCON) 2010 Conference*, 2010, pp. 1121–1126.

[8] R. Ferguson, “The State of Learning Analytics in 2012 : A Review and Future Challenges” *Media*, no. March, 2012.