FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Fostering motivation through Artificial Intelligence techniques in educational serious games

Ana Carolina Ribeiro Moura

Mestrado Integrado em Engenharia Informática e Computação

Orientador: **Rosaldo José Fernandes Rossetti, PhD**

Co-Orientador: **Susana Maria Dessa Aguiar, PhD**

July de 2017

Fostering motivation through AI techniques in educational serious games

Ana Carolina Ribeiro Moura

Mestrado Integrado em Engenharia Informática e Computação

July of 2017

Abstract

The term serious game refers to games that have a bigger impact than entertainment and have been growing over the last years. Serious games are usually developed for a target audience and a target teaching goal. This results in a high cost development for a small audience.

One of the most relevant problems in serious games is the need to adapt and balance the game, this balance is restricted to one person and it cannot be extrapolated to a bigger audience, turning the small audience to just a few people. In order to have an efficient control of learning, it is necessary to understand the latent relation between the cognitive capacity, motivation and performance of each person. The more personalised the material given to each student, the better their learning will be, because the balance is tailored to their needs.

Machine learning, especially reinforcement learning (RL) can be used for automated behaviour generation and it can be applied to an agent that controls the difficulty of a game in an unknown, unsupervised environment. For that reason, the application of algorithms like Q-learning may help on the creation of learning curves in a game. The broad objective of this dissertation is exploring how Artificial Intelligence can help monitor and adapt a game in real time to a player's needs and profile. Specifically, we want to study the state-of-the-art in the context of serious games as well of adaptative games.

There is a broad range of work in serious games and how they are the answer to motivate the students, however one needs to keep a flow state in the student for better results. Some papers also explore how AI can help with these questions but there is no concrete answer to this need and no definite study of how to do it. Nonetheless, the research work helps us define some parameters needs for the success of this dissertation work.

The goal is to create a game with real time adaptation in the area of Mathematics, that can create a reliable profile of any player inside our target audience and adapt to the needs of that profile. Ideally, this will be the next step on e-learning and serious games development as we can expand our audience to a bigger number with the same resources.

We achieved our objectives by implementing an adaptation of the conventional Q-Learning algorithm by creating a double layer Q-Learning with adaptative matrices. This algorithm was the core AI in a proof of concept game which was tested with 79 pupils of the 1st and 2nd grade (7th and 8th grade in Portugal). The tests showed how the AI managed the difficulty throughout time, therefore creating a personalised flow state for each player.

Acknowledgements

I would want to thank my advisors Rosaldo Rossetti and Susana Aguiar with whom I've been collaborating closely towards many of the results presented in this work for sharing with me their expertise in their fields.

To my parents who were my biggest supporters and were always ready to help me with whatever was necessary, always believed in me and showed me unconditional support throughout the years, thank you so much, the reason I'm here today is you.

And a special thanks to David Caminha for always keeping me going, always being there for the good and bad moments and helping me achieve this goal of finishing this journey with success.

To all these people thanks, without you this would not be possible.

Ana Moura

"Simplicity is the ultimate sophistication."

Leonardo da Vinci

Table of Contents

1. Introduction 1
   1.1. Scope and motivation 1
   1.2. Aim and Goals 2
   1.3. Document Structure 3
2. Serious Games and motivation 5
   2.1. Effective learning environment, Flow experience and Motivation 5
   2.2. Motivation Theory 7
   2.3. Serious Games 8
   2.4. Gamification 9
3. Adaptive Game Design 12
   3.1. Player Experience Modelling 12
   3.2. Reinforcement Learning 13
   3.3. Q-Learning 13
4. Methodological approach 18
   4.1. Conceptual Design 18
   4.2. Game Design 19
   4.3. Game Adaptation 20
   4.4. Game Architecture 21
   4.5. Game AI 22
   4.5.1. Q-learning for Difficulty 24
   4.5.2. Q-learning for numbers 27
5. Results 32
   5.1. Type-1 results 33
   5.2. Type-2 results 35
   5.3. Type-3 results 37
   5.4. Type-4 results 39
   5.5. Type-5 results 41
6. Conclusions 44
   6.1. Summary of Research and Findings 44
   6.2. Future work 45
References 48

List of Figures

Figure 1 - EFM model for Educational Game Design (Song & Zhang, 2008) 6
Figure 2- Representation of flow zone 6
Figure 3-Types of Game Thinking and Primary Design Goal (Marczewski, 2013). 9
Figure 4 - Gameplay exemple 19
Figure 5-Gameplay example:right answer 20
Figure 6 - Example of an adaptative game system 21
Figure 7-Game Architecture 21
Figure 8-Normal Q-learning matrix 22
Figure 9-Exploration value Scale 24
Figure 10- Difficulty Q-learning state-actions 24
Figure 11- Chart study of learning discounts 25
Figure 12- Game Cycle 27
Figure 13- Q-learning for Numbers 28
Figure 14- Student A Results 33
Figure 15 - Student B Results 35
Figure 16- Student C Results 36
Figure 17 - Student D Results 37
Figure 18 - Student E Results 38
Figure 19 - Student F Results 40
Figure 20 - Student H Results 41
Figure 21 - Student I Results 41

List of Tables

Table 1 - Student A Difficulty Q-learning matrix 34
Table 2- Student A Numbers Q-learning matrix 34
Table 3 - Student B Q-learning Difficulty matrix 35
Table 4- Student B Q-learning for Numbers matrix 35
Table 5- Extract from Student C Q-learning for Numbers matrix 36
Table 6 - Student C Q-learning Difficulty matrix 36
Table 7- Student D Q-learning 37
Table 8- Student D Q-learning for Numbers matrix 37
Table 9- Student D Self-Assessment 38
Table 10 - Student E Q-learning Difficulty matrix 39
Table 11- Student E Q-learning Numbers matrix 39
Table 12 - Student D Self-Assessment 39
Table 13- Student F Q-learning Difficulty matrix 40
Table 14 - Student F Q-learning for Numbers matrix 40

Acronyms

AI	Artificial Intelligence
EFM	Effective learning environment, Flow experience and Motivation
EV	Exploration Value
PEM	Player experience modelling
RF	Reinforcement Learning
SG	Serious Games

1. Introduction

1.1. Scope and motivation

Learning is one of the main characteristics that separate us from animals and is where the majority of our time is spent. There are many researches and studies on the improvement of the education overall. They seek new ways on how to improve learning and teaching techniques so a student can make the best out of their cognitive ability.

For that reason, we can see with the passing of time a change on teaching ways. The education started with a more cognitive way of teaching, black boards where teachers can introduce new subjects and that the students would repeat until they knew it. However now, the new theories for better learning affirm that a more hands on work is better for the complete knowledge of a new subject. Moreover, they agree that not every student can learn in the same way, they will have different skills to different task and instead of obligating them to study in the same rhythm they should be able to learn at their own pace. (Mouaheb, Fahli, Moussetad and Eljamali 2012). Also, motivation has been a big study subject and how it can improve the learning progress of the students and therefore theories like flow have been gaining more importance in a classroom.

Over several years there have been many attempts and studies in regard to of how games have powerful characteristics to help a student achieve a higher state of motivation and therefore a better knowledge of the subject in question. (Methaneethorn, 2008). For that reason, a new type of game was created: Serious games. Currently there is no formal acceptance of its definitions, however the core point in any definition is that the main focus of it is the pedagogical objective without leaving the realm of games and the enjoyment they bring. (Freitas, S., Earp, J., Ott, M. et al., 2012).

In works like Silva, Almeida, Rossetti Coelho (2013) and Ribeiro, Almeida, Rossetti, A. Coelho, A. L. Coelho, (2012) we can see how serious game can help one to learn diferente types of knowladge as these works focused on evacution plan, showing the broad range of serious games.

These games are accepted as being the next big stage in the way of educating and training students as they can be pioneers to a new way of education, in fact, research states that learning through games is better than the conventional way, since it maintains the attention of the student for longer periods of time without overloading their cognitive system. (Aguilera, 2003; Barbosa A. F. S., Pereira P. N. M., Dias J. A. F. F., and Silva F. G. M. 2014).

Çankaya and Karamete (2009) explain that one of the problem with these games is the insufficient assistance during the progress and the "one size fits all". type of games that don't adapt the game to the student's needs and skills bringing boredom to those more skilled and anxiety to those that lack some knowledge. (Squire, 2003). For that one must go to the roots of the motivation and how it occurs and understand that flow only comes with the balance between an individual's skills and difficulties on the tasks.

Creating a game for every student type is expensive and hardly scalable which means another approach must be done if we want to combine the enjoyment from games and the motivation to overcome new challenges, this means creating a game that can adapt itself to their player profile.

To do that first it is important to understand how can the system understand the type of the player through their inputs and interaction with the game and second the system must learn, while unsupervised, what the flow state of the player is and challenge them according to their own needs.

The integration of a flexible, real time difficulty adjustment can bring the balance needed for each player based on their own Player Experience Model. (Charles, McNeill, McAlister, Black, Moore, Stringer, Kücklich and Kerr n.d.). Reinforcement learning is a technique that allows an agent to learn in an unsupervised environment the best state change according to the reward system implemented. Q-learning is an algorithm with high adaptability to an unknown set of states from which it can learn what is the best next state to go for. By using adapted forms of exploration, we can also create a better set of stages for the agent to choose from based on what are the worst/best areas of knowledge of the player/student. That way there will be a difficulty adjustment based on challenge of the task but also based in areas where the student is less skilful.

1.2. Aim and Goals

Besides the clear motivation to the improvement of education and its social impact, our work is motivated by the ever-growing need of improvement of the education system and game adaptation in the context of serious games. The aim of this dissertation is to explore how Artificial Intelligence can help monitor and adapt a game in real time to a player's needs and profile. However, a few questions arise in the pursuit of such an aim: How to motivate and keep the motivation high in digital games? How can Artificial Intelligence contribute to the adaptation accounting for the student's needs? To answer to these questions the following goals are identified:

- Study state-of-the-art approaches in serious games, game adaptation and reinforcement learning with focus on Q-learning algorithms.
- Create a reliable and realistic profile of any player.
- Create a reliable form of adapting the game difficulty to a student's knowledge.

- Implement a proof-of-concept game endowed with dynamic model creation and real-time adaptation for the management of different levels of difficulty in the field of Mathematics.

The practical expected outcome from this dissertation is a serious game that can be tested on the target audience to better understand if game adaptation in serious games should be the next step in the growing of this area.

1.3. Document Structure

The remainder of this report has the following structure: Chapter 2 and 3 explore the state of the art with the former being the review of important concepts in Serious Games and motivation, whereas the latter the explore game adaptation and Q-learning concepts. Chapter 4 showcases the planned methodology for the construction of the game and algorithm creation. Chapter 5 shows the study of the results. Lastly Chapter 6 closes this report drawing conclusions on the main achievements by giving an overview of the work done and future improvements that should be taken on account if one wants to recreate and improve this project.

2. Serious Games and motivation

The importance of motivation during school education has been recognised and worked as a focal point of students' development and learning process. Goleman (1996) states: "The extent to which emotional upsets can interfere with mental life is no news to teachers. Students who are anxious, angry or depressed don't learn; people who are caught in these states do not take in information efficiently or deal with it well." (p. 78).

According to Goleman, there is an intrinsic relationship between the student's emotions and their classroom performance and willingness to overcome challenges. Positive emotions are considered precursors of good motivation and development. Students who perform well usually have positive emotions and a high motivation to develop their knowledge. Because motivation is increasingly a necessity to overcome problems, it has been repeatedly addressed in research projects in various disciplines.

2.1. Effective learning environment, Flow experience and Motivation

The game is going to follow the model proposed by Song & Zhang (2008), namely effective learning environment, flow experience and motivation (EFM), which is believed to achieve high states of motivation on the players. Such model sugests the following guidelines:

- A challenge should match a player's skill level and the challenge should progress with the learning curve of the player. Every time the player improves there should be a reward for doing so.
- A game should have clear goals and objectives at any given time and should present new sub-goals at the player pace.
- There should be constant and unambiguous feedback on the game progression.
- Player's attention should be guided and not distracted by irrelevant things.
- The player should feel somewhat in control of their own choices.

To achieve a state of **Effective learning** it is required the player to be immersed in the game without many distractions. A constant well-balanced sense of challenge and appropriate tools to accomplish them should be given to the player.

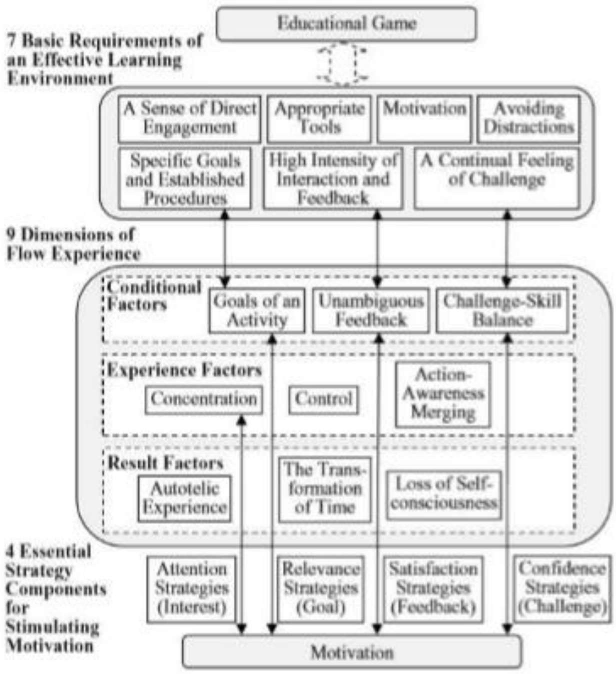


Figure 1 - EFM model for Educational Game Design (Song & Zhang, 2008)

Flow experience is defined by being "an optimal experience resulting in intense engagement, heightened motivation, and receptiveness to information, and diminished perception of time" (Pavlas, 2010) and as shown in Figure 2 is the midpoint between one skill and the challenges proposed to them. If something is more challenging than the skill of the player it may result in anxiety alternatively if the player is too skilful for the challenge may result in boredom.

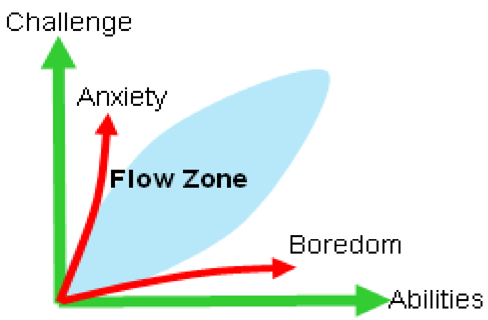


Figure 2- Representation of flow zone

By using these techniques one may give the player a bigger sense of immersion to increase his focus (attention), thus improving their learning process and motivation. (Huang, Huang, and Tschopp 2010).

2.2. Motivation Theory

Since always motivation has been considered a potential tool for better learning and development of students. There have been thorough studies trying to figure out what motivation really is and how to attain it.

According to Weiner (1992) "Motivation is the study of the determinants of thought and action – it addresses why the behaviour is initiated, persists, and stops, as well as what choices are made". In the same way that there are several definitions, there are also many theories of how motivation is achieved. Weiner separates these theories into two types: mechanistic theories and theories based on cognitive approaches.

On the one hand, the mechanistics claim that humans are a kind of machine in which motivation is based on accumulating needs, instincts, and desires. On the other hand, the cognitives believe that motivation is based on thoughts and beliefs and that the human has the choice to feel motivated or not.

Moreover, there are others namely Eccles, and Wigfield, (2002). who divide theories into four types: theories focused on self-efficiency and control; task-focused theories (those focused on intrinsic motivation, self-determination, flow, interest, and goals); theories based on expectations and values (attribution theory, the expectancy-value models and self-worth theory) and theories based on motivation and knowledge (social cognitive theories of self-regulation and motivation and theories of motivation and volition).

Learning is information processing and it seems extremely related to human cognition. Due to this fact, several theories try to relate the cognitive capacity to the motivation in the human being.

As we already mentioned, there are numerous theories in the literature and hence several attempts to categorise them. These theories try to find different factors of why some people are more motivated than others. However, these factors are sometimes intrinsic to their personal nature. (Tüzün, Yilmaz-Soylu, Karakuş, İnal, and Kızılkaya 2009).

Of all the theories, the most accepted one is Self-regulation theory (SRT) that involves processes of guidance for one's own thoughts, behaviours and emotions so that an individual can achieve its objectives. This is the theory that will be focused on this dissertation when it come to the creation of an application. The goal is that the application helps the students guide their thoughts and emotions by striking a balance between the motivation offered, the knowledge of each student and their entertainment.

According to Keller, John M. (1987) ARCS Model of Motivational Design Theory, there are four steps to promoting and sustaining motivation in the learning process: Attention, Relevance, Self-confidence, and Satisfaction.

Attention

Attention can be gained in two ways: perceptual arousal and Inquiry arousal. The former uses the surprise or uncertainty to win the interest of the student while the latter stimulates curiosity by giving new and more challenging questions or problems to be solved. Some methods used are:

- Active participation in which strategies of games or roleplay are adopted to keep students involved with the material or subject;
- In order to reinforce the knowledge apply it in several ways: sonorous, visual.
- Using small doses of humour can affect the student's emotional state by helping to avoid overload his cognitive load.
- Simulating a problem, so that the student can relate to actual experiences.
- Pose problems and questions for the students themselves to solve.

Relevance

Relating what they are learning to practical and real situations helps maintaining a higher motivation in the student. In order to do this, it is necessary to use natural language and credible examples that students are familiar with.

Self-confidence

Keeping a motivated student also goes through giving them the sense of accomplishment. Keeping them confident that they are able to solve exercises and problems with increased degrees of difficulty becomes critical so that they do not give up. For this, one can give help when needed or make the learning curve less steep so that they can adapt.

Satisfaction

A student should have a reward at the end of their endeavours whether be it by the sense of achievement or a praise from someone they look up to. Giving feedback and reinforcement gives the student the feeling of appreciation of their hard work and in turn, makes them want to overcome new challenges.

2.3. Serious Games

Since games appeared they have fascinated the entire population, playing is an enjoyable, goal-directed activity that gives us a sense of satisfaction and hooks our attention for many hours. Since the beginning of their existence, the population is keen to learn in the context of the same to be the best they can be or to beat another person. (e.g. learn the best technique in chess, or check our knowledge with Trivial Pursuit). Boring and complex tasks can quickly turn to something that entertains us for hours at the same time as we learn consciously or unconsciously. (Ade-Ibijola 2013)

The use of games with an educational purpose is an idea that has existed for many years. The earliest recorded is from 3000 years ago, in China, and in Europe there is a record in the 18th century. Nowadays games are being used in all ages, from nursery to graduation and all study areas as they can be used at all levels of cognitive objectives either to review some basic facts that the student isn't too comfortable or to test one's knowledge in a different way.

The design intention in serious games is, therefore, different from conventional games, even though the entertainment variable is present in these games that is not their main purpose nor the focus, nevertheless, it is a required component in serious games. (Gunter, Kenny and Vick, 2008)

2.4. Gamification

There is a misconception and blurred line between Gamification and Serious Games. A serious game, however, is a game with a bigger purpose than entertainment. It has all the elements of a game, looks and feels like a game but has a defined outcome or message that the creators want the user to understand.

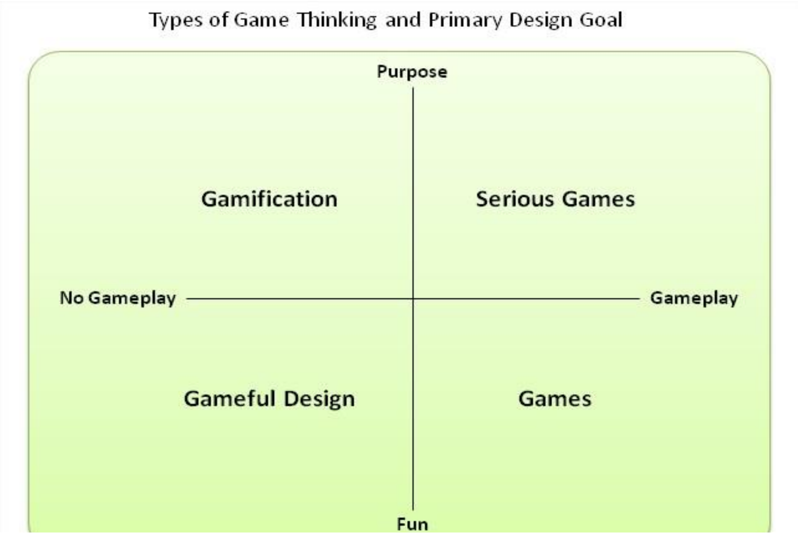


Figure 3-Types of Game Thinking and Primary Design Goal (Marczewski, 2013).

Gamification, however, is the concept of taking elements and ideas from games and applying them to things that are not a game.

We can see gamification applied to serious game in the work of Silva, Almeida, Rossetti and Coelho (2013) being used in the context of evacuation drills.

It is about the psychology of games and people, adding little things like progress bars, points, badges and leaderboards to mundane activities bringing engagement and motivation to the user. These elements bring something that games have by default, the big separation from gamification and gameplay is gameplay. You don't create a game with gamification.

3. Adaptive Game Design

As previously mentioned, one of the big parts of keeping a student motivated is, in fact, the balance between difficulty and his knowledge. By balancing the difficulty of the levels and increasing them in a curve that the student feels comfortable with there is an improvement in their results and their engagement with the learning process. To adapt a game to a user we first need to create a profile of that user, which can be accomplished through a few different ways as mentioned in the work of Georgios N. Yannakakis (n.d.) and explained in this chapter.

This chapter also explores how Reinforcement Learning (RL) techniques can be used to adapt the game experience to the player's skill by allowing the agent to learn from the interaction with the environment through reinforcement and punishment mechanism. (Hocine, Nadia, Gouaich, Abdelkader 2011) This technique will be used to adapt the game to the player needs based on their models.

3.1. Player Experience Modelling

Games have been used for a multiculturally diverse world of gamers. This means that skills, preferences, and knowledge may be widely different between the users and there has been a growing need in the tailoring of the experience to individual players and therefore the task of user modelling and experience-based adaptation within games has become more important and challenging.

Player experience modelling (PEM) is the study and use of AI techniques for the construction of computational models of experience for players. In the following sections we will name a few techniques used to trace the model of the user:

Subjective PEM

One of the most direct ways to develop a profile of the user is to directly ask them about their opinion and build a model based on such data. Subjective player experience modelling it is based on data usually got in questionnaires. However, these data collection methods have limitations as it depends solely on the user feedback which may include self-deception and it depends on the vivid memory of the user.

Objective PEM

Objective PEM traces the bodily alterations occurring during game play, facial expression, posture, and speech may show alterations on the attention and focus level. Monitoring such alterations may assist in recognising the mental state of the player and synthesising new approaches as how to make them do what we want to.

Gameplay-based PEM

The main assumption in Gameplay-based PEM is that the actions and reactions of the player manifest their own state of mind and emotional connection to the game. Their real-time preferences are linked to their cognitive and processing pattern. It is possible to analyse the patterns of the interaction and associate it with a state of mind. Any element derived from the interaction between the player and the environment forms the basis for the gameplay-based PEM.

Features like time spent on a task, number of tries or even in game decisions can be used as a way to map the overall idea of the user profile. Difficulty adjustment can be performed based on the unique profile of the user and create a link between challenge and player satisfaction.

Gameplay-based is the least intrusive PEM approach but it is very prone to errors and it may result in an inaccurate mapping of the player profile and experience.

3.2. Reinforcement Learning

Player Experience Models are needed to understand what is the player state at any given time and how it is mutable during a game but that information it is only viable for studying the changes that occur in one player during a game run. For it to be useful in the context of a serious game the information of the model can be used to define and adjust a game based on the emotional context of a player.

To personalise and tailor the player experience based on their profile model we can use Reinforcement Learning techniques (RL). (Dobrovsky, Borghoffand Hofman 2016). Those techniques allow any agent to gather information by interacting with the environment and learn through positive or negative reinforcement mechanisms. (Hado van Hasselt. 2012).

3.3. Q-Learning

The Q-learning is a RL process that uses the negative and positive reinforcements, the knowledge of the current agent state, the action performed and the observation of the state change arising from the action to learn about an unknown environment in an unsupervised method, meaning it can occur real time in any game with any player to adapt the game. (Russel, Norvig, Peter 2010, Sutton, Richard S.; Barto, Andrew G. 1998).

Q-learning consists of an agent with states **s** from which it can have a set of actions **a.** Performing an action **a** from state **s** creates a reward **Q(s,a)** which can be positive or negative. At every iteration, the Q-values are updated according to Eq.1

$$Q(s,a) \leftarrow Q(s,a)+\alpha[r+\gamma \max Q(s',a')-Q(s,a)] \quad (1)$$

where γ is the discount factor used so the values of Q are finite and α is the learning constant.

By doing the action **a**, the state (**s**) changes (**s'**) and receives a reward **r**. In the state **s'** the best action with the highest reward value will be chosen, represented as **maxQ(s', a')** in the equation above.

Discount factor (γ)

Is the factor that decides the relevance of future rewards. It can be anything from zero to one. The left extreme makes the algorithm "blind" to future outcomes of one choice and it only considers the current reward, while the right extreme will make the algorithm attempt a choice that will grant a higher reward in the long run.

Learning rate(α)

Learning rate or also called step size factor decides how much of a turn over the new information does to old information. A factor of near zero mean that there is no override and by consequence the agent learns nothing, while if the value is one the agent only considers the new acquired information.

Exploration and Exploitation

In a basic Q-learning the agent only learns a Q-function and at any given moment it chooses the best optimal action, however this result may be biased by lacking trials during the learning process. So, to contradict that outcome there should be two things that an agent can do: exploitation and exploration.

Exploitation is using the knowledge it already has found and use it in the current stage (s) and maximize the outcome.

Exploration, on the contrary, is the act of choosing an action that may not result in the best outcome.

In the literature Poole Mackworth (2010) shows several suggested ways to tackle this problem of how to choose between when to explore and when to exploit, we are going to show the less complex and least resource consuming:

"Optimism in the face of uncertainty":

This strategy encourages exploration by creating a false sense of how good the outcomes of each decision can be. To do that we initialize the Q-matrix with high values.

By having higher values in all state-action pairs, the Q-learning will opt more likely to search all possible outcomes, however the agent can be falsely certain that a given state-action pair is good for a long time, even without proofs.

For a fast convergence, the values should be as close to the ending ones as possible, trying to overestimate will make a slow convergence overall.

This strategy is not recommended if the dynamic can change with time, as it only explores in the early beginning and after that it picks the best action.

ε-greedy strategy:

Is a strategy to select the action that maximizes Q[s,a] almost every time but selects a random action ***ε*** of the time, given $0 \leq \varepsilon \leq 1$. ***ε*** can be changed throughout the process and it should, so when the agent is just starting to learn it should be encouraged to explore most of the time and with time it should exploit most of the time.

The problem shown with this strategy is the fact that all actions represent the same importance but the best action. So, if a few actions show good results and many show bad results it spends more effort trying to figure out which of the good actions are more promising rather than putting some effort to explore the many bad actions. This makes this strategy biased but it may be controlled.

A common method is called **Gibbs** or **Boltzmann distribution** where $e^{Q[s,a]/\tau}$ is proportional to the probability of selecting an action a in state s. Meaning that an action is selected with the probability given by

$$(e^{Q[s,a]/\tau})/(\textstyle\sum_a e^{Q[s,a]/\tau}) \qquad (2)$$

where τ is the temperature. A high temperature means that the actions are more likely to be chosen with the same probability and when $\tau \rightarrow 0$, these best action is always the chosen.

More strategies can be found but most of them turn out to be really complex as with neural-networks and the target audience to test this product may not have a computer with enough resources to actually run more complex algorithms. So, in the scope of this dissertation they will be ignored.

4. Methodological approach

This chapter's main concern is to explain the thought behind the development of the application, what and how it is going to be implemented. We will also explain how the research in this paper is connected to the application and how it is used to improve the experience of the student in the serious game and his knowledge of the concepts taught.

Traditionally games only detect the amount of right and wrong answers and badly adapt the levels to the gamer's needs. Our proposal wants to take it a step further and use Player Experience Modelling in depth to better adapt the game and keep the player from getting frustrated or bored. The idea behind this is that if a game is too easy and offers no challenge it bores the player, but if it is too hard it can make the player frustrated and make him give up. With the use of real-time game adaptation it is possible to keep the balance needed to achieve high motivation in all different players.

Moreover, the use of AI and the logs saved from it can help teachers understand what are the main mistakes the student makes and what are their main problematic areas. The solution presented is a concept prototype.

4.1. Conceptual Design

The application to be developed is going to be a proof of concept game focused in learning fractions with AI support for designing and balancing the challenge to the player. Because the fractions module represents a lot of teaching material this game will focus on part of the learning process: Fractions as a part of a whole. Usually students are around the seventh grade (in Portugal) when learning which means they are around the age of 14 which will be considered when designing the game.

The implementation of the following design is going to be in Unity as it is a highly adaptive game engine with possibility to export to many systems.

Assumptions for the prototype:

- Within the scope of this paper we will disregard the influence of the game theme, design and the gender of the student. However, in future development the design and gender of the student should be taken into consideration as they influence their motivation to play as seen on “Gender differences in game activity preferences of middle school children: implications for educational game design" on works by Kinzie, M., & Joseph, D. (2008) or Choi, B. Huang, J. Jeffrey, A. and Baek. (2013)

- The design of the game will be simplistic and 2D as these graphics are more suited for a learning environment and the graphical design is out of scope on the master thesis
- The audio used will be as simple as possible as will all the multimedia aspects so we can avoid unnecessary distractions. Sound will be used as key way of making the student understand if the answer is right or wrong and when a button is clicked.
- The main aim of the prototype is to test if the output of the AI would be useful for creating and analysing the student as well as the adaptation within borders the difficulty within one basic level.

4.2. Game Design

Although as mentioned on the work of Ibrahim and Jaafar (2009) the story and design of a game is important for the success of any game, we are ignoring the design of the game on the student performance the choices made to this game were based on: overall knowledge and common sense of the community as well as the simplicity of implementation.

The story was made to be able to adapt to a bigger game and shows a part of what it can be. A person needs some crystals for a journey and will ask people around for help so he can gather them. In this prototype, the game shows the main character helping an old lady that fell and broke the crystal to pieces, however if the player collects all of them he can create the crystal again.

The game has a basic point and click design in which the player must click on the object that equals the fraction the old lady is telling.

On the first stage, the project will focus on a solid basis of a serious adaptive game, in this first phase we will explore the entry points to monitor player performance and profile creation.

The idea of the game was created based on the work of AguiarRossetti (2016).

Figure 4 - Gameplay exemple

Based on the assumptions the visual part of the game was created as shown in Figure 4. On the text box, we can see the fraction to be answered and spread throughout the forest layout possible answers. If the player is able to answer correctly a piece of crystal will go to the box on the left, as seen in Figure 5.

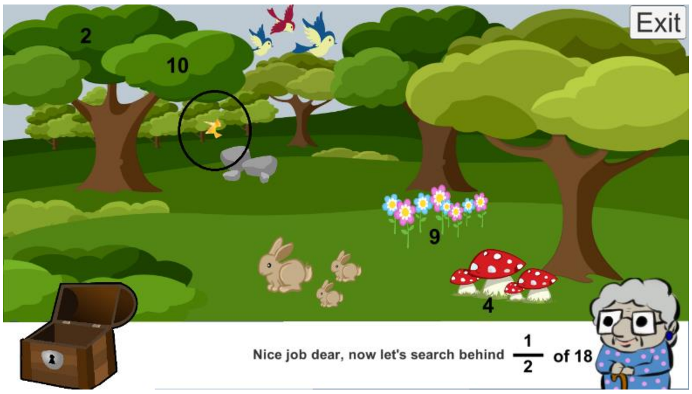


Figure 5-Gameplay example:right answer

This creates a visual stimulation that combined with the audio indicates to the player that the answer was correct. As studied, visual and audio inputs are big factors to the player motivation.

4.3. Game Adaptation

The concept will only consider the **Gameplay-based PEM** of the player in a non-invasive way. The system will track the player's answers and if they are right or wrong. As we want to be able to, in this stage, study the impact of the questions given by the AI to the student's motivation we will disregard other factors like time spent on each question and multiple mistakes.

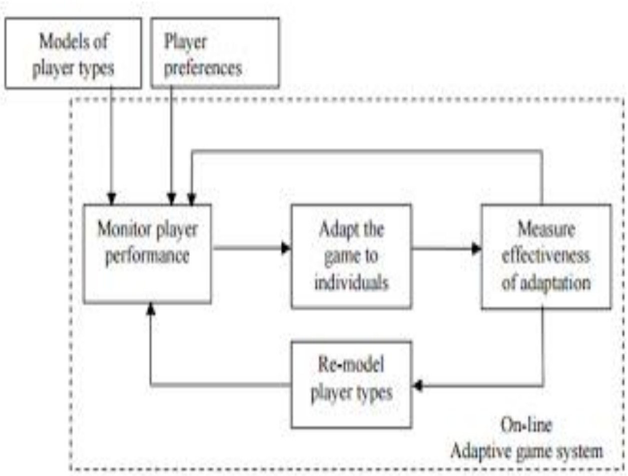


Figure 6 - Example of an adaptative game system

For monitoring the player performance we use a Q-learning matrix that indicates the difficulties of the player, so we gradually can increase or decrease the difficulty as needed without being random. Each iteration of the game more information is collected and a better awareness of the player profile can be created. The prototype would be in a bigger game the first stage of the game adaptation and its main concern is to record as much information as possible and use it in real time.

Further stages would use this information to recreate the levels and adapt them since the beginning without need for learning.

4.4. Game Architecture

The game will have 3 classes as represented in Figure 7 FileManager, GameManager and AI. The grey boxes represent where the inputs of the player are read.

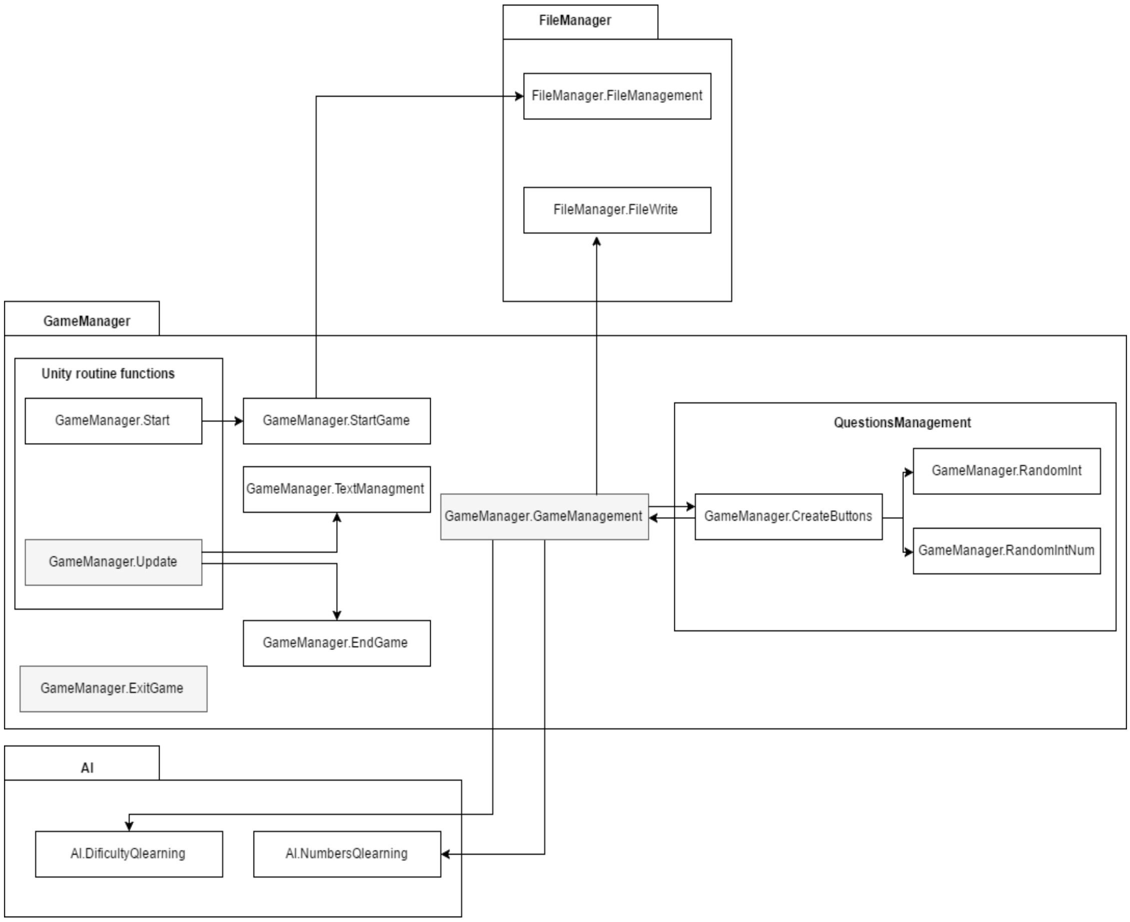


Figure 7-Game Architecture

The FileManager as the name points stores all functions needed to create and safely save the logs created during the game. The GameManager handles all the game logic needed. Also, it is the main class and stores the game cycle, as such it is normal that the Unity routine functions are in here. Finally, the AI class handles, as the name indicates, anything related to AI, it has both Q-learning algorithms and all the data needed for them.

For testing and result analysis the game will be played by a study group, which will include, for more diversity, students from the 6th and 7th grade. With this we hope we can have a reliable sample of the audience to whom this project has relevance. In our study, we want to compare how the AI categorises the student and if it is in agreement with the profile the teacher gives us, furthermore we want to keep track of how the AI evolves during the run phase and if it in fact shows signals of adaptation. In addition, the study group will have a questionnaire on the beginning of the game where they can self-assess on their ease with each type of questions, which will be used to cross compare with the results in the game.

4.5. Game AI

Q-Learning is an Off-Policy algorithm for Temporal Difference learning. It can be proven that given sufficient training under any policy, the algorithm converges with probability 1 to a close approximation of the action-value function for an arbitrary target policy.

However, in this situation we need to be able to choose the next action based on the answers given by the player. The perfect action to be done from one state will change based on the player inputs this means that if the player is struggling with level x we want to choose the best action to lower the level difficulty, however, if the player is acing the same level x we want to choose the best action to increase the level difficulty.

Because of this requirement we had to adjust how the Q-learning works and adapt the table so the information shown is more than one best action but a range of actions based on the output we want to create.

A normal Q-matrix is constituted by values (usually positive) and at each state the best action is chosen based on the higher value in the matrix:

$$Q = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 80 & 0 \\ 0 & 0 & 0 & 64 & 0 & 100 \\ 0 & 0 & 0 & 64 & 0 & 0 \\ 0 & 80 & 51 & 0 & 80 & 0 \\ 64 & 0 & 0 & 64 & 0 & 100 \\ 0 & 80 & 0 & 0 & 80 & 100 \end{bmatrix} \end{array}$$

Figure 8-Normal Q-learning matrix

The use of Q-learning will converge to a neutral state of the best equilibrium to get to a goal, in this case the flow state of a student. This is called exploitation that given a set of options we chose the best to our problem.

However, the student needs to be challenged in order to improve, meaning he needs to leave his comfort zone, by using adapted exploration we can gather more information instead of evaluating one student as a stable line. The adapted exploration will be based on a different approach to the normal Q-learning, as when we are exploring instead of having all actions **a** available we will narrow them to the specified need.

In our AI, the values in the matrix may represent different importance based on what we need to do and we manipulated the values so that we create the output needed but still using the base of the Q-learning algorithm. That being said, we can make a negative number more relevant than a positive or create a range near zero that is more likely to be chosen for the action and next stage.

Because we have the need to challenge a student but also stabilizing the questions difficulty from time to time and the information may be dynamically unstable, we need to address with care the balance between exploration and exploitation. According to the study made of the state of art until the moment and the needs for the project we will use the ε-greedy strategy with adaptation.

In this project two Q-learnings were used and although they are similar we will explain their purpose and differences in two different sections. The first is the Q-learning in control of difficulty adjustment inside the game, this is the Q-learning responsible for agreeing if the student is ready for the next stage inside the game or if it should lower the difficulty. The output of this is then used in the second which chooses which number to give the student based on this information.

Temperature or Exploration Value

The game uses an adaptation of the basic temperature term in the **Boltzmann distribution** for creating different outputs of the algorithms based on the questions streak that the player has been doing. The Exploration Value as the name say the capacity to decide when the algorithm should explore, the further from 0 the more it explores.

- Positive value explores to increase difficulty
- Negative value explores to decrease difficulty
- Near 0 explores to keep same difficulty

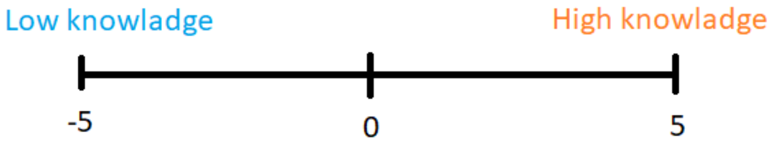


Figure 9-Exploration value Scale

At each iteration the EV is updated based on the answer the player gave, plus 1 if right and minus 1 if it was wrong. This translates in difference enough so that the answer the player gave is valued enough but not that a single answer can change the outcome by too much as it represents a change of one tenth in the scale. The scale is of -5 to +5 as this represents a scale variable enough so that we can adapt the game as it needs. The idea is to represent the basic scale of 1 to 10 where 5 is average but it is translated to 0 so we have the difference between the positive numbers and negative.

In addition, as the temperature is used in an exponent the values cannot be too far from the 0 since the higher the value of the exponent the more important slight changes become. Therefore, by keeping the temperature at lower values we have more room to change it according to the answer given.

4.5.1. Q-learning for Difficulty

This Q-learning has 5 states in total and from each state it can decide among 3 actions. These actions influence what is the next state as we can see in the diagram below. The arrows represent actions and the squares the different states. In our code an action (a) of lower difficulty is equal to the value 0, same difficulty 1 and higher difficulty equals 2.

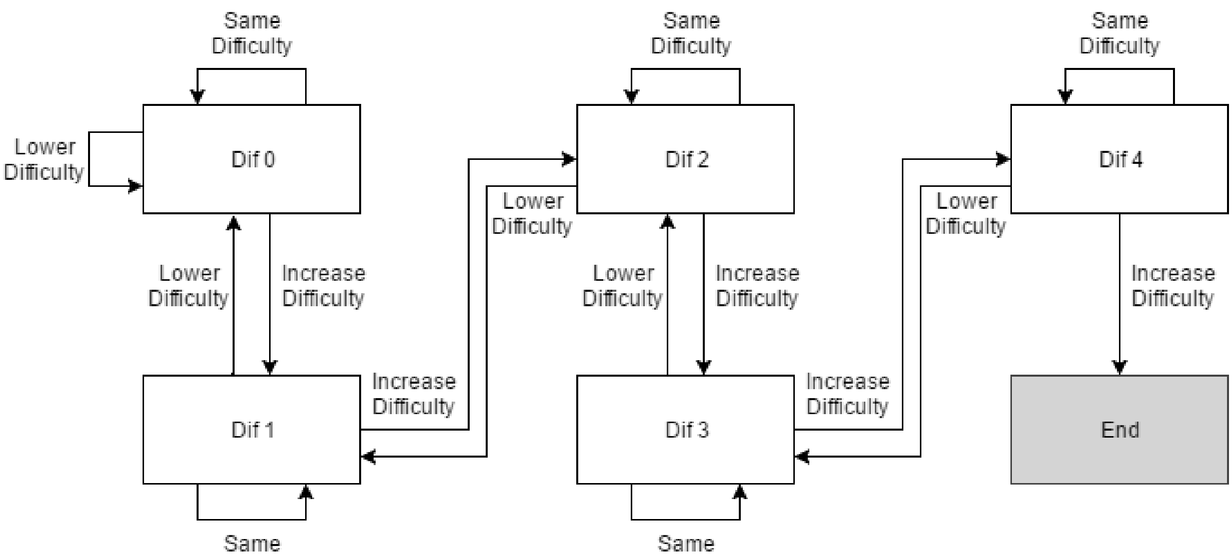


Figure 10- Difficulty Q-learning state-actions

The objective of this game is to maximize the challenge while keeping the balance with the knowledge of the student. In the attempt to maximize when possible the challenge we use formula (Eq 3) which is the inverse of the distance from the current state(si) to the final state (st) plus or minus v, plus if the answer was correct minus if it was wrong.

$$r = 1/(st - si) \pm v \tag{3}$$

At each iteration the value of the reward in any given Q(s,a) is updated with the reward function and its updated the exploration value, which are used to choose the next a that will give us the next state.

As in this Q-learning we want the algorithm to be able to foresee the near future the discount factor is set to a high value 0.8 because it will make it strive for a long-term high reward.

At the same time, we want a high turnover of the information to create a faster learning curve for the agent as we don't want the player to have to play a high number of iterations until the agent learns the needs of the player so we set the learning rate to 0.9 as with our test it shows us the need for a higher override of information.

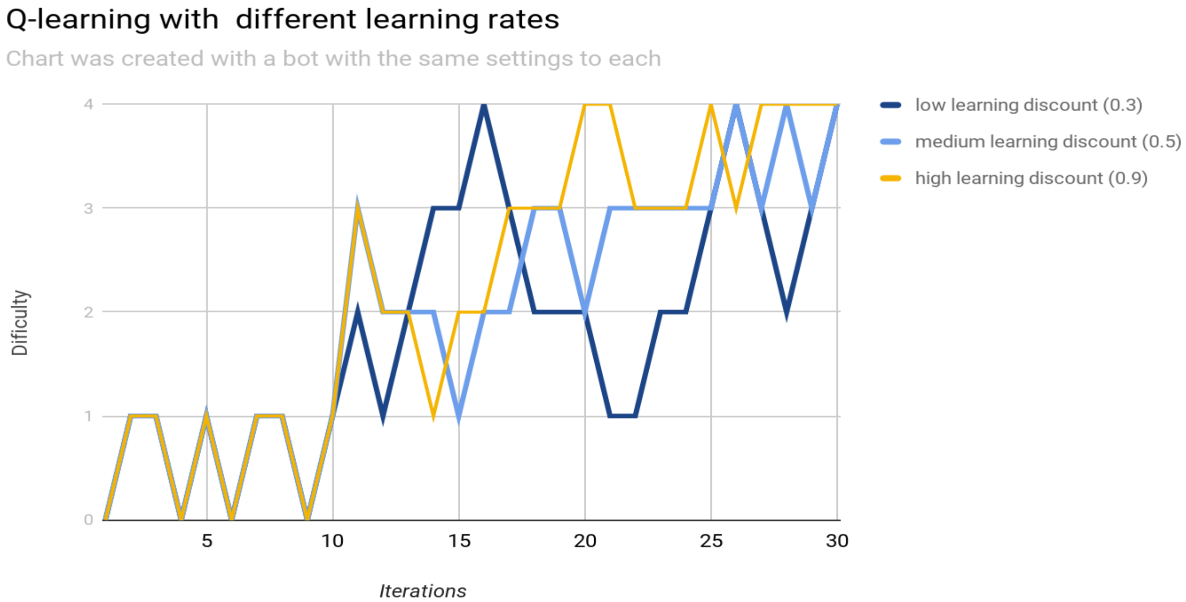


Figure 11- Chart study of learning discounts

For the creation of the graph above all three tests started with a same Q-matrix which was stable at a zero to one difficulty (created by a bot which would always answer wrong when above difficulty 0 and answer right when difficulty was 0.

This matrix was then used with a bot that would answer right to anything below difficulty 4 and answer wrong to anything at difficulty 4.

The graph shows that although all learning discounts will in fact converge for the same result, higher learning discounts adapt quicker to a player and will converge faster for the expected result. For the scope of analysing students quickly and knowing their needs without the need of repeating too many iterations we set the learning rate to a higher value.

Besides using the exploration value to further adapt the algorithm we also used a variable "answerStreak" that exists to make sure that the student truly knows the difficulty he is in before updating the difficulty.

The Q-Learning algorithm has been adapted and implemented and is presented in Algorithm 1.

Algorithm 1: Q-Learning for Difficulty

Require: Load or Initialize Q(s,a) with arbitrary values
  Initialize si to first state
  if si != final state
    lowerDifValue = $e^{\text{QMatrix[si][0] / | -minRangeEv - Ev |}}$
    if (Ev >= 0)
      sameDifValue = $e^{\text{QMatrix[si][1] / | Ev |}}$
    else
      sameDifValue = $e^{\text{QMatrix[si][1] / | - Ev |}}$
    higherDiValuef = $e^{\text{QMatrix[state][2] / | maxRangeEv - Ev |}}$;
    Calculate the probability of each one
    Pick the choice that corresponds to the random generated between 0-1
    if (choice = lowerDif)
      Ev ← Ev+2
      a ← 0
      answer Streak = 0
    else if *(*choice = *sameDif*)
      a ← 1
    else if *(*choice *= higherDif)*
      if(player answered all numbers in this state && answer Streak high)
        Ev ← 1
        a ← 2
      else
        a ← 1
    answerStreak←answerStreak++
    s' ← s + a -1
  Observe player answer
  Q(s,a)←Q(s,a)+α[r(s,a)+γ*max[Q(s',ai] - Q(s,a)]
  s←s'

During the game, the user will interact with the virtual environment through the mouse by clicking on the right answer. The game will keep records of the current iteration, the answer and will provide the Q-learning with a new reward and update the exploration value. Using the adopted policy, the next Q learning values are updated and send to the second Q learning that will choose the number according to the input value. Figure 12 shows the game cycle that happens during a gameplay session.

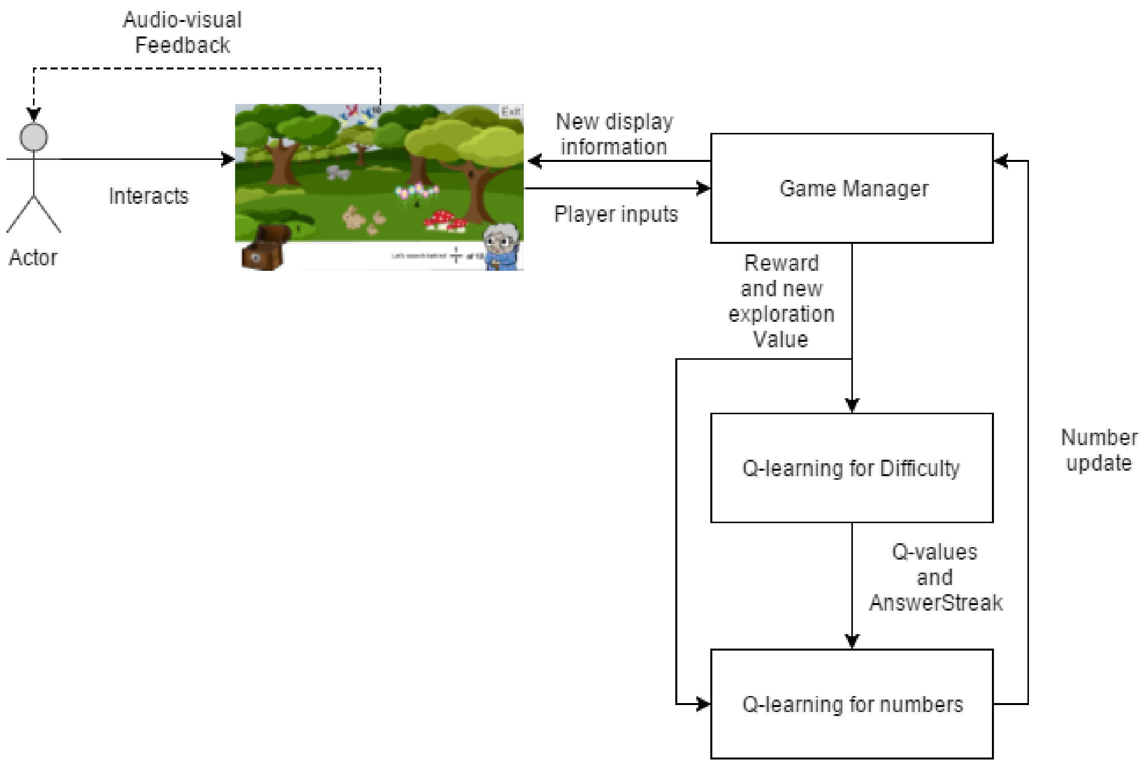


Figure 12- Game Cycle

4.5.2. Q-learning for numbers

As a Q-learning is based on Markov Decision Processes this means there is a table with all possible outcomes, as a way of not increasing the computational resources needed and a logical separation between what is the difficulty and how the numbers compare with each other this project uses two Q-learnings to create the next question.

This division came handy as it can be easily changed for a bigger process, also showing how a multilayer Q-learning can be used to tackle problems where the connection between action state would grow in an exponential way. For simplification of testing we subdivided the numbers into levels of difficulty as we need less iterations to get results. The goal however, is to let the AI decide what is considered hard numbers and easy numbers, in the meantime we used the input of experts (teachers) for definition of the difficulty.

This Q-learning is based on a 10x10 matrix with the states numbered from one to ten and the actions numbered from one to ten also. This means that an action is the transition from one number to another and the action chosen will be the next state as well.

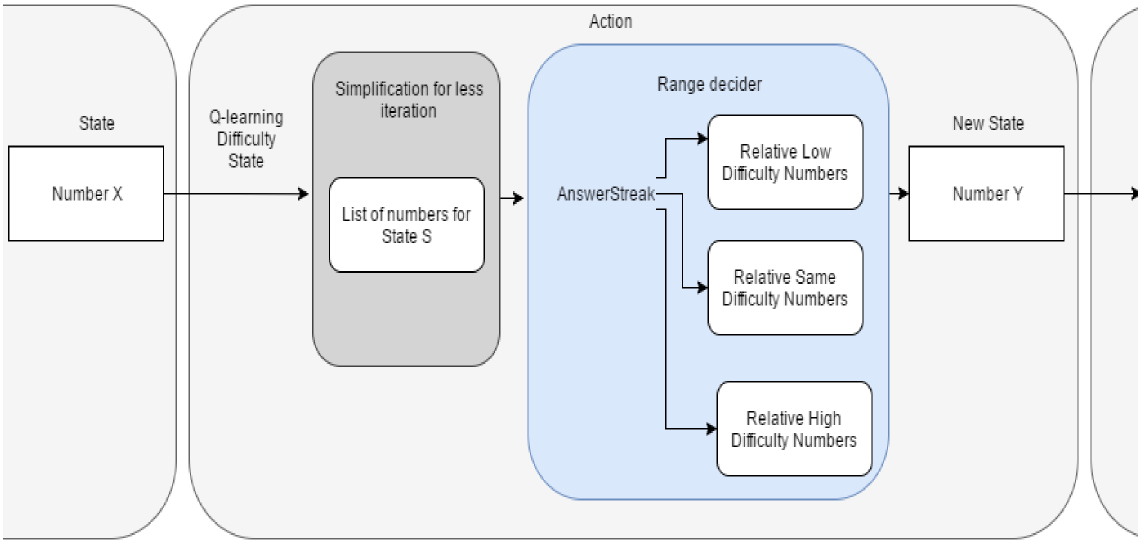


Figure 13- Q-learning for Numbers

In Figure 13 we have represented the Q-learning basis. For simplicity of show we name our initial state number X as it can represent any given number and the new state is named number Y as it can be any given number, even X.

The choice of the action starts by reading which state the Difficulty Q-learning decides is going to be next. Because this is going to be tested in a school and we want to create results quickly, instead of at each state having all 10 numbers we select a given list, made up with the help of teachers, which is representative of the current difficulty. This approach can be ignored in a full game creation as we do not have time constraints with the players.

After that we use the range decider, that internally uses the answer streak from the Difficulty Q-learning which will be 0 if the difficulty will be lower, 1 if the difficulty will be the same and 2 or more if the difficulty will be higher. The range decider will then choose based on that the range of numbers that should be chosen:

- If answer streak is 0 this means we are going to lower the difficulty and we want to choose the easier number in the level.
- If answer streak is 1 this means we are keeping the difficulty (e.g. 1 -> 1), and because we want to keep the same difficulty we choose numbers that are neutral in this level.
- If answer streak is 2 means we want to get a higher difficulty number, so we choose the more difficult numbers in the level

The goal for this Q-learning is to show us the relative difficulty between each number. Since this difficulty is not chained (i.e. 7 may be harder than 5 but it shouldn't take part when comparing the difficulty for 2 and 5) we create this Q-learning short-sighted (i.e. discount factor equals 0) as we don't want it to look at future rewards when learning the relation between two numbers. Such values should be pure and not affected by future occurrences.

For creating this relational matrix we need to decide on what more difficult number will be, in the same range of difficulty or less difficult than the number we are in. We decide to assign a range of difficulty from minus ten to positive ten where minus ten is a much easier number and positive ten is much more difficult number.

In the attempt to create this matrix as closer to reality as possible we change the rewards always based on what was the value before. This let us keep the purity of the myopia of Q-learning and use the rMatrix to further manipulate the values. Equation 4 shows how the reward is calculated where v equals 2.

$$r[s,a] = r[s,a] \pm v \tag{4}$$

As we keep needing a high turnover we use the same learning rate as the difficulty Q-learning based on the test made for its creation.

In Algorithm 2 we present the adaptation of the Q-Learning algorithm used for the Numbers Q-learning:

Algorithm 2: Numbers Q-Learning

Require: Load or Initialize Q(s,a) with arbitrary values
Initialize si to first number
for each cycle
list = AllNumbers[siQlearningDif]
for each number **in** list
if (answerStreak == 0)
if (number < 0)
value_number = $e^{(\text{numberRange+QMatrix[si][number]} - 1)}$
if (answerStreak == 1)
value_number = $e^{(\text{numberRange-|QMatrix[si][number]} - 1|)}$
if (answerStreak >= 2)
if (number > 2)
value_number = $e^{(\text{numberRange-QMatrix[si][number]} - \text{startOfHardNumbers})}$
Calculate the probability of each one
a ← choice that corresponds to the random generated between 0-1
Observe player answer
Q(s,a)←Q(s,a)+α[r(s,a) - Q(s,a)]
s←a

Such adaptation was inspired with the use of temperature to adapt values into being more flexible without overriding the matrix. Based on our "answerStreak" we read the values in different ways because we want to favor different ranges to create flow states more adaptable to the player.

If the "answerStreak" is 0, it means that the player is struggling with the numbers, so what we want is to choose the ones that are easier than the one currently used, the formula 5 is used to create a tendency to choose an easier number but with a Qmatrix value closer to zero, meaning it is not too easy, in this formula we must constrain QMatrix[si][number] < 0.

$$e^{(numberRange+QMatrix[si][number]\ -\ 1)} \quad (5)$$

In the case of wanting to keep the same difficulty of numbers ("answerStreak" equals 1), we want to favor the values closer to zero the formula 6 creates a bigger value for positive numbers closer to zero and decreases for higher or lower numbers. This way we keep exploring the numbers we don't have much information while staying on the same stage of relative difficulty.

$$e^{(numberRange-|QMatrix[si][number]\ -\ 1|)\ /\ 1} \quad (6)$$

In the case of wanting to increase the difficulty of the numbers inside one level of difficulty we use the formula 7 that creates a propensity for picking values around 2, which we consider to be the start of the relatively harder numbers. For insuring that we only choose harder than that value, we constrain QMatrix[si][number] > 2.

$$e^{(numberRange-QMatrix[si][number]\ -\ startOfHardNumbers)} \quad (7)$$

The algorithms update their Q values after each episode, and after updating, the new number is shown to the player and the system observes the user's performance thus improving the information about the student's skill. It is important to note that any given choice is based on probabilities that get fine-tuned with time.

5. Results

As mentioned earlier, the main goal of this study is to manage the motivation of the player with the help of an AI implementation that could adapt the game to each player's needs, while the player is using our application we saved the logs of what is happening in the game which will be used for validating our hypothesis. Hence a study was conducted in a school with our target audience. This study was carried out to test if the AI can in fact adapt to different needs. In this chapter, issues relating to the data analysis will be discussed.

As described we want to validate our approach that AI can be used to track a player model and adapt a game to each player profile, which can be used for further motivational management in a full game without a need to spend resources on the creation of multiple game difficulties. In this case study we want to see if we can assess a player's difficulties for later improvement of the game adaptation.

It was considered during this approach that although this is created to increase motivation, the focus of this study is to create the flow state of each player, and the assumptions were that the game itself is only a tool to test the AI and further improvements should be done in order to use the data from the AI.

During the tests we divided the results into five main categories:

- Type-1: These are students that answered each question with ease and making just a few to no mistakes; this creates a graph with a quick increase in difficulty
- Type-2: These are students that have deep problems with the questions and make just a few to no correct answers, this creates a graph with no difficulty increase.
- Type-3: These are the students that have difficulties but may or may not end the game, this creates a graph with slower difficulty increase or some decreases in the difficulty, it also may stabilize in one difficulty.
- Type-4: These are the students that have a few difficulties but overall know the answers, this creates a graph with slower difficulty increase but no drop in the difficulty
- Type-5: These are the students where the answers given are odd and should be reviewed further as the data collected is contradictory.

The practical part of this study was carried out at Webster's High School, Anugs, Scotland. There were 79 pupils participating in this study. 33 of the second year (8th grade in Portugal) and 46 of the first year (7th grade in Portugal).

On every graph a green dot represents a correct answer and above them is the number being tested, while a wrong answer is represented by a red dot with the number being tested below them. The y axis is the difficulty chosen after the answer in the given iteration. We should note that each graph starts with a red dot (at iteration 0), which is the result of the graph creation and should be ignored.

5.1. Type-1 results

This population is composed of students that are at ease with fractions that may have a few mistakes, however the AI sees they don't need a decrease of difficulty nor a long stay in the difficulty. These test subjects have a fast increase in difficulty and end the game quicker than any other. Furthermore, they are the ones where the flow is the traditional oblique line between knowledge and challenge as when we increase the challenge they have the knowledge to keep going.

This category has 35 of the 79 participants and there are two main divisions inside this category the ones that don't give any wrong answer at all and the ones that do but quickly can overcome it, this may be due to distraction or a lack of judgement for a given number.

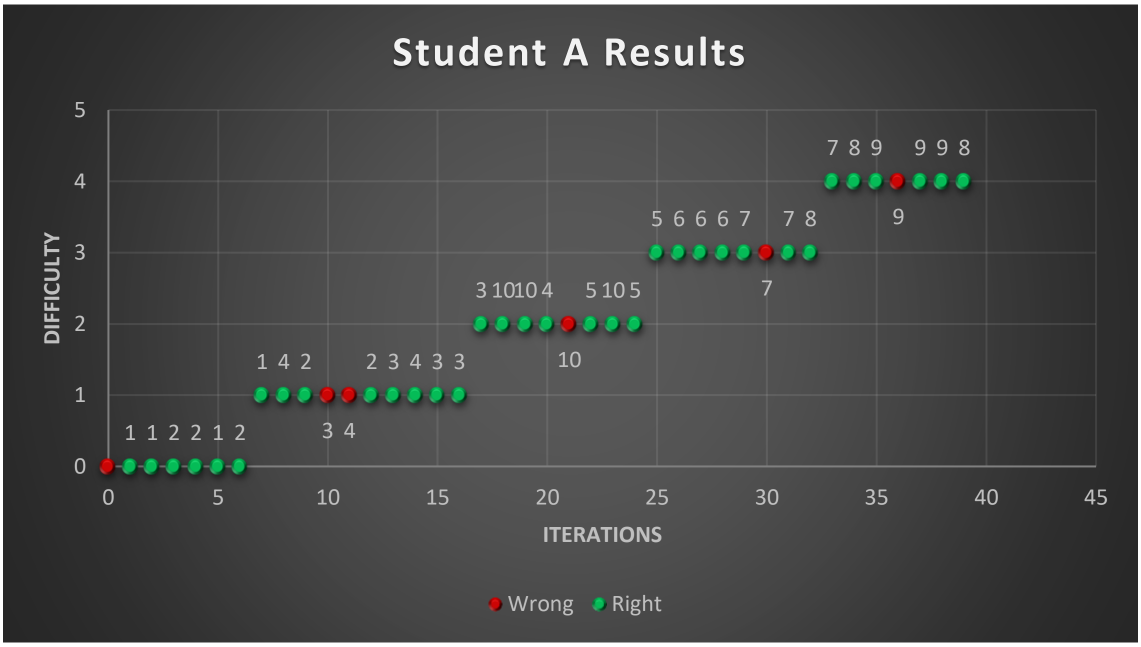


Figure 14- Student A Results

Student A is a student with a good knowledge overall, as we can see there are few mistakes. For each mistake, we can see how the AI functions by looking at the overall answers of the student for that number. There are two main cases we want to further explore: the 10 and the 7 as they work in different ways.

The wrong answer on 10 came after two correct answers on 10 which the AI only chooses to test the student one more time. Since the student answers correctly it is accepted as the student knowing that number. Meanwhile, the seven only had one right answer before the wrong one and therefore is tested two more times before being accepted as known.

After the 45 iterations the student has the given profile tables:

S\A	1	2	3
0	0	2.16	0.45
1	0	2.78	0.52
2	0	3.60	0.67
3	0	5.66	1.12
4	0	5.06	0

Table 1 - Student A Difficulty Q-learning matrix

S\A	1	2	3	4	5	6	7	8	9	10
1	-1.78	-3.76	1.1	-1.69	1.1	1.1	1.1	1.1	1.1	1.1
2	-0.16	-1.69	-3.76	1.1	1.1	1.1	1.1	1.1	1.1	1.1
3	1.1	1.1	-3.76	0.191	1.1	1.1	1.1	1.1	1.1	-1.69
4	1.1	-3.76	-1.69	1.1	1.1	1.1	1.1	1.1	1.1	-1.69
5	1.1	1.1	1.1	1.1	-1.69	1.1	-1.69	1.1	1.1	-1.69
6	1.1	1.1	1.1	1.1	1.1	-1.69	-1.69	1.1	1.1	1.1
7	1.1	1.1	1.1	1.1	1.1	-1.69	0.19	-3.76	1.1	1.1
8	1.1	1.1	1.1	1.1	1.1	1.1	-1.69	1.1	-1.69	1.1
9	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	-1.78	1.1
10	1.1	1.1	1.1	-1.69	0.191	1.1	1.1	1.1	1.1	-1.69

Table 2- Student A Numbers Q-learning matrix

On table one we can see there is no reward for decreasing difficulty as it never was needed, on the other hand the reward for increasing is not really high because once it happened there was no need to redo it so there is not much information for that change of states, however we can see that the higher the difficulty the bigger the Q value for staying in that state, this will help with further runs of the algorithm as it is not short sighted and will tend to increase the difficulty for this student.

On the other hand, the table for the numbers, ignoring the 1.1 has they aren't real values but product of the simplification explained in chapter 4, we can see there isn't a really harder number than other for the student as most of them are negative values meaning they are easier than the one he is answering at the moment. The only positive values can be ignored as they are really close to 0, meaning there is a slight ignorable increase in difficulty.

The other case inside this category represents the students that never gave a wrong answer.

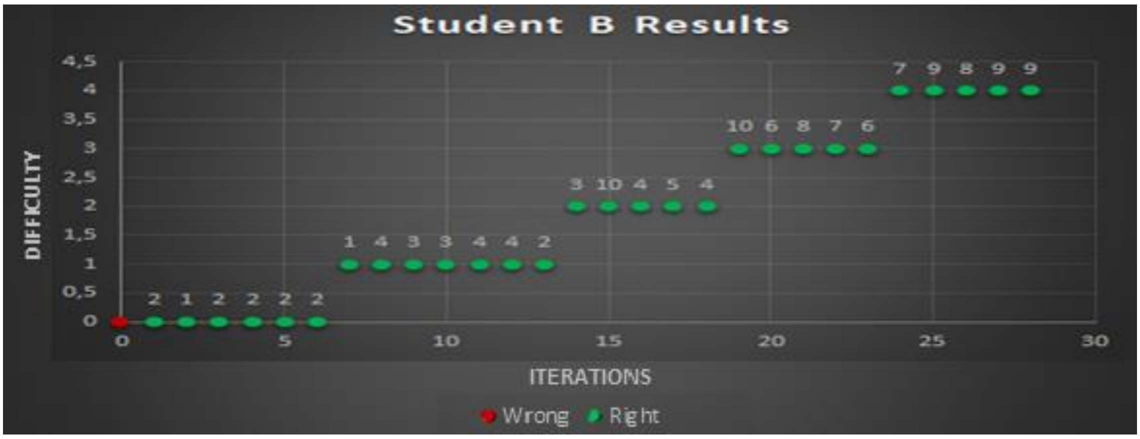


Figure 15 - Student B Results

S\A	1	2	3
0	0	2.16	0.45
1	0	2.52	0.52
2	0	2.35	0.67
3	0	3.92	1.12
4	0	3.92	0

Table 3 - Student B Q-learning Difficulty matrix

S\A	1	2	3	4	5	6	7	8	9	10
1	-1.69	-3.76	1.1	-1.69	1.1	1.1	1.1	1.1	1.1	1.1
2	-3.76	-5.77	-1.69	1.1	1.1	1.1	1.1	1.1	1.1	1.1
3	1.1	1.1	-1.69	-1.69	1.1	1.1	1.1	1.1	1.1	-1.69
4	1.1	-1.69	-1.69	-1.69	-1.69	1.1	1.1	1.1	1.1	-1.69
5	1.1	1.1	1.1	-1.69	1.1	1.1	1.1	1.1	1.1	1.1
6	1.1	1.1	1.1	1.1	1.1	1.1	-1.69	-1.69	1.1	1.1
7	1.1	1.1	1.1	1.1	1.1	-1.69	1.1	1.1	-1.69	1.1
8	1.1	1.1	1.1	1.1	1.1	1.1	-1.69	1.1	-1.69	1.1
9	1.1	1.1	1.1	1.1	1.1	1.1	1.1	-1.69	1.1	1.1
10	1.1	1.1	1.1	-1.69	1.1	-1.69	1.1	1.1	1.1	1.1

Table 4- Student B Q-learning for Numbers matrix

As expected the results of student B aren't very different from student A, as they both have a good performance, the biggest difference is the number of iterations done to end the game.

5.2. Type-2 results

This population is the students on the polar opposite, the students that have very high struggles with fractions and don't have enough knowledge to increase the difficulty. It only represents 2 out of the 79.

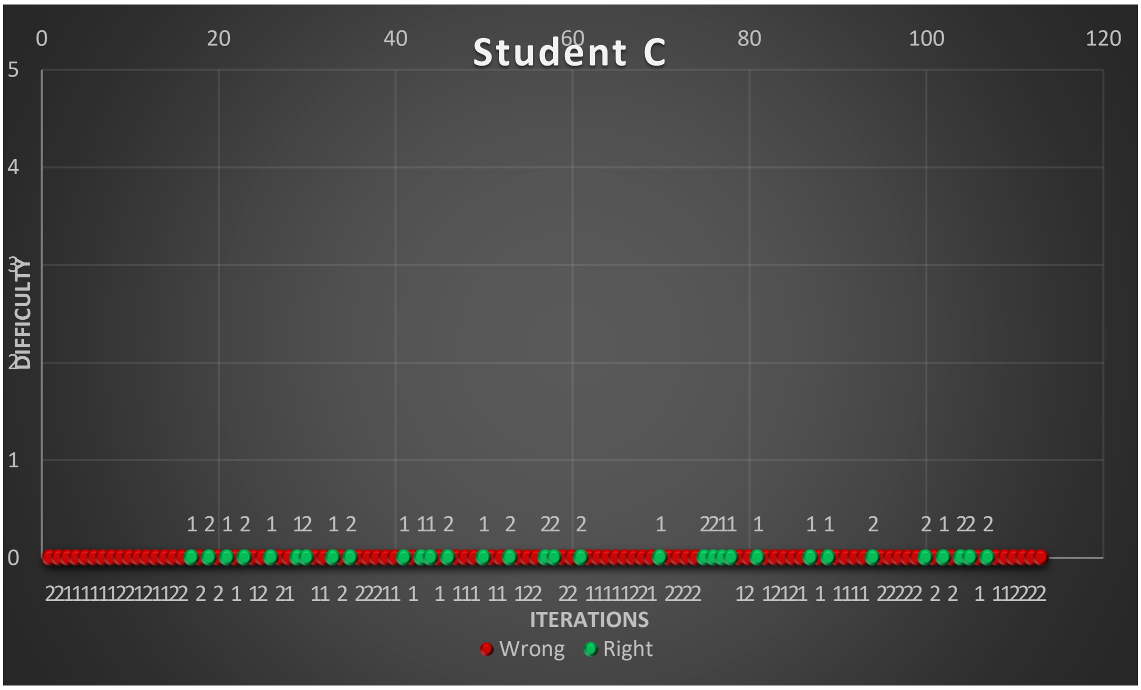


Figure 16- Student C Results

This student shows high difficulties in the first level making the progress impossible, this is a case where an AI is not enough to keep motivation high and progress in the difficulty.

As expected the matrices are fairly simple and only 2 numbers were tested.

S\A	1	2	3
1	10	10	1.1
2	10	10	1.1
3	1.1	1.1	1.1
4	1.1	1.1	1.1
5	1.1	1.1	1.1

Table 5- Extract from Student C Q-learning for Numbers matrix

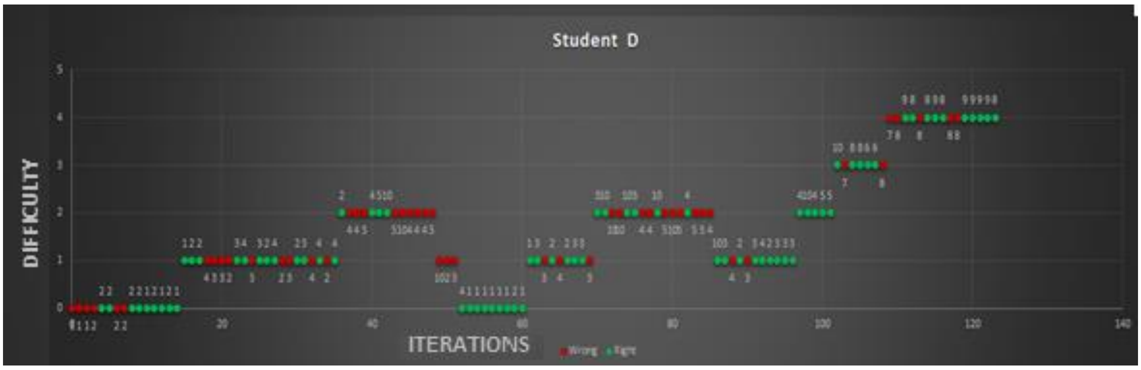
S\A	1	2	3
0	2.202311	2.001706	0
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0

Table 6 - Student C Q-learning Difficulty matrix

The matrix from the Q-learning for Numbers only has data from the relations between one and two and they are all filled as expected with the highest number possible 10.

The matrix from the Q-learning Difficulty has values in decreasing the state 0 to a lower state and maintaining the state this is because we do accept the Q-learning trying to lower the difficulty in the lower state but the outcome is keeping the same difficulty.

5.3. Type-3 results



This population is the students that have big struggles with the fractions and have a really hard time increasing the difficulty of the game and have many decreases of difficulty. In these cases, the AI tends to test the student more and alter the difficulty level more often. These are the students that may or may not be able to overcome the difficulties and end the game. It represents 12 out of the 79.

S\A	1	2	3
0	0.45	3.94	2.73
1	2.76	4.40	3.79
2	3.68	4.87	0.67
3	0	4.67	0.67
4	0	8.31	0

Table 7- Student D Q-learning

S\A	1	2	3	4	5	6	7	8	9	10
1	-7.77	-5.77	-1.69	1.1	1.1	1.1	1.1	1.1	1.1	1.1
2	-7.77	-5.77	-1.81	1.78	1.1	1.1	1.1	1.1	1.1	1.1
3	1.1	-3.77	-5.77	-3.78	1.1	1.1	1.1	1.1	1.1	-1.69
4	-1.69	-3.77	3.79	5.77	2.17	1.1	1.1	1.1	1.1	-5.77
5	1.1	1.1	1.1	-0.18	0.19	1.1	1.1	1.1	1.1	2.17
6	1.1	1.1	1.1	1.1	1.1	-1.69	1.1	1.91	1.1	1.1
8	1.1	1.1	1.1	1.1	1.1	-1.69	1.91	1.781	-5.77	1.1
9	1.1	1.1	1.1	1.1	1.1	1.1	1.1	-3.76	-5.77	1.1
10	1.1	1.91	-1.69	0.191	3.78	1.1	1.91	1.1	1.1	2.17

Figure 17 - Student D Results

Table 8- Student D Q-learning for Numbers matrix

As we can see the tables of this student are more dynamic as more relationships between states were tested. We can analyse that the numbers 4 and 5 are numbers the student struggles with as most of the transitions to 4 and 5 create a positive value. Also, we can see that this kind of student needs more iterations to be able to finish the game.

This student does assess himself with difficulties with the numbers 4 and 5 as it can be seen on his self-assessment, so we can hypothesise that the AI in fact was able to create a profile table with success and it wasn't just a case of random answers.

Number	Knowledge
1	High
2	Medium
3	Medium
4	Low
5	Low
6	Low
7	Low
8	Low
9	Low
10	Medium

Table 9- Student D Self-Assessment

In the case of student D, he can end the game as in his session he had time to do the necessary number of iterations and show the AI sufficient knowledge to pass. On the other hand, the example of student E he shows a really focused problem in the area of numbers 6, 7 and 8. As we can see in the graph below. This subject only had time to answer to 70 game iterations but we can easily assess what are the main problems.

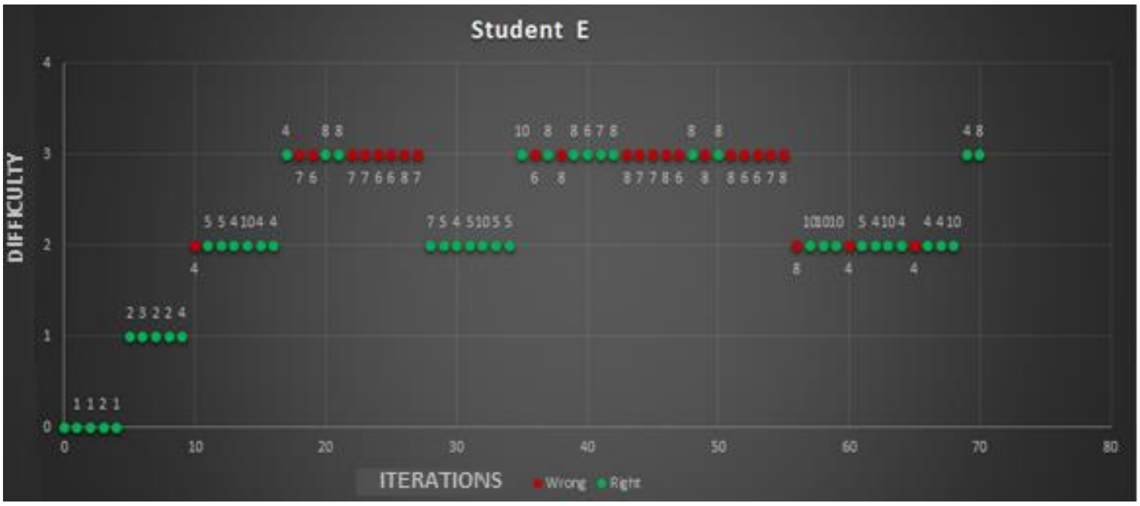


Figure 18 - Student E Results

Student E shows specific problems which translate in an easy to read matrix. Table 9 shows how in level 2 of difficulty (because it has changed more than once contrary to cases given until now) the outcome expected is to keep or increase the difficulty with main focus on increasing and the next level has the decrease and keep of difficulty with positive values. This student in particular can be qualified as a student of high level 2 to level 3 knowledge.

The table 11 give us further information on the needs of this student and as figure 18 shows, the information it gives us is that the student as some difficulties on the 6,7,8 as these are the numbers with positive.

We can further see that the student self-assessment shows that he knows he has difficulties on the number 7 and 8, however the self-assessment on the number 6 is that the knowledge is high but the game logs show us otherwise.

S/A	1	2	3
0	0	1.84	0.45
1	0	1.83	0.07
2	0	6.16	7.79
3	5.16	8.57	0
4	0	0	0

Table 10 - Student E Q-learning Difficulty matrix

S/A	1	2	3	4	5	6	7	8	9	10
1	-5.77	-3.76	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
2	-1.69	-1.69	-1.69	-1.69	1.1	1.1	1.1	1.1	1.1	1.1
3	1.1	-1.69	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
4	1.1	1.1	1.1	-3.78	-5.77	1.1	1.91	-1.69	1.1	-5.77
5	1.1	1.1	1.1	-5.77	-3.76	1.1	1.1	1.1	1.1	-3.76
6	1.1	1.1	1.1	1.1	1.1	3.79	-0.16	-3.78	1.1	1.1
7	1.1	1.1	1.1	1.1	-1.69	3.79	1.81	1.78	1.1	1.1
8	1.1	1.1	1.1	1.1	1.1	1.78	5.77	3.78	1.1	-1.69
9	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
10	1.1	1.1	1.1	-3.78	-1.69	1.91	1.1	1.1	1.1	-3.76

Table 11- Student E Q-learning Numbers matrix

Number	Knowledge
1	High
2	High
3	High
4	High
5	Medium
6	High
7	Low
8	Low
9	Medium
10	High

Table 12 - Student E Self-Assessment

5.4. Type-4 results

This group contains the students that have some struggles with the fractions and have a harder time increasing the difficulty of the game but usually don't have decreases in the difficulty. In these cases, the AI may test the student more times but will likely not decrease the difficulty but tends to test the student more. It represents 28 out of the 79.

Student F is a student that on level 2 got stuck awhile as he gives wrong and right answers intercalated. However, around iteration 50 we can see an increase in right answers on the numbers previously with wrong answers. At this time the difficulty increases. Once again, he has a wrong answer on the first 6 that he then answers two times correctly.

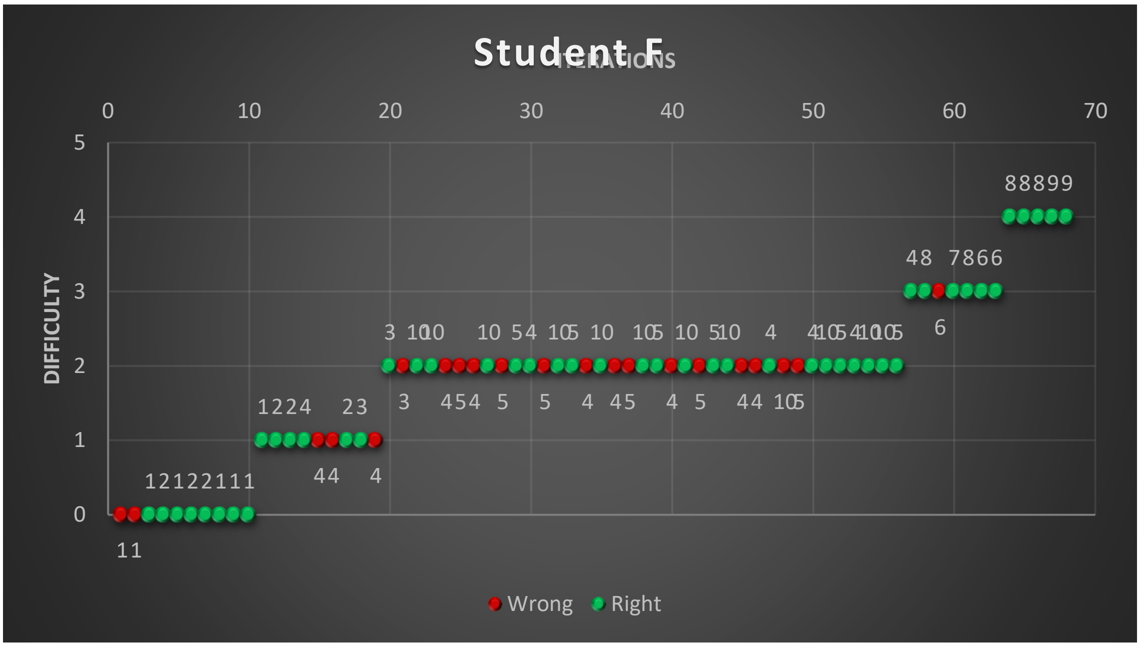


Figure 19 - Student F Results

S/A	1	2	3
0	0	2.64	0.45
1	0	1.99	0.525
2	0	5.69	0.675
3	0	5.09	1.125
4	0	3.92	0

Table 13- Student F Q-learning Difficulty matrix

S/A	1	2	3	4	5	6	7	8	9	10
1	-1.77	-5.77	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
2	-3.76	-3.76	-1.69	-1.69	1.1	1.1	1.1	1.1	1.1	1.1
3	1.1	1.1	1.91	1.91	1.1	1.1	1.1	1.1	1.1	-1.69
4	1.1	-1.69	-1.69	7.77	5.77	1.1	1.1	-1.69	1.1	-7.78
5	1.1	1.1	1.1	-1.77	-3.76	1.1	1.1	1.1	1.1	-5.77
6	1.1	1.1	1.1	1.1	1.1	-1.69	-1.69	-1.69	1.1	1.1
7	1.1	1.1	1.1	1.1	1.1	1.1	1.1	-1.69	1.1	1.1
8	1.1	1.1	1.1	1.1	1.1	0.19	1.1	-3.76	-1.69	1.1
9	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
10	1.1	1.1	1.1	5.77	-1.78	1.1	1.1	1.1	1.1	-3.76

Table 14 - Student F Q-learning for Numbers matrix

This is a case where we can see the power of the AI with the adaptability it can have. The difference between student E and F aren't that further apart as both show a specific problem, but because E shows a steadier streak of wrong answers the difficulty is lowered while F, because his are more intercalated, the AI tries to test the student without changing the difficulty.

5.5. Type-5 results

Student H shows results that are considered random and strange as we can see on figure 20. The answers given by this student seem, in fact, to be the result of random iterations. It only represents 2 out of the 79 participants.

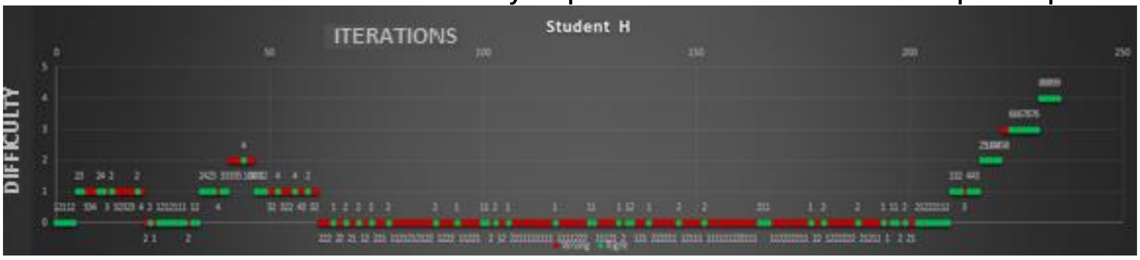


Figure 20 - Student H Results

Figure 20 starts by being quite normal for any student however around iteration 60 there is a tendency for all answer to be wrong which in a normal case we would accept as being a student with difficulties. However, on iteration 200 the student starts by answering correctly to the questions ending the game with most of them correct in only 50 iterations. For this we can acknowledge two different happenings: The student got bored and started answering randomly or at the iteration 200 the player swap with another one.

Although this result is not the best we can see in this gameplay the adaptation of AI to different circumstances.

The other odd case appearing is student I, who shows high difficulties on the early iterations but suddenly overcome, and because as this project was a test to how the AI can create a player profile, no help was created in the game meaning there is no reason for such thing to happen.

However, the difference between student H and I is that student I shows problems in the early begging which can mean a lack of understanding of the game mechanics and not boredom like student H, hence making a good study case if this project should be continued.

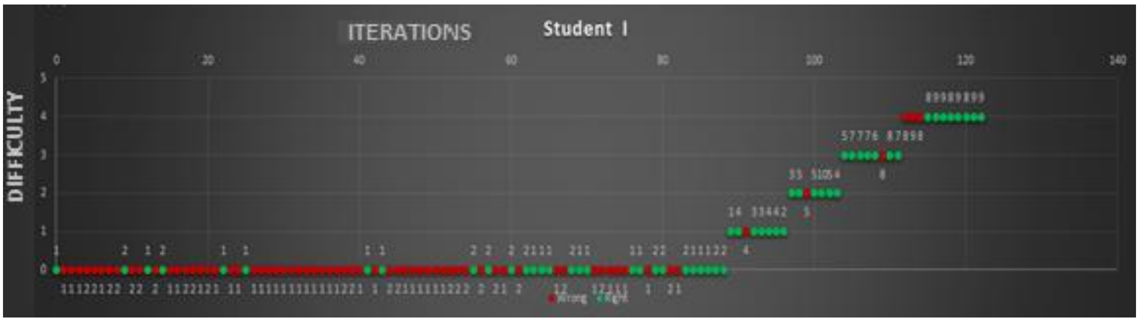


Figure 21 - Student I Results

The test results are considered positive as we can see how the game can adapt itself based on the answers given by any player, it creates a solid player profile close to what a teacher would do based on the same answers. The difference is that the same program was able to create different study environments hence consuming less resources and time. This tool may be able to help a teacher further understand the special needs that each student has without the time consuming individual quizzes.

6. Conclusions

6.1. Summary of Research and Findings

According to the literature, recent research points to the notion that motivation is crucial factor when learning a new subject. One of the key factors for motivation is the notion of flow state which is the creation of a state where the student feels his knowledge is enough to overcome the challenges presented to him while keeping them exciting. (Kiili, Freitas, Arnab and Lainema, T. 2012).

The related research is broad and covers a large number of topics, but by the granularity of the needs of each student there are small and specific approaches to create motivation in any given type of student. The focus of this research is on creating an Artificial Intelligence that was able to handle this granularity and adapt a single game to any given person by creating each player profile thus being able to adapt the challenges presented to the knowledge the player has.

The Q-learning algorithm was used as an approach to create this adaptation but because of the requirements present with this project there was the need to adapt the Q-learning to be more dynamic in our environment, by introducing exploration values and negative numbers and the notion of range inside one matrix we created a more adaptable algorithm with almost the same complexity. This proved us that is possible to create a highly adaptable AI to a player profile with roughly the same resources, and once the rules are made to the Q-learnings the game can be made by anyone that only has to use the results given by them to create the new questions.

While on the creation of this work we acknowledge other contribution that creating an AI in a serious game can make. The teachers of the test population when questioned by the specific difficulties of each student weren't able to give us an answer and gave us more of a qualitative notion of each student, we then knew that with this we could actually give the teachers information that they don't have in the classrooms as they don't keep a log of each answer of each student and only know more or less what their difficulties are.

Also on the test bed there were students that recently changed grade and by using our program the teachers were able to have an insight of what the students, that are going to start, know or not and be able to create more adaptable study plans.

Even though the project was a success and the input we had was positive based on what we expected we know that further work can and should be made for a creation of a true serious game.

We want to use the words from Porayska-Pomsta (2003) as reflection of the work: "The model often combines the theoretical contributions by other researchers in a somewhat simplified and even naive manner which leaves it open to a considerable amount of criticism." (p.299).

However, this does not mean that the work created is to be overlooked as many conclusions and data can be taken in account as they have been proved to work with our test subjects.

By relating the goals of this dissertation with the results of the study and evaluation, the main contribution of this research has been:

- Creating an adaptative game difficulty management
- Adaptation of Q-learning algorithm to be in two layers
- Adaptation of Q-learning algorithm for more dynamic information in one matrix.

It is true that our AI for motivation contains some limitations. Despite its limitations, we considered this implementation throws light on the process of how to create a more adaptable serious game that keeps the students motivated to learn, which we believe is a promising step into creating a learning environment that cares about learners.

6.2. Future work

The work in this dissertation has been tested on the target audience with success. However, there still is much more study to be done to draw more solid conclusions as only 79 students were tested. A bigger sample as well as a post-game questionnaire would be needed to further analyse this work.

Regarding the approach used to create the AI, we consider it to be a positive solution as the approach worked as expected. Furthermore, we tested and analysed with success the use of the two Q-learning algorithms. Also, the fact of being able to have more than one outcome from one Q-matrix and the use of different ranges shows itself useful as it is more dynamic without spending more resources.

Moreover, the method developed will be used in behaviour elicitation tools using serious games such as the works presented in Rossetti, Almeida, Kokkinogenis, Gonçalves (2013) and Rossetti, Oliveira, Bazzan (2007).

Nonetheless, many improvements can and should be considered before starting a project aimed at a finished game. First of all, one type of game does not fit all and this should be taken into consideration. We cannot create one single game enjoyable for all; we can however study what the biggest common factors in the target audience regarding the type of game would be.

Also, there is the need to actually create a game and not a test environment with game components. There should be a storyline, objectives and game mechanics that would create a complete game. We should also point out the necessity of creating help in the game and tools to unstuck the player from a given place as a way of motivation.

Moreover, there are improvements that may be implemented in the AI. We can improve the knowledge of the Numbers Q-learning by using the relativity of the numbers to change the Q-matrix, in other words, if 3 is harder than 2 and 4 is harder than 3 then we can assume that 4 is harder than 2. Furthermore, we can create the difference between difficulties of an easier transition; this mean if we are transitioning from difficulty 4 to 3 we should choose the harder numbers from difficulty 3; and the same applies to increasing the difficulty, i.e. first the easier numbers should be chosen.

Also, we should take into account which is the whole part of the fraction when the student answers as it can give us more information on the troubles the student had. We recommend to further adapt the reward functions to include more than one fixed number, such as time spent or help asked.

Finally, it would be interesting to study the emotions of the players during a game and calculate their arousal and valence to see if a given game creates times of boredom or frustration.

References

Ade-Ibijola, A. O., Adetunmbi, A. O., Alese, B. K. (2013). *A Simple Computer Game for Learning Mathematics*. eleed, Iss. 9.

AguiarRossetti, S. (2016). *A game design to teach National 5 Lifeskills Mathematics to secondary school children: A case study*. Individual project, Department of Teaching and Learning Enhancement, Edinburgh Napier University.

Aguilera, M. and Mendiz, A. (2003). *Video games and education*. Computers in Entertainment, vol. 1, no. 1, article 1.

Barbosa A. F. S., Pereira P. N. M., Dias J. A. F. F., and Silva F. G. M. (2014). *A new methodology of design and development of serious games* International Journal of Computer Games Technology, Article ID 817167, 8 pages.

Çankaya, S. and Karamete A. (2009). *The effects of educational computer games on students' attitudes towards mathematics course and educational computer games,* Procedia—Social and Behavioral Sciences, vol. 1, no. 1, pp. 145–149,

Charles, D., McNeill, M., McAlister, M., Black, M., Moore, A., Stringer, K., Kücklich, J. and Kerr, A. (n.d.). *Player-Centred Game Design: Player Modelling and Adaptive Digital Games*. Computing and Information Engineering and Centre for Media Research, University of Ulster.

Choi, B. Huang, J. Jeffrey, A. and Baek. (2013) *Development of a scale for fantasy state in digital games.* Computers in Human Behavior, vol. 29, no. 5, pp. 1980–1986,

Dobrovsky, A., Borghoff, U. and Hofman, M. (2016). *An Approach to Interactive Deep ReinforcementLearning for Serious Games*. In: 7th IEEE International Conference on Cognitive Infocommunications.

Freitas, S., Earp, J., Ott, M. et al., (2012). *Hot issues in game enhanced learning: the GEL viewpoint.* Procedia Computer Science, vol. 15, pp. 25–31,

Goleman, D. (1996). Emotional intelligence. London [etc.].: Bloomsbury.

Gunter, G. A., Kenny, R. F. and Vick, E. H. (2008) *Taking educational games seriously: using the RETAIN model to design endogenous fantasy into standalone educational games*. Educational Technology Research and Development, vol. 56, no. 5-6, pp. 511–537.

Hado van Hasselt. (2012). *Reinforcement Learning in Continuous State and Action Spaces.* In: Reinforcement Learning: State of the Art, Springer, pages 207-251.

Hocine, Nadia, Gouaich, Abdelkader. (2011) *A Survey of Agent Programming and Adaptive Serious Games.* RR-11013, pp.8.

Huang, W.-H., Huang, W.-Y., and Tschopp J. (2010). *Sustaining iterative game playing processes in DGBL: the relationship between motivational processing and outcome processing.* Computers & Education, vol. 55, no. 2, pp. 789–797.

Ibrahim, R. and Jaafar, A. (2009). *Educational games (EG) Design Framework: Combination of Game Design, Pedagogy and Content Modelling.*

J. F. Silva, J. E. Almeida, R. J. F. Rossetti and A. L. Coelho,(2013) *A serious game for EVAcuation training,* in 2013 IEEE 2nd International Conference on Serious Games and Applications for Health (SeGAH), Vilamoura , pp. 1-6.

J. F. Silva, J. E. Almeida, R. J. F. Rossetti and A. L. Coelho, (2013) *Gamifying evacuation drills,* 2013 8th Iberian Conference on Information Systems and Technologies (CISTI), Lisboa, pp. 1-6.

J. Ribeiro, J. E. Almeida, R. J. F. Rossetti, A. Coelho, A. L. Coelho, (2012) *Towards a Serious Games Evacuation Simulator* in 26th European Conference on Modelling and Simulation, ECMS, Koblenz, Germany, May 29 - June 1, 2012, pp. 697-702.

Kinzie, M., & Joseph, D. (2008). Gender Differences in Game Activity Preferences of Middle School Children: Implications for Educational Game Design. *Educational Technology Research and Development, 56*(5/6), 643-663.

Keller, John M. (1987). *Development and use of the ARCS model of instructional design.* Journal of instructional development 10.

Kiili, K., Freitas, S., Arnab, S. and Lainema, T. (2012). *The Design Principles for Flow Experience in Educational Games.* Procedia Computer Science,15. pp.78-91.

Marczewski, A. (2013). *Gamification: A Simple Introduction. Tips advice and thoughts on gamfication.*

Methaneethorn, J. (2008). *Towards Motivation Modelling within A Computer Game Based Learning Environment: An Empirical Study.* Ph.D. University of Glasgow.

Mouaheb, H., Fahli, A., Moussetad, M. and Eljamali, S. (2012). *The serious game: what educational benefits?.* Social and Behavioral Sciences 46. pp.5502 – 5508.

Pavlas, D. (2010). *A Model of Flow and Play in Game-based Learning: The Impact of Game Characteristics, Player Traits, and Player States.* Ph.D. University of Central Florida.

Poole, D., Mackworth, A. (2010). *Artificial Intelligence: Foundations of Computational Agents.* Cambridge University Press.

R. J. F. Rossetti, J. E. Almeida, Z. Kokkinogenis and J. Gonçalves, (July-Aug. 2013) *Playing Transportation Seriously: Applications of Serious Games to Artificial Transportation Systems*, in IEEE Intelligent Systems, vol. 28, no. 4, pp. 107-112.

R. J. F. Rossetti, E. C. Oliveira, A. L. C. , (2007) *Bazzan, Towards a specification of a framework for sustainable transportation analysis,* in 13th Portuguese Conference on Artificial Intelligence, EPIA, Guimarães, Portugal, pp. 179-190

Russel, Stuart J. l; Norvig, Peter (2010). *Artificial Intelligence: A Modern Approach* (Third ed.). Prentice Hall.

S. Eccles, J. and Wigfield, A. (2002). *Motivational Beliefs, Values, and Goals. Michigan: Annual Reviews.*

Song, M. and S. Zhang, (2008). *EFM: A model for educational game design. In: Lecture Notes in Computer Science*. Edited by S. Link. Springer US. 509 – 517.

Squire, K., (2003). *Video games in education*. International Journal of Intelligent Games & Simulation, vol. 2, no. 1, pp. 49–62.

Sutton, Richard S.; Barto, Andrew G. (1998). Reinforcement Learning: An Introduction. MIT Press.

Tüzün, H., Yilmaz-Soylu, M., Karakuş, T., İnal, Y. and Kızılkaya, G., (2009). *The effects of computer games on primary school students' achievement and motivation in geography learning.* Computers & Education, vol. 52, no. 1, pp. 68–77,

Weiner, B (1992). *Human Motivation, Metaphots, Theories, and Research.* Sage Publications, Inc.

Yannakakis, G. N. (n.d.). *Game AI Revisited.* Center for Computer Games Research. IT University of Copenhagen