



HAL
open science

How to Evaluate Competencies in Game-Based Learning Systems Automatically?

Pradeepa Thomas, Jean-Marc Labat, Mathieu Muratet, Amel Yessad

► To cite this version:

Pradeepa Thomas, Jean-Marc Labat, Mathieu Muratet, Amel Yessad. How to Evaluate Competencies in Game-Based Learning Systems Automatically?. 11th international conference on Intelligent Tutoring Systems, Jun 2012, Chania, Greece. pp.168-173, 10.1007/978-3-642-30950-2_22 . hal-01359548

HAL Id: hal-01359548

<https://hal.science/hal-01359548>

Submitted on 2 Sep 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

How to Evaluate Competencies in Game-Based Learning Systems Automatically?

Pradeepa Thomas, Jean-Marc Labat, Mathieu Muratet, Amel Yessad

Laboratoire d'Informatique de Paris 6
Université Pierre et Marie Curie
4, place Jussieu,
75005 Paris, France

{pradeepa.thomas, jean-marc.labat, mathieu.muratet,
amel.yessad}@lip6.fr

Abstract. Serious games are increasingly used in schools, universities or in vocational training. When they are used in the classroom, teachers often have to deal with the lack of tools for monitoring the students during the game and assessing them after the game. So they often tend to add assessment questionnaires to the fun sequence of “learning by playing”, to ensure that students have learned during the session. Our goal is to enable the teacher to do without this type of questionnaires by providing them an automated tool for monitoring and analyzing the actions performed by learners. The system combines an “expert Petri Net” and a domain and game action ontology. Our first experiment conducted on a sample of fifteen students showed that the diagnostic tool gives relatively close results to those of an online assessment questionnaire proposed by the teacher.

Keywords: Serious games, Game-based learning, Assessment, Petri Nets, User tracking

1 Introduction

The question of learning through serious games is often asked. Much research has been carried out [1], [2], [3]. When the serious games are used in the classroom, teachers often have to deal with the lack of tools for monitoring and assessing students. So they often tend to use assessment questionnaires to ensure that students have learned during the session. This practice interrupts the game dynamics created by game-based learning systems. Our contribution is a tool for teachers to monitor and analyze the progress of the player (from traces of the game). The system uses indicators inspired from Hollnagel’s analysis of human errors [4]. The tool is based on an “expert” Petri net and a domain and game action ontology. Petri Nets are used to model the expert rules of the domain and to diagnose the non-compliance of these rules. The ontology represents the domain concepts and their equivalent in terms of game actions, relations between game actions and between concepts and actions.

The original publication is available at www.springerlink.com

After having highlighted the difficulties of assessing learning in game-based learning systems we explain in detail the algorithms used by showing how the properties and tools of a Petri Net (PN) can be used to label the behavior of the player. We then present the first encouraging results of a comparative evaluation between our system and an assessment using an online questionnaire.

2 Assessment in Game-Based Learning Systems

Several studies have considered the issue of automatic assessment of learning in serious games. Thus, in [5] the authors added a system of state machines to the game, i.e. predefined situations that the teacher wants to watch. The teacher sets the states he wants to trace in the monitoring system. This approach is interesting but requires the teacher a great effort of interpretation to analyze the provided game indicators. In [6], the system compares the student causal graph to the teacher's one and highlights the missing and erroneous link. In [7], the authors use plan space exploration to generate a suitable game play for the learner. The game is automatically adapted to the actions of the player. Our approach has the same objectives as these approaches but uses different techniques and is more interested in labeling the errors performed by the player: the main goal is to assist the teacher in his evaluation of the "learning player".

3 Automatic Monitoring and Game Action Analysis System

3.1 Diagnostic Indicators

Drawing on the work of Hollnagel [8], we defined a classification of actions made or not made by the learner, using the CREAM method (Cognitive Reliability and Error Analysis Method). In the case of error analysis in game-based learning systems, we present in the figure 1 an evaluation of the actions of the player.

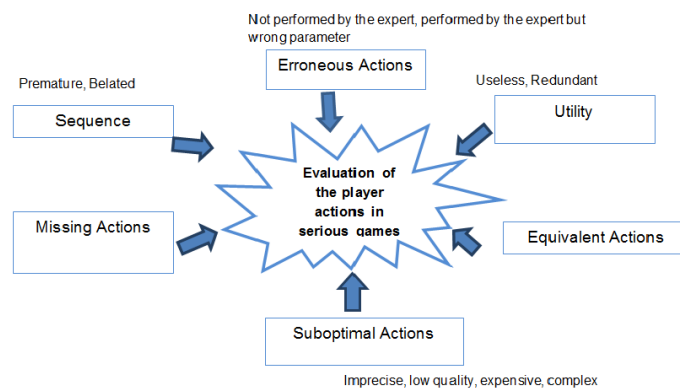


Fig. 1. Evaluation of the player's actions in the game-based learning systems according to the CREAM method

Each label is explained in [9]. For example, suboptimal actions enable progress in the game but the multiplication of these actions by the learner reveals a poor mastery of the field and imperfect skills. The player manages to overcome difficulties without finding the right solution. Belated and premature actions are necessary to resolve the current problem but do not happen in the right sequence. Equivalent actions are not performed by the expert, but produce exactly the same result as those of the expert.

3.2 Combining an Expert Petri Net and a Game Action and Domain Ontology

As detailed in [10] [9], we use a Petri net to follow the progress of the learner step by step. The idea is to analyze every "pedagogically significant" action performed by the player and to label each one according to the headings defined in Figure 1. The transitions of the Petri Net are game actions and the places represent game properties. The Petri net describes the expert behavior in the game: it performs the actions that the expert uses to solve the problem. The Petri net is built using a reverse engineering process on the game engine and by extracting domain rules from experts. Once the network is initialized by marking the places that describe the data of a problem to solve, the Petri Net will list all the solutions, i.e. the graph of the actions leading to an expert resolution of the problem (Petri Net reachability graph). Petri nets have been used in the field of game-based learning systems but rather to design games and to validate and verify the consistency of scenarios [11][12][13]. The ontology of game actions completes the approach. We use it to link game actions to the domain competencies and to represent the equivalence and sub-optimality relations between game actions. Ontologies have been used for the diagnosis of errors in learning systems [14] as well as for knowledge diagnosis in serious games [15].

3.3 Game Actions Labeling Algorithm

The goal of this algorithm is to analyze the actions performed by the player step by step and to compare them with the « expert » Petri net. The system provides the teacher with an overview by presenting the percentage of each label defined in Figure 1. Thus, the list of missing actions allows the teacher to identify blocking points. Moreover, even if a student passes the level, the multiplication of erroneous actions demonstrates a process of trial and error to reach the solution. Finally, belated and premature actions reveal a lack of optimization in the sequence of actions. The player performs the correct actions but not at the most opportune moment.

The diagnosis algorithm works as follows:

1. Expert Petri Net loading and reachability graph calculation
2. Player's traces loading and sub-optimal / equivalent actions research: the ontology is queried first to detect these error categories. These actions are labeled and then replaced by the corresponding expert action.
3. With the reachability graph, identification of :

- (a) Right actions : firable transitions
 - (b) Erroneous actions : transitions that don't appear in the reachability graph
 - (c) Redundant actions: live transitions (available in the expert Petri Net) but all the output places are marked (the player already has the information he requested)
 - (d) Premature actions : live transitions but not firable because prerequisite transitions are missing
 - (e) Missing actions : transitions that appear in the reachability graph but not in the traces of the player
4. Finally, the belated actions are obtained as follows: each time the player performs a non firable transition, the system calculates and stores the expected ones. Thus, when the player performs an "expected action" , it is a belated one.

4 Case Study

4.1 The Game

Ludiville has been developed by KTM Advance, for the "Banque Populaire Caisse d'Epargne" Group. It is designed for fledgling account managers. The goal of the player is to meet the demand of a customer by handling a more or less pre-filled loan file by performing domain linked game actions. One of the particularities of the game is to allow the player to use generic action when he doesn't know what type of information to ask the client for. For instance, "ask for document" action can be used instead of "Pay slips". These generic actions ensure that the learning player does not get frustrated. He can move forward inside the game without being held up by a lack of knowledge about some domain aspects. However, these cards yield fewer points than the specific cards, which reveal *a priori* core competencies.

4.2 The Experiment

The aim of the experiment is to compare the results of the diagnostic tool with that of an online questionnaire. The game has been tested on fifteen Higher National Certificate students. The part of the course on mortgages has not been addressed by the teacher beforehand. After a quick presentation of the game interface, the students played independently for about an hour. They all finished the first level. The traces containing all the actions performed for each client and each attempt were collected in XML files. At the end of the game, students responded to an online questionnaire developed by the teacher, referring to the concepts covered in the game. For example, they were asked to name the key documents to identify the personal characteristics of the client. The questions were classified according to four tabs defined in the game: client, project, loan and finalizing. Students had not been warned initially that they would be assessed at the end of the session. We chose to analyze in detail the last customer case of the first level. This is an assessment case that contains most of the skills used in the previous cases. This validation is considered as the "boss" at the end of a level in "traditional" video games. Some students had to go through several at-

tempts to complete this case. We chose to analyze the latest. Take the average of all trials would have penalized those who started several times. However, the number of attempts has been passed on to the teacher to give a more accurate evaluation.

For each student, an overall average and average per competency was calculated by coding the responses to the questionnaire as follows: right (2 pts), approximate (1pt), no answer (0pt), wrong (-1pt). The diagnostic indicators (right action, too early, too late, sub-optimal, equivalent erroneous) were also related to each sub area of expertise and coded as follows : right (1pt), sub-optimal (1,5 pts), premature and belated (1 pt), erroneous (-0,5pt), missing (1pt).

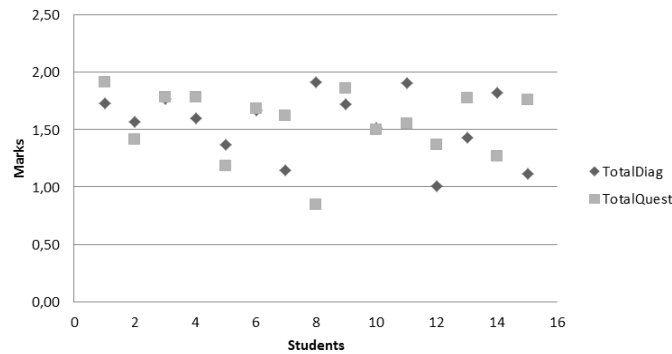


Fig. 2. Results comparison

4.3 Discussion et Perspectives

Comparing the two point clouds in Figure 2 shows that the results coincide for two-thirds of the students. The average is 1.55 for the diagnosis tool while it is 1.56 for the questionnaire. The Wilcoxon signed rank test gives p-value at 0.69. The two series are quite close. For student 8 and student 15, the differences are explained by a misunderstanding in the questionnaire (the same question). For student 8, the importance of the gap is due to the multiplication of non-expert actions. In our tool, we can see hesitations, trials and errors. Thus, at the rating of game actions, students who had thoughtful behavior have clearly an advantage compared with those who have adopted a process of trial and error, by multiplying the attempts. Moreover, it is not because they have increased the errors that they did not finally learn from their mistakes: this explains why their results in the questionnaires are good. We should refine the labeling of non-expert actions in order to isolate those that specifically reveal misconceptions.

5 Conclusion

From several experiments, it will be possible to identify players' behavior patterns using data mining techniques such as clustering. In the rest of our work, we also plan to analyze in detail the various attempts on the same mission: how does the player

adjust his strategy when he starts again a mission? In this regard, the Petri Net-based approach when implemented in real time on a game, allows for automatic and appropriate guidance. Indeed, when the player is blocked because he did not perform an action, the system can send a clue. The authors want to thank the French government who funded this research.

References

1. Johnson, W.L., Wu, S.: Assessing Aptitude for Learning with a Serious Game for Foreign Language and Culture. Proceedings of the 9th international conference on Intelligent Tutoring Systems. pp. 520–529. Springer-Verlag, Berlin, Heidelberg (2008).
2. Virvou, M., Katsionis, G., Manos, K.: Combining Software Games with Education: Evaluation of its Educational Effectiveness. *Education Technology & Society*. 8, (2005).
3. Defreitas, S., Oliver, M.: How can exploratory learning with games and simulations within the curriculum be most effectively evaluated? *Computers & Education*. 46, pp. 249-264 (2006).
4. Hollnagel, E.: Barriers and accident prevention. Ashgate Publishing, Ltd. (2004).
5. del Blanco, Á., Torrente, J., Marchiori, E.J., Martínez-Ortiz, I., Moreno-Ger, P., Fernández-Manjón, B.: Easing assessment of game-based learning with e-adventure and LAMS. Proceedings of the second ACM international workshop on Multimedia technologies for distance learning. pp. 25–30. ACM, New York, NY, USA (2010).
6. Shute, V.J., Masduki, I., Donmez, oktay: Conceptual Framework for Modeling, Assessing and Supporting Competencies within Game Environments. *Science*. 8, pp. 137-161 (2010).
7. Thomas, J.M., Young, R.M.: Annie: Automated Generation of Adaptive Learner Guidance for Fun Serious Games. *IEEE Transactions on Learning Technologies*. 3, pp. 329-343 (2010).
8. Hollnagel, E.: Cognitive reliability and error analysis method: CREAM. Elsevier (1998).
9. Thomas, P., Yessad, A., Labat, J.M.: Petri nets and ontologies : tools for the « learning player » assessment in serious games. Proceedings of the 11th IEEE International Conference on Advanced Learning Technologies. IEEE Computer Society, Athens, Georgia, USA pp. 415-419 (2011).
10. Yessad, A., Thomas, P., Capdevilla, B., Labat, J.-M.: Using the Petri Nets for the Learner Assessment in Serious Games. *Lecture Notes in Computer Science*. pp. 339-348. Springer (2010).
11. Natkin, S., Vega, L.: Petri Net Modelling for the Analysis of the Ordering of Actions in Computer Games. Présenté à 4th International Conference on Intelligent Games and Simulation (GAME-ON 2003) (2003).
12. Araujo, M., Roque, L.: Modeling Games with Petri Nets. *Breaking New Ground: Innovation in Games, Play, Practice and Theory*. , West London, United Kingdom (2009).
13. Brom, C., Šisler, V., Holan, T.: Story Manager in ‘Europe 2045 ’ Uses Petri Nets.
14. Abou Assali, A., Lenne, D., Debray, B.: Case Retrieval in Ontology-Based CBR Systems. In: Mertsching, B., Hund, M., et Aziz, Z. (éd.) *KI 2009: Advances in Artificial Intelligence*. pp. 564-571. Springer Berlin Heidelberg, Berlin, Heidelberg (2009).
15. Conlan, O., Hampson, C., Peirce, N., Kickmeier-Rust, M.D.: Realtime Knowledge Space Skill Assessment for Personalized Digital Educational Games. 2009 Ninth IEEE International Conference on Advanced Learning Technologies (2009). pp. 538–542. IEEE Computer Society, Riga, letvia (2009).