Potentiating Digital Educational Environments Through Data Analytics

Flávio Lima Faria[1](✉), Maitê Gil[2], Eva Oliveira[1], and Cristina Sylla[3]

[1] 2Ai, Polytechnic Institute of Cávado and Avenue, Barcelos, Portugal
[2] Research Centre on Child Studies, Universidade do Minho, Braga, Portugal
[3] Research Centre on Child Studies – ITI/LARSyS, Guimarães, Portugal

Abstract. When we think of educational games or apps, data analysis is not what comes first to our mind. However, a fundamental feature of all types of games/apps is the data. When playing, the system is constantly collecting data from the users, about the current state of the game and making predictions and decisions based on that data. This is what data analytics is all about. Games are an amazing way for scientists and educators to communicate with multiple data sets, as well as a great way for developers to get a wealth of relevant information to inspire the development of new algorithms and theories. The main goal of this investigation is to propose and validate techniques that allow the application of data analysis concepts to improve digital educational environments, especially the ones related to storytelling. In this paper, we discuss the development and implementation process of specific analytical techniques applied to Mobeybou, a set of story apps for children, as a first step to develop a set of guidelines to support, inform and optimize the evaluation of the efficiency of educational games/apps using data analysis.

Keywords: Data analytics · Educational games · Interactive story apps · Children

1 Introduction

This work aims to develop a set of guidelines and techniques to optimize the evaluation of the efficiency of educational tools, such as games and story apps directed to children using data analysis. The goal is not to develop a new data mining or analytics tool, but rather to propose and validate techniques that allow the application of data analysis concepts to improve digital educational environments, especially the ones related to storytelling. With these techniques, we intend, for example, (i) to collect and analyse relevant data of the users and their contexts, in order to understand and optimize the learning process in environments that use educational games, and (ii) to better understand students' learning process. In order to develop this general framework, we started by customizing existing data collection methods and algorithms that we have then implemented in the Mobeybou story apps [1]. In this paper, we present and discuss the development and implementation process of these specific analytical techniques.

E. I. Brooks et al. (Eds.): DLI 2020, LNICST 366, pp. 40–47, 2021.
https://doi.org/10.1007/978-3-030-78448-5_3

2 Data Analytics and Games

Any team responsible for designing a game needs to understand its users, e.g., why do they play, why do they stop playing, how long do they play? These questions can be answered by analyzing data collected during the game play and thus informed design decisions can be made instead of relying on mere assumptions. By integrating data collection into a game, it is possible to collect different types of information about the users: How many people play the game every day, how many times do these people play during a certain period of time, what stages of the game do they play the most or how long do they spend playing each session. In short, data analysis allows designers to focus on the most relevant tasks to improve their games and to make them more attractive to their users [2].

2.1 Data Analytics and Educational Games/Apps

Regarding educational contexts, Agasisti and Bowers (2017) propose that the instruments to be used for data analysis in specific research and to answer the desired questions can be ordered in three different approaches considering its specific purpose: Educational Data Mining (EDM), Learning Analytics (LA) and Academic Analytics (AcAn).

Educational Data Mining uses computational methods to detect patterns and recurrences in large amounts of collected educational data. The main beneficiaries of this approach are researchers, analysts, faculty and tutors. EDM represents a host of education-specific machine learning tools, providing an excellent foundation for our discussion of learning design analytics. EDM is an emerging discipline, concerned with developing methods for exploring large educational data streams [3].

Learning Analytics uses many of the Educational Data Mining analysis models but focuses on teaching and learning activities with great attention to the achievement goals, that is, the results. LA uses data analysis to help researchers understand the factors that are critical to learning. Learners, faculty and tutors can benefit from this data analysis method.

Academic Analytics, on the other hand, focuses on the systematic use of data to generate information that can improve the internal efficiency of operations and management processes in educational institutions. Thus, its use is more focused on the organizational level of education, such as personnel, resources and material management. In our work, we intend to use tools from the first two approaches (EDM and LA) to assist our research, as we believe they are more relevant to analyze the data collected from the Mobeybou users in order to improve the quality of storytelling and the educational aspects of the app.

When applied during the various stages of game development, data analysis methods can provide in-depth insights of learning patterns in digital game environments, as well as inform and facilitate design iterations for optimizing learning and engagement [3]. These insights can inform each stage of the game design and production and be carried out in synchrony with game refinement cycles to fuel iterative, data-driven design for better engagement in learning. In the early stages of development, the definition of the data structure and visualization analysis can be valuable in supporting design, since the discovery of data structures can reveal deeper interaction patterns performed by the user

as the mechanics are solidified in the following stages of development. This predictive approach can support game production by providing learning prediction and behavior detection for adaptability in the game supporting the learning objectives [4].

There are several methods for evaluating in-game learning: "Broad analysis goals (or 'metagoals') common to the expert EDM synopses are visualization, relationship mining, and prediction […]. Visualization involves graphic representations of data; relationship mining looks at associative data patterns; and prediction can project outcomes via algorithms of sequence, probability, and regression. Specific method types include: descriptive visualization, social networks, clustering, association, classification/regression, and pattern mining"[5:2].

A simple example is to measure how many levels a student has completed or to record the last completed level. The benefit of this approach is that it is very straightforward and relatively easy to implement. The analysis of the last challenge a student successfully overcame can be used as an independent assessment variable of the student's performance [6]. Andersen et al. [7] applied a different method for evaluating success in a study of the effects of different tutorial styles on learning gameplay, looking at how well a player was learning while playing a game. This approach works well when the gameplay mechanics are properly aligned with the learning content.

Another method is Evidence Centered Design (ECD) [8], which depends on three main components: a competency model, an evidence, and a task model. The competency model describes the knowledge level intended to be achieved with the system. The evidence model analyzes student's interactions with a system through predefined rules. The evidence model also shows a relationship between the competence model and the observed scores. Finally, the task model provides a framework for describing the tasks a student performs within the system [9]. In this investigation, we focus only on the post-game data collection method, that is, the data is analyzed after the sessions and not in real time. The data collected in the app used as a model is grouped and made available for analyses at 12-h intervals.

This brief panorama of evaluation methods for educational games indicates a diversity of practices. The fact that educational games have been used as an innovative instructional strategy to make learning more effective justifies such diversity, since it is important to systematically evaluate these games and check for evidence of its impact on learning. However, a systematic literature review carried out by Petri and von Wangenheim [10], in which the authors identify and describe seven different approaches to systematically evaluate educational games, has shown that most of the approaches are developed in an ad-hoc manner and do not provide an explicit definition of the study, its execution and data analysis. Therefore, the authors argue that there is a need for research on the definition and operationalization of educational game evaluations. According to Steiner et al.: "A crucial question is how to harness and make sense of game-based user data in an educationally relevant manner" [11:196]. In this sense, the data analytics techniques developed in this study are an effort to obtain more systematic data collection and, consequently, more valid and uniform results.

Although evidence of learning is the main objective when investigating, developing or designing educational games, it is not the only aspect that can be investigated on a

game. Many other important questions may arise during the development of an educational game. Among them, it is important to highlight the following: Are the goals of the game really rewarding the kinds of behaviors that the game wants to encourage? Does the difficulty of the levels progress adequately to the proposed objectives? Answering these questions may affect the interpretation of learning measures and may have implications for the redesign of game mechanics. Gameplay data analysis can be a valuable ally in determining what types of mechanics work, or not, to meet the learning goals. However, the obtained data can only be consistently and assertively analyzed when these learning goals are well defined [12].

The use of data analytics in educational environments can inform both educators and designers. For educators, the data collection can be seen as an iterative cycle of hypothesis formation, testing and improvement. E.g., the educator can understand which learning topics need to be reinforced. In this process, the objective is not only to transform data into knowledge, but also to apply the extracted knowledge to make decisions about how to modify the digital educational environment to improve student's learning. For game designers, data analytics is a way to evaluate the design of applications or games and can help to improve the game design. Therefore, data collection in educational environments can be a powerful ally for assessing the effectiveness of interactive experiences [13].

Our literature review identified a lack of detailed work on data analysis in educational games. This review also indicates that there is a growing interest in evidence-based education, where games and educational applications that aim to improve teaching and learning can be validated with real data obtained from user interactions.

3 Developing Data Analytics Techniques for the Mobeybou Interactive Story Apps

Based on the assumptions presented above, we propose the development of a set of data analytics techniques to be integrated in the Mobeybou apps, a set of interactive story apps directed to pre and primary school children. The integration of these techniques will allow the collection of relevant data of the user's interactions with the story apps that will inform (i) future design decisions and (ii) the evaluation of the user's learning process. In this section, we describe the decisions regarding the development of such data analytics techniques.

The development of techniques for capturing the user's activity data has great relevance for validating the collected information. These techniques should be concise and built to collect and catalog the data in a reliable manner. Based on the above presented understandings, we developed an interaction logging system, a kind of activity map, which, by saving the user's activity log, can provide a database to inform teachers and researchers about which game mechanics and elements are more effective in promoting the user's achievement of the learning objectives.

Several methodological procedures are necessary to accomplish this goal, namely: (i) choosing the data platform in which the study is going to be developed; (ii) defining which data will be collected and recorded considering the learning goals; (iii) customizing the data analytics techniques; (iv) implementing the data analytics techniques; (v) establishing a time window for the collection of this data; (vi) analysing the collected

data; and, finally, (vii) indicating, based on the analysis of the data, paths for both the improvement of the story apps and the development of a set of guidelines and techniques for using data analysis in educational applications or games with a storytelling component.

3.1 Using Unity Analytic as a Data Platform

We choose the Unity Analytics data platform to carry out this investigation since it is a simple but powerful data platform that also provides data recording for projects developed using Unity. Unity Analytics, is the only analysis solution integrated in the Unity engine that does not require the installation of a third-party SDK (Software Development Kit). Unity Analytics allows collecting specific information that helps to adjust the gameplay and offer the best possible experience to the users. To obtain the behavioral data it is necessary to record the events or activities that the user performs within the game. Specific actions or choices carried out by the users are captured. To do this, a series of events are created that are triggered by the user's actions and indicate Unity Analytics to capture certain values. The collected data is generally divided into the following categories: geodemographic data; behavioral data; operational data; funnels; and qualitative data [14].

Along the entire data collection process it is mandatory to assure the anonymity of the users. Therefore, the data collection is merely used, to understand how the game applies to all users, and only relevant data is stored. Some types of data are irrelevant and some data can also be inferred by evaluating interconnected events. Understanding the kinds of data we can access and its meaning is fundamental to convert the collected data into improvements or new features for a project. In the next section, we present the decisions we have made about which data needs to be collected and stored by the data analytics techniques under development.

3.2 Design of Adaptive Data Analysis Techniques

In order to collect relevant data using the Unity Analytics platform, we apply scripts at key points in the story app that serves as a test model. These scripts act as a trigger that records the user's actions within the interactive application.

The first design decision was to map every button of the app's opening interface. The recording of all possible interactions in the opening interface is necessary to get information about how the user navigates through the user interface. Important users' choices within the opening interface that have pedagogical implications are, for instance, related to the following questions: What does the player do first: Read the story, read the glossary or play the integrated game? Does s/he read the story from the beginning or selects a different page to begin the reading? Which page do the users read most frequently? Which page do the users read less frequently? Which sequence of pages do they prefer to read? How many users visit the glossary page?

This kind of data can help researchers and educators to get an overview of the users' reading behaviour, which is relevant both to improve the interactive story application - for instance by redesigning the less visited pages, - and to develop pedagogical approaches

- for example, by focusing on the preferred sequence of pages or on the integration of the glossary into the story.

The second design decision was to map every button of the app's menu interface. This is relevant to record the most selected language or the most chosen character by the users, as well as to access if the users turn off the music or the integrated narration of the story. In the story pages, the data is collected each time the users touch an interaction area and the navigation buttons. The analysis of such interconnected events allows us to infer valuable information such as: How long does it take users to read each page? How many users give up on completing the reading? Do the users do all possible interactions before moving to the next page? What kind of interaction seems to involve users the most? Again, interpreting such data can lead to relevant understandings about the users' reading behaviour and inform design and pedagogical decisions.

We also map data such as the user's life cycle, how many users are online, how many use the app daily and how many times each user interacts with the app. The country in which the user is located is also registered.

The above mentioned aspects are examples of custom events that we use to capture data of the player's behavior, however, and depending on the defined goals, parameters are often changed over time, some other parameters can be added and others removed. This is why simplicity and ease of implementation and management for data analysis techniques for educational games or applications is so important. In our investigation, the developed data analytics techniques were implemented in May 2020, and will collect data during a time window of six months. After that, the collected data will be analysed to understand if the chosen parameters and custom events are effective for informing future design and pedagogical decisions, as well as good enough to support the development of a technique for using data analysis techniques for educational applications or games with a storytelling component.

Concerning the analysis of the collected data, it is important to highlight that it requires new techniques, since interactive learning platforms, such as digital games, have opened up new opportunities to collect and analyze student's data, to map patterns and trends in that data and to test hypotheses about how students learn.

As discussed in Sect. 2.1, there is a lack of well-established and well-founded techniques to analyze the efficiency of educational games and applications. Data collection, for example, is performed in various ways and not all are clearly justified. Although different data may be collected, there is still little guarantee about which data is most suitable for the evaluation of educational games. However, if properly designed and applied, the use of data analytics can be a valuable contribution to studies that aim at evaluating educational games by operationalizing an important phase for the evaluation, that is, the collection of data.

4 Conclusion and Future Work

Along this investigation, we have found evidence that data analysis can be a valuable tool for evaluating educational games. Following the customization of data analytics techniques to implement in story apps for children, a set of guidelines to optimize the evaluation of the efficiency of educational games using data analysis will be collected.

Future work includes: (i) analysis of the data recorded during this investigation in order to determine if the selected data was adequate to answering the questions related to the story apps and their learning goals; (ii) the design of general techniques to determine which data is important to collect in order to answer the desired research questions; (iii) the design of data management techniques that can be adaptable and easily implemented in educational game projects. As a long-term goal, we plan to apply data analysis in educational games in an agile and adaptable way with the aim of providing data that can help to improve teaching in educational games.

Acknowledgments. MoBeyBOU: Moving Beyond Boundaries - Designing Narrative Learning in the Digital Era, has been financed by national funds through the Portuguese Foundation for Science and Technology (FCT)- and by the European Regional Development Fund (ERDF) through the Competitiveness and Internationalisation Operational Program under the reference POCI/01/0145/FEDER/032580.

References

1. Sylla, C., et al.: Designing narrative learning in the digital era. In: Proceedings of the ACM International Conference on Human Factors in Computing Systems, CHI 2019 Extended Abstracts, Glasgow, Scotland, UK, 4–9 May. ACM Press, New York (2019)
2. Tlili, A., Chang, M.: Data Analytics Approaches in Educational Games and Gamification Systems. Springer, Singapore (2019). https://doi.org/10.1007/978-981-32-9335-9. ISBN: 9813293357, 9789813293359
3. Freire, M., Serrano-Laguna, Á., Manero, B., Martinez-Ortiz, I., Moreno Ger, P., Fernández-Manjón, B.: Game Learning Analytics: Learning Analytics for Serious Games (2016). https://doi.org/10.1007/978-3-319-17727-4_21-1
4. Agasisti, T., Bowers, A.: Data Analytics and Decision-Making in Education: Towards the Educational Data Scientist as a Key Actor in Schools and Higher Education Institutions (2018)
5. Owen, V.E.: Using Learning Analytics to Support Educational Game Development: A Data-Driven Design Approach (2015)
6. Delacruz, G.C., Chung, G.K.W.K., Baker, E.L.: Validity Evidence for Games as Assessment Environments (CRESST Report 773), Los Angeles, CA, USA (2010)
7. Andersen, E., Rourke, E.O., Liu, Y., et al.: The impact of tutorials on games of varying complexity. s.l. : Proceedings of the CHI 2012, pp. 59–68 (2012)
8. Mislevy, R.J., Steinberg, L.S., Almond, R.G.: On the structure of educational assessments. Meas. Interdisc. Res. Persp. **1**, 3–62 (2003)
9. Shute, V.J.: Stealth assessment in computer-based games to support learning. In: Tobias, S., Fletcher, J.D. (eds.) Computer Games and Instruction. s.l. : Information Age Publishers, pp. 503–524 (2011)
10. Wangenheim, G.P., Gresse von, C.: How to Evaluate Educational Games: A Systematic Literature Review (2016)
11. Steiner, C., Kickmeier-Rust, M., Albert, D.: Making Sense of Game-based User Data: Learning Analytics in Applied Games (2015)
12. Wang, H., Sun, C.-T.: Game Reward Systems: Gaming Experiences and Social Meanings (2012)

13. Alonso-Fernandez, C., Calvo-Morata, A., Freire, M., Martinez-Ortiz, I., Fernández-Manjón, B.: Applications of data science to game learning analytics data: a systematic literature review. Comput. Educ. 141 (2019) https://doi.org/10.1016/j.compedu.2019.103612
14. Documentation, Unity scripting languages manual – Unity Blog. Unity Technologies Blog. https://docs.unity3d.com/Manual/UnityAnalytics.html